#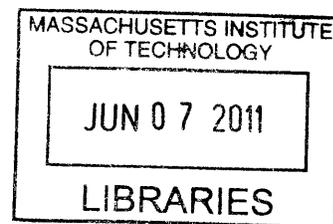 Discovery, Characterization, and Rational Design of the Enzymes Involved in Monoterpene Indole Alkaloid Biosynthesis in Madagascar Periwinkle

by

Lesley-Ann Giddings

B. A. Chemistry
Smith College, 2005

SUBMITTED TO THE DEPARTMENT OF CHEMISTRY IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT THE
MASSACHUSSETTS INSTITUTE OF TECHNOLOGY

ARCHIVES

JUNE 2011

Signature of Author:___

Department of Chemistry
May 11, 2011

Certified by:_

Sarah E. O'Connor
Associate Professor of Chemistry
Thesis Supervisor

Accepted by:____

Robert W. Field
Professor of Chemistry
Chairman, Departmental Committee on Graduate Students

This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:

Chair: _____

_____
John M. Essigmann
Professor of Chemistry and Biological Engineering

Thesis Supervisor:___

_____
Sarah E. O'Connor
Associate Professor of Chemistry

Committee Member:_

_____
Alexander M. Klibanov
Professor of Chemistry and Biological Engineering

DISCOVERY, CHARACTERIZATION, AND RATIONAL DESIGN OF THE ENZYMES INVOLVED IN MONOTERPENE INDOLE ALKALOID BIOSYNTHESIS IN MADAGASCAR PERIWINKLE

by

Lesley-Ann Giddings

ABSTRACT

The chemical diversity found in plants has served as a major source of inspiration to many synthetic and biological chemists. Nature has evolved enzyme active sites to catalyze the synthesis of structurally complex compounds that serve as pharmaceuticals, insecticides, dyes, perfumes, and biofuels. In *Catharanthus roseus*, approximately 130 structurally complex monoterpene indole alkaloids are produced, including the clinically used anti-mitotic drugs, vinblastine and vincristine. The common intermediate to all monoterpene indole alkaloids is strictosidine, the product of an asymmetric Pictet-Spengler condensation of tryptamine and the iridoid monoterpene secologanin. This reaction is catalyzed by the enzyme strictosidine synthase. This thesis describes the use of kinetic isotope effects, the rate dependence on pH, as well as structural and computational data to propose a mechanism by which strictosidine synthase catalyzes the Pictet-Spengler reaction. Notably, the data also shed light on the mechanism of the widely used nonenzymatic reaction. Interestingly, the Pictet-Spenglerase strictosidine synthase belongs to a superfamily of enzymes that have mainly been observed to catalyze ester hydrolysis. Using the β-propeller fold conserved in both strictosidine synthase and the related hydrolase, paraoxonase, rational mutagenesis was used to convert strictosidine synthase into a hydrolase. Intriguingly, during the rational design process, the function of a closely related strictosidine synthase homolog was also functionally characterized as a hydrolase. In addition to reengineering proteins with new catalytic activity, the chemical diversity in plants can also modified using metabolic engineering. However, this approach requires knowledge of the genetic blueprints of the plant to be known. Using the recently released *C. roseus* transcriptome sequencing data along with co-expression analysis, this thesis describes the functional characterization of a new P450 gene involved in metabolizing a key intermediate in the biosynthesis of bioactive bisindole alkaloids. With the functional characterization of this new gene, a combination of gene silencing and synthetic biology techniques will provide a greater understanding on how to "tune" alkaloid biosynthesis in *C. roseus* in order to generate more functionally diverse molecules.

Thesis Supervisor: Sarah E. O'Connor
Title: Associate Professor of Chemistry

Acknowledgements

There are a number of people who have provided me with guidance throughout my life that I would like to acknowledge. I would first like to thank God for surrounding me with positive people and providing me with opportunities to help me reach this place in life. One person who has had a very positive impact on my life is my advisor Sarah O'Connor. Sarah believed in me when I told her I wanted to quit during my first year of graduate school. I had never taken a single biochemistry course before coming to MIT, but she gave me her Voet & Voet *Biochemistry* textbook and said that she would support me if I wanted to stay at MIT. Your brilliance, generosity, mentorship, and patience are what make you an outstanding teacher. You have inspired many and I am truly grateful to have had you as an advisor.

I would like to thank Professors Alexander Klibanov and John Essigmann, my committee members for providing me with constructive feedback on my research. Additionally, I would like to thank my former committee member Professor Daniel Kemp whose positive words helped instill confidence in me, which is needed in a place like MIT.

I am also grateful to the O'Connor lab members for my scientific training, but most importantly for allowing me to be myself. The parties, tennis and volleyball, the lab trips including those to the muddy, all helped me to feel as though I was a part of a family. I would like to especially thank Justin Maresh for teaching me enzyme kinetics, Peter Bernhardt for his many insightful discussions and ideas, one of which is the basis for Chapter 3, Weerawat "Ricky" Runguphan for teaching me how to clone and being a good source of lab entertainment, Weslee Glenn and Aimee Usera for always listening to my random thoughts and being good friends, David Liscombe for teaching me about gene discovery in plants as well as being a good friend, and Jenna Caldwell for being such a hard worker and a great student to work with. I would also like to thank my collaborators, Summer Thyme from Professor David Baker's lab at the University of Washington and Michael Hicks from Professor Patsy Babbitt's lab at UCSF, for their homology modeling and discussions on protein evolution.

I became introduced to the world of natural products thanks to the late John Daly, formerly of the Laboratory of Bioorganic Chemistry at the National Institutes of Health. I am eternally grateful for the support and guidance he gave me, which helped create many opportunities throughout my scientific career thus far. I am also grateful to Cristina Suarez, my undergraduate advisor at Smith College, for giving me research opportunities and guidance throughout my undergraduate studies.

Importantly, I would like to thank my boyfriend Joel for his friendship, love, and support. Thank you for teaching me not worry so much about life and for showing me how to make the most of it. I would also like to thank wonderful family for always making me feel loved even when separated by distance. Lastly, I thank my mother, my number 1 fan, who has always put me first in her life. I dedicate this thesis to you.

Lesley-Ann Giddings

Table of Contents

List of Figures

Chapter 5

List of Tables

Abbreviations

Standard one- and three-letter codes are used for the naturally occurring amino acids. Standard one-letter codes are used for nucleotides.

ABT:	Aminobenzotriazole
AU:	Absorbance units
B:	B-factor
Bis-tris:	Bis-(2-hydroxyethyl)amino-tris(hydroxymethyl)methane
bp:	Base pairs
BLAST:	Basic Local Sequence Alignment Search Tool
Br:	Bromo
C:	Carbon
Ca:	Calcium
$CaCl_2$:	Calcium chloride
CD:	Circular dichroism
$CDCl_3$	Deuterated chloroform
cDNA:	Complementary deoxyribonucleic acid
Cl:	Chloro
Cr:	*Catharanthus roseus*
CSC:	Cell suspension cultures
CYP:	Cytochrome P450
DCM:	Dichloromethane
DMSO:	Dimethylsulfoxide
DNA:	Deoxyribonucleic acid
EDTA:	Ethylenediaminetetraacetic acid
ESI MS:	Electrospray ionization mass spectrometry
EST:	Expressed sequence tag
F:	Structure factor
Fig.:	Figure
FT-ICR-MS:	Fourier transform ion cyclotron resonance mass spectrometry
Glc:	Glucose
H:	Hydrogen
HCl:	Hydrochloric acid
HEPES:	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC:	High performance liquid chromatography
HSQC:	Heteronuclear single quantum correlation
Hz:	Hertz
I:	Intensity of the reflection
IPTG:	Isopropyl-β-D-thiogalactopyranoside
KIE:	Kinetic isotope effect
K_m:	Michaelis-Menten constant
LB:	Luria-Bertani

LC-MS:	Liquid chromatography mass spectrometry
Me:	Methyl
MeO:	Methoxy
MS:	Mass spectrometry
m/z:	Mass-to-charge ratio
NAA:	1-Napthaleneacetic acid
NaCl:	Sodium chloride
NADPH:	Nicotinamide adenine dinucleotide phosphate
NaOH:	Sodium hydroxide
NMR:	Nuclear magnetic resonance
PCR:	Polymerase chain reaction
pNPAc:	p-Nitrophenyl acetate
R_{cryst}:	Residual function calculated for 95% of data
Re:	Recombinant
R_{free}:	Residual function calculated for randomly selected data
R_{merge}:	A measure of agreement of reflections within a data set
RMSD:	Root mean square deviation
RNA:	Ribonucleic acid
RNAi:	Ribonucleic acid interference
Rs:	*Rauvolfia serpentina*
SDS-PAGE:	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TBBL:	5-Thiobutyl butyrolactone
TOF:	Time of flight
Tris:	Tris(hydroxymethyl)aminomethane
UPLC:	Ultra performance liquid chromatography
UCSF:	University of California, San Francisco
UV:	Ultraviolet
V_{max}:	Maximum velocity when substrate concentration approaches infinity

CHAPTER 1

BACKGROUND AND SIGNIFICANCE

1.1 Natural products

Natural products are organic molecules produced by living organisms found throughout nature. These small molecules, which are also referred to as secondary metabolites, are not typically considered to be essential for the growth, development, and reproduction of the organism, nor are they produced under all environmental conditions or physiological stages [1]. In many organisms, natural products are released as chemical responses to protect the host or to allow the host to communicate with the environment. The structural complexity that Nature has evolved in secondary metabolism has been of great interest to chemists and biologists for decades. Notably, between 1981 and 2006, of all small molecule new chemical entities, 34 % were classified as natural products or semisynthetic derivatives [2, 3] (Fig. 1.1). Because natural products have evolved to exhibit biological activity, these compounds have successfully served as templates for the development of pharmaceuticals, dyes, flavors, repellants, material protectants, perfumes, herbicides, and bioenergy sources [4-6]. Furthermore, other applications for natural products can be determined by studying their interactions with the proteins they target in their native host [7]. While a complete understanding of the mechanisms of action of these compounds is ideal, natural products continue to serve as a rich source of novel structures with unique bioactivities.

topotecan
anticancer
source: semisynthesis from camptothecin
from *Camptotheca acuminata*
plant

artemesinin
antiparasitic
source: *Artemesia annua*
plant

resveratrol
anticancer, antinflammatory,
and hypotensive agentagent
source: *Vitis vinifera* and many other plants
plant

digoxin
antiarrythmic
source: *Digitalis lanata*
plant

vancomycin
antibiotic
source: *Amycolatopsis orientalis*
soil bacteria

epoprostenol
inhibits platelet aggregation
sources: *mammalian tissues*
and *Plexaura homomalla* (coral)
made synthetically

mupiricin
antibiotic
source: *Pseudomonas fluorescens*
bacteria

Figure 1.1 Examples of natural products and derivatives that are clinically approved drugs.

14

The wide array of stereocenters, rings, and oxygenation patterns found in natural products serves as a major source of inspiration for organic chemists. Nature has provided a template for chemists to discover new reactions to generate these interesting synthetic targets [8]. Synthetic efforts are critical, as the isolation of a lead compound from a complex mixture of natural products can be challenging and often fails to provide enough material for full structural and biological characterization. Many synthetic research groups are interested in developing new methodologies and catalysts for regio- and stereo-selective syntheses and semi-syntheses of natural products and their derivatives [9, 10]. Additionally, chemists have constructed libraries of compounds using "diversity oriented synthesis", a process whereby efficient synthetic routes to combinatorial libraries are based on structural motifs found in natural products [11]. High-throughput screens have been developed to identify bioactive small molecules produced in these libraries [12, 13]. However, these synthetic approaches have their drawbacks. Synthesis is labor- and time-intensive as well as expensive. While the use of combinatorial chemistry may yield libraries of new pharmacophores for drug development, natural products are so skeletally diverse that new synthetic strategies need to be developed to provide easy access to many natural product analogs. Most importantly, identifying and subsequently deconvoluting the biological target of the newly synthesized small molecules remains to be the bottleneck in high-throughput methods aimed at creating new molecules with enhanced bioactivity [12].

Although synthetic methods have improved significantly in terms of providing access to natural and unnatural products, the structural complexity of all natural products cannot be achieved via synthesis. As the genetic blueprints for the biosynthesis of many natural products become available, metabolic engineering has proven to be a powerful tool for the generation of natural

products and derivatives, especially in plants [14]. While the biosynthetic machinery required to assemble natural products must be known, studying nature to identify, characterize, and manipulate enzymes from these biosynthetic pathways can serve as a useful tool for increasing chemical diversity and finding unique bioactivities.

Natural products have been isolated from microorganisms, marine life, insects, animals such as venomous snakes, as well as from plants, which is the focus of my thesis. Although less than 15 % of higher plant species having been explored rigorously for bioactivity, plants harbor over 200,000 known chemical entities that are, for the most part, classified as secondary metabolites [14-16]. Notably, plants are still a major source of drugs as evidenced by the 91 plant-derived compounds that have been in clinical trials since 2007 [15, 16]. Plant natural products can be divided into four major groups based on biosynthetic origins: (1) terpenoids, (2) glucosinolates and cyanoglucosides, (3) alkaloids, and (4) phenylpropanoids and other phenolic compounds. The building blocks of these natural products are acetyl/malonyl coenzyme A, shikimic acid, mevalonic acid, 1-deoxyxylulose 5-phosphate, and a variety of proteogenic and non-proteogenic amino acids (Fig. 1.2). These precursors are derived from primary metabolic pathways such as glycolysis, the pentose phosphate cycle, and photosynthesis. Unlike the clustered, and therefore more tractable, biosynthetic genes that direct the synthesis of natural products in microorganisms, genes involved in the synthesis of plant natural products are typically scattered throughout the genome, complicating the gene discovery process. However, the recent advances in molecular biology, classic biochemical techniques, sequencing of genomes and transcriptomes, and enzyme engineering have all provided major insight on natural product biosynthesis in plants.

acetyl coenzyme A

mevalonic acid

shikimic acid   deoxyxylulose-5-phosphate

malonyl coenzyme A
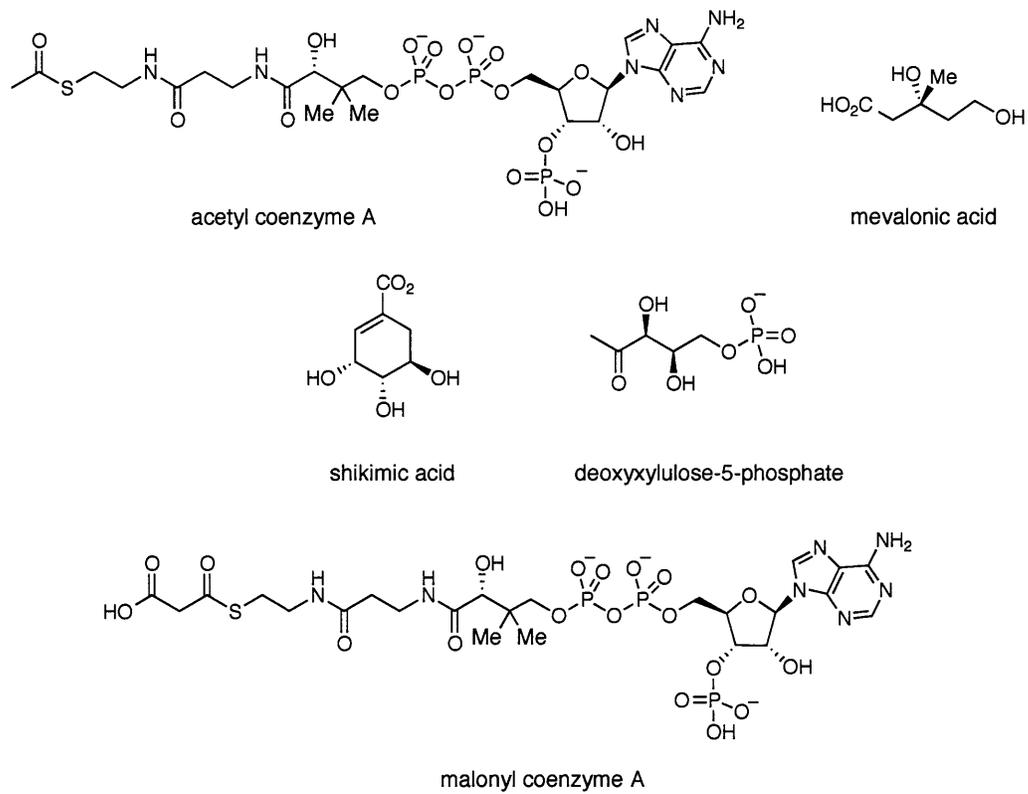
Figure. 1.2 Examples of some of the building blocks of natural products.

1.2 Monoterpene indole alkaloid biosynthesis

The monoterpene indole alkaloids are derived from tryptophan (shikimate pathway), isopentenyl diphosphate (IPP, 1-deoxyxylulose 5-phosphate pathway), and dimethylallyl diphosphate (DMAPP, 1-deoxyxylulose 5-phosphate pathway). These alkaloids have been identified in eight different plant families, but are most commonly found in Apocyanceae, Loganiaceae, and Rubiaceae families. In the Apocyanceae family, *Catharanthus roseus*, also referred to as the Madagascar periwinkle, produces approximately 130 monoterpene indole alkaloids. Many of these compounds have important bioactivities, such as the anti-cancer properties of vinblastine and vincristine (Fig. 1.3). In recent years, major progress has been made to identify the cDNAs in *C. roseus* that encode the biosynthetic enzymes responsible for the production of monoterpene indole alkaloids.

ajmalicine
(hypotensive agent)

serpentine
(topoisomerase II inhibitor)

ajmaline
(antiarrhythmic agent)

strychnine
(convulsant, pesticide)

R=Me   vinblastine
R=CHO vincristine
(anticancer agents)
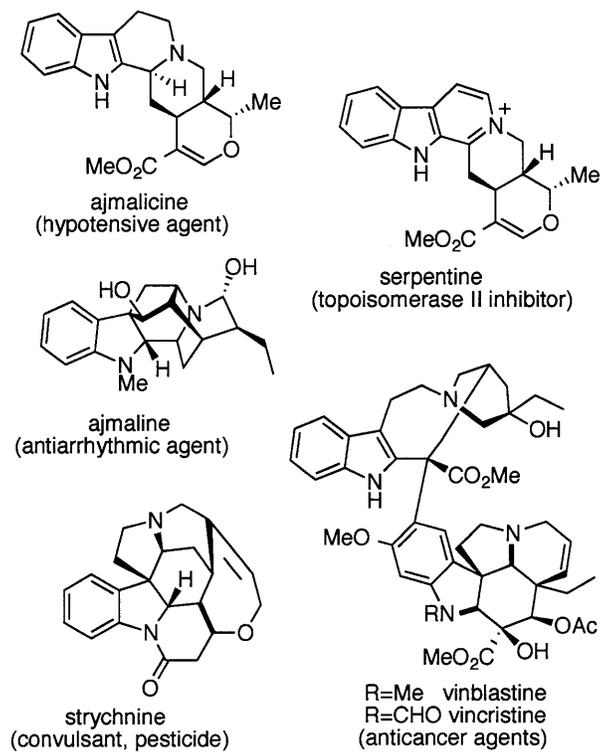
Figure 1.3 Bioactive monoterpene indole alkaloids found in *C. roseus*.

19

The first committed step in monoterpene indole biosynthesis is the condensation of tryptamine and secologanin. Tryptamine is formed by the pyridoxal-dependent decarboxylation of L-tryptophan by tryptophan decarboxylase (TDC) (Fig. 1.4 A) [17]. Recently, Runguphan and coworkers have shown that upon RNAi-induced silencing of tryptophan decarboxylase *in planta*, alkaloid production is eliminated and can be rescued by the addition of tryptamine and analogs [18]. Interestingly, overexpression of TDC did not result in an increase in monoterpene alkaloids [19, 20]. Feeding studies with ^{3}H-labeled monoterpenes suggest that secologanin is derived from geraniol and the downstream biosynthetic intermediates 10-hydroxygeraniol, iridodial, and iridotrial [21, 22]. To form secologanin, geranyl diphosphate synthase (GPPS) condenses one unit of IPP and one unit of DMAPP in a head-to-tail fashion to form geranyl diphosphate (Fig. 1.4 B) [23]. Based on sequence similarity to the known *Arabidopsis thaliana* GPPS, a cDNA has been identified for GPPS in *C. roseus*; however, this cDNA has not been functionally characterized. Geraniol synthase (GS) then converts geranyl diphosphate to geraniol. A cDNA encoding this enzyme has been isolated and characterized from *Ocimum basillicum* [24] and *Cinnamomum tenuipilum* [25]. The isolation of the cDNA encoding geraniol synthase from *C. roseus* has not yet been reported [26]. Geraniol is then hydroxylated at the 10-position by the P450 geraniol-10-hydroxylase (G10H, CYP76B6) (Fig. 1.3). The cDNA that encodes G10H has been isolated and characterized *in planta* and *in vitro* via heterologous expression in yeast [27]. The G10H enzymatic product subsequently undergoes cyclization and hemiacetal formation to form an iridotrial intermediate. The oxidation of the iridotrial intermediate to a carboxylic acid and subsequent glucosylation, hydroxylation, and esterification lead to the formation of loganin. The order of these enzymatic steps is only speculative. Of the enzymes in this section of the pathway, 7-deoxyloganin hydrolase (DL7H) [28] has been partially purified, and a cDNA that

encodes *S*-adenosyl methonine-dependent loganic acid O-methyl transferase (LAMT) has been isolated, cloned for heterologous expression in *Escherichia coli*, and functionally characterized [29]. Loganin then undergoes oxidative cleavage via the heme-dependent P450 secologanin synthase (SLS, CYP72A1) to form secologanin. The cDNA that encodes SLS has been identified and heterologously expressed in *E. coli* for *in vitro* biochemical characterization [28].

A. tryptamine biosynthesis

L-tryptophan   tryptamine
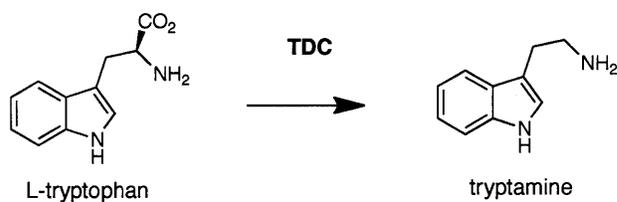
B. secologanin biosynthesis

DMAPP   IPP

geraniol   geranyl diphosphate

iridotrial
(hemiacetal form)   loganin   secologanin
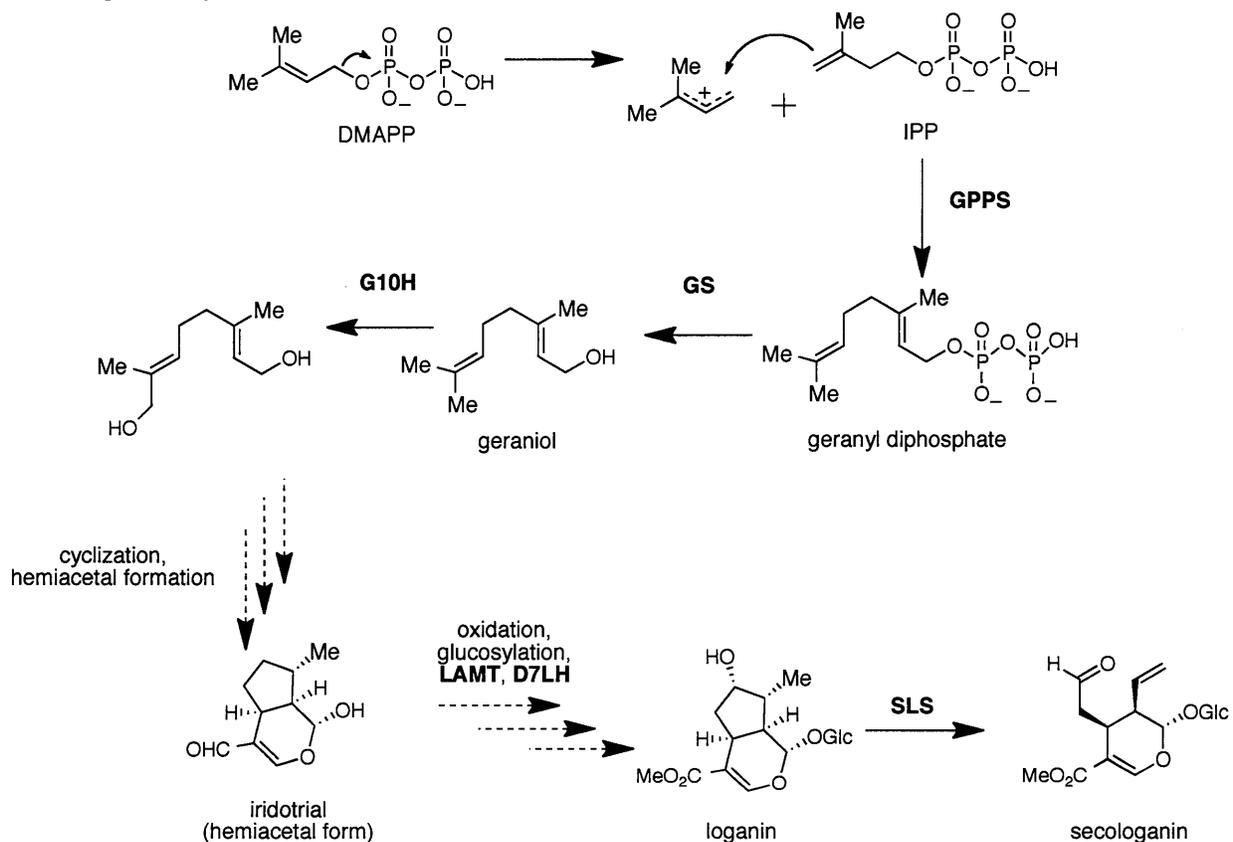
Figure 1.4 Biosynthesis of tryptamine (A) and secologanin (B), precursors to monoterpene indole alkaloids in *C. roseus*. The dashed arrows represent steps for which no enzyme has been identified.

Tryptamine and secologanin undergo a diasteroselective Pictet-Spengler condensation reaction to form (*S*)-strictosidine, a transformation catalyzed by strictosidine synthase (STR, Fig. 1.5) [30, 31]. STR has been isolated and cloned from other monoterpene indole alkaloid-producing plants such as *Ophiorrhiza pumila* [32] and *Rauvolfia serpentina* [33], the source of STR for crystallographic studies [34, 35]. Strictosidine is the common intermediate for all monoterpene indole alkaloids. The other epimer, 3-(*R*)-vincoside, has not been isolated in plants. Notably, when *in vitro* assays of STR or downstream enzymes and various synthetic stereoisomers of condensed tryptamine and desvinylsecologanin were performed, the enzymes demonstrated stringent substrate preference for substrates with the *S* stereochemistry at C3 [36].

Once formed, strictosidine is then deglucosylated by strictosidine-β-D-glucosidase (SGD) to form a hemiacetal species. The cDNA of SGD has been isolated, cloned, and heterologously expressed in *E. coli* for *in vitro* characterization [37]. SGD has also been identified in other plants, such as *R. serpentina* [38, 39]. After deglucosylation, the reactive hemiacetal undergoes several allylic isomerizations, and the biosynthetic pathway diverges (Fig. 1.5). Several reduction and rearrangement products are formed from deglycosylated strictosidine, including dehydrogeissoschizine and cathenamine (Fig 1.5). Importantly, several medicinally important compounds arise from the divergence of the biosynthetic pathway such as ajmalicine (hypotensive agent), serpentine (topoisomerase II inhibitor), yohimbine (used for treatment of erectile dysfunction) [40]. While most of the dedicated enzymes that control this branchpoint remain unknown, some studies suggest that enzyme-catalyzed reductive and oxidative chemistries are involved [41-46].
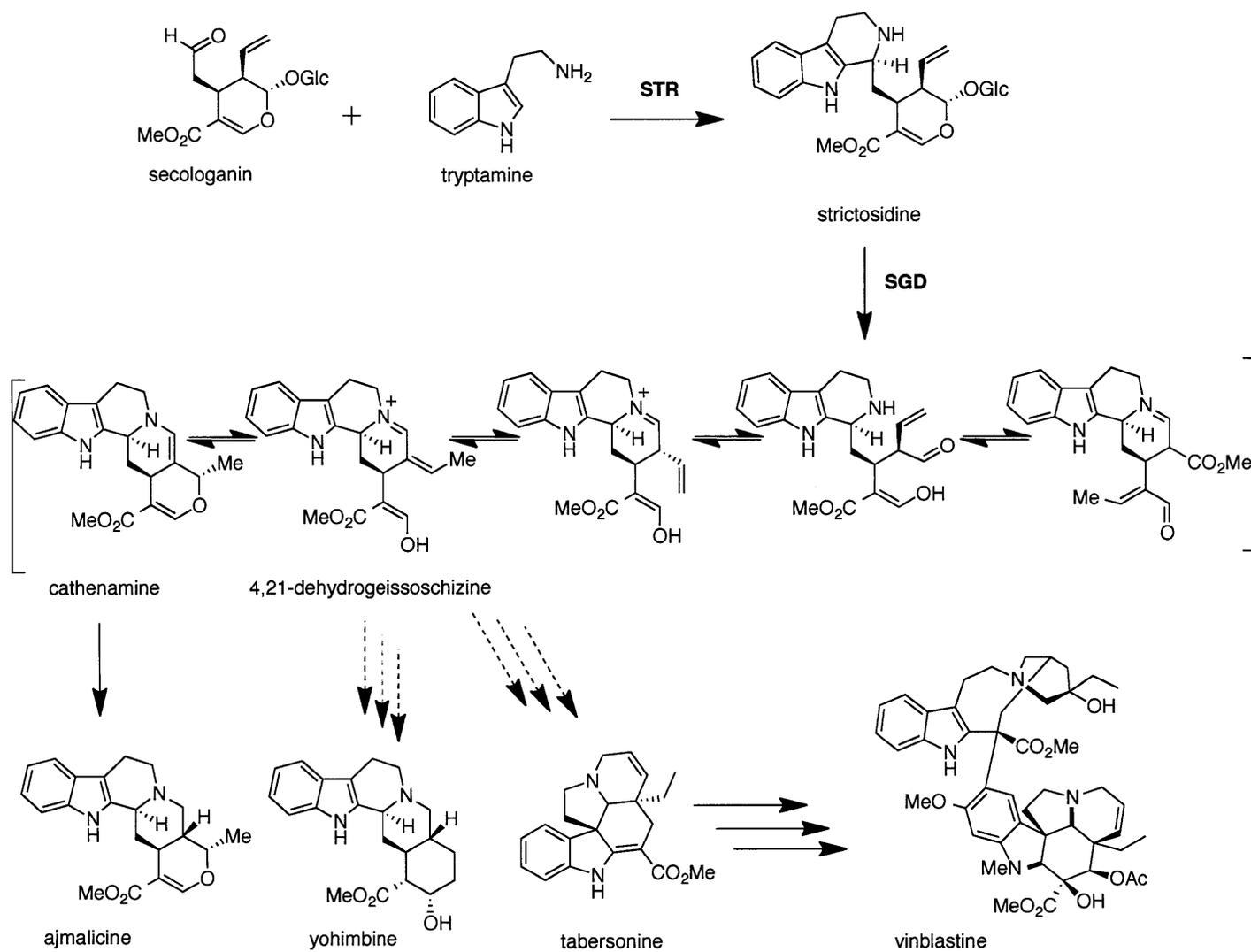
Figure 1.5 The first committed steps in monoterpene indole biosynthesis yield a variety of structurally diverse monoterpene indole alkaloids. Multiple arrows are used to denote several intermediate steps. The dashed arrows represent steps for which no enzyme has been identified.

Vindoline biosynthesis is the best characterized pathway, with five of the six predicted enzymes required to catalyze the transformation from tabersonine to vindoline having been cloned (Fig. 1.6). Notably, all intermediates involved in conversion of tabersonine to vindoline are produced in the aerial parts of differentiated *C. roseus* plants and are not found in roots (Fig. 1.6) [47]. The specific localization of these enzymes has guided the search for the genes involved in vindoline biosynthesis. For example, some enzymes are highly expressed in the early stages of leaf development in specialized idioblast and laticifer cell types [48, 49] or in the epidermis of young leaves [50]. Enzymes from these tissues can be purified to homogeneity and sequenced to identify the corresponding cDNA. The localization of these enzymes can also help deconvolute sequencing of *C. roseus* cDNA libraries prepared from whole organs. Additionally, the expression of the transcripts involved in the late stages of vindoline biosynthesis is light regulated in young leaves, which is also helpful in the identification of cDNA transcripts [51].

Starting with tabersonine, the NADPH-dependent tabersonine-16-hydroxylase (T16H, CYP71D12) hydroxylates the 16-position of the indole moiety to produce 16-hydroxytabersonine [52]. The cDNA for T16H has been isolated, cloned, and heterologously expressed in *E. coli* for *in vitro* biochemical studies [53]. A cloned and characterized *S*-adenosyl methionine-dependent 16-*O*-hydroxytabersonine-16-*O*-methyl transferase (16OMT) then methylates 16-hydroxytabersonine to form 16-methoxytabersonine [54]. *In vitro* studies on 16OMT were conducted using *E. coli* as a heterologous host for protein expression [50].

The only uncharacterized enzymatic step in vindoline biosynthesis is the hydration of the 2,3-double bond of 16-methoxytabersonine to 16-2,3-dihydro-3-hydroxytabersonine. Following

hydration of 16-methoxytabersonine, the *S*-adenosyl methionine-dependent *N*-methyl transferase (NMT) then transfers a methyl group to the indole nitrogen to form desacetoxyvindoline [55]. Interestingly, this enzyme has been recently identified as a homolog to tocopherol *C*-methyl transferases involved in vitamin E biosynthesis [56]. A cDNA for the NMT has been isolated, and *E. coli* was used as a heterologous host for *in vitro* biochemical characterization [56]. The last two steps in vindoline biosynthesis involve a hydroxylation at the C4-position and a subsequent acetylation. The first step requires a 2-oxyglutarate-dependent diooxgenase, desacetoxyvindoline 4-hydroxylase [57], followed by an acetyl coenzyme A-dependent deacetylvindoline-*O*-acetyl transferase [51]. cDNAs for both these enzymes have been identified, cloned, and heterologously expressed in *E. coli* for *in vitro* biochemical characterization. Once vindoline is made, the alkaloid catharanthine can dimerize with vindoline to ultimately form the bisindole alkaloids, such as vinblastine and vincristine, in mature leaves [46, 58, 59].

Figure
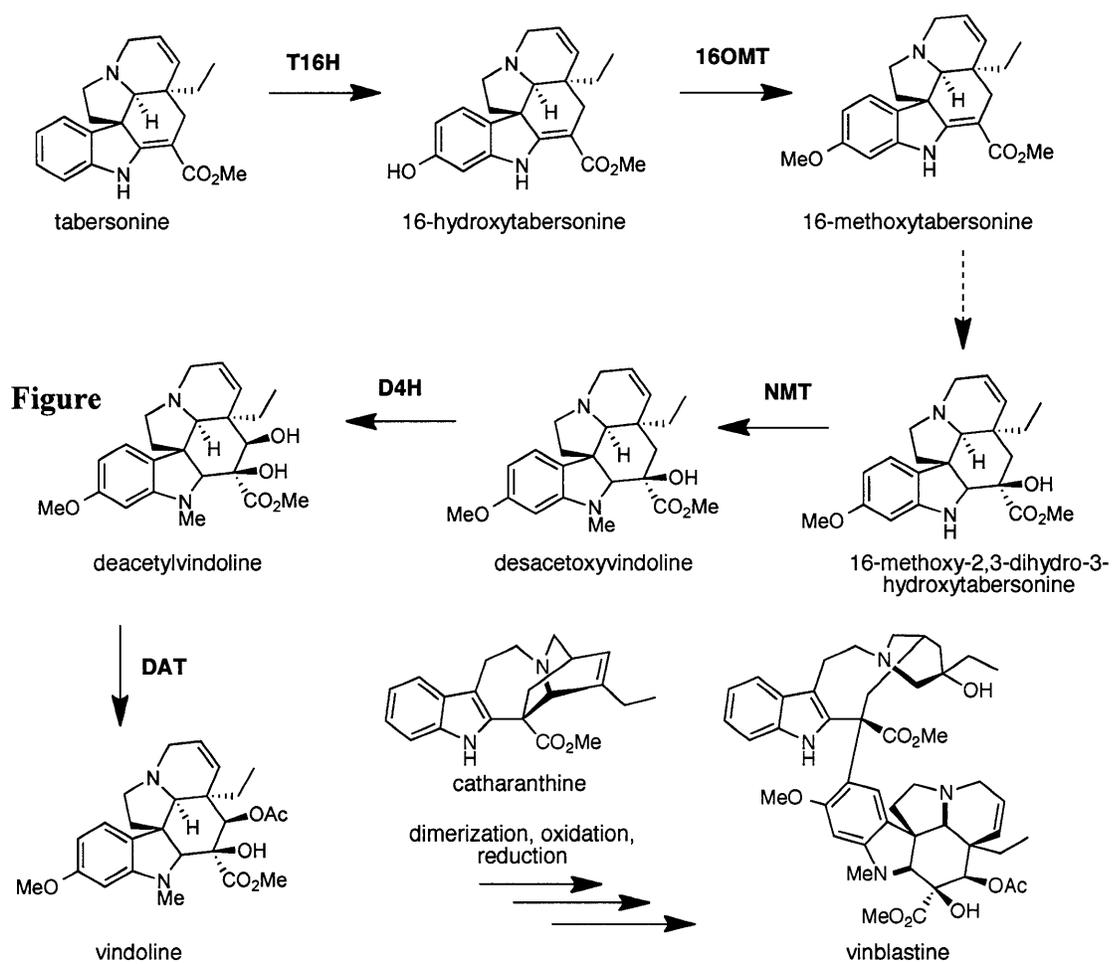
1.6 Biosynthesis of vindoline, precursor to bisindole alkaloids such as vinblastine.

1.3 Why study monoterpene indole alkaloid biosynthesis in *C. roseus*?

Many enzymes involved in monoterpene indole alkaloid biosynthesis in *C. roseus* that remain to be identified catalyze complex and unique biochemical transformations, making this an attractive metabolic system to explore. Significant progress has been made despite the complexity of the system, which contains an intricate network of natural products that are synthesized and regulated in tissue-, organ-, and development-specific ways. In the past 15 years, the addition of elicitors [60] has been demonstrated to increase the levels of secondary metabolites in plants, which has aided in the elucidation of alkaloid biosynthesis in *C. roseus*. The plant stress hormone methyl jasmonate has been found to upregulate the production of monoterpene indole alkaloids and enhance enzyme activity during the germination stage of *C. roseus* and *Cinchona* seedlings [61]. This increase in enzyme production can be used to facilitate identification of enzymes involved in alkaloid biosynthesis for functional characterization and to determine where these enzymes are localized *in planta*.

Elicitation has also been used to determine the flux of alkaloids *in planta*. The jasmonate-responsive transcription factors, ORCA2 and ORCA3, have been identified using transfer DNA tagging approaches and, in *C. roseus*, upon overexpression of these genes an increase in alkaloid production has been observed [62, 63]. Interestingly, G10H is not regulated by ORCA2 and ORCA3, but is methyl jasmonate-responsive, indicating that other jasmonate-responsive transcription factors exist. This detail can potentially be exploited to study G10H biosynthesis, which is especially attractive since many of those steps leading to G10H production are unknown.

Additionally, the compartmentalization of alkaloids, such as those involved in vindoline

biosynthesis, can be further exploited using elicitors such as methyl jasmonate to study to

metabolites made in specific plant tissues. For example, while tabersonine is converted to

vindoline in aerial parts of the plants, tabersonine is metabolized to produce different alkaloids in

the roots of *C. roseus* [64] (Fig. 1.7). The biosynthesis of the alkaloids produced in the roots is

highly susceptible to methyl jasmonate elicitation, and can thereby be used to aid in the

identification of the enzymes involved in monoterpene indole alkaloid biosynthesis [65] (Fig.

1.7).
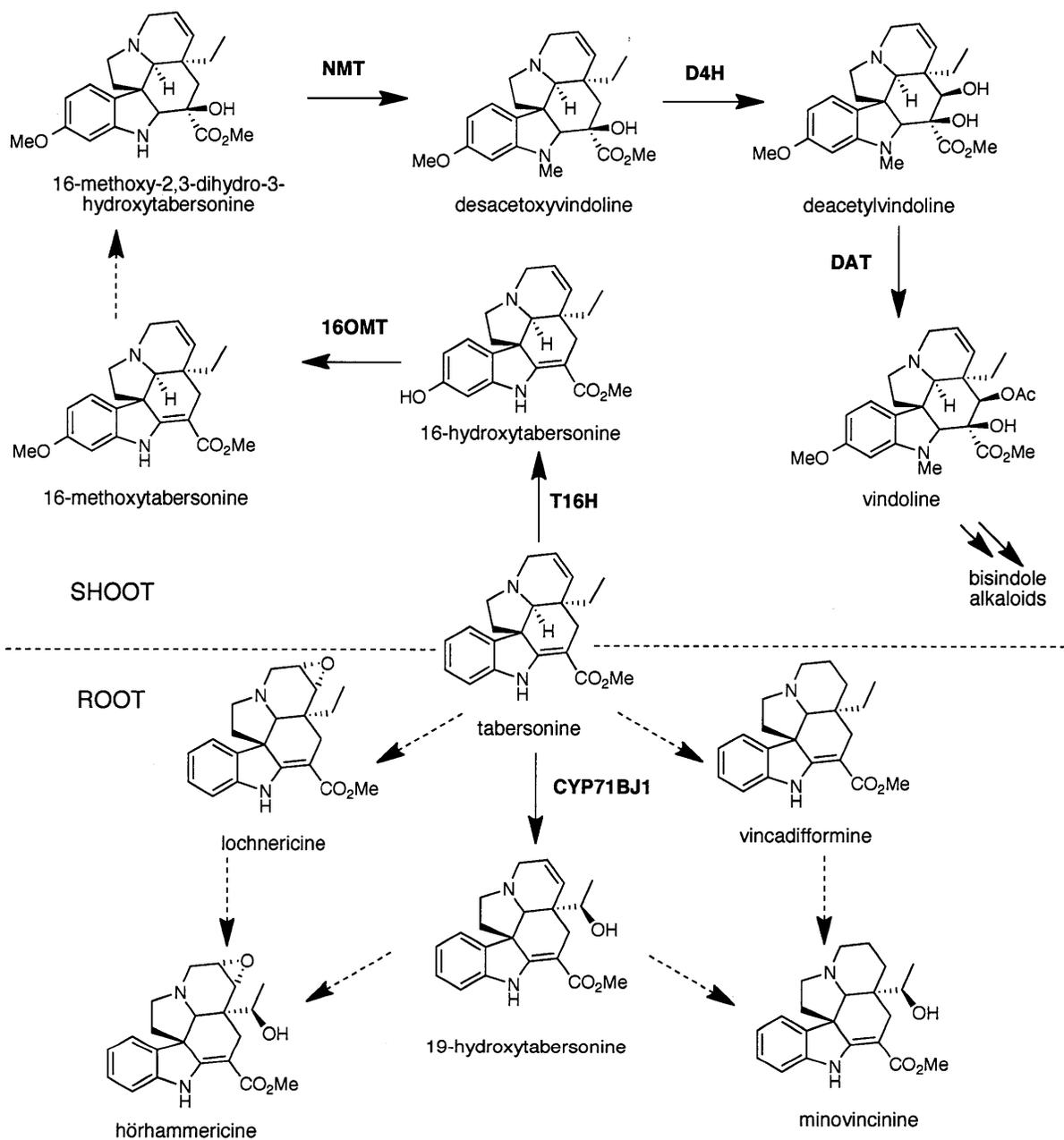
Figure 1.7 Example of tabersonine-derived alkaloids produced in different organs of *C. roseus*. Tabersonine is converted to vindoline in the idioblasts and laticifers of leaves and stems. In the roots and under certain conditions within cell cultures, tabersonine can be transformed to hörhammericine, lochnericine, and minovincine. The dashed arrows represent steps for which no enzyme has been identified.

30

The study of plant natural products has also benefited by the discovery that heterologous genes can be transformed into plants via the soil bacteria *Agrobacterium tumefaciens and A. rhizogenes*. Dicotyledenous *C. roseus* plants infected with *A. rhizogenes* bacterium produce "hairy roots" (Fig. 1.8 A) at the site of infection, which can be propagated in plant growth media. The chemical stability and fast growth that can be achieved in hairy root culture have provided biochemists with a robust system to study alkaloid biosynthesis and its regulation. It takes only 2-3 weeks to obtain gram quantities of hairy roots (Fig. 1.8 B) that accumulate high levels of alkaloids. Other plant tissue such as seedlings (Fig. 1.8 C), used to study alkaloids produced in more specialized tissues, such as the leaves, take longer to grow.

Precursor-directed biosynthesis in hairy roots has been used to assess the bottlenecks of monoterpene indole alkaloid biosynthesis in *C. roseus* as well as to probe the substrate specificities of the enzymes involved. Additionally, feeding substrates to hairy roots can guide mutagenesis studies that may ultimately aid in reengineering the pathway enzymes to produce more structurally diverse alkaloids. For example, McCoy [66, 67] and coworkers reported that STR does not turnover 5-and 6-substituted tryptamine analogs. However, with the identification and *Agrobacterium*-mediated transformation of the V214M STR mutant [67], Runguphan [68] and co-workers were able to engineer the production of more structurally diverse molecules in hairy root cultures.

Most importantly, with the development of sequencing technology, an increasing amount of sequence data will become available for plant genomes and transcriptomes. To date, there are 11 plants for which the entire genome has been sequenced [69], all of which have been published

31

within the past 11 years. The transcriptomes of hundreds of plants, including *C. roseus,* have also been reported. These data will enable the unknown steps in alkaloid synthesis to be characterized, which could lead to the reconstitution of parts of plant secondary metabolism in faster growing organisms, such as *E. coli* or yeast (*Saccharomyces cerevisiae*). Examples of yeast reconstitution have been reported for portions of the plant pathways for artemisinin [70] and benzoisoquinoline [71].

Finally, secondary metabolic pathway enzymes most likely arose from primary metabolism [72, 73], suggesting that studying the intricate chemistries catalyzed by plant natural product biosynthetic enzymes may provide insight on how classes of enzymes may be evolutionary related. Secondary metabolites outnumber primary metabolites by an order of magnitude, and with roughly 300,000 documented species of higher plants, a large piece of unknown chemical terrain remains to be explored [14].

A.

Hairy roots at *A. rhizogenes* infection site

B.

C.

Figure 1.8 Examples of *A. rhizogenes*-infected seedlings (A), hairy root culture (B), and differentiated *C. roseus* seedlings (C).

1.4. Research goals and thesis overview

A. <u>Elucidation of the mechanism of a Pictet-Spengler catalyzing enzyme.</u> The Pictet-Spengler reaction is a key reaction in the synthesis of tetrahydro-β-carboline and tetrahydroisoquinoline alkaloids [74]. In organic synthesis, acidic conditions are used to generate a mixture of enantiomeric products from the condensation of a β-ethyl arylamine and an aldehyde, and the mechanism of product formation is not clearly understood [74]. However, Nature has evolved a "Pictet-Spenglerase," strictosidine synthase (STR), to catalyze the asymmetric synthesis of strictosidine, the common tetrahydro-β-carboline intermediate for all monoterpene indole alkaloids. Chapter 2 focuses on the use of biochemical techniques, such as kinetic isotope effects, crystallographic studies, and theoretical calculations to propose a mechanism for the STR-catalyzed diastereoselective synthesis of strictosidine. The results of this research have provided a plausible mechanism for STR acid-base catalysis [31].

B. <u>Use of rational design to convert strictosidine synthase into a hydrolase.</u> STR belongs to the lactonohydrolase/paraoxonase superfamily, which is comprised of a number of hydrolases including paraoxonase and C20orf3, as well as other proposed hydrolases [75]. The unique "Pictet-Spenglerase" catalytic activity of STR suggests that this enzyme may be an outlier in this protein superfamily [76]. The crystal structures of paraoxonase [77] and STR [35] reveal that both enzymes share a similar protein fold. Chapter 3 describes how the β-propeller fold is exploited to convert STR into a hydrolase by using the sequences of closely related STR homologs that do not display Pictet-Spenglerase activity as a guide for rational mutagenesis. The results of this work reveal that one of the STR homologs from *Vitis vinifera* is a hydrolase. Moreover, using the sequence of this *V. vinifera* hydrolase as a target, STR from *C. roseus* was

34

converted into a hydrolase. In addition to the discovery and functional characterization of a new protein, the data also demonstrate that the β-propeller architecture can indeed be used as scaffold to introduce new activities within the lactonohydrolase/paraoxonase family. As such, there may be a possible evolutionary link between STR and some of the hydrolases found in plants and higher eukaryotes [72].

C. <u>Discovery and characterization of a unique P450 involved in monoterpene indole alkaloid biosynthesis</u>. Many steps in monoterpene indole alkaloid biosynthesis remain to be characterized at the enzymatic and genetic level. Tabersonine, the precursor to the clinically used bisindole alkaloids, can undergo many oxygenation reactions. Some of these transformations are speculated to be P450-dependent, but the cDNA of these enzymes remain unknown. With the recent release of the *C. roseus* transcriptome sequence data, Chapter 4 describes how hierarchical clustering and co-expression analyses enabled the identification of putative P450 enzymes involved in the metabolism of tabersonine. A whole cell assay was designed to screen for P450 activity using tabersonine as a substrate. In this work, CYP71BJ1, the first member of a new P450 family, was demonstrated to hydroxylate both lochnericine and tabersonine alkaloids at the 19-postion. Similar methods can be used to screen additional candidate P450 genes that oxidize tabersonine, such as lochnericine-6,7-epoxidase.

1. 5 References

1.   Dewick, P.M., *Medicinal natural products: A biosynthetic approach.* 3rd ed 2009, West Sussex, United Kingdom: John Wiley & Sons, Ltd.

2.   Zhu, Y.-Z., et al., *Natural products: Essential resources for human survival* 2007, Singapore: World Scientific.

3.   Newman, D.J. and G.M. Cragg, *Natural products as sources of new drugs over the past 25 years.* J. Nat. Prod., 2007. **70**(3): p. 461-477.

4.   Hanson, J.R., *Natural products: Secondary metabolites* 2003, Cambridge, UK: The Royal Society of Chemistry.

5.   Raja, P.B. and M.G. Sethuraman, *Natural products as corrosion inhibitor for metals in corrosive media — A review* Material Letters, 2008. **62**(1): p. 113-116.

6.   Ragauskas, A.J., et al., *The path forward for biofuels and biomaterials.* Science, 2006. **311**(5760): p. 484-489.

7.   Gershenzon, J. and N. Dudareva, *The function of terpene natural products in the natural world.* Nature Chemical Biology, 2007. **3**: p. 408-414.

8.   Walsh, C.T. and M.A. Fischbach, *Natural products 2.0: Connecting genes to molecules.* J. Am. Chem. Soc., 2010. **132**(8): p. 2469-2493.

9.   Chen, K. and P.S. Baran, *Total synthesis of eudesmane terpenes by site-selective C–H oxidations.* Nature, 2009. **459**: p. 824-828.

10.   Davie, E.A., et al., *Asymmetric catalysis mediated by synthetic peptides.* Chem. Rev., 2007. **107**(12): p. 5759-5812.

11.   Shang, S. and D.S. Tan, *Advancing chemistry and biology through diversity-oriented synthesis of natural product-like libraries* Current Opinion in Chemical Biology, 2005. **9**(3): p. 248-258.

12.   Chan, J.N.Y., C. Nislow, and A. Emili, *Recent advances and method development for drug target identification* Trends in Pharmacological Sciences, 2010. **31**(2): p. 82-88.

13.   Kingston, D.G.I., *Modern natural products drug discovery and its relevance to biodiversity conservation.* J. Nat. Prod., 2011. **74**(3): p. 496-511.

14.   Wu, S. and J. Chappell, *Metabolic engineering of natural products in plants; tools of the trade and challenges for the future.* Curr. Opin. Biotech., 2008. **19**(2): p. 145-152.

15.   Li, J.W.-H. and J.C. Vederas, *Drug discovery and natural products: End of an era or an endless frontier?* Science, 2009. **325**(5937): p. 161-165.

16.   Saklani, A. and S.K. Kutty, *Plant-derived compounds in clinical trials.* Drug Disc. Today, 2008. **13**(3-4): p. 161-171.

17.   Leete, E., *Biogenesis of Rauwolfia alkaloids. II. Incorporation of tryptophan into serpentine and reserpine.* Tetrahedron, 1961. **14**(1-2): p. 35-41.

18.   Runguphan, W., J. Maresh, and S.E. O'Connor, *Silencing of tryptamine biosynthesis for production of nonnatural alkaloids in plant culture.* Proc. Natl. Acad. Sci. USA, 2009. **106**(33): p. 13673-13678.

19.   Canel, C., et al., *Effects of over-expression of strictosidine synthase and tryptophan decarboxylase on alkaloid production by cell cultures of Catharanthus roseus.* Planta, 1998. **205**(3): p. 414-419.

20. Whitmer, S., *Aspects of terpenoid indole alkaloid formation by transgenic cell lines of Catharanthus roseus overexpressing tryptophan decarboxylase and strictosidine synthase*, 1999, Lelden University.

21. Uesato, S., S. Matsuda, and H. Inouye, *Mechanism for iridane skeleton formation from acyclic monoterpenes in the biosynthesis of secologanin and vindoline in Catharanthus roseus and Lonicera morrowii.* Chem. Pharm. Bull., 1984. **32**(4): p. 1671-1674.

22. Banthorpe, D.V., B.V. Charlwood, and M.J.O. Francis, *The biosynthesis of monoterpenes.* Chem. Rev., 1972. **72**(2): p. 115-155.

23. Oudin, A., et al., *The iridiod pathway in Catharanthus roseus alkaloid biosynthesis.* Phytochem Rev, 2007. **6**(2-3): p. 259-276.

24. Iijima, Y., et al., *The biochemical and molecular basis for the divergent patterns in the biosynthesis of terpenes and phenylpropenes in the peltate glands of three cultivars of basil.* Plant Physiol., 2004. **136**(3): p. 3724-3736.

25. Yang, T., et al., *A geraniol-synthase gene from Cinnamomum tenuipilum.* Phytochemistry, 2005. **66**(3): p. 285-293.

26. Simkin, A.J., *Rapport a'activite studium 2008-1010*, 2010, Université François Rabelais: Tours. p. 1-21.

27. Collu, G., et al., *Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis.* FEBS Lett., 2001. **508**(2): p. 215-220.

28. Irmler, S., et al., *Indole alkaloid biosynthesis in Catharanthus roseus: new enzyme activities and identification of cytochrome P450 CYP72A1 as secologanin synthase.* Plant J., 2000. **24**(6): p. 797-804.

29.   Murata, J., et al., *The leaf epidermome of Catharanthus roseus reveals its biochemical specialization.* Plant Cell, 2008. **20**(3): p. 524-542.

30.   Mizukami, H., et al., *Purification and properties of strictosidine synthase from Catharanthus roseus cultured cells.* Biochemistry, 1979. **18**(17): p. 3760-3763.

31.   Maresh, J.J., et al., *Strictosidine synthase: mechanism of a Pictet-Spengler catalyzing enzyme.* J. Am. Chem. Soc., 2008. **130**(2): p. 710-723.

32.   Yamazaki, Y., et al., *Camptothecin biosynthetic genes in hairy roots of Ophiorrhiza pumila: Cloning, characterization and differential expression in tissues and by stress compounds.* Plant Cell Physiol., 2003. **44**(4): p. 395-403.

33.   Treimer, J.F. and M.H. Zenk, *Strictosidine synthase from cell cultures of Apocynaceae plants.* FEBS Lett., 1979. **97**: p. 159-162.

34.   Ma, X.Y., et al., *Crystallization and preliminary X-ray crystallographic analysis of strictosidine synthase from Rauvolfia: the first member of a novel enzyme family* Biochimica et Biophysica Acta, 2004. **1702**(1): p. 121-124.

35.   Ma, X., et al., *The structure of Rauvolfia serpentina strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins.* The Plant cell, 2006. **18**(4): p. 907-920.

36.   Bernhardt, P., N. Yerkes, and S.E. O'Connor, *Bypassing stereoselectivity in the early steps of alkaloid biosynthesis.* Organic and Biomolecular Chemistry, 2009. **7**: p. 4166-4168.

37.   Scott, A.I., S.L. Lee, and W. Wan, *Indole alkaloid biosynthesis: Partial purification of "ajmalicine synthetase" from Catharanthus roseus.* Biochem. Biophys. Res. Comm., 1977. **75**(4): p. 1004-1009.

38.   Barleben, L., et al., *Expression, purification, crystallization, and preliminary X-ray analysis of strictosidine glucosidase, an enzyme initiating biosynthetic pathways to a unique diversity of indole alkaloid skeletons.* Biochimica et Biophysica Acta, 2004. **1747**(1): p. 89-92.

39.   Barleben, L., et al., *Molecular architecture of strictosidine glucosidase: The gateway of the biosynthesis of the monoterpenoid indole alkaloid family.* Plant Cell, 2007. **19**(9): p. 2886-2897.

40.   O'Connor, S.E. and J. Maresh, *Chemistry and biology of monoterpene indole alkaloid biosynthesis.* Nat. Prod. Rep., 2006. **23**(4): p. 532-547.

41.   Hemscheidt, T. and M.H. Zenk, *Partial purification and characterization of a NADPH dependent tetrahydroalstonine synthase from Catharanthus roseus cell suspension cultures.* Plant Cell Rep., 1985. **4**(4): p. 216-219.

42.   Stöckigt, J., A. Pfitzner, and P.I. Keller, Tetrahedron Lett., 1983. **244**: p. 2485.

43.   Stöckigt, J., *Indirect involvement of geissoschizine in the biosynthesis of ajmalicine and related alkaloids.* J. Chem. Soc. Chem. Comm., 1978: p. 1097-1099.

44.   Stöckigt, J., *Indirect involvement of geissoschizine in the biosynthesis of ajmalicine and related alkaloids.* J. Chem. Soc. Chem. Comm., 1978(24): p. 1097-1099.

45.   Pfitzner, A. and J. Stöckigt, *Partial purification and characterization of geissoschizine dehydrogenase from suspension cultures of Catharanthus roseus.* Phytochemistry, 1982. **21**(7): p. 1585-1588.

46.   Sottomayor, M., et al., *Peroxidase and the biosynthesis of terpenoid indole alkaloids in the medicinal plant Catharanthus roseus (L.) G. Don.* Phytochemistry Rev., 2004. **3**(1-2): p. 159-171.

47.   Murata, J. and V. De Luca, *Localization of tabersonine 16-hydroxylase and 16-OH tabersonine-16-O-methyltransferase to leaf epidermal cells defines them as a major site of precursor biosynthesis in the vindoline pathway in Catharanthus roseus.* Plant J., 2005. **44**(4): p. 581-594.

48.   De Luca, V. and A.J. Cutler, *Subcellular localization of enzymes involved in indole alkaloid biosynthesis in Catharanthus roseus.* Plant Physiol., 1987. **85**(4): p. 1099-1102.

49.   St-Pierre, B., F. Vazquez-Flota, and V. De Luca, *Multicellular compartmentation of Catharanthus roseus alkaloid biosynthesis predicts intercellular translocation of a pathway intermediate.* Plant Cell, 1999. **11**(5): p. 887-900.

50.   Levac, D., et al., *Application of carborundum abrasion for investigating the leaf epidermis: molecular cloning of Catharanthus roseus 16-hydroxytabersonine-16-O-methyl transferase.* Plant J., 2008. **53**(2): p. 225-236.

51.   St-Pierre, B., et al., *The terminal O-acetyltransferase involved in vindoline biosynthesis defines a new class of proteins responsible for coenzyme A-dependent acyl transfer.* Plant J., 1998. **14**(6): p. 703-713.

52.   St-Pierre, B. and V. De Luca, *A cytochrome P-450 monooxygenase catalyzes the first step in the conversion of tabersonine to vindoline in Catharanthus roseus.* Plant Physiol., 1995. **109**(1): p. 131-139.

53.   Schröder, G., et al., *Light-induced cytochrome P450-dependent enzyme in indole alkaoid biosynthesis: tabersonine 16-hydroxylase.* FEBS Lett., 1999. **458**(2): p. 97-102.

54.   Cacace, S., et al., *A flavonol O-methyltransferase from Catharanthus roseus performing two sequential methylations.* Phytochemistry, 2003. **62**(2): p. 127-137.

55.   Dethier, M. and V. De Luca, *Partial purification of an N-methyltransferase involved in vindoline biosynthesis in Catharanthus roseus.* Phytochemistry, 1993. **32**(3): p. 673-678.

56.   Liscombe, D.K., A.R. Usera, and S.E. O'Connor, *Homolog of tocopherol C methyltransferases catalyzes N methylation in anticancer alkaloid biosynthesis.* Proc. Natl. Acad. Sci. USA, 2010. **107**(44): p. 18793-18798.

57.   Vazquez-Flota, F., et al., *Molecular cloning and characterization of desacetoxyvindoline-4-hydroxylase, a 2-oxoglutarate dependent-dioxygenase involved in the biosynthesis of vindoline in Catharanthus roseus (L.) G. Don.* Plant Mol. Biol., 1997. **34**(6): p. 935-948.

58.   Sottomayor, M., F. DiCosmo, and A. Barceló, *The fate of the catharanthine and vindoline during the peroxidase-mediated enzymatic synthesis of α-3',4'-anhydrovinblastine.* Enz. Microbiol. Tech., 1997. **21**(7): p. 543-549.

59.   Sottomayor, M., et al., *Purification and characterization of α-3',4'-anhydrovinblastine synthase (peroxidase like) from Catharanthus roseus (L.) G. Don.* FEBS Lett., 1998. **428**(3): p. 299-303.

60.   Rijhwani, S.K. and J.V. Shanks, *Effect of elicitor dosage and exposure time on biosynthesis of indole alkaloids by Catharanthus roseus hairy root cultures.* Biotechnol. Prog., 1998. **14**(3): p. 442-449.

61.   Aerts, R.J., et al., *Methyl jasmonate vapor increases the developmentally controlled synthesis of alkaloids in Catharanthus and Cinchona seedlings.* Plant J., 1994. **5**(5): p. 635-643.

62.   Memelink, J., R. Verpoorte, and J.W. Kinje, *ORCAnisation of jasmonate-responsive gene expression in alkaloid metabolism.* Trends in Plant Science, 2001. **6**(5): p. 212-219.

63.   van der Fits, L. and J. Memelink, *ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism.* Science, 2000. **289**(5477): p. 295-297.

64.   Laflamme, P., B. St-Pierre, and V. De Luca, *Molecular and biochemical analysis of a Madagascar periwinkle root-specific minovincinine-19-hydroxy-O-acetyltransferase.* Plant Physiol., 2001. **125**(1): p. 189-198.

65.   Morgan, J.A. and J.V. Shanks, *Inhibitor studies of tabersonine metabolism in C. roseus hairy roots.* Phytochemistry, 1999. **51**(1): p. 61-68.

66.   McCoy, E., M.C. Galan, and S.E. O'Connor, *Substrate specificity of strictosidine synthase.* Bioorg. Med. Chem. Lett., 2006. **16**(9): p. 2475-2478.

67.   Bernhardt, P., E. McCoy, and S.E. O'Connor, *Rapid identification of enzyme variants for reengineered alkaloid biosynthesis in periwinkle.* Chem. Biol., 2007. **14**(8): p. 888-897.

68.   Runguphan, W., X. Qu, and S.E. O'Connor, *Integrating carbon-halogen bond formation into medicinal plant metabolism.* Nature, 2010. **468**(7322): p. 461-464.

69.   Nelson, D.R. and D. Werck-Reichhart, *A P450-centric view of plant evolution.* Plant J., 2011. **66**(1): p. 194-211.

70.   Ro, D.-K., et al., *Production of the antimalarial drug precursor artemisinic acid in engineered yeast.* Nature, 2006. **446**: p. 940-943.

71.   Hawkins, K.M. and C.D. Smolke, *Production of benzylisoquinoline alkaloids in Saccharomyces cerevisiae.* Nat. Chem. Biol., 2008. **4**(9): p. 564-573.

72.   Khersonsky, O. and D.S. Tawfik, *Enzyme promiscuity: a mechanistic and evolutional perspective.* Annu. Rev. Biochem., 2010. **79**: p. 471-505.

73.   Siegler, D.S., *Plant secondary metabolism* 1998, Norwell, Massachusetts: Kluwer

Academic Publishers. 1-15.

74.   Cox, E.D. and J.M. Cook, *The Pictet-Spengler condensation: A new direction for an old

reaction.* Chem. Rev., 1995. **95**(6): p. 1797-1842.

75.   Kobayashi, M., et al., *Lactone-ring-cleaving enzyme: genetic analysis, novel RNA

editing, and evolutionary implication.* Proc. Natl. Acad. Sci. USA, 1998. **95**(22): p.

12787-12792.

76.   Hicks, M.A., et al., *The Evolution of New Function in Strictosidine Synthase-like

Proteins*, 2011: San Francisco. p. 1-28.

77.   Harel, M., et al., *Structure and evolution of the serum paraoxonase family of detoxifying

and anti-atherosclerotic enzymes.* Nature structural & molecular biology, 2004. **11**(5): p.

412-419.

CHAPTER 2

MECHANISM OF STRICTOSIDINE SYNTHASE: A PICTET-SPENGLER

CATALYZING ENZYME

2.1 Introduction

The Pictet-Spengler reaction produces tetrahydro-β-carboline and tetrahydroisoquinoline compounds that are essential in the (bio)synthesis of bioactive natural products. The structural complexity that can be achieved using this reaction makes the Pictet-Spengler reaction an important method for alkaloid synthesis [1]. A number of synthetic research groups have focused on developing asymmetric catalysts to control the stereochemical outcome of Pictet-Spengler reactions in the synthesis of complex molecules [2, 3]. However, Nature has evolved an enzyme with an active site built for asymmetric synthesis at pH 7. Biochemists are interested in studying the Pictet-Spengler reaction catalyzed by biosynthetic enzymes to understand how Nature developed the machinery to catalyze such intricate chemistry.

The Pictet-Spengler reaction occurs in several steps [1, 4]. First, an electron-rich β-arylethylamine such as tryptamine condenses with an aldehyde to form an iminium intermediate (Fig. 2.1, steps 1-3). Second, the aryl amine attacks the electrophilic iminium (the cyclization step) to yield a positively charged intermediate (Fig. 2.1, step 4), which is then deprotonated to yield two possible enantiomers of a β-carboline product. Intriguingly, after iminium formation, Pictet-Spengler reactions that utilize indole amine substrates can proceed either by an attack of C2 on the indole moiety to yield the 6-membered ring intermediate (Fig. 2.1, step 4) or by an attack of C3 to yield a spiroindolenine intermediate (Fig. 2.1, step 4a) that can undergo a 1,2-alkyl shift (Fig. 2.1, step 4b) to form the product. Both C2 and C3 of the indole moiety of the aryl amine are nucleophilic. According to Baldwin's rules [5], forming the 6-membered ring

intermediate via a 6-endo ring closure is favorable, while 5-endotrig cyclization to form

the spiroindolenine is unfavorable. However, evidence exists for both mechanisms in

solution [6-9], and at the outset of these studies, the predominant mechanism of the

nonenzymatic or the enzymatic Pictet-Spengler reaction was not entirely clear.
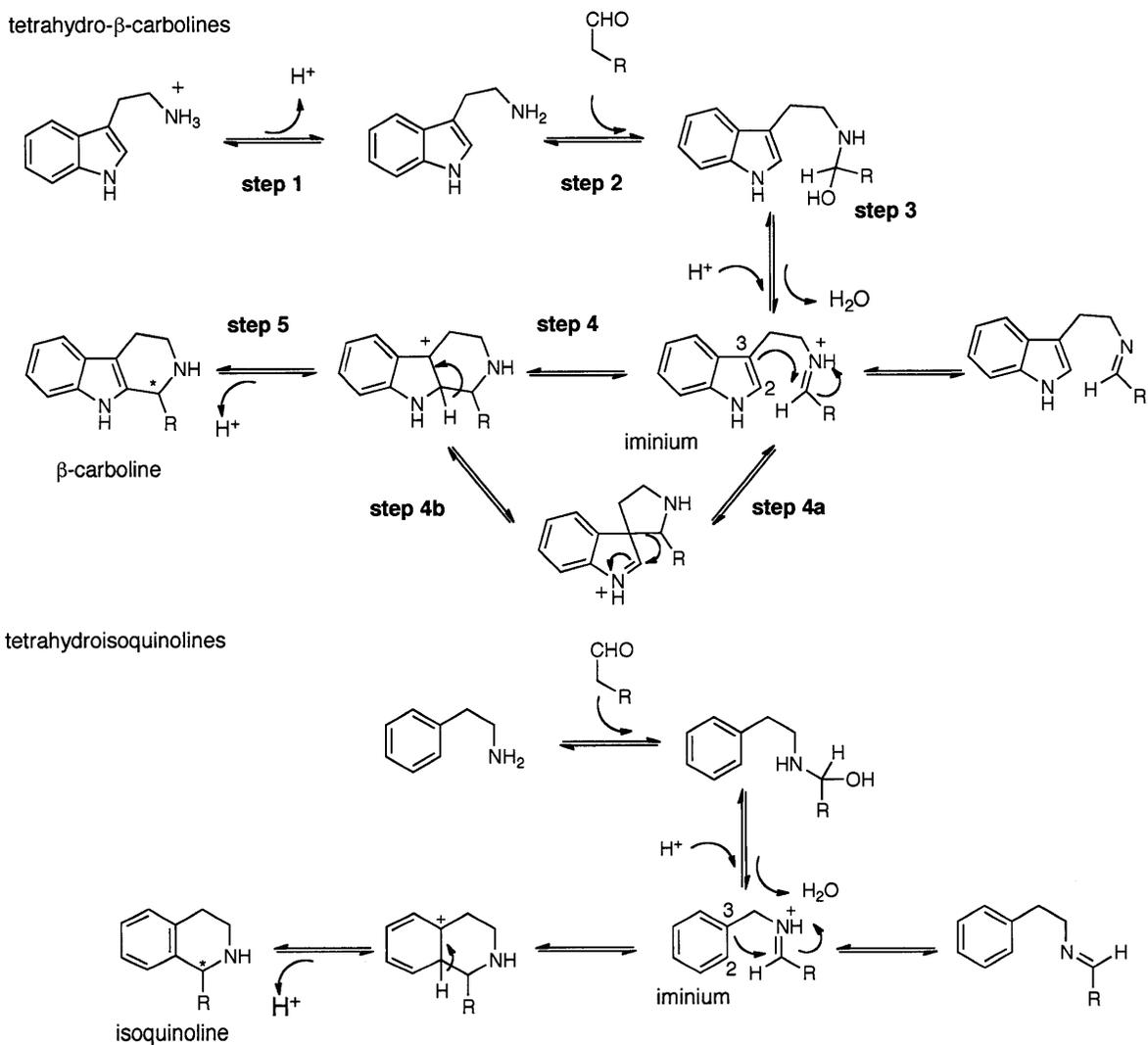
Figure. 2.1 Mechanism of the Pictet-Spengler reactions that utilize an aldehyde and either indole amine and phenethylamine substrates to form tetrahydro-β-carbolines and tetrahydroisoquinolines, respectively.

To date, only four enzymes are known to catalyze the Pictet-Spengler reaction: deacetylipecoside synthase [10], deacetylisoipecoside synthase [10], norcoclaurine synthase (NCS) [11], and strictosidine synthase (STR) [12] (Fig. 2.2). These enzymes have been isolated from several plant biosynthetic pathways [10, 13-16]. Deacetylipecoside and deacetylisoipecoside synthases catalyze a Pictet-Spengler condensation of dopamine and secologanin to form (*R*)-deacetylipecoside and (*S*)-deacetylisoipecoside, respectively (Fig 2.2 B). These enzymes form the common precursor for a large family of tetrahydroisoquinoline alkaloids. NCS catalyzes a condensation of dopamine and 4-hydroxyphenylacetaldehyde to form (*S*)-norcoclaurine, the precursor to over 5000 benzoisoquinoline alkaloids (Fig. 2.2 C). Lastly, STR [17, 18] catalyzes the first-committed step in the biosynthesis of thousands of monoterpene indole alkaloids, a structurally diverse family of natural products that includes pharmaceutically valuable compounds such as vinblastine, vincristine, and camptothecin [19, 20] (Fig. 2.2 A). This chapter focuses on understanding the mechanism by which STR converts tryptamine and the iridoid monoterpene secologanin to form the product (*S*)-strictosidine (Fig. 2.2), the common biosynthetic precursor to all monoterpene indole alkaloids.
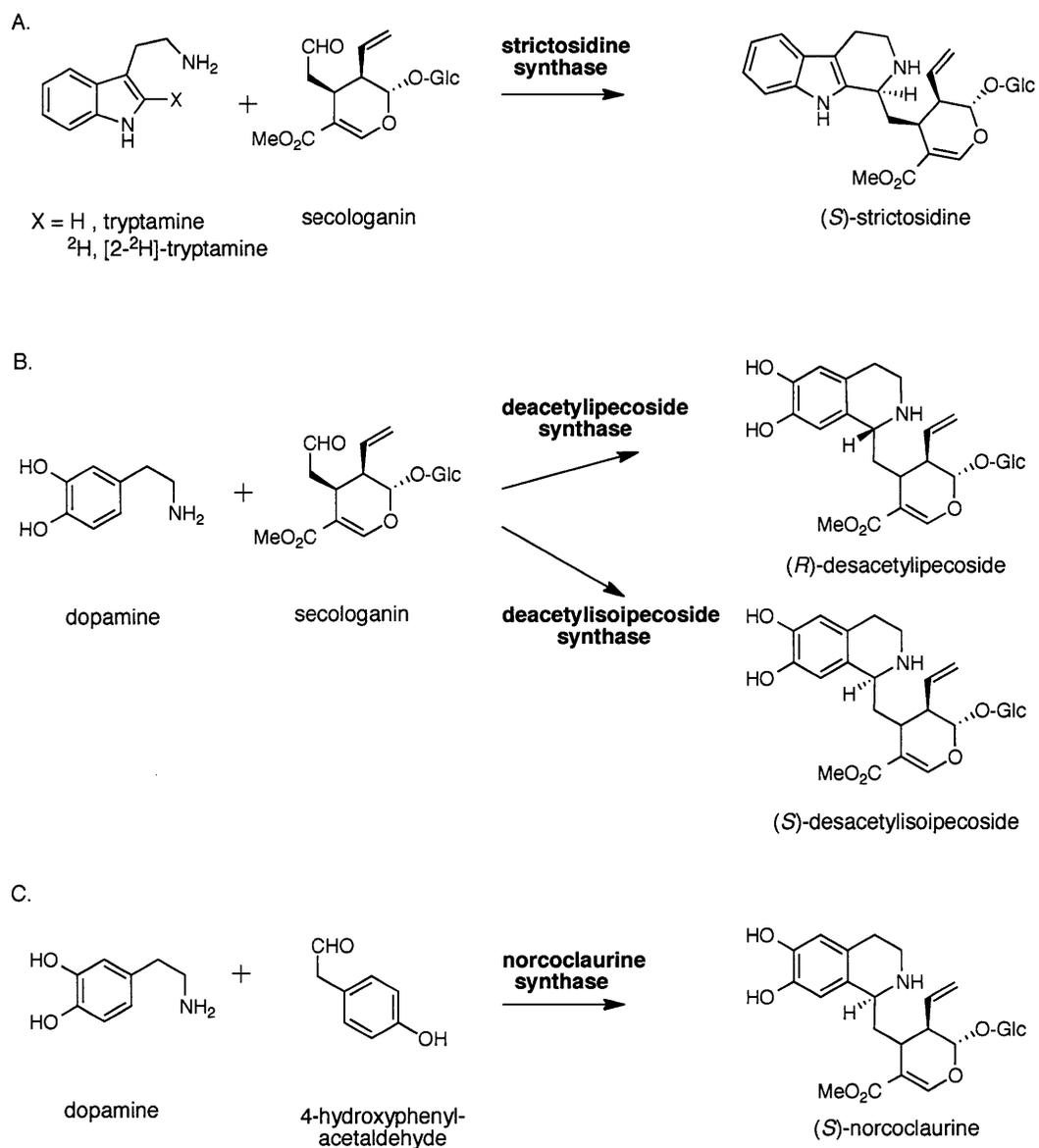
Figure 2.2 Pictet-Spenglerases found in nature. A) Strictosidine synthase. B) Deacetylipecoside and deacetylisoipecoside synthases. C) Norcoclaurine synthase.

At the time of this study, little was known about the structures and mechanisms of the Pictet-Spengler-catalyzing enzymes. STR was the first Pictet-Spenglerase to be identified and was isolated over 30 years ago from *Catharanthus roseus* and *Rauvolfia serpentina* (89 % sequence homology to the enzyme from *C. roseus*). Several steady-state kinetic analyses had been conducted on STR to understand the enzyme's mechanism of catalysis [21, 22]. Finally in 2006, STR from *Rauvolfia serpentina* was co-crystallized in the presence of both secologanin (2FPC.pdb) and tryptamine (2FPB.pdb) substrates [23]. These structures provided the first insights into the orientation of substrate binding and enzymatic mechanism. The STR crystal structure revealed the presence of three ionizable residues, Tyr151, His307, and Glu309 (*R. serpentina* numbering) in the enzyme active site (Fig. 2.3) [23]. Site-directed mutagenesis of Tyr151 to phenylalanine did not alter the catalytic activity significantly, suggesting that the ionizable hydroxyl group does not play a key role in catalysis [23]. His307 appeared to be involved in binding to the glucose moiety of secologanin as evidenced by the crystal structure and the significant increase in secologanin K_m after mutation of this residue to alanine [23]. Site-directed mutagenesis of Glu309 resulted in a 900-fold decrease in catalytic activity, supporting the involvement of this residue in catalysis [23].
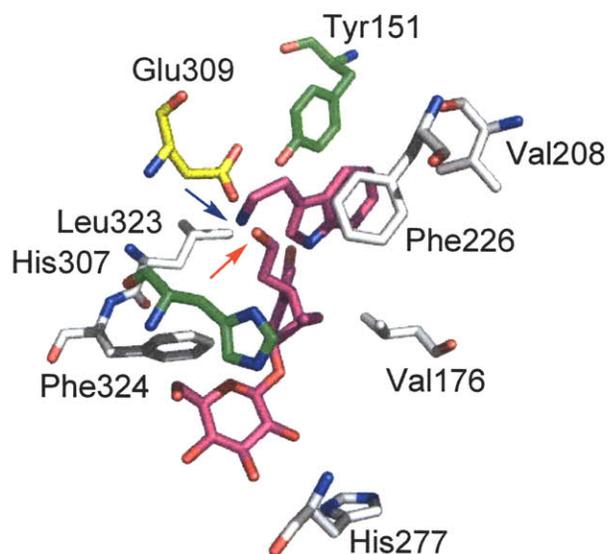
Figure 2.3 STR crystal structure highlighting select amino acids that comprise the active site, which binds tryptamine and secologanin. The crystal structure of STR with tryptamine (2FPB.pdb) is overlaid with structure of secologanin (2FPC.pdb). Both substrates are shown in pink. The amine of tryptamine is highlighted with a blue arrow, and the aldehyde of secologanin is highlighted with a red arrow. Surrounding polar residues are shown in green, and the key Glu309 residue is shown in yellow. Nonpolar residues are shown in gray (Numbering is from *R. serpentina* enzyme).

Using kinetic isotope effects (KIEs), pH rate profiles, and STR structural data, we proposed a mechanism for enzyme catalysis. The rate-dependence on pH and KIE data suggest the involvement of both an acid-catalyzed step presumably involved in iminium formation (Fig. 2.1, step 3) and a base-catalyzed step involved in the final deprotonation step (Fig. 2.1, step 5). Based on primary KIEs measured for enzyme assays between the pH range of 3.10 to 9.19, we propose that an active site glutamate residue, previously implicated by site directed mutagenesis [23], acts as a general acid and a base that initially assists in formation of an iminium species (step 3, Fig. 2.1), and then acts to deprotonate C2 to form strictosidine (step 5, Fig. 2.1). Additionally, *ab initio* and crystallographic studies with a potential transition state mimic provided insight into the nature of substrate binding and the productive transition states involved in the reaction. Specifically, *ab initio* calculations provided insight into step 4 of Fig. 2.1 using tryptamine and acetaldehyde as a model substrates for the Pictet-Spengler reaction. Notably, *ab initio* calculations suggest that formation of the spiroindolenine intermediate shown in Fig. 2.1 (step 4a) is nonproductive, and a 1,2-alkyl shift of this intermediate (Fig. 2.1, step 4b) does not occur.

2.2 Results

2.2.1 Observed isotope effects of the maximum rate, V

Isotope effects of the enzymatic reaction catalyzed by STR from *C. roseus* were measured with tryptamine and [2-^{2}H]-tryptamine to determine whether the isotopically sensitive step is the rate-controlling. STR was expressed and purified as reported by McCoy et al [24]. I collaborated with Dr. Justin Maresh, who prepared tryptamine and [2-

^{2}H]-tryptamine (Fig. 2.1) substrates following parallel synthetic procedures to ensure that any trace impurities that might affect the reaction rate would be present in both labeled and unlabeled samples. Upon incubating purified STR with isolated secologanin [24] and tryptamine/ [2-^{2}H]-tryptamine at 30 °C, differences in the rate of hydrolyzed strictosidine (hydrolysis occurred because assays were quenched with 2 M NaOH) formation were observed by HPLC (Fig. 2.4) [25]. At saturating concentrations of both substrates tryptamine/ [2-^{2}H]-tryptamine and secologanin, a primary deuterium isotope effect of 2.67 ± 0.13 on the maximum rate, ^DV, was observed [26]. For the purpose of this chapter, isotope effects on limiting macroscopic rate constants V_{max} and V_{max}/K_m are referred to as ^DV and $^D(V/K_x)$, respectively (the subscript "x" indicates the varied substrate). The maximum rate, V, expressed under saturating substrate conditions, is controlled by the rate constants of all steps after substrate binding up to and including the release of product [27]. As such, the isotope effect on V is essentially dependent on the chemical reaction and product release. A primary KIE on ^DV indicates that the isotopic sensitive step is rate-controlling. Since we observed a significant primary KIE for ^DV, the data suggest that the chemical steps of the Pictet-Spengler reaction, namely deprotonation of the 6-membered intermediate (Fig. 2.1, step 4) and not product release, are largely rate-controlling.
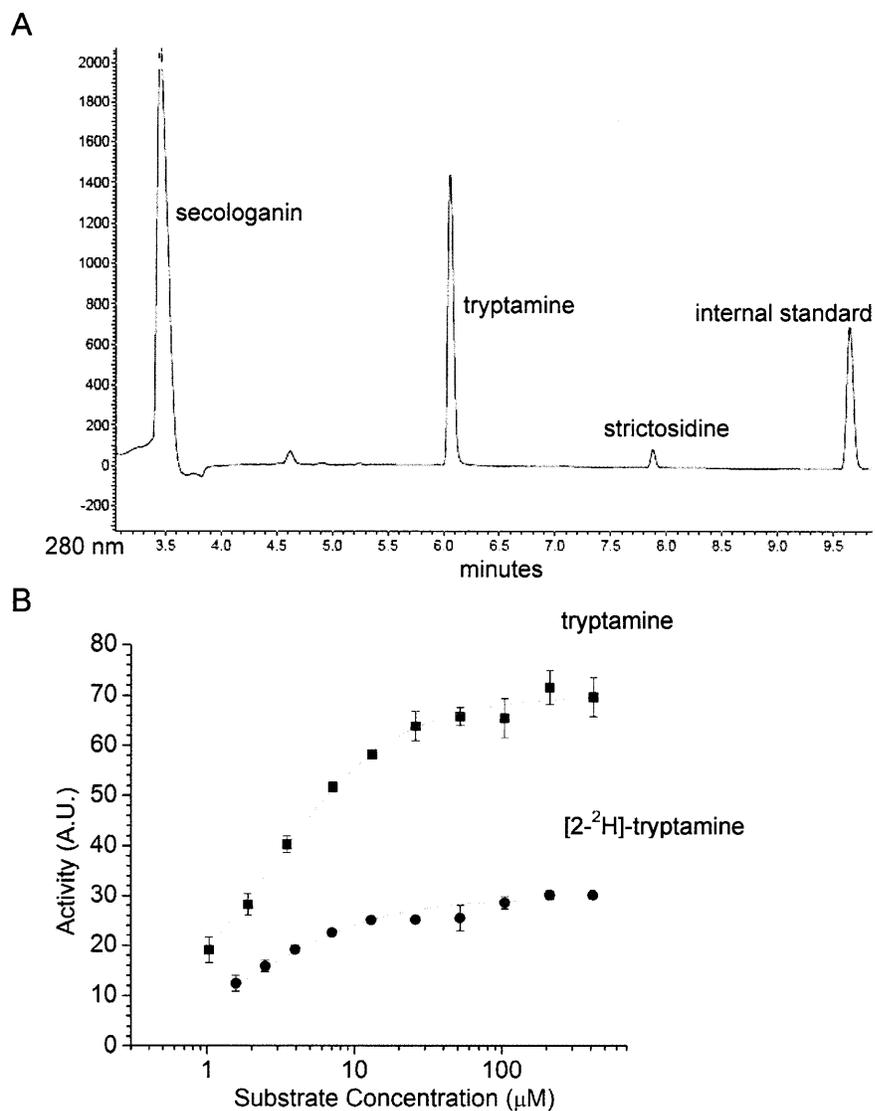
Figure 2.4 A) Representative HPLC trace for STR-catalyzed reaction of tryptamine with secologanin to yield strictosidine. As a result of quenching the assays with NaOH, the methyl ester of secologanin and strictosidine have been hydrolyzed. B) Representative kinetic data for tryptamine and [2-^{2}H]-tryptamine (2.5 mM secologanin, 80 nM STR, 30°C) obtained using this assay. The error bars represent the 95 % confidence of the standard deviations of 4 experiments.

2.2.2 KIE V/K

At limiting substrate concentrations, V/K is the pseudo second order rate constant. The magnitude of isotope effects on V and V/K can be used to estimate the relative rates at which the substrates dissociate from the enzyme [27]. A smaller, but significant primary KIE was observed at limiting concentrations of either tryptamine/[2-^{2}H]-tryptamine or secologanin (^{D}(V/K$_{tryp}$) = 1.87 ± 0.37 and ^{D}(V/K$_{sec}$) = 1.45 ± 0.28). When the varied substrate forms products faster than it dissociates from the enzyme-substrate complex, there are external commitments to catalysis, which lower the isotopic effect on V/K. As such, tryptamine is most likely released at a faster rate than secologanin from the enzyme-secologanin-tryptamine ternary complex [27]. The magnitude of ^{D}(V/K) isotope effects also provides information about order of substrate binding for multi-substrate enzymes. In an ordered mechanism, ^{D}(V/K) approaches unity if the varied substrate binds to the enzyme first [27]. Our data show primary KIE values greater than unity were measured (^{D}(V/K$_{tryp}$) = 1.87 ± 0.37 and ^{D}(V/K$_{sec}$) = 1.45 ± 0.28), suggesting that the order of substrate binding is random. The random binding of substrates is consistent with structural data obtained of STR co-crystallized with tryptamine or secologanin, data revealing that either substrate can bind to STR independently of one another [23].

Additionally, isotope effects on V and V/K can be used to describe the relationship between the Michealis constant (K$_m$) and the dissociation constant K$_D$ [28]. Using the equation derived from Klinman et al. (equation 2.1) [28], previously measured K$_m$ values,

and the isotope effects on V/K, the K_D for secologanin and tryptamine were estimated to be $K_{D\,tryp} = 2$ μM ($K_{m\,tryp} = 4$ μM) and $K_{D\,sec} = 22$ μM ($K_{m\,sec} = 50$ μM), respectively.

$$\frac{^DV - 1}{^D(V/K) - 1} = K_m/K_D \quad \text{(Equation 2.1)}$$

2.2.3 Rate dependence on pH

The effect of pH on V can be informative about the mechanism of an enzyme, especially in this case where the chemical reaction, rather than for example product release, is rate-controlling. Residues involved in acid or base catalysis, product release, substrate binding, or maintaining the structural conformation of an enzyme can be titrated by changing the pH of the assay conditions. Since, under steady-state conditions, the enzyme-substrate complex dominates, we assumed that V is not sensitive to substrate binding. Conformational changes may occur after both substrates are bound and can contribute to V. However, upon examination of STR crystal structures with and without substrates, the crystal structures appear to be the same, suggesting that this enzyme does not undergo any obvious enzyme conformational changes. Furthermore, the relatively large KIE measured on V led us to assume that V is primarily sensitive to the protonation states of groups involved in catalysis, and not product release. To verify this, STR was assayed with secologanin (at a concentration 50 fold above K_m) and tryptamine/[2-^{2}H]-tryptamine (concentration of 100 K_m) over a pH range of 3.10 – 9.19 at 30 °C (Fig. 2.5). The data best fit a diprotic model (equation 2.2), which was used to estimate the pKa values of two ionizable species to be 4.70 and 8.28. When the pH profiles of assays with

57

tryptamine and [2-^{2}H]-tryptamine were compared, a large primary isotope effect was still

observed from pH 4.6 to pH 9.19, indicating that the chemical step, and not product

release, is rate-controlling. The decrease in ^{D}V at acidic pH may be due to either rate-

controlling product release or another mechanism. This decrease prevents the exact

assignment of the pKa at acidic pH; however, the significant primary isotope effect that is

still observed indicates that the titrated residue is a catalytic species. Based on the pH

profile, we estimated the pKa of the residue to be 4.6. Loss of activity was observed at

extreme pH; however, circular dichroism (CD) spectroscopy suggested that the protein

secondary structure was unchanged at basic and acidic pH values, though CD

spectroscopy does not provide a detailed picture of enzyme structure.

Based on previously reported site-directed mutagenesis of the three ionizable residues

(Tyr151, His307, and Glu309) in the STR active site, we attributed the pKa of 4.60 to

Glu309, which was previously implicated to be critical for enzyme catalysis (a 900-fold

decrease in activity was observed when this residue was mutated to alanine) [23]. The

assignment of the residue responsible for the pKa of 8.28 was not clear. The pKa of the

hydroxyl group of tyrosine in solution is 10.2 [29]. However, even if the active site

environment modulated the pKa of this phenolic moiety to 8.28, mutation of Tyr151 to

phenylalanine resulted in only a small decrease in k_{cat} of 78 min^{-1} (Tyr151) to 58 min^{-1}

(Tyr151Phe), indicating that this residue is not involved in acid-base catalysis [23]. The

pKa of histidine in solution ranges from 5.5-7. Mutation of His307 to alanine resulted in

a 40-fold decrease in enzyme activity. This residue could be involved in catalysis, but the

mutation of Glu309 to alkaline resulted in a more significant decrease in rate (900-fold)

and, most importantly, the His307 is positioned further away from the primary amine of tryptamine (4 Å) and the aldehyde in secologanin (5 Å) to be directly involved in catalysis [23]. Based on the STR crystal structures, there are no other ionizable residues within the active site that are positioned near the reactive moieties of tryptamine and secologanin. Furthermore, there does not appear to be an ordered water molecule in the active site that could account for this pH dependence of the enzymatic rate. Although the pKa of tryptamine in solution is 10.2 [29, 30], the pKa value can shift in a hydrophobic environment such as an enzyme active site, within a range of 7.5-10.5 [30]. Therefore, the pKa of the basic species could be the tryptamine substrate. However, since the rate-controlling step occurs after cyclization (Fig. 2.1, step 4), the deprotonation of tryptamine may not be observed in the pH dependence of V. The iminium that forms during the reaction (Fig. 2.1, step 3) could also be the ionizable species displaying the pKa of 8.28. The pKa of an iminium is between 7 and 8 [31]. If the iminium is deprotonated, it is no longer electrophilic, and then cyclization cannot occur. Since structural analyses did not reveal any other obvious residues that have a basic pKa of 8.28, our best hypothesis is that the either the tryptamine substrate or iminium intermediate is the basic species observed in the pH rate profile.
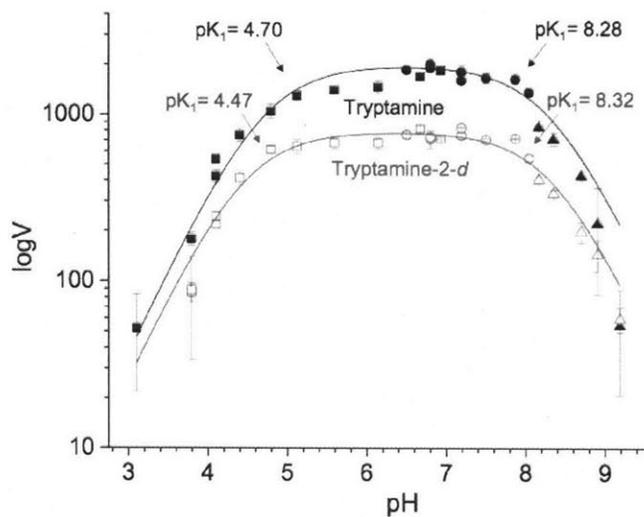
Figure 2.5 pH profile of enzymatic activity under saturating (V_{max}) substrate conditions. Tryptamine is represented by filled symbols; [2-^{2}H]-tryptamine is represented by open symbols. Circles, squares, triangles represent citrate, phosphate and borate buffer conditions, respectively. The error bars represent the 95 % confidence of the standard deviation of three experiments.

2.2.4 Co-crystallization of STR with an inhibitor

The Pictet-Spengler reaction of indole amine substrates has been proposed to proceed through a spiroindolenine intermediate resulting from attack from the C3-position of the indole (Fig. 2.1, step 4a and 4b) on to the carbon of the iminium, which can undergo a 1,2-alkyl shift and subsequent deprotonation (Fig. 2.1, step 5) to form product [9]. However, there is also evidence that a direct attack on the iminium carbon by C2 on the indole (Fig. 2.1, step 4) contributes to the mechanism [8, 32, 33]. It is not clear what mechanism is favored for the enzyme. Crystallographic studies could be used to demonstrate how the iminium is oriented relative to C2 and C3 during the cyclization step. However, the iminium intermediate is not stable enough for co-crystallization with *Rs*STR. To examine step 4 in detail, a potential transition state mimic (Fig. 2.6 A) was synthesized by Dr. Anne Friedrich via a reductive amination of the iminium intermediate formed during a Pictet-Spengler condensation between tryptamine and secologanin. This reductive amination product has a flexible sp^3 hybridized amine linkage between the indole and secologanin, which, once bound in the enzyme active site, could possibly model the conformation of the transition state leading from the iminium to the cyclized state in Fig. 2.1, step 4. When assayed with *Cr*STR (10 nM), tryptamine (20 µM), and secologanin (80 µM), the reductive amination product was found to be a potent inhibitor of *Cr*STR with an estimated IC_{50} of 3 ± 0.5 nM. The kinetic data was fit to a sigmoidal logistic curve for the estimation of IC_{50} values, and the error represents the standard deviation of three individual experiments.

The inhibitor readily co-crystallized with *Rs*STR, and a crystal structure of the complex was obtained at 3 Å resolution. Drs. Joachim Stöckigt, Santosh Panjikar, and Elke Loris completed this work. The best fit to the electron density suggested that the indole of the inhibitor did not match the orientation observed when tryptamine alone was bound to *Rs*STR (2FPB.pdb) (Fig. 2.6) [23]. The data demonstrate that tryptamine can adopt different binding modes whereas the secologanin moiety of the inhibitor overlaid well with the previously reported enzyme-secologanin complex (2FPC.pdb). The electron density of the inhibitor did not show a productive cyclization state since C2 (4.78 Å) and C3 (3.86 Å) of the inhibitor were not positioned close to the electrophilic iminium carbon (Fig. 2.6). We therefore concluded that the rigidity of the sp^2 hybridized iminium moiety appears to critical for productive binding that can lead to cyclization, and that the more flexible sp^3 hybridized amine of the inhibitor can adopt nonproductive conformations. Unfortunately, these structural studies were not able provide insight into the mechanism of the cyclization step (Fig. 2.1, step 4).
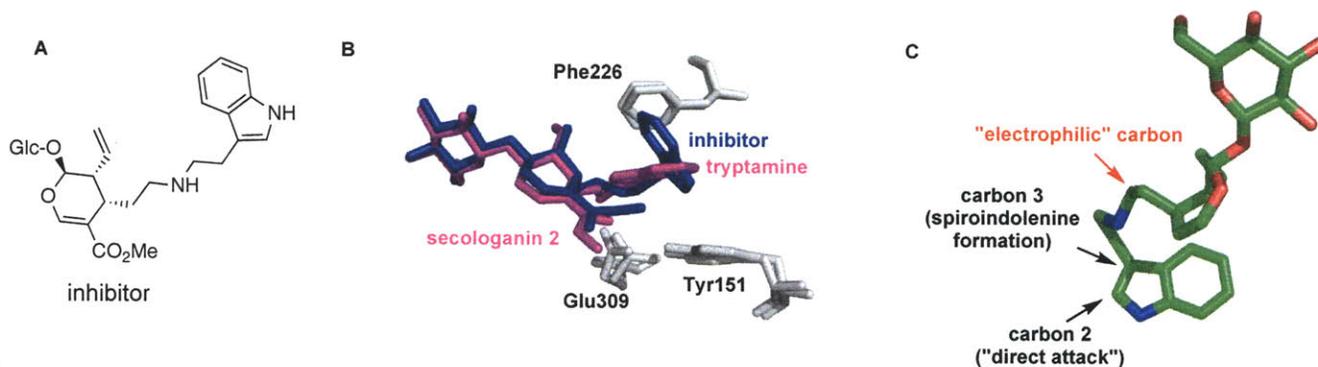
Figure 2.6. A) Chemical structure of the inhibitor. B) Overlay of the inhibitor (in blue) co-crystallized with *Rs*STR compared with *Rs*STR co-crystallized with tryptamine (in magenta, from 2FPB.pdb) and secologanin (in magenta, from 2FPC.pdb). The indole moiety of the inhibitor does not overlay with tryptamine. Surrounding amino acids are shown in gray for comparison. C) Detail of the structure of the inhibitor as it crystallized in the STR active site. The nucleophilic carbons 2 and 3 (highlighted by black arrows) are not positioned close to the electrophilic carbon of the iminium species (highlighted by red arrow). The methyl ester of the inhibitor is omitted for clarity.

2.2.5 Theoretical calculations

Since co-crystallization with a transition state mimic did not provide insight on the orientation of the iminium species in the cyclization step, we turned to theoretical modeling of transition states on the pathway to step 4 in Fig. 2.1. Dr. Baron Peters from the lab of Dr. Bernhardt Trout used *ab initio* calculations to model the transition states of the spiroindolenine and the 6-membered ring intermediates from the iminium species. The free energies of intermediates and transition states (Fig. 2.6) were computed relative to the trans form of the iminium species, the lowest calculated energy ground state structure. To simplify the time required for computational analysis, calculations were performed with the iminium species formed from the condensation of tryptamine and acetaldehyde instead of secologanin. Eight different transition states were found for the cyclization at both C2 (Fig. 2.1, step 4) and C3 of the indole (Fig. 2.1, step 4a). These transition states link either the trans- or cis-iminium species with 5- or 6-membered ring intermediates via attack of C2 or C3 of the indole with the methyl group (from acetaldehyde) oriented either "over" or "away" from the indole during the cyclization step (Fig. 2.7). Fig. 2.7 shows the free energies of the stable species that could potentially form along the reaction coordinate with the lowest energy transition states required to form this network of species. Some transition states of higher energies have been omitted because they provide redundant connections in the network. On the left axis, transition states can be accessed from the trans-iminium species and the corresponding reaction timescales are shown on the right axis.
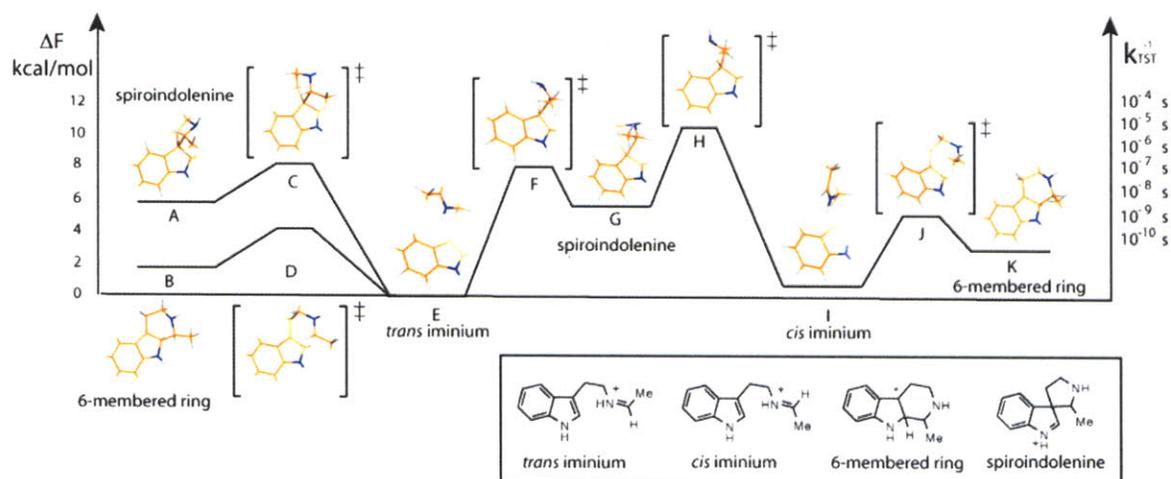
Figure 2.7 Free energy ΔF minima (represented by A, B, E, G, I, and K) and transition states (represented by C, D, F, H, and J) plotted to show the free energy landscape. Trans and cis refer to isomers of the iminium species. The timescales on the right correspond to the time for passage through the network from the trans-iminium state.

This computational analysis indicated that a direct attack from C2 of the indole onto the electrophilic carbon of the iminium (both cis and trans forms) to form the 6-membered ring intermediate is orders of magnitude faster than forming a 5-membered ring spiroindolenine. Previous mechanistic studies of the Pictet-Spengler reaction proposed that 1,2-alkyl shifts occurred after the formation of 5-membered ring/spiroindolenine intermediates to form the 6-membered ring intermediate, which is deprotonated to form product(s) [33]. However, searches for the transition state between the spiroindolenine and 6-membered ring intermediates repeatedly resulted in transition states leading in the reverse direction to re-form the iminium species, as evidenced by a shift in the electron density on nitrogen to the adjacent carbon to reform the sp^2-iminium. A transition state that invoked breaking the C-C (instead of the C-N bond) in the spiroindolenine to form the 6-membered ring transition state could be found in the computational search, but the energy of this transition state was calculated to be high (approximately 30 kcal mol^{-1}), suggesting that a 1,2-shift involving a C-C bond cleavage also does not occur. Therefore, the lower energy pathway is from the iminium to directly form the 6-membered ring intermediate. Even if the spiroindolenine formed, the 1,2-alkyl shift required to convert this intermediate to the strictosidine product would not occur.

2.3 Discussion

The Pictet-Spengler reaction, performed under acidic conditions, is widely used in the total synthesis of alkaloids. Interestingly, Nature has evolved CrSTR, an asymmetric enzyme catalyst, which at physiological pH forms (S)-strictosidine, the central biosynthetic intermediate for thousands of plant monoterpene indole alkaloids. This chapter describes the use of KIEs, rate dependence on pH, and crystallographic studies to provide insight on the mechanism by which STR catalyzes the Pictet-Spengler reaction.

Surprisingly, a primary isotope effect on V and V/K was observed when STR is assayed with tryptamine/[2-^{2}H]-tryptamine and secologanin, indicating that the final deprotonation is the rate-controlling step of the reaction (Fig. 2.1, step 5). This observation was initially difficult to rationalize, since loss of a proton at C2 on the indole results in rearomatization, which is expected to be an energetically favorable process. We propose a mechanistic rationale to address this observation. To regain the aromaticity of the indole that is lost during the cyclization step (Fig. 2.1, step 4), the indole can undergo either deprotonation (Fig. 2.1, step 5) or the reverse reaction to re-form the iminium (Fig. 2.1, step 3) intermediate [34]. If the rate of the reverse reaction to form the iminium is faster than the rate of deprotonation to form the strictosidine product, the final deprotonation step will be rate-controlling. Notably, Dr. Justin Maresh observed a primary KIE for the nonenzymatic Pictet-Spengler reaction using tryptamine and propanal as substrates [12], indicating that the rate-controlling step of the nonenzymatic Pictet-Spengler reaction is also the final deprotonation step.

In the STR-catalyzed reaction, the pH dependence of V (Fig. 2.5) revealed the pKa values of residues involved in acid-base catalysis. The acid-catalyzed step is involved in forming the iminium intermediate (Fig. 2.1, step 3). The base-catalyzed steps are involved in the initial deprotonation of the tryptamine substrate (Fig. 2.1, step 1) and the final deprotonation of the six-membered intermediate to form strictosidine (Fig. 2.1, step 5). The STR pH profile suggests that a residue with a pKa of 4.6 is involved in catalysis, which is close to the expected pKa of a carboxylate moiety. We propose that the Glu309 residue most likely acts as a general acid catalyst in STR-catalyzed reactions for three reasons: (1) the measured pKa of 4.6 is in the pKa range of an active-site glutamate, (2) the Glu309Ala mutant is less active by almost three orders of magnitudes (900-fold), indicating that Glu309 is a catalytic residue, and (3) the Glu309 residue is closely positioned to the aldehyde moiety and primary amine of secologanin and tryptamine, respectively, involved in catalysis. Conversely, for the basic pH dependence observed in the pH profile, we can only speculate which ionizable groups are involved in base-catalysis. Mutagenesis of Tyr151, an obvious candidate based on the pKa of the ionizable residue, does not cause a significant loss in activity when mutated to phenylalanine, suggesting that Tyr151 is not involved in catalysis. The only other possible species that can have pKa values close to 8.27 are the protonated tryptamine substrate or the iminium intermediate. By studying the pH dependence of V/K, the pKa of the free tryptamine substrate that the enzyme binds can be approximated, which would provide information that could possibly be used to rule out the possibility of tryptamine substrate being observed in the pH dependence of V. However, the low sensitivity of the HPLC assays at extreme pH and low substrate concentrations prevented these data from being measured.

Based on the position of Glu309 in the crystal structure, and the lack of other available ionizable active site residues, we also propose that Glu309 carries out the key general base functions; namely that Glu309 deprotonates tryptamine and also performs the final deprotonation step.

In the Pictet-Spengler reaction, the main driving force for cyclization is iminium formation (Fig. 2.1, step 3). This pH dependent process proceeds by the addition of an amine to a carbonyl to form a carbinolamine, followed by the loss of water to form an imine (Fig. 2.1, steps 2-3). Mechanistic studies of enzymatic imine (or iminium) formation have been previously reported for enzymes such as aldolases, acetoacetonate decarboxylase, pyridoxal and pyruvate containing decarboxylases, and dehydratases [35, 36]. Enzymes that catalyze iminium formation are proposed to have a catalytic residue, often a glutamic acid [36, 37] or an ordered water molecule [38] that protonates the carbinolamine to catalyze the formation of iminium species. In STR, there is no visible ordered active-site water molecule but Glu309 is positioned perfectly relative to the aldehyde and amine to be involved in an acid-catalyzed step such as iminium formation. Based on the pH dependence and structural data, we propose that Glu309 is involved in the acid-catalyzed step of iminium formation.

After iminium formation, co-crystallization studies of STR with an inhibitor suggested that the active site alone cannot orient the tryptamine in the correct position to undergo electrophilic addition since a non-productive conformation of the flexible transition state mimic in the STR crystal structure (Fig. 2.6) was observed. The sp^2 hybridized iminium

intermediate must also be properly positioned in a conformation that promotes cyclization to form the correct stereoisomer (Fig. 2.1, step 4). As a result of the non-productive conformation of the potential transition state mimic, we could not speculate if an attack by C2 or C3 of the indole onto the electrophilic carbon of iminium produced strictosidine in the STR active site. Other transition state analogs with different linkages between tryptamine and secologanin can be synthesized and co-crystallized with STR to establish a more vivid picture of the transition state to form (S)-strictosidine. Since co-crystallization studies with a bisubstrate inhibitor did not reveal a productive transition state, *ab initio* calculations were used to determine the energy landscape of a model Pictet-Spengler reaction. Calculated geometries of both theoretical intermediates and transition states, along with the corresponding free energy calculations suggested that the transition state to form the spiroindolenine (Fig. 2.1, step 4b) intermediate is higher in energy than the transition state on the pathway to forming the six-membered ring intermediate via a direct attack by C2 on the indole. Additionally, after forming the spiroindolenine intermediate, no transition state could be found for the 1,2-alkyl shift that is also proposed to occur to form the 6-membered ring intermediate, which is deprotonated to form product. Based on these computational studies, it appears that the spiroindolenine intermediate is non-productive.

After the cyclization step, we propose the newly deprotonated Glu309 could also catalyze the rate-limiting deprotonation step. Dr. Justin Maresh observed that a carboxylate ion can act as a base in the nonenzymatic Pictet-Spengler reaction [12]. Since there are no other ionizable residues in the active site and an ordered water molecule does not appear

to be present, the closely positioned Glu309 could deprotonate of the six-membered ring intermediate to form strictosidine.

To summarize, we propose a mechanism by which STR catalyzes a Pictet-Spengler reaction where tryptamine enters the active site in a protonated state (pKa 8.27) under the conditions of the assay (pH 7), and transfers its proton to Glu309, supplying the proton for general acid catalysis (Fig. 2.8). The primary amine of tryptamine becomes a better nucleophile to attack the aldehyde moiety of secologanin, resulting in rapid generation of a carbinolamine species (Fig. 2.8). Protonated Glu309 could then act as the general acid and protonate the carbinolamine, and catalyze the formation of the iminium species through loss of water (Fig. 2.8). During cyclization, we assume that the enzyme holds the transition state in the appropriate conformation to achieve the correct diastereomer. Lastly, Glu309 is positioned to deprotonate the 6-membered ring intermediate to form strictosidine.
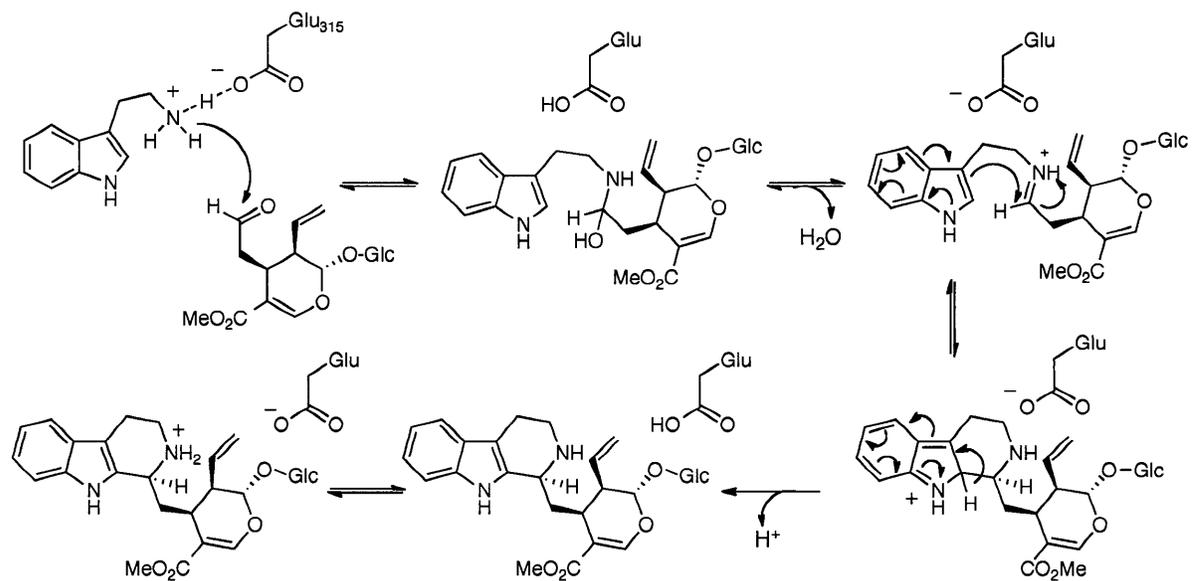
Figure 2.8 Proposed mechanism of STR acid/base catalysis involving Glu309 and protonated tryptamine (Numbering is from *C. roseus* enzyme).

The amino acid sequence and three-dimensional structure of an enzyme can reveal clues regarding the exact nature of the biochemical mechanism. Out of the four Pictet-Spenglerases, only the genes that encode norcoclaurine synthase (NCS) and STR are known. Consequently, we can only speculate about the mechanisms by which these two enzymes catalyze the Pictet-Spengler reaction. Interestingly, these enzymes do not share any sequence homology, suggesting that they may be convergently evolved from different ancestors [23]. Mechanistic studies on the Pictet-Spengler reaction catalyzed by NCS utilizing 4-hydroxyphenylaldehyde and dopamine/[3,5,6-^{2}H]-dopamine (Fig. 2.2 C) revealed that rearomatization of phenethylamine substrates to form tetrahydroisoquinolines is also rate-controlling, since a primary isotope effect was observed in this enzymatic system [11]. Therefore, although the enzyme sequences of NCS and STR are substantially different, the final deprotonation step is rate-controlling in both enzymatic mechanisms. However, since isotope effects were not measured for the nonenzymatic Pictet-Spengler reaction condensation that occurs between 4-hydroxyphenylaldehyde and dopamine/[3,5,6-^{2}H]-dopamine, we cannot speculate that the final deprotonation step of the nonenzymatic reaction is also rate-controlling.

In 2009, a crystal structure of NCS co-crystallized with both dopamine and a nonreactive 4-hydroxybenzaldehyde substrate mimic was solved at 2.1 Å resolution [39] (Fig. 2.9). The structural data revealed the presence of four ionizable residues in the active site: Lys122, Asp141, Glu110, and Tyr 108. Lys122 appears to be closely positioned (2.6 Å) to the carbonyl oxygen of the aldehyde moiety of 4-hydroxybenzaldehyde (Fig. 2.9). Upon mutating Lys122 to alanine, the enzyme became completely inactive, indicating

that NCS utilizes a different enzymatic mechanism than STR; no active site lysine residue is found in the STR structure. Further studies, such as mass spectrometry analysis of enzyme incubated with the 4-hydroxyphenylaldehyde substrate that has undergone sodium cyanoborohydride reduction to reduce possible linkages formed between 4-hydroxyphenylaldehyde and Lys122 or crystal structures of NCS and transition state mimics are required to illuminate the specific catalytic role of Lys122. Another closely positioned residue is Glu110, which is 2.7 Å away from the C5 atom of dopamine, which attacks the iminium to form the tetrahydroisoquinoline. When Glu110 is mutated to alanine the catalytic activity drops more than 80-fold, also suggesting that this residue is involved in catalysis. These data highlight a similarity between NCS and STR, as both enzymes have active site glutamates that appear to be involved in catalysis. Glu110 could catalyze a proton abstraction to form norcoclaurine, analogous to the role that Glu309 plays in STR. Understanding the mechanism of other Pictet-Spenglerases will reveal how Nature modifies its chemistry for catalyzing reactions with β-arylethylamine versus β-phenethylamine substrates.
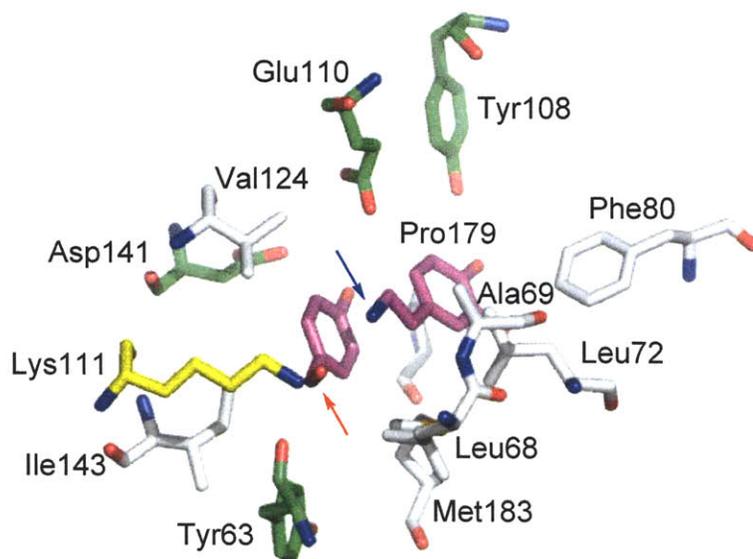
Figure 2.9 NCS *Thalictrum flavum* crystal structure (2VQ5.pdb) highlighting select amino acids that in the active site that binds dopamine and 4-hydroxyphenylaldehyde. Both substrates are shown in pink. The amine of dopamine is highlighted with a blue arrow, and the aldehyde of 4-hydroxyphenylaldehyde is highlighted with a red arrow. Surrounding polar residues are shown in green, and the key Lys111 residue is shown in yellow. Nonpolar residues are shown in gray.

2.4 Experimental Methods

2.2.1 Chemicals, general methods, and analytical techniques

The substrates [2-^{2}H]-tryptamine and tryptamine were synthesized by Dr. Justin Maresh

as previously reported [12]. The bisubstrate inhibitor was synthesized by Dr. Anne

Friedrich [12]. With the exception of secologanin, all other reagents were purchased from

Sigma-Aldrich. Secologanin was isolated as previously described [24]. HPLC separations

were carried out on a Beckman Coulter System Gold 125 HPLC equipped with a model

168 photodiode array detector, Hibar RT 250-4 LiChrosorb C18 column (Merck), and

home-built column heater set to maintain 30 °C.

UPLC and MS analyses were performed in tandem on an Acquity Ultra Performance

BEH C18 column with a 1.7 mm particle size, 2.1 x 100 mm dimension, which was

coupled to a Micromass LCT Premier TOF Mass Spectrometer by Waters Corporation

(Milford, MA) with electrospray ionization source. The capillary and sample cone

voltages were 3000 V and 30 V, respectively for MS analyses. The source and

desolvation temperatures were 100 °C and 300 °C, respectively. The cone and desolvation

gas flow were 60 and 800 L hr^{-1}.

2.2.2 Enzymatic assays

*Cr*STR was expressed and the enzyme activity was measured via an HPLC assay

employing previously reported conditions [24]. The concentration of *Cr*STR was

estimated from the known extinction coefficient at 280 nm. Enzyme (80 nM) was

incubated with secologanin in 100 mM sodium phosphate buffer, pH 7.0 at 30°C. An

internal standard, 1-naphthaleneacetic acid (NAA, 30 μM), was added to assays and was shown to have no effect on the rate of the reaction. Assays were initiated by the addition of tryptamine. ^{D}(V/K) values were measured with one substrate varied and the other substrate held at a concentration at least 50-fold higher than the K_m. The range of varied substrate concentrations spanned 2.5 orders of magnitude with several data points below the expected K_m. When the concentration of secologanin concentration was held at 2.5 mM, tryptamine was varied (1 μM-400 μM). Conversely, when tryptamine was held at 1 mM, secologanin concentrations was varied (3.13 μM-1.6 mM). The concentrations of tryptamine and secologanin were verified from their measured extinction coefficients. Extinction coefficients were measured by creating a standard curve on HPLC. The extinction coefficient for tryptamine at 280 nm was determined to be 3.345 x 10^{-7} AU M^{-1} mm^{-1}. The extinction coefficient of secologanin was determined by HPLC from partial conversion to an iminium by reaction of 2,4-dinitrophenyl hydrazide, and the iminium product was found to have an extinction coefficient of 4.064 x 10^{-7} AU M^{-1} mm^{-1} in pH 5 acetic acid buffer.

Assays were quenched by the addition of 0.1 volume equivalents of 2.0 M NaOH to denature the enzyme and prevent any background Pictet-Spengler reaction. Under these quench conditions, the methyl ester of both secologanin and strictosidine were hydrolyzed and remaining side products were not observed. Initial rates were obtained from five time points. Quenched aliquots (74 μL) of the reaction were directly injected onto an analytical HPLC using a solvent gradient of 22 % to 67 % acetonitrile in 0.1 % aqueous trifluoroacetic acid. The absorbance of tryptamine, hydrolyzed strictosidine, and

NAA was measured at 228 nm and 280 nm. Peak areas of hydrolyzed strictosidine were integrated and normalized to the internal standard. Initial rates were determined from the slope of the linear fit of five normalized data points for each concentration of varied substrate. Kinetic parameters were estimated from the initial rate data direct fit to appropriate forms of the Michealis-Menten equation [40] by non-linear least squares fit using Origin 7.0 software (OriginLab Corp, Northampton, MA). The accuracy of initial rates is diminished at concentrations significantly below the K_m for tryptamine because of the low sensitivity of the HPLC assay. As a result, there is greater error in reported V/K values than in the V values. The isotope effects on V and V/K were determined by dividing the V or V/K values obtained from the Origin software for natural abundance tryptamine by the values obtained for $[2-^2H]$-tryptamine. The errors were estimated from the 95 % confidence of the standard deviations of three experiments. Although the isotopic substitution of $[2-^2H]$-tryptamine was 90.42 %, reported KIE values were not corrected since the applied isotopic corrections for the direct measurements in this study were not found to be significant.

2.2.3 pH rate profiles

The effect of pH on enzymatic activity was observed by incubating secologanin (2.5 mM), 5 nM STR, NAA internal standard, and buffer (varying from pH 3.1-9.7) at 30°C. The assays were initiated by the addition of tryptamine or $[2-^2H]$-tryptamine (0.3 mM). Two different buffer systems were used to control for potential buffer effects on enzyme activity. The effective pKa (pKa') for the experimental conditions was calculated for each buffering agent from the Debye-Hückel equation as described by Ellis [41] using A = 0.5161 for 30 °C [42]. A set of constant composition buffers with a constant ionic

strength (I = 0.31) were prepared from a mixed buffer system of acetic acid (pKa' =

4.64), Tris (pKa' = 6.32), and Bis-Tris (pKa' = 8.00) following the procedure of Ellis

[41] ranging pH 3.7 to 8.76 (pH at 30 °C: 3.70, 4.20, 4.62, 5.06, 5.52, 5.89, 6.89, 6.37,

6.83, 7.29, 7.81, 8.27, 8.76) and ionic strength of 0.2. Alternatively, overlapping constant

ionic strength buffers were also used: citric acid (pKa'$_1$ = 2.97, pKa'$_2$ = 4.61, pKa'$_3$ =

6.25) ranging pH 3.1 to 9.46 (pH at 30 °C: 3.10, 3.42, 3.79, 4.10, 4.40, 4.79, 5.12, 5.59,

6.14, 6.67, and 6.93), potassium phosphate buffers (pKa'$_2$ = 6.72) covered an overlapping

pH range of 6.50 to 8.04 (pH at 30 °C: 6.50, 6.80, 7.19, 7.50, 7.87, and 8.04), and borate

(pKa' = 9.04) buffers were used for assays over a pH range of 8.17 to 9.19 (pH at 30 °C:

8.17, 8.35, 8.70, 8.92, and 9.19).

Initial rates from the slope of the linear fit of five data points were plotted on a

logarithmic scale. Using Origin 7.0 software, the data best fit a diprotic model shown in

equation 2.2. Both buffer conditions yielded equivalent results; identical pKa values were

calculated from both buffer systems.

$$V = \frac{V_{max}}{1 + \dfrac{H}{K_1} + \dfrac{K_2}{H}} \qquad \text{(equation 2.2)} \quad \cdot$$

2.2.4. Circular dichroism measurements

CD measurements were obtained on an Aviv Instruments, Inc. Circular Dichroism

Spectrometer Model 202 from 195-250 nm. Each wavelength step was 1 nm and the

averaging time for each wavelength was 6 seconds. Three CD spectra were obtained and

averaged for each sample at 30 °C. Scans of pure solvent were used for a baseline, which

was subtracted from the protein measurements. All samples were degassed. Spectra were obtained of STR (14 μM) in a 50 mM phosphate buffer with 10 % glycerol at 30 °C. The pH was adjusted to pH 3 or pH 9 (as evidenced by pH paper) by adding either concentrated HCl or NaOH to the assay.

2.2.5 *Ab initio* calculations

Geometry optimizations and vibrational frequencies were computed by Dr. Baron Peters using the 6-31G* basis set and the PW1PW91 density functional provided by Gaussian 03 software (Gaussian, Inc., Wallingford, CT) [43] as previously described [12].

2.2.6 Inhibition kinetics of *Cr*STR with an inhibitor

The reductive amination product of the iminim formed during the Pictet-Spengler condensation of tryptamine and secologanin was determined to be a potent inhibitor by Dr. Anne Freidrich. Inhibition kinetics were measured by incubating *Cr*STR (10 nM) with tryptamine (20 μM), secologanin (80 μM), varying concentrations of the inhibitor, and NAA (60 μM), in pH 7.0 sodium phosphate buffer (100 mM) at 30 °C as previously reported [12].

2.2.7 Co-crystallization of *Rs*STR with an inhibitor

Using previously described procedures, the *Rs*STR gene was expressed by Drs. Joachim Stöckigt, Santosh Panjikar, and Elke Noris [44]. The X-ray data were collected using synchrotron radiation at the X11 beamline of the European Molecular Biology Laboratory in Hamburg, Germany. The complete data were collected to 3.0 Å resolution. The data were processed as previously described [12] using DENZO and scaled using

SCALEPACK [45]. The data collection and refinement statistics are shown in the Table 2.1. The atomic coordinates have been deposited into the Protein Data Bank (2VAQ.pdb).

Table 2.1 Data collection and refinement statistics of the STR-inhibitor complex

wavelength (Å)	0.8148
space group	R3
Unit Cell (Å)	
	a=b=150.0
	c=121.7
total reflections	119018
unique reflections	20319
mosaicity	1.0
resolution (Å)	20-3.00
completeness (%)	100 (100)[a]
$I/\sigma(I)$	18.1 (3.6)
R_{merge} (%)[b]	10.1 (52.9)
resolution (Å)	20-3.00
R_{cryst} / R_{free} (%)[c]	18.7/24.7
average B (Å) for protein	48.6
average B (Å) for inhibitor	52.4
average B (Å) for water	30.5
Number of atoms	
non-hydrogen	4896
water	14
RMSD	
bond (Å)	0.019
angles (deg)	1.94

[a]The values in parentheses correspond to the last resolution shell.

[b]$R_{merge} = \Sigma_{hkl}\Sigma_i|I_i$ (hkl) $- < I(hkl) >|/\Sigma_{hkl}\Sigma_i< I_i(hkl) >$, where $< I(hkl) >$ is the average intensity over symmetry equivalent reflections.

[c]R_{cryst} (R_{free}) $= \Sigma_{hkl}||F_o(hkl)| -|F_c(hkl)||/\Sigma_{hkl}|F_o(hkl)|$, where F_o and F_c are observed and calculated structure factors, respectively.

2.5 References

1. Cox, E.D. and J.M. Cook, *The Pictet-Spengler condensation: A new direction for an old reaction.* Chem. Rev., 1995. **95**(6): p. 1797-1842.

2. Taylor, M.S. and E.N. Jacobsen, *Highly enantioselective catalytic acyl-Pictet-Spengler reactions.* J. Am. Chem. Soc., 2004. **126**(34): p. 10558-10559.

3. Seayad, J., A.M. Seayad, and B. List, *Catalytic asymmetric Pictet-Spengler reaction.* J. Am. Chem. Soc., 2006. **128**(4): p. 1086-1087.

4. Pictet, A. and T. Spengler, *Formation of isoquinoline derivatives by the action of methylal on phenylethylamine, phenylalanine and tyrosine.* Ber. Dtsch. Chem. Ges., 1911. **44**: p. 2030-2036.

5. Anslyn, E.V. and D.A. Dougherty, *Modern physical organic chemistry* 2006, Sausalito: University Science Books. 1104.

6. Bailey, P.D., *Direct proof of the involvement of a spiro intermediate in the Pictet-Spengler reaction.* J. Chem. Res., Synopses, 1987: p. 202-203.

7. Ibaceta-Lizana, J.S., et al., *Electrophilic substitution in indoles. 13. The synthesis and rearrangement of 2-deuteriospiro[cyclopentane-3'-indolenine].* J. Chem. Soc. Perkin Trans. II, 1987(9): p. 1221-1226.

8. Kowalski, P. and J.L. Mokrosz, *Structure and spectral properties of β-carbolines. Part 9. New arguments against direct rearrangement of the spiroindolenine intermediate into β-carboline system in the Pictet-Spengler cyclization. An MNDO approach.* Bulletin des Societes Chimiques Belges, 1997. **106**(3): p. 147-149.

9.   Ungemach, F. and J.M. Cook, *The spiroindolenine intermediate. A review.* Heterocycles, 1978. **9**: p. 1089-1119.

10.   De-Eknamkul, W., N. Suttipanta, and T.M. Kutchan, *Purification and characterization of deacetylipecoside synthase from Alangium lamarckii Thw.* Phytochemistry, 2000. **55**(2): p. 177-181.

11.   Luk, L.Y.P., et al., *Mechanistic studies on norcoclaurine synthase of benzylisoquinoline alkaloid biosynthesis: An enzymatic Pictet-Spengler reaction.* Biochemistry, 2007. **46**(35): p. 10153-10161.

12.   Maresh, J.J., et al., *Strictosidine synthase: mechanism of a Pictet-Spengler catalyzing enzyme.* J. Am. Chem. Soc., 2008. **130**(2): p. 710-723.

13.   Samanani, N., D.K. Liscombe, and P.J. Facchini, *Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis.* Plant J., 2004. **40**(2): p. 302-313.

14.   Minami, H., et al., *Functional analysis of norcoclaurine synthase in Coptis japonica.* J. Biol. Chem., 2007. **282**(9): p. 6274-6282.

15.   Mizukami, H., et al., *Purification and properties of strictosidine synthase from Catharanthus roseus cultured cells.* Biochemistry, 1979. **18**(17): p. 3760-3763.

16.   Treimer, J.F. and M.H. Zenk, *Purification and properties of strictosidine synthase, the key enzyme in indole alkaloid formation.* Eur. J. Biochem., 1979. **101**(1): p. 225-233.

17. McKnight, T.D., et al., *Nucleotide sequence of a cDNA encoding the vacuolar protein strictosidine synthase from Catharanthus roseus.* Nuc. Acids Res., 1990. **18**(16): p. 4939.

18. Kutchan, T.M., et al., *The cDNA clone for strictosidine synthase from Rauvolfia serpentina. DNA sequence determination and expression in Escherichia coli.* FEBS Lett., 1988. **237**(1-2): p. 40-44.

19. van Der Heijden, R., et al., *The Catharanthus alkaloids: pharmacognosy and biotechnology.* Curr. Med. Chem., 2004. **11**(5): p. 607-628.

20. O'Connor, S.E. and J. Maresh, *Chemistry and biology of monoterpene indole alkaloid biosynthesis.* Nat. Prod. Rep., 2006. **23**(4): p. 532-547.

21. Pfitzner, U. and M.H. Zenk, *Homogeneous strictosidine synthase isoenzymes from cell suspension cultures of Catharanthus roseus.* Planta Med., 1989. **55**(6): p. 525-530.

22. de Waal, A., A.H. Meijer, and R. Verpoorte, *Strictosidine synthase from Catharanthus roseus: purification and characterization of multiple forms.* Biochem. J., 1995. **306**(Pt 2): p. 571-580.

23. Ma, X., et al., *The structure of Rauvolfia serpentina strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins.* The Plant cell, 2006. **18**(4): p. 907-920.

24. McCoy, E., M.C. Galan, and S.E. O'Connor, *Substrate specificity of strictosidine synthase.* Bioorg. Med. Chem. Lett., 2006. **16**(9): p. 2475-2478.

25. Parkin, D.W., *Methods for the determination of competitive and noncompetitive kinetic isotope effects,* ed. P.F. Cook 1991, Boca Raton: CRC Press. 269-290.

26.   Northrop, D.B., *Minimal kinetic mechanism and general equation for deuterium isotope effects on enzymic reactions: uncertainty of detecting a rate-limiting step.* Biochemistry, 1981. **20**(14): p. 4056-4061.

27.   Cook, P.F. and W.W. Cleland, *Mechanistic deductions from isotope effects in multireactant enzyme mechanisms.* Biochemistry, 1981. **20**(7): p. 1790-1796.

28.   Klinman, J.P. and R.G. Matthews, *Calculation of substrate dissociation constants from steady-state isotope effects in enzyme-catalyzed reactions.* J. Am. Chem. Soc., 1985. **107**(4): p. 1058-1060.

29.   Perrin, D.D., *Dissociation constants of organic bases in aqueous solution* 1965, London: Butterworths.

30.   Cleland, W.W., *Advances in enzymology and related areas of molecular biology*, ed. A. Meister. Vol. 45. 1977, New York: John Wiley & Sons. 320-348.

31.   Frey, P.A. and A.D. Hegeman, *Enzymatic reaction mechanisms* 2007, New York: Oxford University Press.

32.   Casnati, G., A. Dossena, and A. Pochini, *Electrophilic substitution in indoles. Direct attack at the 2-position of 3-alkylindoles.* Tetrahedron Lett., 1972. **13**(52): p. 5277-5280.

33.   Kowalski, P., A.J. Bojarski, and J.L. Mokrosz, *Structure and spectral properties of β-carbolines. 8. Mechanism of the Pictet-Spengler cyclization: an MNDO approach.* Tetrahedron, 1995. **51**(9): p. 2737-2742.

34.   Zollinger, H., *Hydrogen isotope effects in aromatic substitution reactions.*, in *Advances in Physical Organic Chemistry*, V. Gold, Editor 1964, Academic Press: London. p. 163-196.

85

35.   Hupe, D.J., *Chapter 8 Enzyme reactions involving imine formation.* New Compr. Biochem., 1984. **6**: p. 271-301.

36.   Lorentzen, E., et al., *Mechanism of the Schiff base forming fructose-1, 6-bisphosphate aldolase: structural analysis of reaction intermediates.* Biochemistry, 2005. **44**(11): p. 4222-4229.

37.   St-Jean, M., et al., *High resolution reaction intermediates of rabbit muscle fructose-1,6-bisphosphate aldolase: substrate cleavage and induced fit.* J. Biol. Chem., 2005. **280**(29): p. 27262-27270.

38.   Fullerton, S.W.B., et al., *Mechanism of the Class I KDPG aldolase.* Bioorg. Med. Chem., 2006. **14**(9): p. 3002-3010.

39.   Llari, A., et al., *Structural basis of enzymatic (S)-norcoclaurine biosynthesis.* J. Biol. Chem., 2009. **284**(2): p. 897-904.

40.   Northrop, D.B., *Fitting enzyme-kinetic data to V/K.* Anal. Biochem., 1983. **132**(2): p. 457-461.

41.   Ellis, K.J. and J.F. Morrison, *Buffers of constant ionic strength for studying pH-dependent processes.* Meth. Enz., 1982. **87**: p. 390-405.

42.   Holtzhauer, M., *Basic methods for the biochemical lab* 2006, Springer-Verlag: Heidelberg, Germany.

43.   Tomasi, J. and M. Persico, *Molecular interactions in solution: An overview of methods based on continuous distributions of the solvent.* Chem. Rev., 1994. **94**(7): p. 2027-2094.

44.   Ma, X.Y., et al., *Crystallization and preliminary X-ray crystallographic analysis of strictosidine synthase from Rauvolfia: the first member of a novel enzyme family* Biochimica et Biophysica Acta, 2004. **1702**(1): p. 121-124.

45.   Otwinowski, Z. and W. Minor, *Processing of X-ray diffraction data collected in oscillation mode.* Methods Enzymol., 1997. **276**: p. 307-326.

CHAPTER 3

REENGINEERING STRICTOSIDINE SYNTHASE: CONVERSION OF A PICTET-

SPENGLERASE INTO A HYDROLASE

3.1 Introduction

Biocatalysts are used for a wide range of practical applications, including the generation of energy sources, bioser sor applications, food preservatives, starch processing, diagnostics, therapeutics, and biomaterials [1-3]. Natural enzymes are often used as biocatalysts since these enzymes can catalyze up to almost 10^{20}-fold rate accelerations with high regio- and stereoselectivity [4]. While the construction of a protein *de novo*, with a defined catalytic function is the holy grail of protein engineering, it is important to note that an exhaustive search of all sequence space for even a small protein with new catalytic activity is impossible. Considering that there are 20 different amino acids for each position in the protein, the sequence space of a 100 residue protein, which is a relatively small protein, is 20^{100} [5]. Current strategies sidestep these limitations by utilizing experimental and computational methods to narrow the sequence space when designing new proteins [1]. Directed evolution is a highly successful experimental method, which mimics Darwinian evolution in a test tube to create a protein with a unique property without knowing any structural or mechanistic information. Iterative rounds of random mutagenesis and/or gene recombination, followed by high throughput screening or selection for a desired property, and gene amplification are performed until a desired property or function is found [6].

When sequence, mechanistic, and/or structural information for a protein is known, rational design can be a viable strategy to create enzymes with altered properties [6, 7]. Given the relatively small number of structural scaffolds compared to the numerous diverse catalytic functions found in nature [8-10], methods to "mix and match" proteins

with similar folds can be a powerful "rational" strategy to create new catalytic activities. The resulting "hybrid", which for the purposes of this chapter is defined as an enzyme with structural elements from more than one enzyme, can be constructed in a number of different ways. Single-point mutations based on the structures of existing or homologous enzymes [11-13], introduction of secondary-structural elements [14, 15], domain-swapping [16-19], and fusions [20] between two enzymes with different activities have each been used to construct hybrid enzymes [7, 10, 20]. These rational design approaches have been utilized extensively in the study and design of a number of proteins; the design of zinc fingers with novel DNA recognition [21-23] and transcription factors [24, 25] have been particularly successful.

Parallels can be drawn between the process of creating hybrid enzymes and the natural evolutionary mechanisms that generate new enzymes. Presumably, when enzymes within the same superfamily evolve in nature, new catalytic activities arise from changes in catalytic motifs while a similar protein scaffold is maintained [26]. Enzymes involved in natural product biosynthesis provide a rich source for understanding the evolution of new enzyme functions from pre-existing scaffolds since many enzymes of secondary metabolism are speculated to be derived from primary metabolism [27-29]. For example, Trapp et al conducted a phylogenetic analysis of the amino acid sequences of 33 plant terpene synthases and found that terpene synthases involved in primary metabolism, such as kaurene synthase B and copalyl diphosphate synthase, have become specialized by gene duplication and the loss of introns and domains to produce a superfamily of terpene synthases that are involved in secondary metabolism [28]. Additionally, new and diverse

catalytic activities within enzyme superfamilies such as class I and II terpene synthases have been speculated to have emerge from evolutionary conserved protein folds [8]. To explore whether homologo is enzymes evolve from pre-existing scaffolds, this chapter describes the introduction o a different catalytic activity into the highly conserved β-propeller scaffold in the lactonohydrolase/paraoxonase superfamily. While most of the enzymes within this superfamily are believed to catalyze a metal-dependent hydrolysis reaction, a small subset of these enzymes catalyze a metal-independent Pictet-Spengler reaction [30]. This superfamily appeared to be an excellent system for exploring how Nature could use a single protein scaffold to catalyze two distinct chemical reactions. Moreover, we envisioned that by exploring protein design efforts with this superfamily, we could potentially better understand whether or not enzymes that catalyze the Pictet-Spengler reaction evolved from preexisting hydrolases.

The lactonohydrolase/paraoxonase superfamily is composed of lactonases, hydrolases, and Pictet-Spenglerases found in bacteria, mammals, and plants. With over 2,500 sequences, the majority of the members of this large superfamily have yet to be functionally uncharacterized [31]. Five representative members of this superfamily, however, have been functionally and structurally characterized: human senescence marker protein 30 [32, 33], drug responsive protein-35 [34], diisopropyl fluorophosphatase [35, 36], paraoxonase (PON1) [37, 38] and strictosidine synthase (STR) [39, 40]. Human senescence marker protein 30 catalyzes a calcium/zinc-dependant hydrolysis of various carbohydrate lactones and organophosphates [33]. Similarly, drug responsive protein-35 from *Staphylococcus aureus* possesses calcium-dependent

lactonase activity [34]. Diisopropyl fluorophosphatase from the squid head ganglion of *Loligo vulgaris* catalyzes calcium-dependent hydrolysis of organophosphates whereas the calcium-dependent human PON1 also catalyzes the hydrolysis organophosphates, as well as aryl esters and lactones (Fig. 3.1 A). STR, described in Chapter 2, catalyzes a Pictet-Spengler condensation reaction between secologanin and tryptamine to form strictosidine, a central biosynthetic intermediate for all monoterpene indole alkaloids [41-43] (Fig. 3.1 B). Protein from cognate cDNAs for STR from *Catharanthus roseus* [43], *Rauvolfia Serpentina* [44], and *Ophiorriza pumila* [45] have been functionally characterized. Moreover, recent structural data indicate that all of five representative members of the lactonohydrolase/paraoxonase superfamily, all with varying degrees of sequence homology, share a common 6-bladed β–propeller fold (select examples shown in Fig. 3.1 C) [33, 35, 37, 39, 46]. Despite these structural similarities, STR seems to be an outlier as this enzyme catalyzes a metal-independent stereoselective Pictet-Spengler condensation, a reaction fundamentally different from the metal-dependent hydrolase chemistry exhibited by the other known family members.

Members of the lactonohydrolase/paraoxonase superfamily, often annotated as STR homologs, have been found in many species, such as *Vitis vinifera* and *Arabidopsis thaliana* plants, that do not produce strictosidine-derived alkaloids [47, 48]. Since metabolic analysis of the host plants suggest that it is unlikely that these STR homologs exhibit Pictet-Spengler activity, we and other researchers have speculated that these homologs instead exhibit the lactonohydrolase or paraoxonase activity displayed by the other known members of this superfamily [30]. Moreover, these STR homologs have also

been found in mammals [49, 50]. Notably, the human homolog of STR, C20orf3 does not exhibit Pictet-Spengler activity [51], but does display weak calcium-dependent hydrolase activity [52]. Ilhan et al reported that C20orf3 isolated from liver (HepG2) and pancreatic (Rin-5F) cells hydrolyzes phenyl acetate and β-napthyl acetate. Furthermore, sequence analysis indicates that STR and its homologs are weakly similar to the calcium-dependent hydrolase PON1 [37, 38, 53]. Using the crystal structure of PON1,

We converted *C. roseus* STR into a calcium-dependent hydrolase by incorporating calcium-binding residues. Additionally, we introduced three spheres of mutations in STR based on conserved residues from other STR homologs that are speculated to catalyze hydrolysis. Remarkably, one of the strictosidine synthase homologs from *V. vinifera*, which is homologous to STR, was functionally characterized as a hydrolase. The structural similarities and the capacity to mutate one catalytic activity to another suggest an evolutionary link between these highly distinct classes of enzymes, and illustrate the plasticity of the β-propeller scaffold.
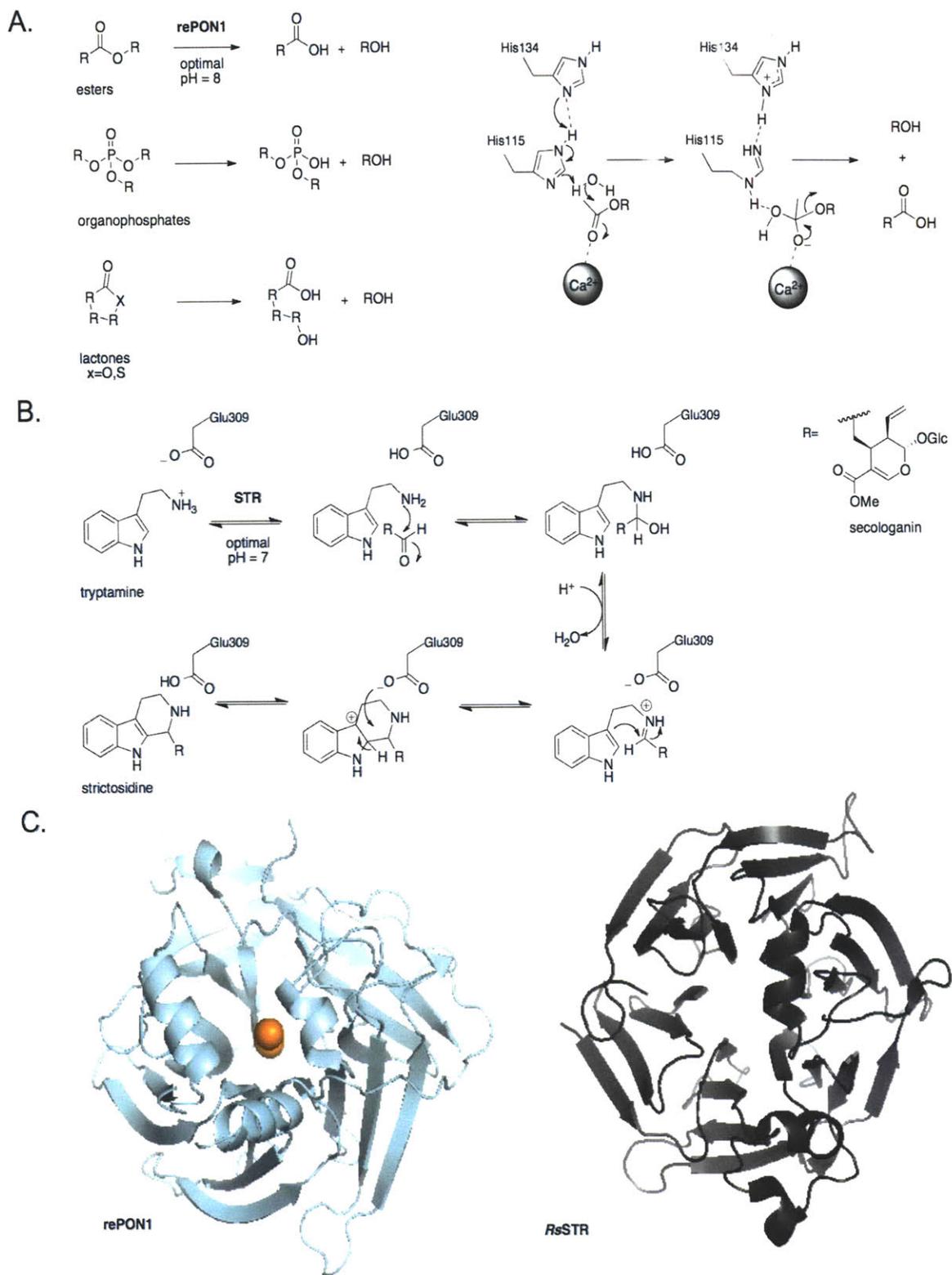
Figure 3.1 A) Mechanism of human serum PON1. B) Mechanism of *C. roseus* STR. C) Crystal structures of recombinant PON1 (1V04.pdb) with calcium ions shown in orange and *Rauvolfia serpentina* STR (2FP8.pdb).

3.2 Results

3.2.1 Design of CrSTR hybrid proteins

To rigorously explore the relationship among the STR superfamily members, Michael Hicks, a graduate student in the laboratory of Professor Patsy Babbitt at UCSF, performed a global computational sequence comparison of proteins that belong to the lactonohydrolase/paraoxonase superfamily. A large subset of proteins, including the small number of characterized hydrolases, was found to contain conserved metal-binding residues [30]. Moreover, the sequences of homologs from *Vitis vinifera* and *Arabidopsis thaliana* that also contained these metal-binding residues were found to be closely related to STR. With this knowledge, we collaborated with graduate student Summer Thyme from the laboratory of Professor David Baker at the University of Washington to explore whether the residues responsible for hydrolase activity (and metal-binding) of closely related STR homologs could be incorporated into STR. We predicted that these substitutions would result in a change of activity from a Pictet-Spenglerase to a hydrolase. We chose to use amino acid sequences of closely related STR homologs that do not display Pictet-Spenglerase activity instead of using the PON1 sequence; the sequence of the structurally characterized PON1 is too distant to be used effectively to map mutations onto STR (12 % sequence identity). Also, PON1 had to be evolved before reproducible heterologous expression could be obtained [54]. Several of the hydrolases in this superfamily are insoluble or unstable *in vitro*; hence, using STR as a template seemed to be the more productive starting point to attempt to convert the Pictet-Spenglerase into a hydrolase, rather than to begin with a hydrolase and convert it to a Pictet-Spenglerase.

Thyme and Baker used BLASTp [55] to mine the GenBank database for STR homologs that were closest in sequence to validated Pictet-Spenglerases from *C. roseus* (*Cr*STR), *Rauvolfia serpentina*, and *Ophiorrhiza pumila*. Homologs from *Vitis vinifera* (CAO67974.1, CAN77945.1, CAO67963.1, CAO66499.1) and *A. thaliana* (NP_177542.1), proteins of notable similarity (37-49 % sequence identity), appeared to be closest in sequence identity. The close similarity of STR with the STR homolog, CAN77945.1, from *V. vinifera* was also noted in Hick's and Babbitt's computational network analysis [30]. Extensive metabolome analysis of the *V. vinifera* and *A. thaliana* plants indicated that no strictosidine or strictosidine-derived alkaloids are produced, suggesting these homologs do not function as Pictet-Spenglerases. Moreover, STR requires a key glutamate residue for catalysis [39, 56], and this key catalytic residue is missing in each of these genes. Sequence analysis of these *A. thaliana* and *V. vinifera* homologs indicated that there were conserved residues that were not found in any of the validated Pictet-Spenglerases (Fig. 3.2). Notably, the five residues that bind the catalytic calcium in PON1 align with the putative metal-binding residues in *A. thaliana* and *V. vinifera* homologs (Fig. 3.2). Summer Thyme also created homology models of the *V. vinifera* homologs (CAO67974.1, CAN77945.1), which were based on the crystal structures of STR from *Rauvolfia serpentina* (*Rs*STR, 2FP8.pdb) and evolved PON1 (1V04.pdb). These homology models were used to determine which residues in *Cr*STR should be mutated to the conserved residues amongst the homologs in order to introduce hydrolytic activity and metal-binding. Summer Thyme made three spheres of mutations extending outward from the putative active site to create three new hybrid proteins named protein 1 (P1), protein 2 (P2), and protein 3 (P3) (Fig. 3.3). Wild type *Cr*STR already

contains one of the conserved metal-binding residues (Fig. 3.2). The P1 hybrid contained

four metal-binding residues, which are conserved amongst the *A. thaliana* and *V. vinifera*

STR homologs (Fig. 3.3). P2 and P3 *Cr*STR hybrids contained five metal-binding

residues since PON1 and the STR homolog CAO66499.1 contain a fifth metal-binding

residue (Fig. 3.2).

```
emb|AAR95984.1|  rePON1           -MAKLIALTLLGMGLALFRNHQSSYQTRLNALREVQPVELPNCNLVKGIETGSEDMEILP  59
emb|CAO67964.1|  V. vinifera      ------------------------------------------LPTP------TIGPEALAFDC  15
emb|CAN77945.1|  V. vinifera      ------------------------------------------LPSP------TVGPEAIAFDY  15
emb|CAO67963.1|  V. vinifera      ------------------------------------------LPPN------VTGPEALAFDR  15
ref|NP_177542.1|A. thaliana       ------------------------------------------VPET------RSGPEAFAFDS  15
emb|CAO66499.1|  V. vinifera      ------------------------------------------LPSR------VSGPESIAFDC  15
emb|X62334.1|   R. serpentina STR MAKLSDSQTMA--LFTVFLLFLSSSLALSS--PILKEILIEAP------SYAPNSFTFDS  50
emb|X61932.1|     C. roseus   STR MANFSESKSMMAVFFMFFLLLLSSSSSSSSSSSPILKKIFIESP------SYAPNAFTFDS  54


emb|AAR95984.1|  rePON1           NGLAFISSGLKYPGIKSFNPNSPGKILLMDLNEEDPTVLELGITGSKFDVSSFNPHGIST 119
emb|CAO67964.1|  V. vinifera      SGAGPYAS-VADGRVLKWQAESAGFVDFTVASPSRSKQLCDGSSDPAKEPTCGRPLGIG-  73
emb|CAN77945.1|  V. vinifera      TGAGPYAS-VADGRVLKWLDASAGFVDFAFISPSRSKKLCDGSTDPALEPTCGRPLGLG-  73
emb|CAO67963.1|  V. vinifera      LGGGPYTG-VSDGRVLKYGGPSAGFTDFAYTTPTRSKAVCDGTTDPNSGPTCGRPLGVG-  73
ref|NP_177542.1|A. thaliana       TGKGFYTG-VSGGKILKYL-PETGYVDFAQITESSNSSWCDGTIGTALAGRCGRPAGIA-  72
emb|CAO66499.1|  V. vinifera      NGDGPYTG-ISDGRILKWQGSKHGWKEFAITSPFRIPKFCDGSLNPAMEQVCGRPLGLK-  73
emb|X62334.1|   R. serpentina STR TNKGFYTS-VQDGRVIKYEGPNSGFVDFAYASPYWNKAFCENSTDAEKRPLCGRTYDIS- 108
emb|X61932.1|     C. roseus   STR TDKGFYTS-VQDGRVIKYEGPNSGFTDFAYASPFWNKAFCENSTDPEKRPLCGRTYDIS- 112


emb|AAR95984.1|  rePON1           FTDEDNAMYLLVVNHPDAKSTVELFKFQEEEKSLLHLKTIRHKLLPNLNDIVAVGPEHFY 179
emb|CAO67964.1|  V. vinifera      FNNKTGDLYIADAYYG---LFVVGPDGGRATQLATEAEGVPFRFLNAVDVDQETGIVYFT 130
emb|CAN77945.1|  V. vinifera      FNYRTVDLYIADAYHG---LNVVGPKDGRIIQLATAAEGVPFLFLNAVDVDQETGIVYFT 130
emb|CAO67963.1|  V. vinifera      FNNLTGQLYIADAYSG---LLVVGSNGGLATPVATTAEGVPFTRFNGLDVDQLTGNVYFT 130
ref|NP_177542.1|A. thaliana       FNEKTGDLYVADAPLG---LHVISPAGGLATKITDSVDGKPFKFLDGLDVDPTTGVVYFT 129
emb|CAO66499.1|  V. vinifera      FNEAKCDLYIADAYFG---LLVVGRNGGVAKQVAISAEGVPFRFTNALDIDQNTGVVYFT 130
emb|X62334.1|   R. serpentina STR YNLQNNQLYIVDCYYH---LSVVGSEGGHATQLATSVDGVPFKWLYAVTVDQRTGIVYFT 165
emb|X61932.1|     C. roseus   STR YDYKNSQMYIVDGHYH---LCVVGKEGGYATQLATSVQGVPFKWLYAVTVDQRTGIVYFT 169


emb|AAR95984.1|  rePON1           GTNDHYFLDPYLQSWEMYLGLAWSYVVYYSPS--EVRVVAEGFDFANGINISPDGKYVYI 237
emb|CAO67964.1|  V. vinifera      DASARFQRR-EFQN-AVLAGDMTGRLMKYDPRTKQVTVLLRGLGLAVGVAINKDGSFVLV 188
emb|CAN77945.1|  V. vinifera      DASARFQRR-EFML-AVQTGDMTGRLMKYDPRTQEVTVLLRGLGGAGGVTISKDGSFILV 188
emb|CAO67963.1|  V. vinifera      DASSVYELR-YYRFYSIENNDASGRLLKYDPSTKQVTVLIRGLSGPAGAAVSRDGSFVLV 189
ref|NP_177542.1|A. thaliana       SFSSRFSP--IQVLIALGLKDATGKLYKYDPSTKVVTVLMEGLSGSAGCAVSSDGSFVLV 187
emb|CAO66499.1|  V. vinifera      DTSTIFQRW-AYAI-AMQTGDKTGRLLKYDPRSKEVTVLLRGLSFSNGVALSKDKDFVLV 188
emb|X62334.1|   R. serpentina STR DVSTLYDD--RGVQQIMDTSDKTGRLIKYDPSTKETTLLLKELHVPGGAEVSADSSFVLV 223
emb|X61932.1|     C. roseus   STR DVSSIHDDSPEGVEEIMNTSDRTGRLMKYDPSTKETTLLLKELHVPGGAEISADGSFVVV 229


emb|AAR95984.1|  rePON1           AELLAHKIHVYEKHANWTLTPLKSLDFNTLVDNISVDPETGDLWVGCHPNGMKIFFYDSE 297
emb|CAO67964.1|  V. vinifera      SEFIATRIQRYWLRGPKANTSELFLKPTGTPDNIKRN-ARGEFWVAAN----IGAEMAAA 243
emb|CAN77945.1|  V. vinifera      TEFVTNRIQRFWLKGPKANTSELFLKPPGTPDNIKRN-VRGEFWVAVN----IGAG-TAV 242
emb|CAO67963.1|  V. vinifera      SEFIANRTQKFWLRGPKANTSELFFTFQGRPDNIKTS-ITDTFWVAVN----IGKSVPTT 244
ref|NP_177542.1|A. thaliana       SQFTKSNIKRYWIKGPKAGSSEDFTNSVSNPDNIKRIGSTGNFWVASV----VNKIIVPT 243
emb|CAO66499.1|  V. vinifera      TETTTAKVTRYWLQGQKSQLSDTFTQLVGCPDNIQRN-IHGEFWAQN----NCGR-PEL 242
emb|X62334.1|   R. serpentina STR AEFLSHQIVKYWLEGPKKGTAEVLVKIPN-PGNIKRN-ADGHFWVSSSEE-LDGNMHGRV 280
emb|X61932.1|     C. roseus   STR AEFLSNRIVKYWLEGPKKGSAEFLVTIPN-PGNIKRN-SDGHFWVSSSEE-LDGGQHGRV 286


emb|AAR95984.1|  rePON1           NPPASEVLRIQNILTEEPKVTQVYAENGTVLQGSTVASVYKGKLLIGTVFHKALYCEL-- 355
emb|CAO67964.1|  V. vinifera      APLGLRVSEEGKILQVVAFDTGDITRTISEVHEYN-GALYVGSLALPFVGVFTF------ 296
emb|CAN77945.1|  V. vinifera      LPSGLRLSEEGKVLQVVAFGTGDIPKTISEVQEYY-RALYIGSLALPFVGVYRS------ 295
emb|CAO67963.1|  V. vinifera      VPTGQRMDAHGNVLQTVNFEAEYGSTMISEVQGRGEIFLYVGSRDASYVGVYT------- 297
ref|NP_177542.1|A. thaliana       NPSAVKVNSNGEVLQTIPLKDKFGDTLLSEVNEFE-GNLYIGTLTGPFAGILKL------ 296
emb|CAO66499.1|  V. vinifera      KVISIKLNKEGKIMEELSEDFG----PLSEVQEKD-NDLWLGSVLLSYIGMLN------- 290
emb|X62334.1|   R. serpentina STR DPKGIKFDEFGNILEVIPLPPPFAGEHFEQIQEHD-GLLYIGTLFHGSVGILVYD---KK 336
emb|X61932.1|     C. roseus   STR VSRGIKFDGFGNILQVIPLPPPYEGEHFEQIQEHD-GLLYIGSLFHSSVGILVYDDHDNK 345


emb|AAR95984.1|  rePON1           --------
emb|CAO67964.1|  V. vinifera      --------
emb|CAN77945.1|  V. vinifera      --------
emb|CAO67963.1|  V. vinifera      --------
ref|NP_177542.1|A. thaliana       --------
emb|CAO66499.1|  V. vinifera      --------
emb|X62334.1|   R. serpentina  STR GNSFVSSH 344
emb|X61932.1|     C. roseus    STR GNSYVSS- 352
```

Figure 3.2 Amino acid sequence alignment of recombinant PON1 and *A. thaliana* and *V. vinifera* STR homologs with validated Pictet-Spenglerases, *Cr*STR and *Rs*STR. Sphere 1 (blue), sphere 2 (purple), and sphere 3 (green) mutations used to generate P1, P2, and P3 are annotated. Putative metal-binding residues are in red and known calcium-binding residues in PON1 are in bolded red. The catalytic glutamate involved in STR Pictet-Spengler condensation is highlighted in yellow. The G258 insertion that was made in sphere 2 mutations is not listed. The alignment was generated by ClustalW [46].
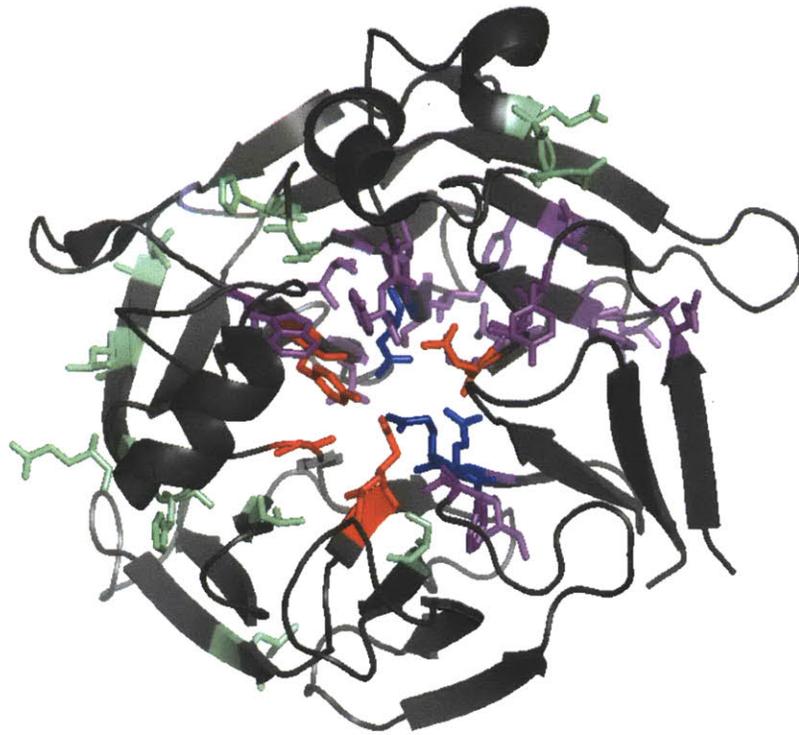
Figure 3.3 *Rauvolfia serpentina* STR (2FP8.pdb) with 3 spheres of mutations stemming outward from the active site. The spheres are annotated in different colors (sphere 1 = blue, sphere 2 = purple, sphere 3 = green). The putative metal-binding residues in sphere 1 are shown in red. One of the five metal-binding residues already exists in wild type *C. roseus*, three metal-binding residues were introduced in sphere 1 mutations, and the fifth metal-binding residue was introduced in sphere 2 mutations.

3.2.2 Cloning, expression, and purification of CrSTR hybrids and STR homolog from V. vinifera

Dr. Peter Bernhardt used a synthetic gene assembly procedure to create hybrid proteins P1, P2, and P3, and cloned each of these genes into a pET28b (+) vector for protein expression in *Escherichia coli* BL21 DE3 cells. For positive controls that exhibit hydrolase activity, the closely related STR homolog with putative hydrolase activity from *V. vinifera* (CAN77945.1) was cloned into pET32b (+) for expression in *E. coli* Rosetta DE3 cells, and the PON1 G2E6 variant was cloned into pET32b (+) for expression in *E. coli* Origami DE3 cells. Tawfik and coworkers [57] have shown that the pET32 vector facilitates the expression of the relatively insoluble PON1 protein in *E. coli*. While recombinant Pictet-Spenglerases have generally been produced in good yields [40, 58], the STR homolog from *V. vinifera* was difficult to express, similar to paraoxonases [59]. After lysis, most of the expressed protein from *V. vinifera* remained associated with the *E. coli* cell pellet. To maximize protein yield, the STR homolog *V. vinifera* and the three CrSTR hybrid proteins were solubilized in 0.1 % tergitol and purified from *E. coli* using affinity chromatography with all buffers containing calcium (1 mM). CrSTR was purified in the same manner to serve as a negative control for hydrolase activity. For a positive control, the G2E6 variant of PON1 was expressed and purified according to Aharoni et al [57]. Protein purity was assessed by SDS-PAGE analysis (Fig. 3.4).

3.2.3 Activity and substrate specificity assays of CrSTR hybrids and STR homolog (CAN77945.1) from V. vinifera

The appropriate substrate to assess the hydrolase activity of the purified proteins was difficult to choose; the physiological substrate(s) of PON1 is not known with certainty, and the substrates of the putative hydrolases in human, V. vinifera, and other species have not been investigated. The proteins were assayed with p-nitrophenyl acetate (pNPAc) for two reasons: (1) PON1 accepts this substrate and (2) hydrolysis of this substrate produces a yellow color that can be readily and quantitatively measured at 405 nm on a UV-visible spectrophotometer [38]. Biochemical assays of the STR homolog from V. vinifera (CAN77945.1) and CrSTR hybrids with pNPAc (0.5 mM, 1 mM, 1.6 mM, 2.1 mM, 2.6 mM, and 3.2 mM) gratifyingly resulted in the accumulation 4-nitrophenol (Fig. 3.5 A). While background hydrolysis of pNPAc at pH 8 is observed, STR homolog from V. vinifera and CrSTR hybrids displayed rates of hydrolysis above background levels. P1, P2, and P3 hybrids exhibited similar rates of hydrolysis. Variations in the rates of hydrolysis for P1 prevented a quantitative analysis of the differences in rates between P1, P2, and P3 hybrid proteins. Assays of CrSTR and purified extracts of E. coli expressing empty vectors [pET28a (+) and pET32b (+)] with pNPAc served as negative controls and displayed negligible hydrolase activity with levels similar to the background pNPAc hydrolysis. PON1 was also assayed with pNPAc as a positive control for hydrolase activity (Fig. 3.5 B) [38]. PON1 shows almost 100-fold higher activity with pNPAc than P3, which is expected since PON1 has been subjected to extensive directed evolution to improve its solubility and catalytic performance [56].

To explore the substrate scope of the STR homolog from *V. vinifera*, which we now refer to as the *V. vinifera* hydrolase, and the most active *Cr*STR hybrid P3, these proteins were assayed with 5-thiobutyl butyrolactone (TBBL) and phenyl acetate substrates. However, the initial rates of hydrolysis were equal to the rates of spontaneous hydrolysis of these substrates. A positive control, PON1, was assayed with TBBL and phenyl acetate substrates and hydrolysis products were observed [38, 60], suggesting that the STR homolog from *V. vinifera* and P3 do not share a similar substrate scope with PON1 and other substrates need to be explored with this subset of enzymes. P3 and the *V. vinifera* hydrolase were also assayed for Pictet-Spenglerase activity but these enzymes failed to convert tryptamine and secologanin to strictosidine. LC-MS analysis indicated that strictosidine (m/z 531) was only produced in assays with *Cr*STR (Fig. 3.6). Although P1 and P2 are not shown in Fig. 3.6, these hybrids also did not produce strictosidine as evidenced by LC-MS analysis.
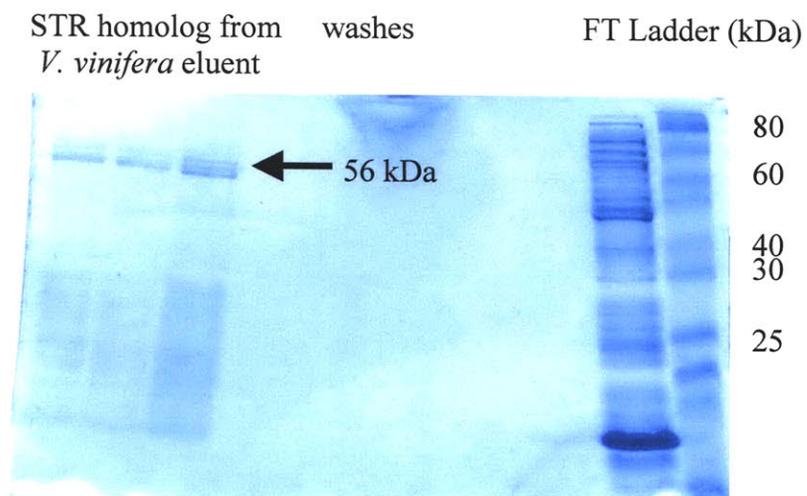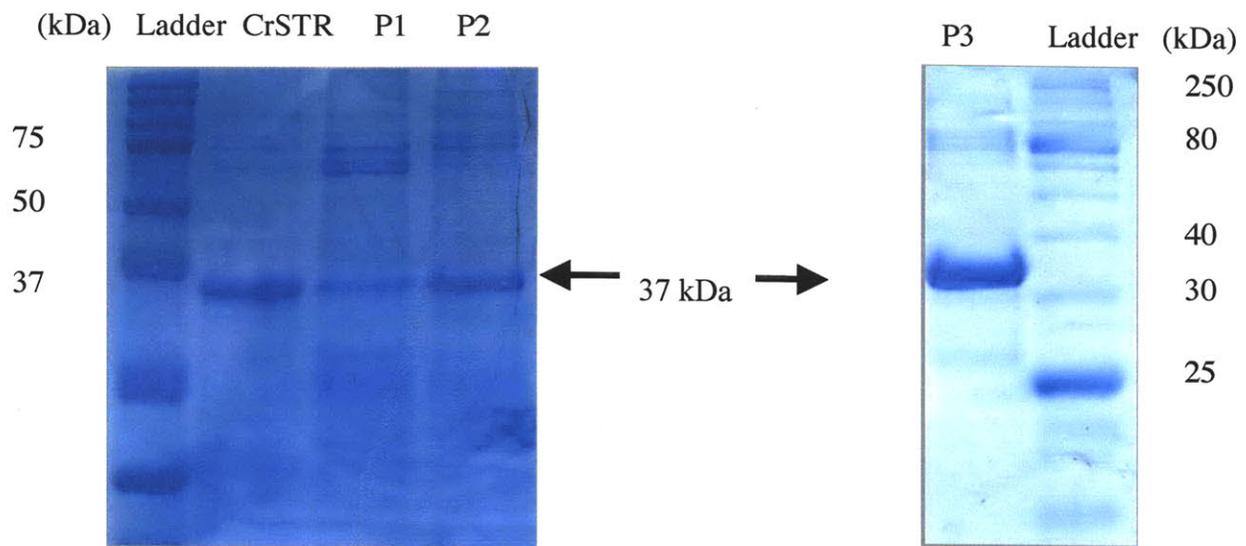
Figure 3.4 SDS-PAGE gels of purified *Cr*STR, P1, P2, P3, and the STR homolog from *V. vinifera*.

A.
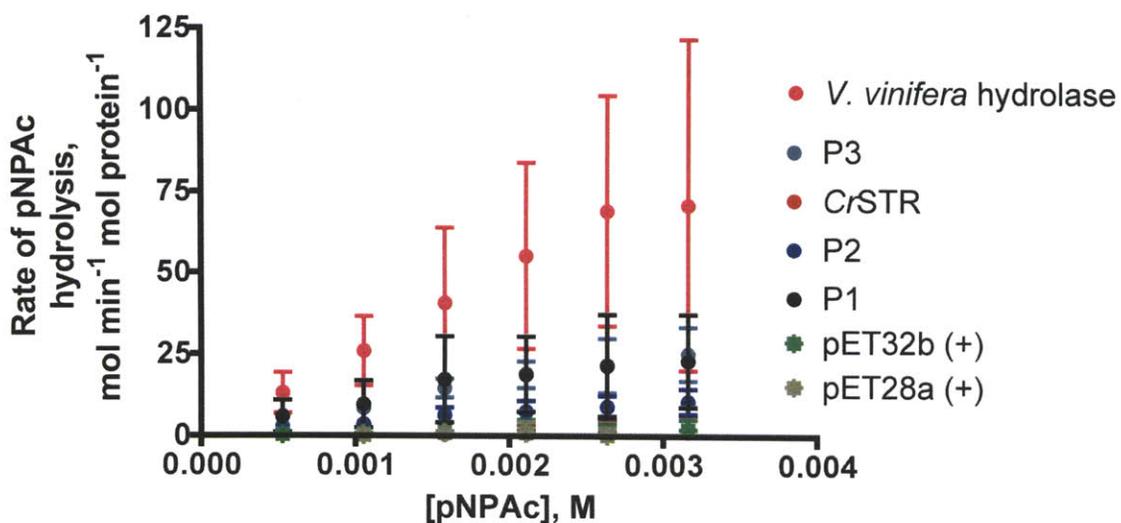
B.
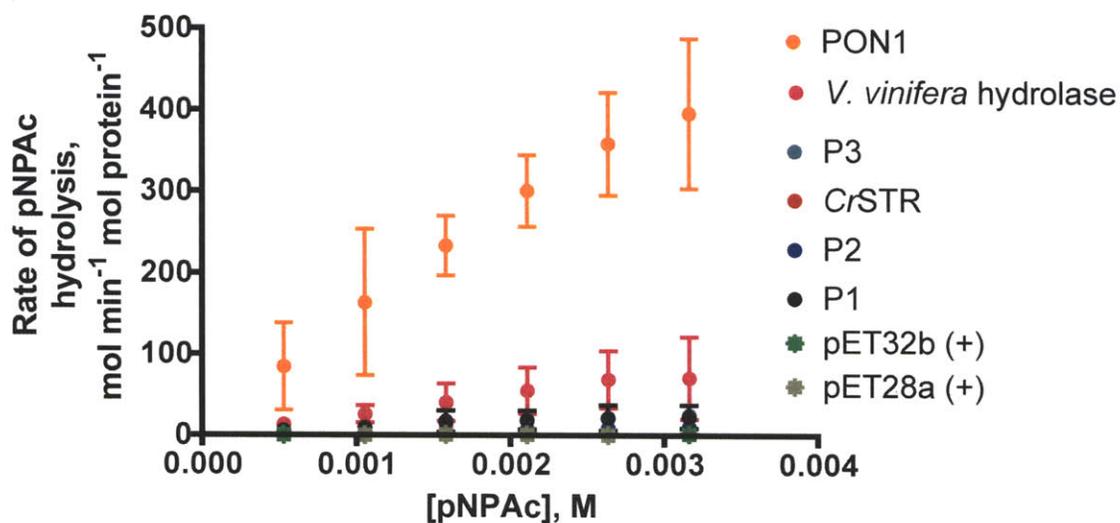
Figure 3.5 A) Hydrolysis of pNPAc by STR homolog from *V. vinifera* (*V. vinifera* hydrolase), P1, P2, P3, empty pET32b (+), empty pET28a (+), and *Cr*STR at pH 8. B) Hydrolysis of pNPAc by PON1 at pH 8 compared to *V. vinifera* hydrolase, P1, P2, P3, and *Cr*STR. The rate of background pNPAc hydrolysis has been subtracted from rates obtained for each concentration of substrate. The error bars represent the 95 % confidence of the standard deviation from three experiments.
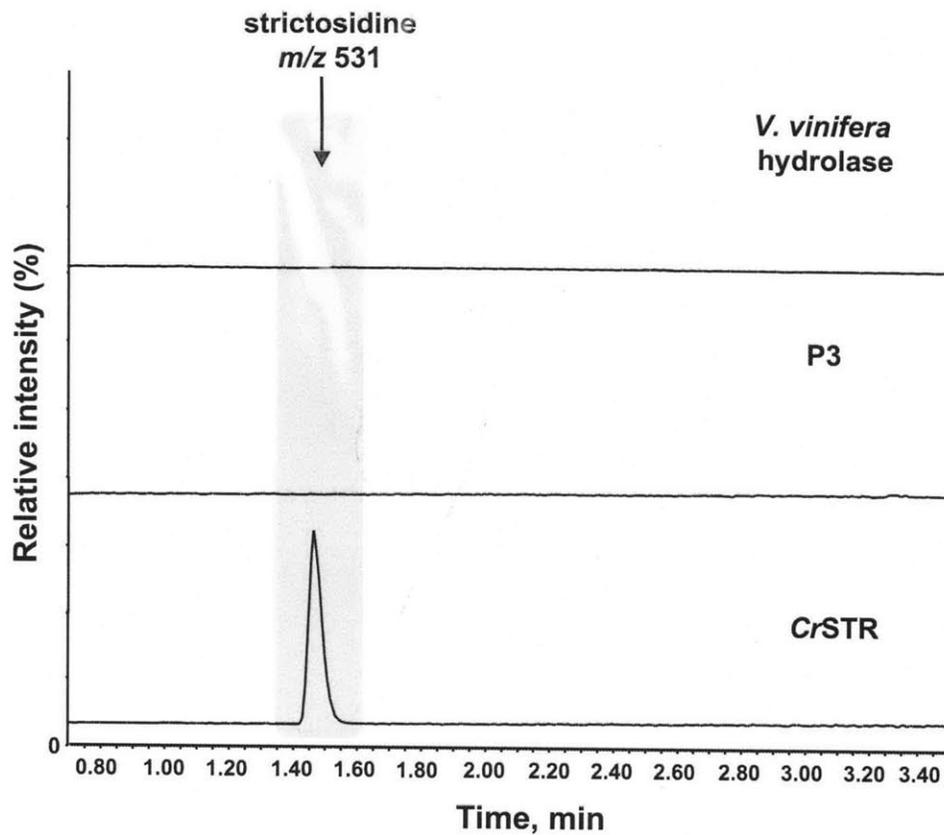
Figure 3.6 LC-MS selected ion chromatograms of strictosidine (*m/z* 531) produced by assays of tryptamine and secologanin incubated with either *V. vinifera* hydrolase, P3, or *Cr*STR at pH 7.

3.2.4 Steady-state kinetics of CrSTR hybrids and V. vinifera hydrolase

The high K_m values for CrSTR hybrids and the V. vinifera hydrolase with pNPAc (Fig. 3.5), along with the limited substrate solubility of pNPAc, precluded the measurement of initial rates in the concentration range of 2- or 3-fold greater than the apparent pNPAc K_m. The data did not facilitate an acceptable fit of the initial rate data to the Michealis-Menten equation. As a result, steady-state kinetic constants for all proteins were estimated by fitting the data to the Lineweaver-Burk plot. Steady-state kinetic analyses of the V. vinifera hydrolase assayed with pNPAc revealed that this enzyme exhibited a V_{max} of 220 ± 140 mol min^{-1} mol protein^{-1}, K_m of 8 ± 3 mM, and V_{max}/K_m of 29 ± 20 mol min^{-1} mol protein^{-1} M^{-1}. The best representative CrSTR hybrid hydrolase, P3, was assayed with pNPAc and estimated to have a V_{max} of 90 ± 40 mol min^{-1} mol protein^{-1}, K_m of 20 ± 9 mM, and V_{max}/K_m of 5 ± 4 mol min^{-1} mol protein^{-1} mM^{-1}. Steady-state analyses of PON1 were also conducted to compare the kinetic parameters to other hydrolases. PON1 was assayed with pNPAc and this enzyme exhibited a V_{max} of 640 ± 160 mol min^{-1} mol protein^{-1}, K_m of 3 ± 0.6 mM, and V_{max}/K_m of 250 ± 90 mol min^{-1} mol protein^{-1} mM^{-1}. The errors were derived from the 95 % confidence of the standard deviation of two experiments from assays with the V. vinifera hydrolase, P3, and PON1.

3.2.5 Assessing metal-binding capabilities of CrSTR hybrids

P1, P2, and P3 CrSTR hybrids were predicted to contain a calcium metal-binding site (Fig. 3.2). Based on previous mechanistic studies on PON1, we hypothesized that calcium was required for P1, P2, and P3 hydrolase activity [53, 57]. To verify this, P3, the hybrid that demonstrated the most consistent expression and activity profiles, was

exchanged into a buffer containing 10 mM EDTA and assayed with varying

concentrations of pNPAc in 100 mM pH 8 HEPES buffer containing 96 mM NaCl and 10

mM EDTA (Fig. 3.7 A). A reduction in hydrolase activity comparable to the rate of

background pNPAc hydrolysis was observed when P3 was assayed in the presence of 10

mM EDTA, suggesting that hydrolysis is metal-dependent (Fig. 3.7 A). Additionally,

when 4 mM calcium inhibitor, terbium chloride, was added to assays with P3, a decrease

of almost 80 % in hydrolase activity was observed (Fig. 3.7 B), also suggesting that P3

binds calcium.

A.
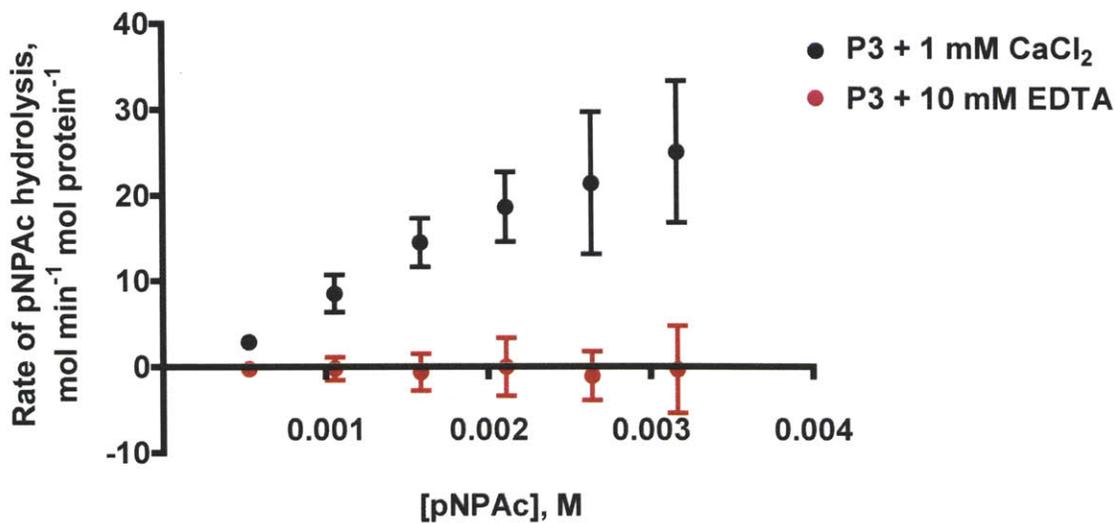
B.
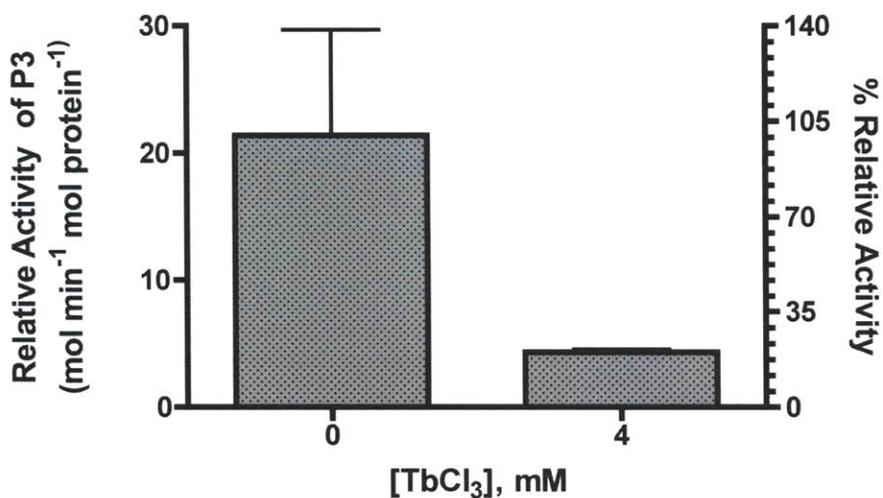
Figure 3.7 A) P3 hydrolysis of pNPAc in the presence of 1 mM $CaCl_2$ (black circles) and 10 mM EDTA (red circles). The rate of background pNPAc hydrolysis has been subtracted from these data. Red and black error bars represent the 95 % confidence of the standard deviation of 2 and 3 experiments, respectively. B) Inhibition of P3 by 4 mM calcium inhibitor, terbium chloride, in the presence of 1mM $CaCl_2$ and 2.6 mM pNPAc at pH 8. Error bars represent the 95 % confidence of the standard deviation of 3 experiments.

To validate that the five metal-binding residues in P3 (E17, N128, N189, D234, N234) coordinate the calcium required for the hydrolysis of pNPAc as predicted, site-directed mutagenesis was used to mutate these residues to alanine residues. However, when the E17A/N189A/D234A/N234A mutant was assayed with pNPAc, the same rate of hydrolase activity was observed compared to wild type P3 hydrolase activity. In human serum PON1, single mutations to convert metal-binding residues (E52A, D168A, or D268A) to alanines resulted in a more than 95 % decrease in hydrolase activity compared to wild type activity when these mutants were incubated with substrates such as phenyl acetate and paraoxon [53]. Moreover, site-directed mutagenesis was used to generate a E75A/D286A/N287A mutant of the *V. vinifera* hydrolase. This triple mutant also displayed levels of hydrolysis activity similar to wild type levels, suggesting that other residues in *Cr*STR hybrids and the *V. vinifera* hydrolase are involved in metal-binding.

3.2.6 Mutations to introduce Pictet-Spenglerase activity into the *V. vinifera* hydrolase

With the successful conversion of *Cr*STR into a hydrolase based upon conserved residues in non-Pictet-Spenglerase STR homologs, the same strategy was used to attempt to convert the *V. vinifera* hydrolase into a Pictet-Spenglerase. A synthetic gene was obtained such that the conserved residues in the *V. vinifera* hydrolase were converted to those residues in *Cr*STR, which included the catalytic glutamate required for Pictet-Spenglerase activity. The gene was cloned into pET32b (+) and transformed into *E. coli* Rosetta DE3 cells for protein expression. After protein purification, the mutant was incubated with tryptamine and secologanin to assay for Pictet-Spenglerase activity. LC-MS analysis indicated that no strictosidine (m/z 531) product (Fig. 3.8) was formed.

In light of the failed attempts to eliminate hydrolysis activity in the *V. vinifera* hydrolase

by mutating putative metal-binding residues systematically to alanine, the *V. vinifera*

mutant that does not catalyze the Pictet-Spengler reaction was assayed with pNPAc to

confirm that there are other potential metal-binding residues involved in catalyzing

hydrolysis. Assays containing 53 nM of the *V. vinifera* mutant in 100 mM HEPES buffer

containing 1 mM $CaCl_2$ and 2.6 mM NaCl were initiated by the addition of pNPAc (0.5

mM, 1 mM, 1.6 mM, 2.1 mM, 2.6 mM, and 3.2 mM). Surprisingly, the *V. vinifera* mutant

hydrolyzed pNPAc at a rate of hydrolysis slightly faster than the wild type *V. vinifera*

hydrolase (Fig. 3.9). Since *Cr*STR lacks hydrolase activity we assumed that residues

mutated to convert the *V. vinifera* hydrolase to *Cr*STR would not participate in

hydrolysis. However, these results further support the need for additional residues to be

mutated in order to identify the metal-binding residues involved in catalyzing hydrolysis.
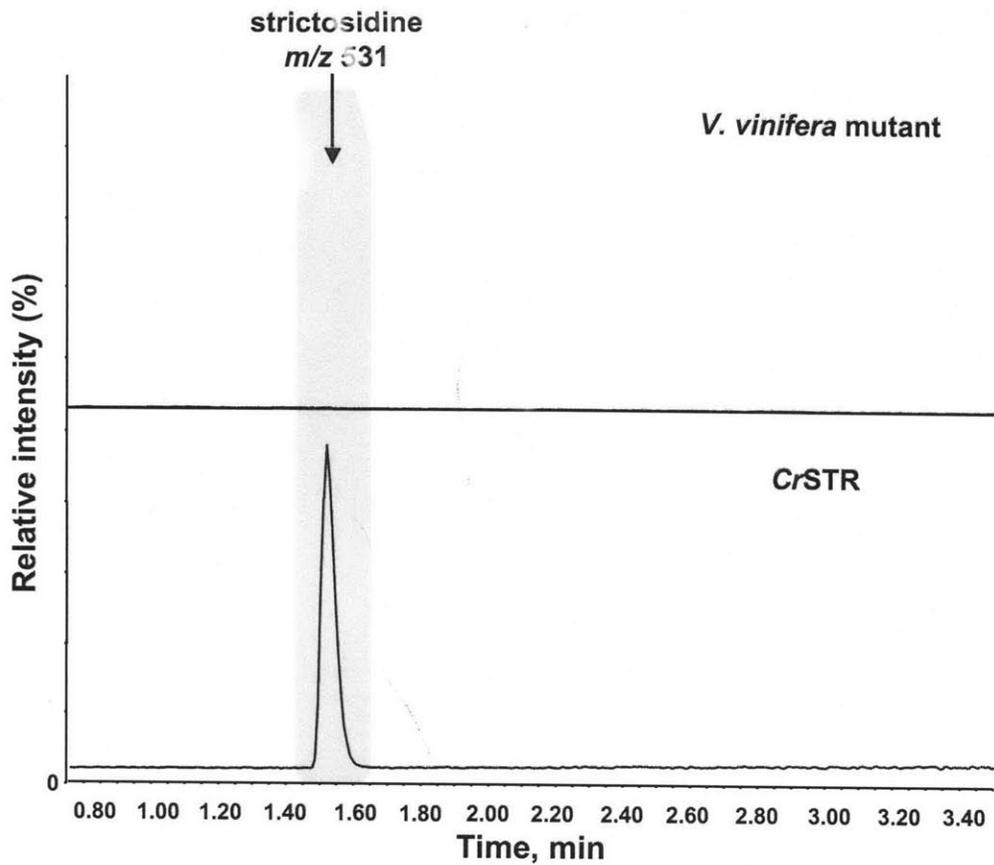
Figure 3.8 LC-MS selected ion chromatograms of strictosidine (*m/z* 531) produced by assays of tryptamine and secologanin incubated with either the *V. vinifera* mutant or *Cr*STR at pH 7.
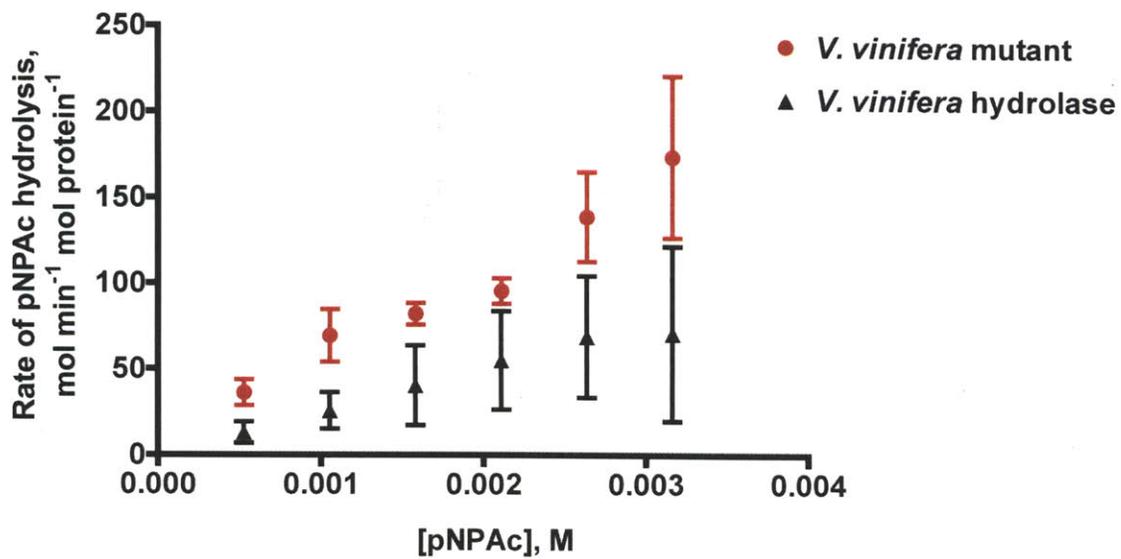
Figure 3.9 Hydrolysis of pNPAc by *V. vinifera* hydrolase and mutant at pH 8. Error bars represent the 95 % confidence of the standard deviation of three experiments.

3.3 Discussion

Protein engineering has undergone major advances with the development of techniques such as directed evolution and rational design [61-63]. While the current trend has been moving away from using rational design alone and more towards combining rational approaches with stochastic methods such as directed evolution [64], this chapter highlights how the rational design of a homologous protein scaffold can be extremely valuable for generating proteins with alternate catalytic activity, and as a framework for understanding how protein evolution may have occurred. There are many successful examples of the use of rational design to introduce new catalytic activities into homologous protein folds found in secondary metabolism. For example, plant sesquiterpene synthases such as γ-humulene synthase [65], tobacco-5-epi-aristolchene synthase [66], and henbane premnaspirodiene synthase [66], all with highly conserved structural and mechanistic features, have been redesigned by rational mutagenesis to yield novel protein activities. This chapter describes the introduction of conserved mechanistic and structural features of closely related STR homologs into STR in an effort to convert a Pictet-Spenglerase into a hydrolase. However, unlike the reengineered sesquiterpene synthases, the mechanistic features of the Pictet-Spenglerase and hydrolase enzymes used to create STR hybrids are quite different. The substrates for these two distinct types of enzymes have diverse functionalities and the chemistries that takes place to form products are also significantly different. Furthermore, in one reaction a metal co-factor is required. This makes the conversion of STR into a hydrolase a unique achievement with potentially significant evolutionary implications.

The lactonohydrolase/paraoxonase superfamily comprises over 2,500 proteins, a few of which have been functionally characterized [31]. Members of this superfamily, in addition to STR, include mammalian paraoxonase [67], a lactonohydrolase from *Fusarium oxysporum* AKU3702 [31], mammalian senescence marker 30 (regucalcin) [33], a gluconolactonase from *Zymomonas mobilis* [68], the human strictosidine synthase homolog C20orf3 [52], diisopropyl fluorophosphatase from *Loligo vulgaris* [36], and the lactonase drug responsive protein 35 from *Staphylococcus aureus* [34]. Unlike STR, all of these proteins have been reported to hydrolyze a variety of substrates such as gluconolactones, lactones, paraoxon, and phenyl acetate.

In the work described in this chapter, we have identified the closely related STR homolog from *V. vinifera* (CAN77945.1) as having hydrolase activity, which can now be added to the subset of characterized enzymes in the lactonohydrolase/paraoxonase superfamily. *V. vinifera* is known to produce phytoalexins and anthocyanins via a prenylpropanoid biosynthetic pathway [69-71] but there is no evidence for the biosynthesis of alkaloids derived from a Pictet-Spengler reaction. This is consistent with the lack of Pictet-Spenglerase activity exhibited by the *V. vinifera* hydrolase *in vitro*. Interestingly, the *V. vinifera* hydrolase is closely related to *Cr*STR (38 % sequence identity). The *V. vinifera* hydrolase lacks the catalytic glutamate required for Pictet-Spenglerase activity [39] and contains residues conserved in the characterized hydrolases that are involved in metal-binding (Fig. 3.2). A similar protein, sharing 35 % sequence identity to the *V. vinifera* hydrolase, was identified in *A. thaliana* (NP_177541.1). *E. coli* lysates containing overexpressed *A. thaliana* (NP_177541.1) protein were reported to hydrolyze the methyl

ester of secologanin when incubated with secologanin and tryptamine [47]. However, these proteins were not properly characterized for the following three reasons: (1) Kibble et al did not purify the protein, (2) there were no controls to account for the presence of endogenous *E. coli* esterases, and (3) the methyl ester of secologanin can easily be hydrolyzed, which can provide misleading results when detecting hydrolase activity. When we assayed the *V. vinifera* hydrolase with tryptamine and secologanin, a secologanic acid product was not observed.

Using conserved residues (including putative metal-binding residues) from closely related STR homologs from *V. vinifera* and *A. thaliana*, and the *Rs*STR and the G2E6 variant of PON1 crystal structures, three hybrid proteins (P1, P2, and P3) based on *Cr*STR were constructed. Each of these hybrids exhibited hydrolase activity with pNPAc. Since the *Cr*STR hybrids displayed similar levels of hydrolysis, quantitative correlations could not be drawn between the rate of hydrolysis and the increase in sphere 1-3 mutations that were introduced into *Cr*STR. While P1 displayed hydrolase activity above the levels of *Cr*STR and background pNPAc hydrolysis, the activity ranges lower than or equal to the activity of the P3 hybrid. P3 was the most consistent and well-behaved hybrid to assay. As such, kinetic parameters were only reported for P3. We suspect that the mutations on the surface of P3 stabilizing the protein structure, similar to the *V. vinifera* hydrolase. A Lineweaver-Burk reciprocal plot of P3 initial rates estimated a high K_m value for pNPAc, suggesting that this is also a non-ideal substrate for P3. The rate of pNPAc hydrolysis with P3 was lower than assays with the *V. vinifera* hydrolase, which exhibited a rate of pNPAc hydrolysis almost 10 times higher than P3. Additional

mutations may be able to improve the rate of pNPAc hydrolysis of P3 such that the rate is equal to that of the *V. vinifera* hydrolase. Furthermore, as the physiological roles and substrates for these STR/PON1 homologs become clearer, we may be able to further optimize hydrolase activity. Subjecting these proteins to directed evolution may also improve the hydrolase activity as has been achieved with paraoxonase.

Pictet-Spenglerases utilize a glutamate residue (Figs. 3.1 B and 3.2) to provide a general acid/base during the course of the reaction mechanism [56]. Therefore, it is not surprising that P1, P2 and P3 do not catalyze the Pictet-Spengler reaction since the key glutamate residue was mutated to an alanine residue in the first sphere mutations. On the other hand, it is difficult to predict a mechanism for hydrolases since, even among the hydrolase members of this superfamily, the substrates, the identities of the metal, and the details of the mechanisms vary. For example, in PON1, a His-dyad, along with the active site Ca^{2+}, has been proposed to activate a water molecule to hydrolyze lactone substrates (Fig. 3.1 A) [72]. However, the STR homologs in this study (CAO67974.1, CAN77945.1, CAO67963.1, CAO66499.1, and NP_177542.1), as well as the human homolog C20orf3, only contain one histidine residue aside from the hexahistidine tag required for affinity purification. As such, we speculate that closely related STR homologs such as the *V. vinifera* hydrolase and P3 may have a somewhat different mechanism of catalysis to PON1. Notably, mutations made to change 4 of the 5 putative calcium-binding residues did not eliminate hydrolase activity, in contrast to a similar experiment performed by Josse et al after making single mutations to metal-binding residues in PON1 [53]. This further suggests that there are important differences between PON1 and the STR-like

hydrolases within this superfamily. Furthermore, our unsuccessful attempt to convert the

V. vinifera hydrolase into a Pictet-Spenglerase by mutating the conserved residues in *V.*

vinifera hydrolase to residues in *Cr*STR, resulted in a slightly more active hydrolase

compared to the wild type *V. vinifera* hydrolase. This was an unexpected result as, again,

this *V. vinifera* mutant contained only two of the four metal-binding residues found in the

wild type *V. vinifera* hydrolase. These results lead us to believe that there may be another

metal-binding site that is not obvious from sequence alignments and homology modeling,

and that the mechanisms of pNPAc hydrolysis by the *V. vinifera* hydrolase and *Cr*STR

could be quite different from PON1. Additional mutations are required to further explore

these mechanistic questions further. This will be key to understanding the roles certain

residues play in hydrolysis to draw hypotheses about the evolution of STR within the

lactonohydrolase/paraoxonase superfamily.

With the current data we cannot draw definitive conclusions about how STR may have

evolved to become a hydrolase, but we can speculate about how differences in sequences

may result in different substrate scopes. PON1 is found in mammalian liver tissue, which

harbor products of fatty acid oxidation, and its physiological substrate is speculated to be

a lactone [38, 73]. Similar to PON1, the STR homolog C20orf3 is also expressed in

human liver [50]; however studies performed in our group have shown that C20orf3 does

not exhibit lactonase activity with the TBBL substrate. C20orf3 has only been shown to

hydrolyze aryl esters such as phenyl acetate and β-naphthyl acetate [52]. Similarly, both

the *V. vinifera* hydrolase and *Cr*STR hybrids did not exhibit lactonase activity with

TBBL. While we cannot yet speculate about the physiological substrate of the *V. vinifera*

hydrolase, gene knockouts and metabolomic analysis may provide more information about potential substrates as vell as help elucidate the role of the *V. vinifera* hydrolase *in planta*.

It is intriguing to speculate that, given the remarkable similarity between STR and STR homologs from *V. vinifera* (CAO67974.1, CAN77945.1, CAO67963.1, CAO66499.1) and *A. thaliana* (NP_177542.1) there may be an evolutionary link between Pictet-Spenglerases and this class of hydrolases. Presumably, Nature evolved a large family of hydrolases, many of which are promiscuous in terms of substrate specificity, to produce a selective protein that catalyzes the highly diverse Pictet-Spengler reaction. The *V. vinifera hydrolase* (CAN77945.1), while definitively exhibiting hydrolase activity, could potentially be a "missing link" or intermediate state in the evolution of Pictet-Spengler activity from the hydrolase. Additional structural and mechanistic studies of these STR synthase homologs should provide a better understanding of how Nature presumably evolved Pictet-Spenglerases into hydrolases in other plants and higher eukaryotic kingdoms.

3.4 Experimental Methods

3.4.1 Chemicals

5-Thiobutyl butyrolactone (TBBL) was a generous gift from Daniel Tawfik (Weizmann Institute of Science, Israel). Secologanin was isolated as previously described [40]. All chemicals were purchased from Sigma Aldrich unless otherwise noted.

3.4.2 General methods and analytical techniques

HPLC separations were performed on a Lichrosorb reverse phase column (Select B, 25 cm x 4.0 mm column, 5 μm particle size) using 10-90 % acetonitrile: water (0.1 % trifluoroacetic acid) over 11 minutes and a flow rate of 1 mL min^{-1}. A Varian Cary 50 Bio UV/Visible Spectrophotometer equipped with a Cary 50 microplate plate reader was used to measure hydrolysis products in colorimetric assays. UPLC and MS analyses were performed in tandem on an Acquity Ultra Performance BEH C18 column with a 1.7 mm particle size, 2.1 x 100 mm dimension, which was coupled to a Micromass LCT Premier TOF Mass Spectrometer by Waters Corporation (Milford, MA) with electrospray ionization source. Analytes were separated using a 10-50 % acetonitrile: water (0.1 % formic acid) over 5 minutes and flow rate of 0.5 mL min^{-1}. For MS analyses, the capillary and sample cone voltages were 3,000 V and 30 V, respectively. The source temperature was 100 °C and the desolvation temperature was 300 °C. The cone and desolvation gas flow were 60 L hr^{-1} and 800 L hr^{-1}, respectively. All reported sequence identities were obtained by using the Fold and Function Assignment 03 server (http://ffas.ljcrf.edu/ffas-cgi/cgi/ffas.pl) [74-76].

3.4.3 Predictions of mutations to convert *Cr*STR into a hydrolase

Michael Hicks, a graduate student in the laboratory of Professor Patsy Babbitt at UCSF, performed a global computational sequence comparison of proteins that belong to the lactonohydrolase/paraoxonase superfamily and found a large subset of proteins, including closely related STR homologs from *Vitis vinifera* and *Arabidopsis thaliana*, were found to contain conserved metal-binding residues [30]. Summer Thyme, a graduate student from the laboratory of Professor David Baker at the University of Washington, used an alignment of the *V. vinifera* and *A. thaliana* homologs (Fig. 3.2), homology models of two of the four *V. vinifera* homologs (CAO67974.1, CAN77945.1), the *Rs*STR crystal structure (2FP8.pdb), and the crystal structure of PON1 (1V04.pdb) to replace residues that were thought to be important for Pictet-Spenglerase activity in *Cr*STR with conserved residues (and metal-binding) in *V. vinifera* and *A. thaliana* STR homologs that are potentially important for introducing hydrolase activity. Three spheres of mutations were made in *Cr*STR stemming from the active site outward to create P1, P2, and P3 STR hybrids.

3.4.3 Gene cloning

C-terminal hexahistidine tagged *Cr*STR in pET28a (+) with flanking NcoI and XhoI restriction sites was provided by Peter Bernhardt. Primers containing mutations to convert C-terminal *Cr*STR into a hydrolase were ordered from Integrated DNA technologies (San Diego, CA) and gene assembly was used to amplify full-length genes. NcoI and XhoI restriction sites were introduced by PCR for standard directional cloning into pET28a (+). Mutations that arose from the gene assembly process were corrected

using the Strategene (Santa Clara, CA) Quickchange Kit. Wild type *Cr*STR, P1, P2, and

P3 constructs were subsequently transformed into *E. coli* BL21 (DE3) cells for protein

expression. The PCR-Blunt II-TOPO vector containing the gene of the STR homolog

from *V. vinifera* (CAN77945.1) was a generous gift from Patsy Babbitt (UCSF). NcoI

and XhoI restriction sites were introduced by PCR for standard directional cloning into

pET32b (+) to create an N-terminal hexahistine tagged a thioredoxin fusion protein. The

V. vinifera mutant (putative Pictet-Spenglerase) was synthesized and cloned into pET32b

(+) by Genscript (Piscataway, NJ) to create and N-terminal hexahistidine tagged

thioredoxin fusion protein. Both *V. vinifera* constructs were transformed into *E. coli*

Rosetta DE3 cells for protein expression. Recombinant PON1 (variant G2E6) in pET32b

(+) was a gift from Dan Tawfik (Weizmann Institute of Science, Israel) and transformed

into *E. coli* Origami B DE3 cells. All liquid and solid media were supplemented with 1

mM calcium chloride.

3.4.4 DNA sequences of *Cr*STR hybrids and *V. vinifera* mutant in pET32b (+)

P1:

```
ATGGGCAGCCCGATTCTGAAGAAAATCTTTATTGAGTCTCCATCTTATGCACCAGAAGCCTTC
ACGTTTGATTCTACGGATAAAGGCTTTTATACGTCTGTCCAAGATGGCCGCGTTATCAAGTAC
GAGGGTCCAAATTCTGGTTTTACCGATTTTGCGTATGCGAGCCCATTTTGGAACAAAGCGTTCT
GCGAAAACAGCACCGATCCGGAAAAGCGCCCGTTATGCGGCCGTACCTACGGCATCAGCTAC
GACTACAAGAACTCTCAGATGTATATTGTGGATGGTCACTATCATCTGTGCGTCGTGGGCAAA
GAAGGCGGCTACGCGACCCAACTGGCGACGAGCGTGCAGGGCGTGCCATTCAAGTGGCTGAA
TGCGGTTACGGTTGATCAGCGTACGGGTATTGTTTACTTTACGGATGTCAGCAGCATTCATGAT
GATTCTCCAGAAGGTGTCGAGGAGATTATGAACACCTCTGATCGCACCGGTCGCTTAATGAAA
TATGATCCATCTACCAAAGAAACCACCCTGCTGTTAAAAGAACTGCATGTTCCAGGTGGCGCG
GAAATCTCTGCCGACGGTAGCTTTGTCGTCGTTGCCGAGTTTCTGTCTAATCGCATTGTCAAAT
ATTGGTTAGAAGGTCCGAAAAAGGGCAGCGCCGAATTTCTGGTGACGATCCCGAACCCAGAT
AACATTAAACGTAATAGCGACGGCCACTTCTGGGTTAGCAGCTCTGAGGAATTAGATGGTGGT
CAGCACGGTCGTGTCGTCAGCCGCGGTATTAAGTTCGATGGTTTTGGTAACATTTTACAGGTT
ATCCCGCTGCCACCACCGTATGAAGGTGAACACTTCAGCGAAATTCAAGAGCATGACGGTTTA
CTGTATATCGGTTCTCTGTTCCACAGCAGCGTTGGTATCCTGGTTTACGATGACCACGATAATA
AGGGCAATAGCTACGTTTCTTCTCTCGAGCACCACCACCACCACCAC
```

P2:

ATGGGCAGCCCGATTCTGAAGAAAATCTTTATTCCGTCTCCATCTGTGGGCCCAGAAGCCTTC
GCGTTTGATTCTACGGACAAAGGCCCGTATACGTCTGTCCAAGATGGCCGCGTTCTGAAGTAC
GAGGGTCCAAATTCTGGTTTTACCGATTTTGCGTATGCGAGCCCATTTTGGAACAAAGCGTTCT
GCGAAAACAGCACCGATCCGGAAAAGCGCCCGTTATGCGGCCGTCCGCTGGGCATCGGCTTT
GACTACAAGAACTCTCAGATGTATATTGCGGATGGTCACTATCATCTGTGCGTCGTGGGCAAA
GAAGGCGGCTACGCGACCCAACTGGCGACGAGCGTGCAGGGCGTGCCATTCAAGTTTCTGAA
TGCGGTTGATGTTGATCAGCGTACGGGTATTGTTTACTTTACGGATGTCAGCAGCATTCATGAT
GATTCTCCAGAAGGTGTCGAGGAGATTATGAACACCTCTGATCGCACCGGTCGCTTAATGAAA
TATGATCCATCTACCAAAGAAACCACCCTGCTGTTAAAAGAACTGCATGTTCCAAATGGCGCG
GAAATCTCTGCCGACGGTAGCTTTGTCGTCGTTGCCGAGTTTCTGTCTAATCGCATTGTCAAAT
ATTGGTTAGAAGGTCCGAAAAAGGGCAGCGCCGAATTTCTGGTGACGATCCCGGGCAACCCA
GATAACATTAAACGTAATAGCGACGGCCACTTCTGGGTTAGCAGCTCTGAGGAATTAGATGGT
GGTCAGCACGGTCGTGTCGTCAGCCGCGGTATTAAGTTCGATGGTTTTGGTAACATTTTACAG
GTTATCCCGCTGCCACCACCGTATGAAGGTGAACACATTAGCGAAGTGCAAGAGCATGACGG
TTTACTGTATATCGGTTCTCTGTTCCACAGCAGCGTTGGTATCCTGGTTTACGATGACCACGAT
AATAAGGGCAATAGCTACGTTTCTTCTCTCGAGCACCACCACCACCACCAC

P3:

ATGGGCAGCCCGATTCTGAAGAAAATCTTTATTCCTTCTCCATCTGTGGGCCCAGAAGCGTTC
GCGTTTGATTCTACGGGCAAAGGCCCGTATACGTCTGTCCAAGATGGCCGCGTTCTGAAGTAC
GAGGGTCCAAATTCTGGTTTTACCGATTTTGCGTATGCGAGCCCATTTTGGAACAAAGCGTTCT
GCGATGGCAGCACCGATCCGGAAAAGCGCCCGTTATGCGGCCGTCCGCTGGGCATCGGCTTTG
ACTACAAGAACTCTCAGATGTATATTGCGGATGCGCACTATGGCCTGTGCGTCGTGGGCAAAG
AAGGCGGCTACGCGACCCAACTGGCGACGAGCGTGCAGGGCGTGCCATTCAAGTTTCTGAAT
GCGGTTGATGTTGATCAGCGTACGGGTATTGTTTACTTTACGGATGTCAGCAGCATTCATGAT
GATTCTCCAGAAGGTGTCGAGGAGATTATGAACACCTCTGATCGCACCGGTCGCTTAATGAAA
TATGATCCATCTACCAAAGAAGTGACCGTGCTGTTAAAAGGCCTGAGCGTTCCAAATGGCGCG
GAAATCTCTGCCGACGGTAGCTTTGTCGTCGTTAGCGAGTTTCTGTCTAATCGCATTGTCAAAT
ATTGGTTAGAAGGTCCGAAAAAGGGCAGCGCCGAATTTTTTGTGACGATCCCGGGCAACCCA
GATAACATTAAACGTAATAGCGACGGCCACTTCTGGGTTGCGAGCTCTGAGGAATTAGATGGT
GGTCAGCACGGTCGTGTCGTCAGCCGCGGTATTAAGTTCGATGGTTTTGGTAACATTTTACAG
GTTATCCCGCTGCCACCACCGTATGAAGGTGAACACATTAGCGAAGTGCAAGAGCATGACGG
TTTGCTGTATATCGGTTCTCTGTTCCACAGCAGCGTTGGTATCCTGGTTTACGATGACCACGAT
AATAAGGGCAATAGCTACGTTTCTTCTCTCGAGCACCACCACCACCACCAC

V. vinifera mutant (proposed Pictet-Spenglerase):

ATGAGCGATAAAATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGAC
GGGGCGATCCTCGTCGATTTCTGGGCAGAGTGGTGCGGTCCGTGCAAAATGATCGCCCCGATT
CTGGATGAAATCGCTGACGAATATCAGGGCAAACTGACCGTTGCAAAACTGAACATCGATCA
AAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGCTGCTGTTCAAAAA
CGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCG
ACGCTAACCTGGCCGGTTCTGGTTCTGGCCATATGCACCATCATCATCATCATTCTTCTGGTCT
GGTGCCACGCGGTTCTGGTATGAAAGAAACCGCTGCTGCTAAATTCGAACGCCAGCACATGG
ACAGCCCAGATCTGGGTACCGACGACGACGACAAGGCCATGGAAGAAGGTGGTCTGGGCTTT
GACGGTTGTCGCCTGGTCCAAGAAGATGAAGGTCTGACGTTCGGTATTAGCACGGGCTCGGGC
CACCTGAGCTCTCGTGGCCTGGAAAGCCTGCTGGAAACCGCACTGCTGGCTCATGTGCTGTTT
ATTATCCCGAACATGTGGTATCTGTCCCACGTGGTTCTGCGCCTGAAACTGGAATCACCGACC
TACGCGCCGAATGCCATTACCTTTGATTATACGGACGCGGGCTTCTACGCATCGGTGGCTGAT

GGTCGTGTTATTAAATGGCTGGATGCAAGCGCTGGTTTTGTTGACTTTGCATTCATCAGTCCGA
GCCGTAGCAAAAAACTGTGCGAAAACTCTACCGATCCGGCTCTGGAACCGACCTGTGGCCGT
ACGTATGACCTGAGTTATAACTACCGCACGGTCGATCTGTATATTGTGGACGGTTACCATCAC
CTGAATGTCGTGGGCCCGAAAGATGGTCGTATTATCCAGCTGGCAACCGCAGCAGAAGGCGT
GCCGTTTCTGTGGCTGTATGCTGTGACCGTTGATCAAGAAACGGGTATCGTTTACTTCACGGA
CGCGAGTGCCCGTTTTCAGCGTCGCGAATTTATGCTGGCCGTCCAAACCGGCGATATGACGGG
TCGTCTGATGAAATACGACCCGCGCACCCAGGAAACCACGCTGCTGCTGCGTGAACTGCATGG
CGCGGGCGGTGTTACCATTTCGAAAGATGGTAGCTTTATCCTGGTTGCCGAATTTGTCACGAA
CCGTATCCAACGCTTTTGGCTGAAAGGCCCGAAAGCAAATACCTCCGAACTGCTGCTGAAACC
GCCGGGCACGCCGGGTAACATTAAACGTAATGTGCGCGGTGAATTTTGGGTCTCTGTGAATAT
CGGTGCAGGCACCGCAGTTCTGCCGTCTGGCCTGCGCCTGAGTGAAGAAGGTAAAGTGCTGC
AGGTTGTCGCCTTTGGCACCGGTGATATTCCGAAAACGTTCGAACAGATCCAAGAATATTATC
GTGCCCTGTATATCGGTTCGCTGGCTCTGCCGTTTGTGGGTGTCTATCGCTCC

3.4.5 Expression of hydrolases

Overnight cultures were grown at 37 °C in sterile LB-broth containing 1 mM CaCl$_2$ with

the appropriate antibiotic selection. Cultures (500 mL) of wild type CrSTR, P1, and P2,

containing 1 mM CaCl$_2$ and 50 μg/mL of kanamycin were inoculated with overnight

cultures (1:100 dilution), and grown at 37 °C until the optical density at 600 nm reached

0.5-0.75. After cultures were chilled at 4 °C for 30 min, protein expression was induced

by the addition of 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG). Cells were

harvested by centrifugation after 18 hours of protein expression at 18°C and stored at

-80°C. Cultures (500 mL) of P3 containing 1 mM CaCl$_2$ and 50 μg/mL of kanamycin

were inoculated 1:100 fold with overnight cultures, and grown at 30 °C until an optical

density at 600 nm of 0.5-0.75 was reached. Protein expression was induced with 1 mM

IPTG and after 4 hours the cells were harvested by centrifugation and stored at -80 °C.

Cultures (500 mL) of the wild type $V.$ $vinifera$ hydrolase containing 1 mM CaCl$_2$ and 100

μg/mL of ampicillin and 34 μg/mL of chloroamphenicol were inoculated with 1:100 fold

overnight culture, and grown at 37 °C until the optical density at 600 nm of 0.5-0.75 was

reached. Protein expression was induced with 1 mM IPTG, and after 6 hours the cells

were harvested by centrifugation and stored at -80 °C. The *V. vinifera* mutant was expressed the same manner, however, the cells were harvested after only 2 hours. PON1 was expressed as previously reported by Aharoni et al [57].

3.4.6 Purification of hydrolases

PON 1 was purified following the procedure reported by Aharoni et al [57]. All other hydrolases were lysed by sonication in lysis buffer (pH 8) containing 50 mM HEPES, 1 mM $CaCl_2$, 300 mM NaCl, 10 mM imidazole, 10 % glycerol, 0.1 mg/mL lysozyme, and 0.4 µg/mL leupeptin and pepstatin protease inhibitors. The lysate was then incubated in 0.1 % tergitol for 2.5 hours at 4 °C. After centrifugation at 15,808 x g, the supernatant was incubated with 0.01 % pre-equilibrated Ni-NTA resin suspension for 1 hour before the flow-through was collected. The resin was washed with one column volume of lysis buffer and 2 column volumes of wash buffer containing 50 mM HEPES, 1 mM $CaCl_2$, 300 mM NaCl, 20 mM imidazole, 10 % glycerol, and 0.1 % tergitol. The resin was then washed with increasing ratios of imidazole and the histadine-tagged proteins were eluted in pH 8 buffer containing 50 mM HEPES, 1 mM $CaCl_2$, 300 mM NaCl, 20 mM imidazole, and 10 % glycerol. The eluent was concentrated in Amicon Ultra centrifugal filter units by Millipore (Billerica, MA) and buffer exchanged using 50 mM HEPES buffer containing 162 mM NaCl, 1 mM $CaCl_2$, and 10 % glycerol. The final protein concentration was determined using a bichinchoninic acid assay by Pierce (Rockford, IL).

3.4.7 Hydrolase activity with p-nitrophenyl acetate and 5-thiobutyl butyrolactone

A stock of 300 mM p-nitrophenyl acetate (pNPAc) was prepared in HPLC-grade methanol and diluted for enzyme assays. Colorimetric assays to detect the formation of p-nitrophenolate at 405 nm were prepared in a MICROTEST 96 well plate from Becton Dickinson Labware (Franklin Lakes, NJ) with a final volume of 250 μL containing either 530 nM (CrSTR, P1, P2, and P3) or 53 nM (PON1, V. vinifera hydrolase, and mutant) of protein in 100 mM HEPES buffer containing 1 mM $CaCl_2$ and 2.6 mM NaCl. After pre-equilibration at room temperature for 10 minutes, the assays were initiated by the addition of pNPAc (2.6 mM). Kinetic parameters were measured using a range of pNPAc concentrations (0.5 mM, 1 mM, 1.6 mM, 2.1 mM, 2.6 mM, and 3.2 mM). Time-points were chosen such that the rate of product formation was linear, to ensure accurate measure of initial rates (between 3 and 30 minutes after initiation of the reaction). The measured pathlength of 0.79 cm and p-nitrophenol extinction coefficient of 18,000 M^{-1} cm^{-1} were used to convert the absorbance units of p-nitrophenol into concentrations by the Beer-Lambert law. Kinetic parameters were estimated by fitting the data to the Lineweaver-Burk plot since limited substrate solubility did not enable an acceptable fit to the Michaelis-Menten equation (maximal substrate concentration was less than the 2-3 x K_m).

To determine lactonase activity, a 220 mM TBBL stock was prepared in HPLC-grade acetonitrile and diluted to 125 mM. Colorimetric assays to detect hydrolase activity at 412 nm were prepared in a MICROTEST 96 well plate with a final volume of 250 μL containing either 530 nM (CrSTR, P1, P2, and P3) or 53 nM (PON1, V. vinifera

hydrolase, and mutant) of protein in 100 mM HEPES buffer containing 1 mM $CaCl_2$ and 2.6 mM NaCl. After pre-equilibration at room temperature for 10 minutes, the assays were initiated by the addition of 21 µM TBBL.

3.4.8 Hydrolase activity with phenyl acetate

HPLC assays were used to detect the hydrolysis of phenyl acetate. Assays were prepared with a final volume of 250 µL containing either 530 nM (CrSTR, P1, P2, and P3) or 53 nM (PON1, V. vinifera hydrolase, and V. vinifera mutant) of protein and 480 µM 1-naphthaleneacetic acid internal standard in 100 mM HEPES buffer containing 1 mM $CaCl_2$ and 125 mM NaCl. After pre-equilibration at room temperature for 10 minutes, assays were initiated by the addition of 2.6 mM phenyl acetate and 20 µL of the assay was quenched with an equal volume of HPLC-grade methanol every 20-30 minutes. Quenched assays were clarified by centrifugation for 5 minutes in a microcentrifuge, and products were analyzed by HPLC at 270 nm.

3.4.9 Assays to detect Pictet-Spenglerase activity

To detect Pictet-Spenglerase activity, assays were prepared in a final volume of 100 µL containing 234 nM protein, and 200 µM tryptamine in 100 mM pH 7 phosphate buffer. Assays were initiated by the addition of 1.2 mM secologanin and incubated at 30 °C overnight. Ten percent of the assay volume was quenched with HPLC-grade methanol, clarified by centrifugation for 5 minutes in a microcentrifuge, and analyzed by LC-MS using selected ion monitoring at the mass of the strictosidine product, m/z 531.

3.4.10 Inhibition of P3 pNPAc hydrolysis in the presence of EDTA and calcium inhibitor, terbium chloride

To determine if P3 is metal-dependent, P3 was buffered exchanged into a pH 8 buffer containing 50 mM HEPES, 133 mM NaCl and 10 mM EDTA, and 10 % glycerol. Protein concentration was measured using the bichinchoninic assay by Pierce (Rockford, IL) before assays were prepared. Colorimetric assays were prepared in the same manner as described in section 3.4.7, however, the pH 8 buffer used in these assays contained 100 mM HEPES, 96 mM NaCl, and 10 mM EDTA.

Calcium inhibition assays with terbium chloride were prepared by buffer exchanging P3 into a pH 8 buffer containing of 4 mM terbium chloride, 50 mM HEPES, 6.5 mM NaCl, and 10 % glycerol. After the protein concentration was measured, colorimetric assays were prepared by incubating 530 nM P3, 4 mM terbium chloride, in 100 mM HEPES pH 8 buffer containing 1 mM $CaCl_2$ and 2.6 mM NaCl. After pre-equilibration at room temperature for 10 minutes, assays were initiated by the addition of 2.6 mM pNPAc.

3.4.11 Generation of mutants to eliminate hydrolase activity in P3 and *V. vinifera* hydrolases

The primers listed in Table 3.1 were used to mutate metal-binding residues in P3 and the *V. vinifera* hydrolase. Mutations were generated using the Quickchange Kit by Strategene (Santa Clara, CA).

Table 3.1 Primers for the amplification of full-length genes (mutations in bold)

Primer name	Primer sequence 5' to 3'
*V. vinifera*E75Aforward	CAACCGTGGGTCCGG**C**G**G**CTATCGCGTTTG
*V. vinifera*E75A reverse	CAAACGCGATAGC**C**G**C**CGGACCCACGGTTG
*V. vinifera*N182Aforward	GCCGTTCCTGTTCCTGG**C**GGCGGTTGACGTTGACC
*V. vinifera*N182Areverse	GGTCAACGTCAAC**C**GCCGCCAGGAACAGGAACGGC
*V. vinifera*D286A+N287Afor	CCGCCGGGTACCCCTGC**GGCGA**TCAAACGTAACGTTCGTGG
*V. vinifera*D286A+N287Arev	CCACGAACGTTACGTTTGAT**CGCCG**CAGGGGTACCCGGCGG
P3E17Aforward	CCATCTGTGGGCCCAGC**G**GCGTTCGCGTTTG
P3E17Areverse	CAAACGCGAACGC**C**GCTGGGCCCACAGATGG
P3N128Aforward	GGCGTGCCATTCAAGTTTCTGG**C**GGCGGTTGATGTTG
P3N128Areverse	CAACATCAACCGC**C**GCCAGAAACTTGAATGGCACGCC
P3N189Aforward	GGCCTGAGCGTTCCAG**C**GGGCGCGGAAATCTC
P3N189Areverse	GAGATTTCCGCGCC**C**GCTGGAACGCTCAGGCC
P3D234A+N235Afor	GACGATCCCGGGCAACCCAGC**GGCG**ATTAAACGTAATAGCGAC
P3D234A+N235Areverse	GTCGCTATTACGTTTAAT**CGCCG**CTGGGTTGCCCGGGATCGTC

3.5 References

1. Glasner, M.E., J.A. Gerlt, and P.C. Babbitt, *Mechanisms of protein evolution and their application to protein engineering.* Advances in enzymology and related areas of molecular biology, 2007. **75**: p. 193-239.

2. Toscano, M.D., K.J. Woycechowsky, and D. Hilvert, *Minimalist active-site redesign: teaching old enzymes new tricks.* Angew. Chem., 2007. **46**(18): p. 3212-3236.

3. Bommarius, A.S. and B.R. Riebel, *Biocatalysis* 2004, Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA. 1-18.

4. Radzicka, A. and R. Wolfenden, *A proficient enzyme.* Science, 1995. **267**(5194): p. 90-93.

5. Floudas, C.A., et al., *Advances in protein structure prediction and de novo protein design: a review.* Chemical engineering science, 2006. **61**(3): p. 966-988.

6. Grunwald, P., *Biocatalysis: biochemical fundamentals and applications* 2009, London: Imperial College Press. 777-829.

7.   Nixon, A.E., M. Ostermeier, and S.J. Benkovic, *Hybrid enzymes: manipulating enzyme design.* Trends in Biotechnology, 1998. **16**(6): p. 258-264.

8.   Anantharaman, V., L. Aravind, and E.V. Koonin, *Emergence of diverse biochemical activities in evolutionarily conserved structural scaffolds of proteins.* Curr. Opin. Chem. Biol., 2003. **7**(1): p. 12-20.

9.   Chothia, C., *One thousand families for the molecular biologist.* Nature, 1992. **357**(6379): p. 543-544.

10.   Orengo, C.A., D.T. Jones, and J.M. Thorton, *Protein superfamilies and domain superfolds.* Nature, 1994. **372**(6507): p. 631-634.

11.   Wu, G., et al., *Convergent evolution of Trichomonas vaginalis lactate dehydrogenase from malate dehydrogenase.* Proc. Natl. Acad. Sci. USA, 1999. **96**(11): p. 6285-6290.

12.   Little, D. and R.B. Croteau, *Alteration of product formation by directed mutagenesis and truncation of the multiple-product sesquiterpene synthases δ-selinene synthase and γ-humulene synthase.* Arch. Biochem. Biophys., 2002. **402**(1): p. 120-135.

13.   Arnold, F., *Combinatorial and computational challenges for biocatalyst design.* Nature, 2001. **409**(6817): p. 253-257.

14.   David, R., et al., *Artificial chemokines: combining chemistry and molecular biology for the elucidation of interleukin-8 functionality.* J. Am. Chem. Soc., 2008. **130**(46): p. 15311-15317.

15.   Cordes, M.H.J., et al., *An evolutionary bridge to a new protein fold.* Nat Struct Biol., 2000. **7**(12): p. 1129-1132.

16.   James, C.L. and R.E. Viola, *Production and characterization of bifunctional enzymes. Domain swapping to produce new bifunctional enzymes in the aspartate pathway.* Biochemistry, 2002. **41**(11): p. 3720-3725.

17.   Back, K. and J. Chappell, *Identifying functional domains within terpene cyclases using a domain-swapping strategy* Proc. Natl. Acad. Sci. USA, 1996. **93**(13): p. 6841-6845.

18.   Hansen, E.H., et al., *Substrate specificities of family 1 UGTs gained by domain swapping.* Phytochemistry, 2009. **70**(4): p. 473-482.

19.   Koide, S., *Generation of new protein functions by nonhomologous combinations and rearrangements of domains and modules.* Curr. Opin. Biotech., 2010. **20**(4): p. 398-404.

20.   Gilbert, M., et al., *The synthesis of sialylated oligosaccharides using a CMP-Neu5Ac synthetase/sialyltransferase fusion.* Nat. Biotech., 1998. **16**(8): p. 769-772.

21.   Pabo, C.O., E. Peisach, and R.A. Grant, *Design and selection of novel CYS2HIS2 zinc finger proteins.* Annu. Rev. Biochem., 2001. **70**: p. 313-340.

22.   Segal, D.J. and C.F. Barbas, III. , *Custom DNA-binding proteins come of age: polydactyl zinc-finger proteins.* Curr. Opin. Biotechnol., 2001. **12**(6): p. 632-637.

23.   Jantz, D., et al., *The design of functional DNA-binding proteins based on zinc finger domains.* Chem. Rev., 2003. **104**(2): p. 789-800.

24.   Gonçalves, M.A.F.V., et al., *Transcription factor rational design improves directed differentiation of human mesenchymal stem cells into skeletal myocytes* Molecular Therapy, 2011: p. 1-11.

25.   Mandell, J.G. and C.F. Barbas, III., *Zinc finger tools: custom DNA-binding domains for transcription factors and nucleases* Nuc. Acids Res., 2006. **34**: p. W516-W523.

26.   Gerlt, J.A. and P.C. Babbitt, *Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprafamilies.* Annu. Rev. Biochem., 2001. **70**: p. 209-246.

27.   Siegler, D.S., *Plant secondary metabolism* 1998, Norwell, Massachusetts: Kluwer Academic Publishers. 1-15.

28.   Trapp, S.C. and R.B. Croteau, *Genomic organization of plant terpene synthases and molecular evolutionary implications.* Genetics, 2001. **158**(2): p. 811-832.

29.   Khersonsky, O. and D.S. Tawfik, *Enzyme promiscuity: a mechanistic and evolutional perspective.* Annu. Rev. Biochem., 2010. **79**: p. 471-505.

30.   Hicks, M.A., et al., *The Evolution of New Function in Strictosidine Synthase-like Proteins*, 2011: San Francisco. p. 1-28.

31.   Kobayashi, M., et al., *Lactone-ring-cleaving enzyme: genetic analysis, novel RNA editing, and evolutionary implication.* Proc. Natl. Acad. Sci. USA, 1998. **95**(22): p. 12787-12792.

32.   Kondo, Y., et al., *Senescence marker protein 30 functions as gluconolactonase in l-ascorbic acid biosynthesis, and its knockout mice are prone to scurvy* Proc. Natl. Acad. Sci. USA, 2006. **103**(15): p. 5723-5728.

33.   Chakrabortyi, S. and B.J. Bahnson, *Crystal structure of human senescence marker protein 30: Insights linking structural, enzymatic, and physiological functions.* Biochemistry, 2010. **49**(16): p. 3436-3444.

34.   Tanaka, Y., et al., *Structural and mutational analyses of Drp35 from Staphylococcus aureus: a possible mechanism for its lactonase activity.* J. Biol. Chem., 2006. **282**(8): p. 5770-5780.

35.   Scharff, E.I., et al., *Crystal structure of diisopropylfluorophosphatase from Loligo vulgaris.* Structure, 2001. **9**(6): p. 493-502.

36.   Hoskin, F.C.G. and R.J. Long, *Purification of a DFP-hydrolyzing enzyme from squid head ganglion.* Arch. Biochem. Biophys., 1972. **150**(2): p. 548-555.

37.   Harel, M., et al., *Structure and evolution of the serum paraoxonase family of detoxifying and anti-atherosclerotic enzymes.* Nature structural & molecular biology, 2004. **11**(5): p. 412-419.

38.   Khersonsky, O. and D.S. Tawfik, *Structure-reactivity studies of serum paraoxonase PON1 suggest that its native activity is lactonase.* Biochemistry, 2005. **44**(16): p. 6371-6382.

39.   Ma, X., et al., *The structure of Rauvolfia serpentina strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins.* The Plant cell, 2006. **18**(4): p. 907-920.

40.   McCoy, E., M.C. Galan, and S.E. O'Connor, *Substrate specificity of strictosidine synthase.* Bioorg. Med. Chem. Lett., 2006. **16**(9): p. 2475-2478.

41.   Stöckigt, J. and M.H. Zenk, *Strictosidine (isovincoside): the key intermediate in the biosynthesis of monoterpenoid indole alkaloids.* J. Chem. Soc. Chem. Comm., 1977(18): p. 646-648.

42.   Treimer, J.F. and M.H. Zenk, *Purification and properties of strictosidine synthase, the key enzyme in indole alkaloid formation.* Eur. J. Biochem., 1979. **101**(1): p. 225-233.

43.   McKnight, T.D., et al., *Nucleotide sequence of a cDNA encoding the vacuolar protein strictosidine synthase from Catharanthus roseus.* Nuc. Acids Res., 1990. **18**(16): p. 4939.

44.   Kutchan, T.M., et al., *The cDNA clone for strictosidine synthase from Rauvolfia serpentina. DNA sequence determination and expression in Escherichia coli.* FEBS Lett., 1988. **237**(1-2): p. 40-44.

45.   Yamazaki, Y., et al., *Camptothecin biosynthetic genes in hairy roots of Ophiorrhiza pumila: Cloning, characterization and differential expression in tissues and by stress compounds.* Plant Cell Physiol., 2003. **44**(4): p. 395-403.

46.   Thompson, J.D., D.G. Higgins, and T.J. Gibson, *Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice.* Nuc. Acids Res., 2004. **22**(22): p. 4673-4680.

47.   Kibble, N.A.J., et al., *Phylogenetic analysis and functional characterisation of strictosidine synthase-like genes in Arabidopsis thaliana.* Functional Plant Biology, 2009. **36**(12): p. 1098-1109.

48.   Gatto, P., et al., *Ripening and genotype control stilbene accumulation in health grapes.* J. Agric. Food Chem., 2008. **56**(24): p. 11773-11785.

49. Fabbri, M., et al., *Animal and plant members of a gene family with similarity to alkaloid synthesizing enzymes.* Biochem. Biophys. Res. Comm., 2000. **271**(1): p. 191-196.

50. Morita, M., et al., *Genomic construct and mapping of gene for CMAP and identification of a proximal novel gene, BSCv.* Genomics, 2000. **67**(1): p. 87-91.

51. Hatzios, S., *Human alkaloid biosynthesis: Chemical inducers of Parkinson's disease,* in *Chemistry* 2005, Massachusetts Institute of Technology: Cambridge.

52. Ilhan, A., et al., *Localization and characterization of the novel protein encoded by C20orf3.* Biochem. J., 2008. **414**(3): p. 485-495.

53. Josse, D., et al., *Identification of residues essential for human paraoxonase (PON1) arylesterase/organophosphatase activities.* Biochemistry, 1999. **38**(9): p. 2816-2825.

54. Aharoni, A., et al., *Directed evolution of mammalian paraoxonases PON1 and PON3 for bacterial expression and catalytic specialization.* Proc. Natl. Acad. Sci. USA, 2003. **101**(2): p. 482-487.

55. Altschul, S.F., et al., *Basic local alignment search tool.* J. Mol. Biol., 1990. **215**(3): p. 403-410.

56. Maresh, J.J., et al., *Strictosidine synthase: mechanism of a Pictet-Spengler catalyzing enzyme.* J. Am. Chem. Soc., 2008. **130**(2): p. 710-723.

57. Aharoni, A., et al., *Directed evolution of mammalian paraoxonases PON1 and PON3 for bacterial expression and catalytic specialization.* Proc. Natl. Acad. Sci. USA, 2004. **101**(2): p. 482-487.

58. Roessner, C.A., et al., *Purification of an indole alkaloid biosynthetic enzyme, strictosidine synthase, from a recombinant strain of E. coli.* Prot. Exp. Pur., 1992. **3**(4): p. 295-300.

59. Otto, T.C., et al., *Purification and characterization of functional human paraoxonase-1 expressed in Trichoplusia ni larvae.* Chemico-Biological Interactions, 2010. **187**(1-3): p. 388-392.

60. Khersonsky, O. and D.S. Tawfik, *Chromogenic and fluorogenic assays for the lactonase activity of serum paraoxonases.* ChemBioChem, 2006. **7**(1): p. 49-53.

61. Nanda, V. and R.L. Koder, *Designing artifical enzymes by intuition and computation.* Nat Chem., 2010. **2**(1): p. 15-24.

62. Bolon, D.N., C.A. Voigt, and S.L. Mayo, *De novo design of biocatalysts.* Curr. Opin. Chem. Biol., 2002. **6**(2): p. 125-129.

63. Lutz, S., *Beyond directed evolution---semi-rational protein engineering and design.* Curr. Opin. Biotech., 2010. **21**(6): p. 734-743.

64. Otten, L.G., F. Hollmann, and I.W.C.E. Arends, *Enzyme engineering for enantioselectivity: from trial-and-error to rational design?* Trends Biotech., 2010. **28**(1): p. 46-54.

65. Yoshikuni, Y., T.E. Ferrin, and J.D. Keasling, *Designed divergent evolution of enzyme function.* Nature, 2006. **440**(7087): p. 1078-1082.

66. O'Maille, P.E., et al., *Quantitative exploration of the catalytic lanscape separating divergent plant sesquiterpene synthases.* Nature Chemical Biology, 2008. **4**(10): p. 617-623.

67. Draganov, D.I. and B.N. La Du, *Pharmacogenetics of paraoxonases: a brief review.* Naunyn-Schmiedeberg's Arch. Pharmacol., 2004. **369**(1): p. 78-88.

68. Kanagasundaram, V. and R. Scopes, *Isolation and characterization of the gene encoding gluconolactonase from Zymomonas mobilis.* Biochimica Biophys Acta., 1992. **1171**(2): p. 198-200.

69. Yamakawa, T., et al., *Formation and identification of anthocyanins in cultured cells of Vitis sp.* Agric. Biol. Chem, 1983. **47**(5): p. 997-1001.

70. Jeandet, P., et al., *Phytoalexins from the vitaceae: biosynthesis, phytoalexin gene expression in transgenic plants, antifungal activity, and metabolism.* J. Agric. Food Chem., 2002. **50**(10): p. 2731-2741.

71. Jaillon, O., et al., *The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla.* Nature, 2007. **449**: p. 463-467.

72. Khersonsky, O. and D.S. Tawfik, *The histidine 115-histidine 134 dyad mediates the lactonase activity of mammalian serum paraoxonases.* J. Biol. Chem., 2006. **281**(11): p. 7649-7656.

73. Draganov, D.I., et al., *Human paraoxonases (PON1, PON2, and PON3) are lactonases with overlapping and distinct substrate specificities.* J Lipid Res., 2005. **46**(6): p. 1239-1247.

74. Jaroszewski, L., et al., *FFAS03: a server for profile--profile sequence alignments.* Nuc. Acids Res., 2005. **33**: p. W284-W288.

75. Jaroszewski, L., L. Rychlewski, and A. Godzik, *Improving the quality of twilight-zone alignments.* Protein Science, 2000. **9**(8): p. 1487-1496.

76.   Rychlewski, L., et al., *Comparison of sequence profiles. Strategies for structural predictions using sequence information.* Protein Science, 2000. **9**(2): p. 232-241.

CHAPTER 4

A STEREOSELECTIVE HYDROXYLATION STEP OF ALKALOID BIOSYNTHESIS

BY A UNIQUE CYTOCHROME P450 IN *CATHARANTHUS ROSEUS*

4.1 Introduction

Cytochrome P450-dependent enzymes (P450s) play a key role in the development and survival of plants [1-3]. Adaptation to terrestrial conditions, the production of lignin for structural reinforcement, pollination, and long range signaling for defense against herbivores and pests are processes that each utilize biochemical pathways in plants. P450s are involved in these pathways since oxygen is often utilized in building structurally complex molecules. These enzymes participate in myriad biosynthetic pathways, some of which lead to the production of fatty acids, terpenoids, phenylpropanoids, cyanogenic glucosides, glucosinolates, and alkaloids [4]. P450s catalyze hydroxylation, epoxidation, oxidation, deamination, dehalogenation, and many other reactions [5]. In oxidation reactions, P450s initiate their chemistry by converting a resting state iron (III) protoporphyrin-IX complex to a reactive oxoiron (IV) porphyrin cation radical intermediate (Fig. 4.1). Specifically, these heme-containing enzymes use a reductase partner to transfer an electron from an NAD(P)H cofactor to reduce the ferric P450 to a ferrous state (Fig. 4.1, step 1) [6]. Molecular oxygen binds to the ferrous heme (Fig. 4.1, step 3), which is subsequently reduced by another electron. Protonation and heterologous cleavage (Fig. 4.1, step 4) forms a reactive iron-oxo complex (Fig. 4.1, step 5) that can catalyze oxygen transfer to a substrate to form the oxygenated product complex (Fig. 4.1, step 6) [4, 6].
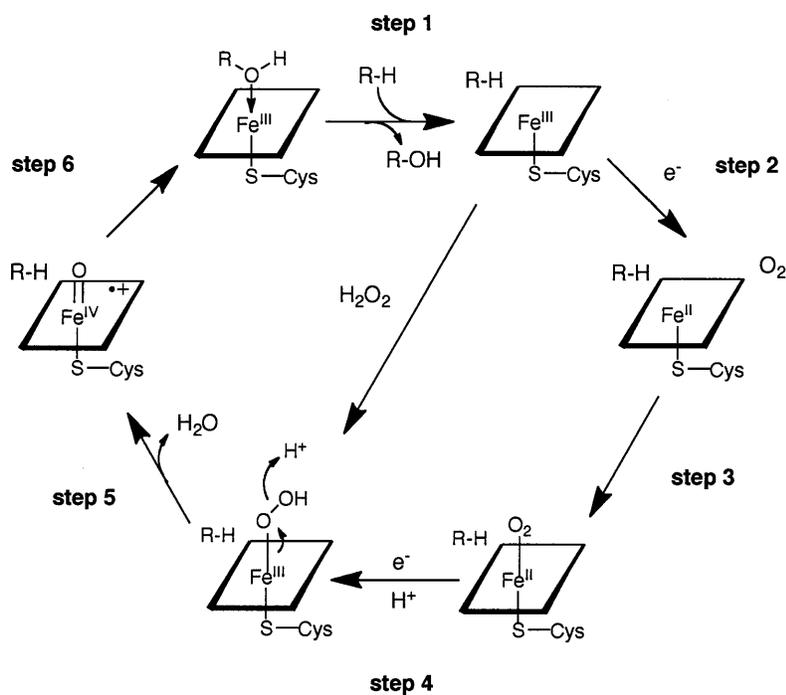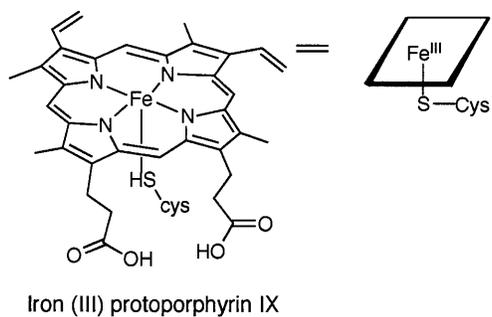
Iron (III) protoporphyrin IX

Figure 4.1 Catalytic cycle for oxygen activation and transfer by P450s.

Plants have an unusually large number of P450 genes compared to prokaryotes and other eukaryotic organisms. P450 genes in the plant genome are estimated to comprise approximately 1 % of total gene complements for sequenced plant species compared to the 0.1-0.5 % of the gene complements found in human, mouse, *Takifugu rubripes*, *Anopheles gambiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Ciona intestinalis*, and *Ciona savignyi* genomes [2, 7]. In fact, P450s have evolved to be the largest enzyme superfamily in plant metabolism. However, the catalytic functions of most plant P450s are unknown because of insufficient knowledge of the pathways and the chemistries these enzymes employ. For example, there are 256 annotated P450s in the model plant *Arabidopsis thaliana* (www.arabidopsis.org/info/genefamily/p450.jsp), and less than 20 % of those genes have been associated with a specific biochemical function [8]. Plant P450s are typically highly substrate-specific enzymes that catalyze regio- and stereo-selective transformations [7, 9]. While P450 genes possess high sequence identity to each other, the lack of correlation between primary structure and catalytic function makes it challenging to identify a P450 that catalyzes a specific biosynthetic transformation [3].

Catharanthus roseus is the sole source of vinblastine and vincristine, alkaloids used clinically to treat leukemia, Hodgkin's lymphoma, and other cancers [10]. The biosynthetic pathway leading to these medicinally important bisindole alkaloids remains to be fully understood. Although alkaloid biosynthesis in *C. roseus* involves several biotransformations known or speculated to be P450-dependent [11], only five *C. roseus* P450s have been functionally-characterized: flavonoid 3',5'-hydroxylase, cinnamate 4-

hydroxylase, geraniol 10-hydroxylase (G10H), secologanin synthase (SLS), and

tabersonine 16-hydroxylase (T16H) [12-16]. These P450 genes were functionally

characterized using various techniques, including "reverse genetics", a strategy in which

proteins are purified and partially sequenced. The partial protein sequence is then used to

screen a cDNA library for corresponding genes. These traditional approaches have

proven to be effective for the functional characterization of some P450 genes, but these

methods are labor-intensive and time-consuming [3]. Recently, a large-scale co-

expression analysis was performed to predict the functional characterization of orphan

P450 genes in *Arabidopsis thaliana* [17, 18]. Co-expression analysis assumes that genes

of the same biosynthetic pathway are similarly regulated on a transcriptional level. The

correlation between tissue-specific gene expression and biochemical function proves to

be extremely valuable for generating hypotheses about specialized gene function.

This chapter describes the use of a *C. roseus* EST collection generated by Illumina

transcriptome sequencing and *in silico* transcriptome analysis to create a corresponding

expression profile in order to predict the role of orphan P450s involved in alkaloid

biosynthesis. In *C. roseus*, the alkaloid tabersonine can be converted into the bisindole

precursor vindoline in aerial organs or to 19-*O*-acetylhörhammericine in roots (Fig. 4.1)

[10, 19]. Although hydroxylation at the 16-position of tabersonine leads to vindoline, a

putative P450 6,7-epoxidase can convert tabersonine to lochnericine [20-22]. Similarly, a

putative P450 hydroxylase can convert tabersonine to 19-hydroxytabersonine [23]. Both

lochnericine and 19-hydroxytabersonine are proposed intermediates in 19-*O*-

acetylhörhammericine biosynthesis (Fig. 4.2). Moreover, the only gene for which a

cognate DNA has been isolated in 19-*O*-acetylhörhammericine biosynthesis is minovincinine 19-hydroxy-*O*-acetyltransferase (MAT) [19]. MAT catalyzes the *O*-acetylation of the 19-hydroxyl group of minovinicinine and hörhammericine to form echitovenine and 19-*O*-acetylhörhammericine, respectively. Using hierarchical clustering of *C. roseus* gene expression data, we identified putative P450 genes that clustered with MAT. A whole cell yeast assay was developed to express and functionally characterize these P450 gene candidates to complement *in silico* transcriptome analyses. Gratifyingly, yeast cell cultures expressing one of these P450 gene candidates, CYP71BJ1, produced a hydroxylated product in the culture medium when supplemented with tabersonine. CYP71BJ1 is the first member of a new plant P450 subfamily, and appears to be a tabersonine/lochnericine 19-hydroxylase.
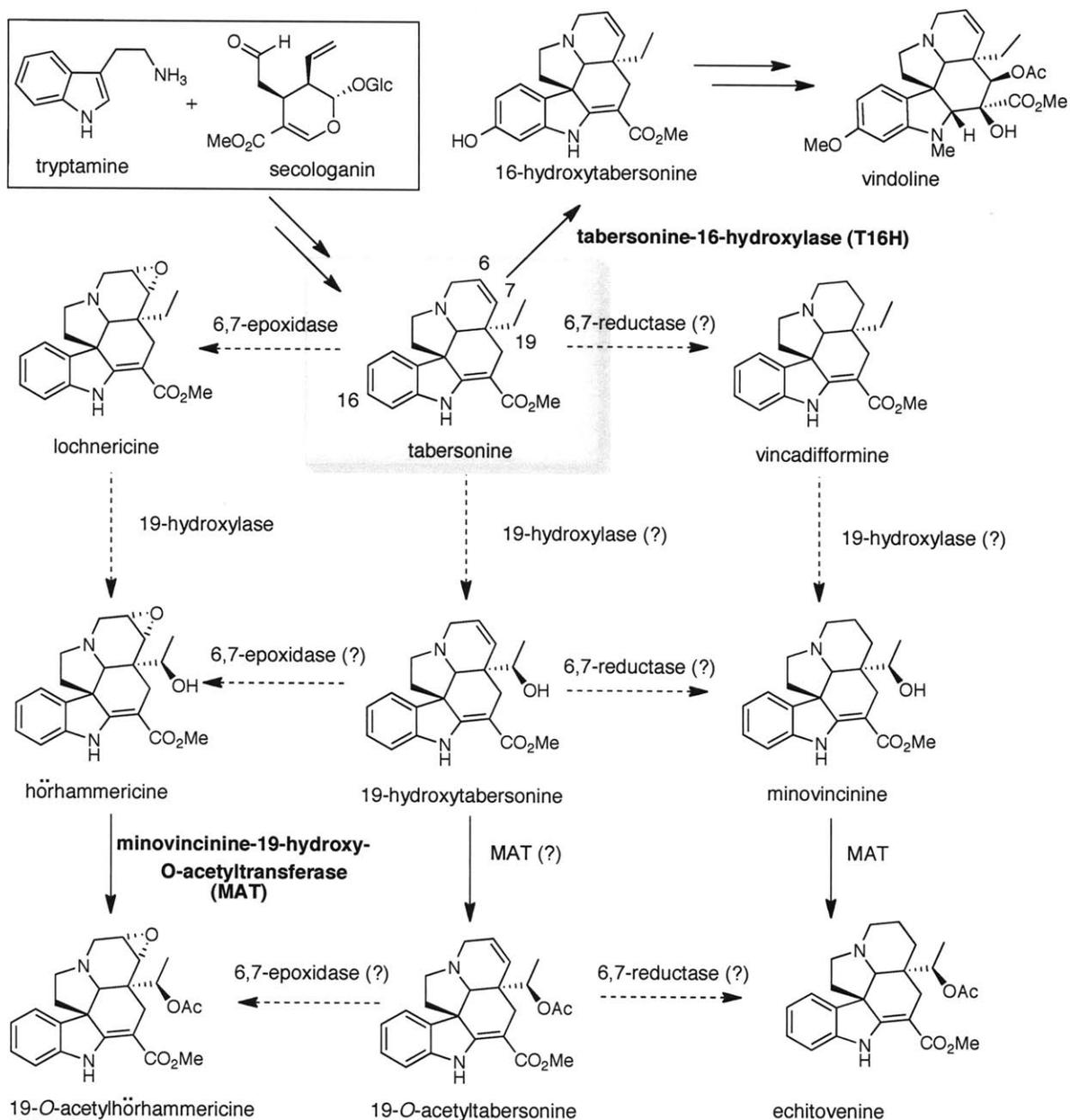
Figure 4.2 Proposed biosynthesis of tabersonine-related alkaloids in *Catharanthus roseus*. Enzymes that have been biochemically characterized are in bold. The dashed arrows represent steps for which no enzyme has been identified. The question marks represent enzymes that may be involved in turning over a particular substrate.

144

4.2 Results

4.2.1 Co-expression analysis using MAT as bait

Dr. David Liscombe performed hierarchical clustering with co-expression analysis of *C. roseus* transcripts to identify P450 genes that clustered with MAT and may be involved in the oxygenation of tabersonine. Expression profiles of known *C. roseus* alkaloid biosynthetic genes [tryptophan decarboxylase, G10H, loganic acid methyltransferase, SLS, strictosidine synthase (STR), strictosidine glucosidase (SGD), T16H, 16-hydroxy-16-*O*-methoxytabersonine (16OMT), 2,3-dihydro-3-hydroxytabersonine-*N*-methyltransferase, desacetoxyvindoline-4-hydroxylase (D4H), deacetylvindoline acetyltransferase, and MAT] that co-expressed with putative P450 genes were shown as a heat map, the portion showing the clusters that we focused on in this study is presented in Fig. 4.3 [10, 19, 24]. Contigs 92197, 80887, 63935, 91544, 89777 all represented partial ORFs. Two partial length transcripts (contigs 88716 and 87898) were clustered with MAT (Fig. 4.3) [19]. Fortunately, these partial ORFs were found to represent the same gene transcript by manual sequence analysis and could be assembled to form the full-length ORF encoding CYP71BJ1. Two other apparent full-length ORFs (41747 and CYP81Z1) were present in the cluster and were also selected for further characterization. Dr. David Liscombe cloned gene candidates from methyl jasmonate-elicited *C. roseus* seedling cDNA into the yeast expression vector pYeDP60. The candidate P450 proteins were then expressed in *Saccharomyces cerevisiae* WAT11 cells, a yeast strain optimized for plant P450 protein expression [25].

SDLG

MeJ + MeJ CSC

Minovincinine 19-O-acetyltransferase (MAT)

CYP71BJ1

92197 (*partial ORF*)
Minovincinine 19-O-acetyltransferase (MAT)
80887 (*partial ORF*)
Desacetoxyvindoline 4-hydroxylase (D4H)
Locus_41747 GenBank: AJ295719

Strictosidine glucosidase (SGD)

Strictosidine synthase (STR)
63935 (*partial ORF*)
91544 (*partial ORF*)
89777 (*partial ORF*)
38939 CYP81Z1

0   50   100
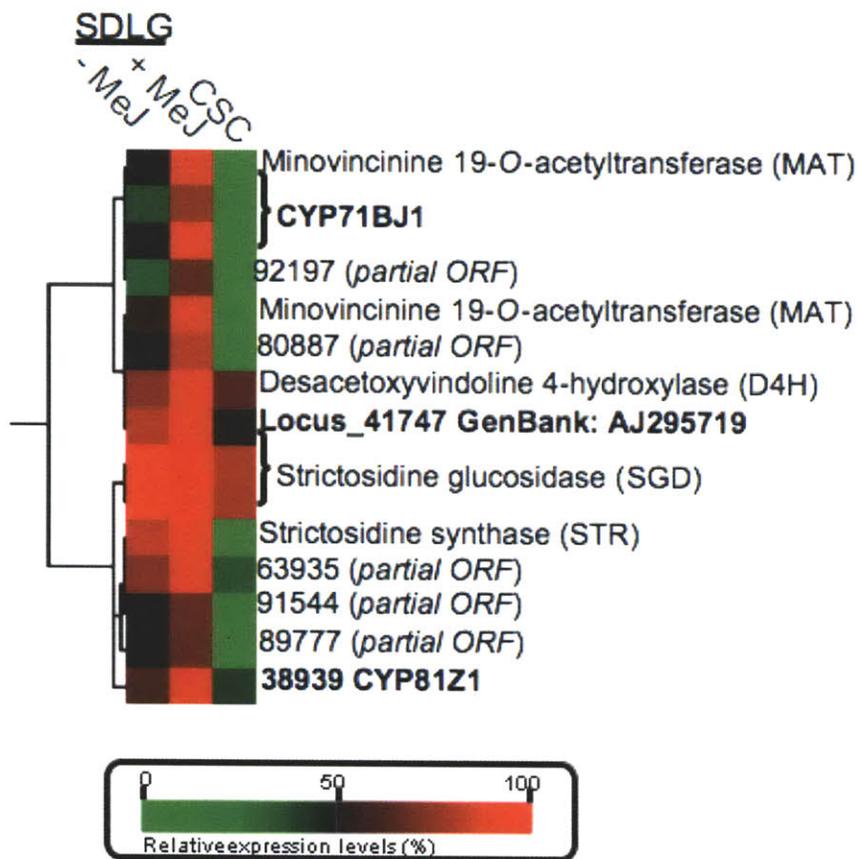
Relative expression levels (%)

Figure 4.3. Section of hierarchical cluster analysis of putative P450 genes (full-length in bold) co-expressed with MAT. This heat map was generated by the Treeview program [26].

4.2.2 Development of yeast whole cell assay

Using the *S. cerevisiae* WAT11 strain harboring the integrated *Arabidopsis thaliana* P450 reductase *ATR1* and the T16H construct, we developed a whole cell assay to detect P450 activity with tabersonine. Yeast cultures expressing T16H were supplemented with 147 μM of tabersonine, and 16-hydroxytabersonine was detected in culture media after 5-10 days (varies between colonies, Fig. 4.4). To determine if other tabersonine-derived alkaloid intermediates could be observed in culture media, two well-characterized enzymes in *C. roseus* vindoline biosynthesis, T16H and 16OMT, were reconstituted in *S. cerevisiae*. Using a different selection marker from pYeDP60 plasmid, 16OMT was cloned into the pYES3-CT plasmid with a tryptophan selection marker. Both plasmids were transformed into yeast. Cultures (5 mL) were grown, protein expression was induced, and the media was supplemented with 147 μM tabersonine. Five days later, selected ion monitoring via LC-MS analysis of the yeast culture media showed a peak with a mass (m/z 353) consistent with that of 16-methoxytabersonine. Direct comparison of the peaks showed that the enzymatic product peak was more hydrophobic than the tabersonine starting material (Fig. 4.4). No enzymatic product was observed in cultures harboring the empty pYES3-CT plasmid that were supplemented with tabersonine.

A.
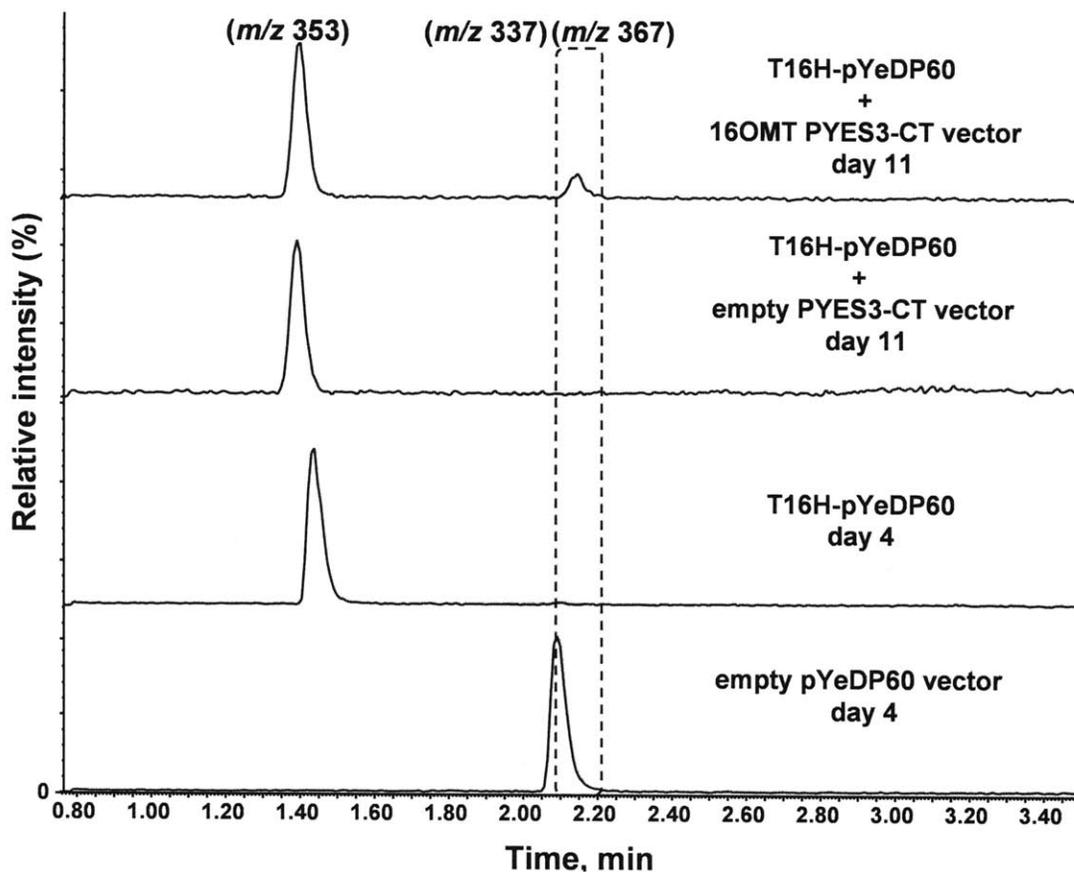
B.
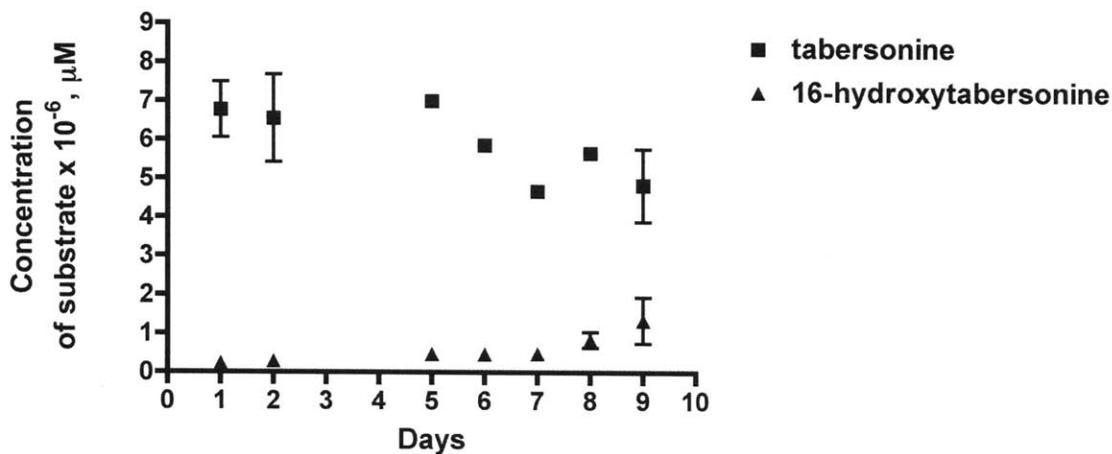
Figure 4.4 A) Selected ion chromatograms of tabersonine (*m/z* 337), 16-hydroxytabersonine (*m/z* 353), and 16-methoxytabersonine (*m/z* 367) in extracted media from yeast cultures expressing either empty pYeDP60 vector (4 days), T16H-pYeDP60 (4 days), T16H-pYeDP60 and empty pYES3-CT vector (11 days), or T16H-pYeDP60 and 16OMT-pYES3-CT (10 days) after being supplemented with tabersonine. B) The accumulation of 16-hydroxytabersonine in the media of yeast cultures expressing T16H nine days after being supplemented with tabersonine.

4.2.3 Protein expression and activity assay

The whole cell yeast assay described in section 4.2.2 was used to screen the candidate P450s CYP71BJ1, 41747, and CYP81Z1, which may be involved in the oxygenation of tabersonine. Yeast cultures that expressed CYP71BJ1 were supplemented with tabersonine, which was completely converted to a more hydrophilic compound with a mass consistent with hydroxylation or epoxidation within 24 hours (Fig. 4.5). Moreover, the CYP71BJ1 product eluted at a retention time distinct from that observed for the T16H product, 16-hydroxytabersonine. No enzymatic product was observed when tabersonine was incubated with yeast cultures harboring the empty pYeDP60 plasmid (Fig. 4.5) or those expressing the other candidate P450s, 41747 and CYP81Z1.

4.2.4 Structural characterization of CYP71BJ1 product

We isolated and purified milligram quantities of the CYP71BJ1 product from 1 L of yeast culture supplemented with 147 µM tabersonine for structural characterization. The enzymatic product was identified as (R)-19-hydroxytabersonine, as evidenced by high resolution mass spectrometry, ^{1}H NMR, ^{13}C NMR, COSY (Fig. 4.6), and HSQC NMR. The ^{1}H NMR spectrum shows a doublet at 1.55 pm and analysis of the COSY spectrum confirmed that this peak represents the proton in the 19-hydroxyl group. The COSY spectrum shows a crosspeak between the doublet at 1.55 ppm corresponding to the proton in the 19-hydroxyl group and the doublet at 2.78 ppm corresponding to the proton at C19 (Fig. 4.6). Furthermore, we observed a crosspeak between the protons at C18 (0.88 ppm) and the proton at C19 (Fig. 4.6) confirming that there is only one proton at C19.
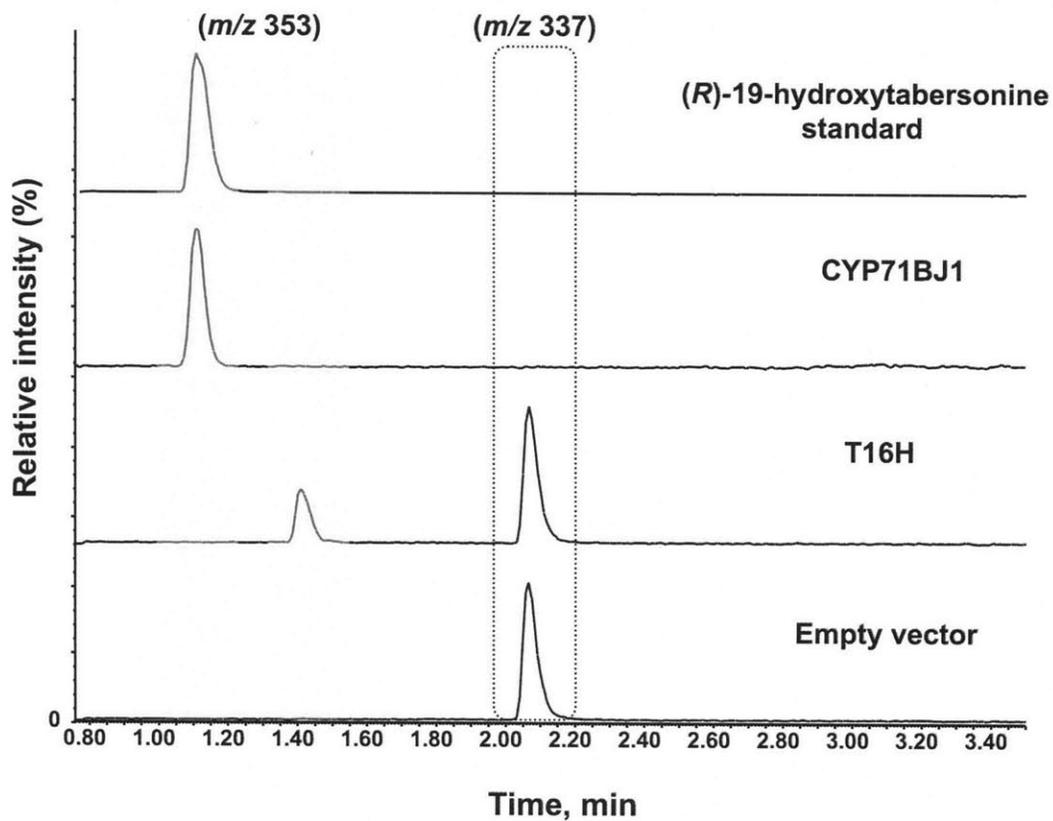
Figure 4.5 Selected ion chromatograms of extracted media from yeast cultures expressing CYP71BJ1, T16H, and empty pYeDP60 vector.
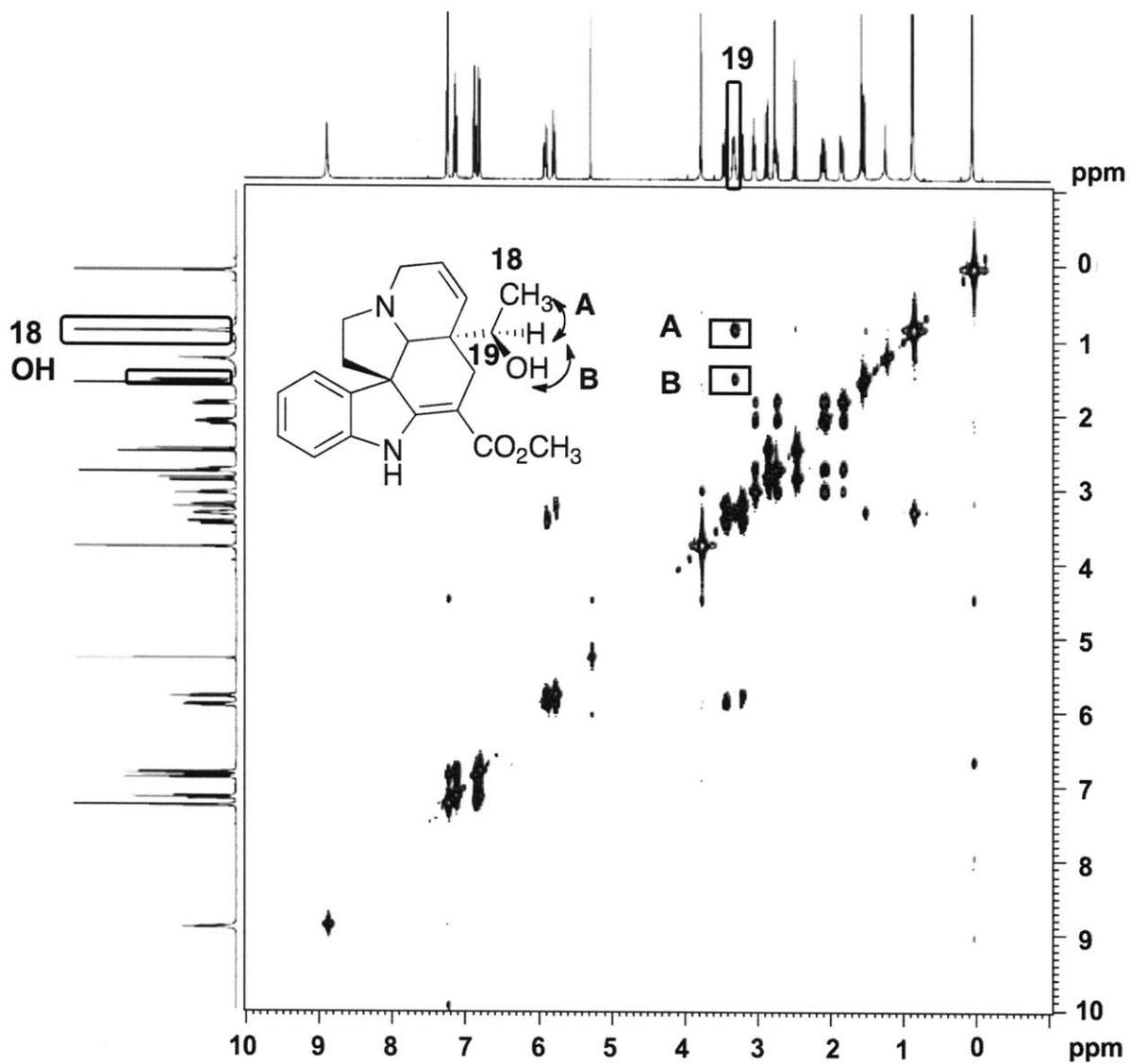
Figure 4.6 COSY of CYP71BJ1 product, (*R*)-19-hydroxytabersonine. The boxed crosspeaks show the correlation between the 19-proton and the C18 methyl group and the proton in the 19-hydroxyl group.

4.2.5 Substrate scope of CYP71BJ1

Microsomes were isolated from yeast cultures expressing CYP71BJ1 and assayed with a variety of alkaloids with different monoterpene indole alkaloid skeletons to determine the substrate scope of the enzyme. Neither lochnericine nor hörhammericine alkaloids are commercially available but both are known to be present in hairy root cultures and in the roots of mature *C. roseus* plants [22, 23, 27]. We isolated small quantities of lochnericine, a potential physiological substrate for CYP71BJ1 (Fig. 4.1), from hairy roots. Although we could not obtain quantities sufficient for NMR characterization, the isolated standard exhibited the expected exact mass and UV signature [28]. Moreover, CYP71BJ1-enriched microsomes converted the isolated compound to a product with an exact mass and UV signature consistent with hörhammericine (Fig. 4.6) [27, 28]. Assays lacking NADPH cofactor or CYP71BJ1-enriched microsomes served as negative controls. Of all substrates tested (Fig. 4.7), CYP71BJ1 only turned over tabersonine and lochnericine. To explore the substrate preference of this enzyme further, we assayed CYP71BJ1-enriched microsomes with a 1:1 mixture of tabersonine and lochnericine (each substrate had a final concentration of 55 μM). After 15 hours of incubation, LC-MS analysis indicated that while both substrates were turned over by the enzyme, all of the lochnericine substrate was consumed, while most of the tabersonine substrate was not converted to product, suggesting that lochnericine is the preferred substrate (Fig. 4.8).
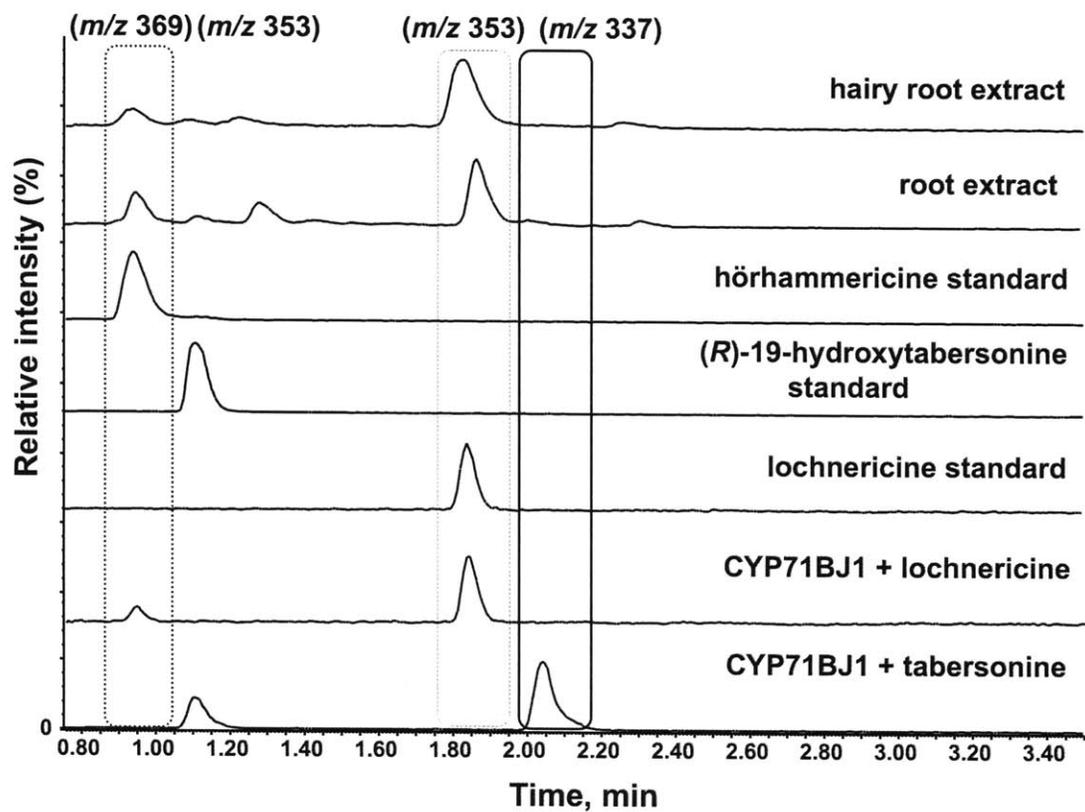
Figure 4.7 Selected ion chromatograms of CYP71BJ1-enriched microsomes that accept tabersonine and lochnericine substrates to produce products that coelute with (*R*)-19-hydroxytabersonine and hörhammericine authentic standards.

yohimbine   catharanthine   vindolinine   norharmane

dihydrotabersonine   16-hydroxytabersonine   16-methoxytabersonine
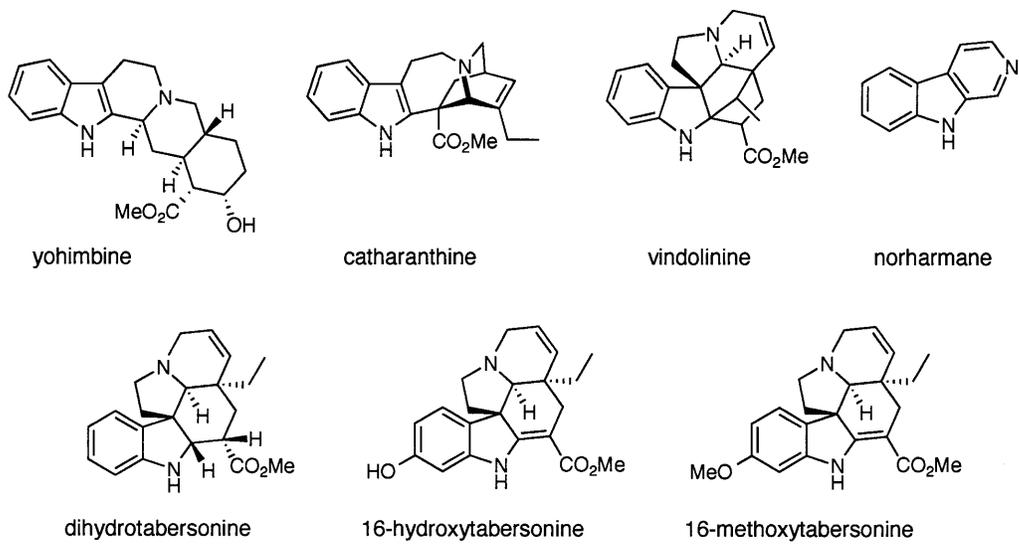
Figure 4.8 Substrates that were not accepted by CYP71BJ1-enriched microsomes.
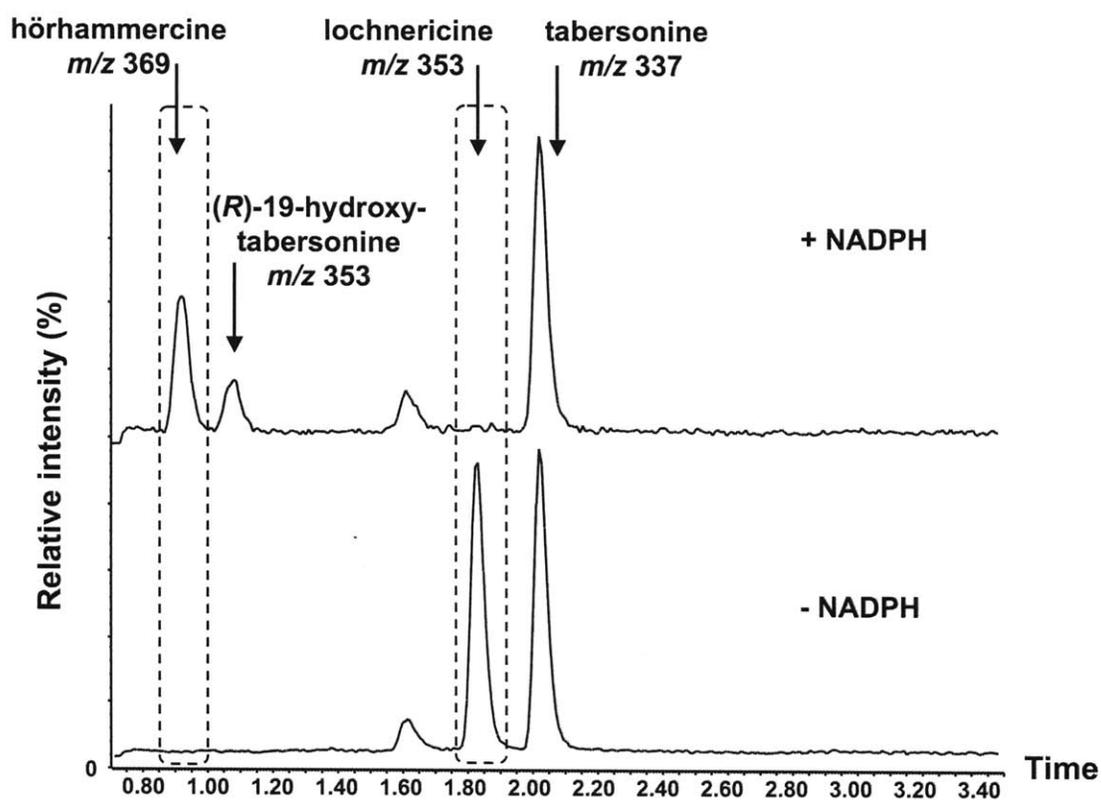
Figure 4.9 Selected ion chromatograms of the formation of (R)-19-hydroxytabersonine (m/z 353) and hörhammericine (m/z 369) from assays of CYP71BJ1-enriched microsomes incubated with a 1:1 mixture of tabersonine (m/z 337) and lochnericine (m/z 353) in the presence and absence of NADPH.

4.2.6 Kinetic analysis and oxygenase inhibition of CYP71BJ1

Steady-state kinetic analysis of CYP71BJ1-enriched microsomes assayed with tabersonine revealed that CYP71BJ1 had a K_m of 300 ± 50 nM and an apparent V_{max} of 5.0 ± 0.2 mM min^{-1} mg^{-1} (Fig. 4.10 A). We also observed reduced activity when increasing concentrations of the aminobenzotriazole (ABT), a P450 monooxygenase suicide inhibitor [29], was added to CYP71BJ1 assays containing 9.2 μM tabersonine. P450 activity decreased to less than 1 % in the presence of 6.4 mM ABT (Fig. 4.10 B).

A.
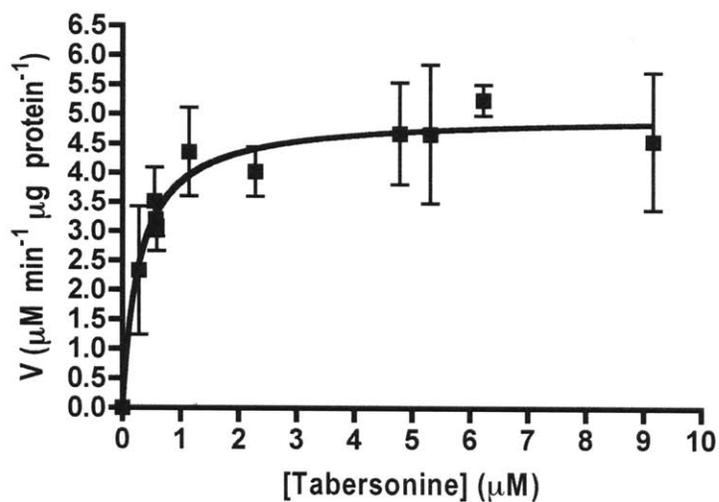
B.
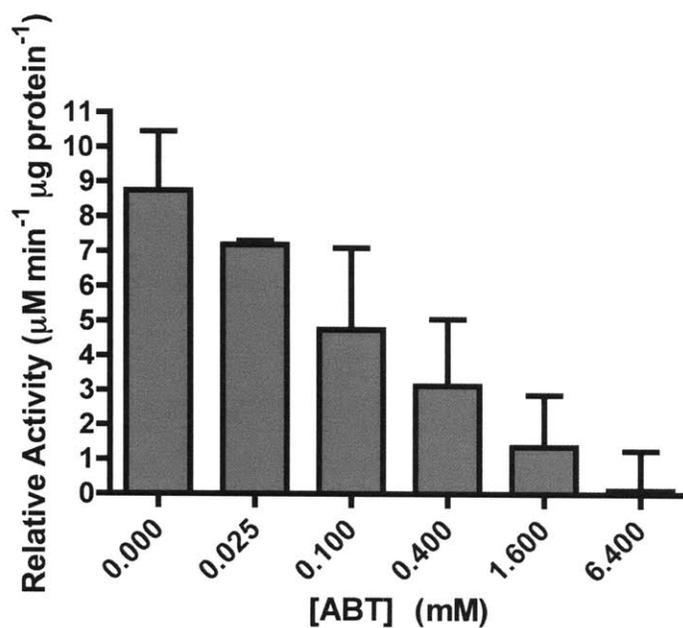
Figure 4.10 A) Steady-state kinetics of CYP71BJ1-enriched microsomes with tabersonine obtained from linear initial rates. B) ABT-monooxygenase inhibition of CYP71BJ1-enriched microsomes with 9.2 μM tabersonine substrate.

4.3 Discussion

This chapter describes how the co-expression of genes serves as a useful tool for predicting the function of orphan P450s. Hierarchical clustering of the *C. roseus* transcriptome revealed three putative P450 genes that cluster with MAT, a gene involved in the acetylation of the 19-position of hörhammericine or minovincinine (Fig. 4.2). These P450s gene candidates may catalyze the oxygenation of tabersonine, a common intermediate in 19-*O*-acetyl hörhammericine biosynthesis. Importantly, apart from *in silico* analyses of the *C. roseus* transcriptome to develop hypotheses about the function of P450 candidates, a simple assay was developed to identify P450s that turn over tabersonine.

Current strategies used to identify and functionally characterize plant P450s employ a combination of methods such as gene expression patterns [18], substrate binding assays [30], *in silico* modeling of active sites [31, 32], mutations [33], and gene silencing [32]. The membrane-bound nature of P450s makes these enzymes difficult to express in heterologous hosts, and the iron-heme cofactor requires suitable redox environments to elucidate their function(s). Yet, in the past 20 years, a number of *in vitro* methods have been developed to express P450s in *S. cerevisiae*, baculovirus, and *Escherichia coli* [25, 34, 35]. After protein expression, affinity purification may be non-trivial, and the isolation of multiple P450-enriched microsomes can be tedious. To avoid problems with P450 expression and purification, we focused on developing a whole cell assay to detect a specific P450 activity using an expression system optimized for P450 expression.

A common cytochrome P450 expression vector, pYeDP60, was used to express T16H in *S. cerevisiae* WAT11, which harbors an integrated plant P450 reductase redox partner. Twenty-four hours after tabersonine was added to the media of yeast cultures, the 16-hydroxytabersonine enzymatic product was detected in the media. To determine if other tabersonine-derived products could be detected in the media, an important detail for the functional characterization of other orphan P450s that turnover tabersonine-derived alkaloids, T16H and 16OMT were co-expressed in yeast. Days after the culture media was supplemented with tabersonine, 16-methoxytabersonine was detected in the media demonstrating that this whole cell assay is useful for detecting tabersonine-derived products.

Hörhammericine biosynthesis is proposed to proceed either through a 6,7-epoxidation of tabersonine to yield lochnericine, which is subsequently hydroxylated at the 19-position [22]. Alternatively, tabersonine can be hydroxylated at the 19-position to yield 19-hydroxytabersonine, which can then undergo a 6,7-epoxidation [22, 23]. Biochemical evidence strongly suggests that both 6,7-epoxidase and the 19-hydroxylase are P450 enzymes [36, 37]. Using the whole cell assay, out of 3 candidate P450 genes CYP71BJ1 was indeed found to catalyze the 19-hydroxylation of tabersonine and lochnericine to form 19-hydroxytabersonine and hörhammericine, respectively.

The CYP71BJ1 enzymatic product was isolated from large-scale whole cell assays for complete structural characterization. ^{1}H NMR chemical shifts of the enzymatic product are consistent with those of the previously reported (*R*)-19-hydroxytabersonine [38, 39].

Additionally, COSY correlations between the proton at C19 and the proton in the 19-hydroxyl group support the formation of a hydroxylated product. Since both (*R*)-19-hydroxytabersonine and hörhammericine share *R*-stereochemistry at C19, we conclude that the 6,7-epoxide in the lochnericine substrate does not direct the stereochemistry of the CYP71BJ1-catalyzed reaction. As such, the order of biosynthetic steps involved in hörhammericine biosynthesis is still ambiguous since CYP71BJ1 accepts lochnericine and tabersonine *in vitro*. The kinetics of CYP71BJ1 were measured only with tabersonine since limited amounts of lochnericine could be isolated from *C. roseus* hairy root cultures. To examine the substrate preference of CYP71BJ1 qualitatively, we set up a competition assay where microsomes were incubated with an equal ratio of lochnericine and tabersonine substrate. Although both substrates were turned over, lochnericine was completely consumed within 15 hours, while most of tabersonine was not turned over, suggesting that lochnericine is the preferred substrate. Morgan et al observed a concurrent reduction in hörhammericine and accumulation of lochnericine in *C. roseus* hairy roots after treatment with the P450 suicide inhibitor ABT, providing further support for lochnericine being the precursor to hörhammericine [37]. The ABT-sensitivity of CYP71BJ1 (Fig. 4.10 B) also provides additional support for the involvement of CYP71BJ1 in hörhammericine biosynthesis *in vivo*, favoring a biosynthetic route via lochnericine. Further studies to disrupt the function of CYP71BJ1 in plant tissue, such as gene silencing in hairy roots via RNA interference, might provide further insight into the physiological role of CYP71BJ1.

The substrate scope of CYP71BJ1 seems to be controlled partially by the presence of the 2,3-double bond, as evidenced by the observation that dihydrotabersonine is not an accepted substrate. Importantly, the vindoline precursors 16-hydroxytabersonine and 16-*O*-methoxytabersonine are also not hydroxylated by CYP71BJ1. Since hydroxylation at the 16-position is the first step in the conversion of tabersonine to vindoline, we speculate that the substrate specificity of CYP71BJ1 may play a role in preventing the 19-hydroxylation of intermediates destined for vindoline biosynthesis. Silencing the gene encoding CYP71BJ1 will provide further insight on the flux of tabersonine throughout *C. roseus*, which will be useful in determining how to increase the metabolic flux towards the commercially valuable vindoline-derived bisindole alkaloids. However, other enzymes, such as the tabersonine 6,7-epoxidase, need to be identified and characterized to be able to efficiently improve the metabolic flux of tabersonine towards bisindole alkaloids. With the *C. roseus* transcriptome now available, the functional characterization of orphan genes involved in alkaloid biosynthesis will hopefully increase, thereby providing better understanding of how the biosynthesis of vindoline and 19-*O*-acetylhörhammericine are regulated *in vitro* and *in planta*.

4.4 Experimental methods

4.4.1 Chemicals, general methods, and analytical techniques

Tabersonine was a generous gift from Viresh Rawal (University of Chicago, Chicago, Il). Unless otherwise noted, all other chemicals were obtained from Sigma-Aldrich. *C. roseus* hairy root cultures were grown according to Morgan et al and *C. roseus* seedlings were grown and elicited with methyl jasmonate as previously reported [25, 37].

UPLC and MS analyses were performed in tandem on an Acquity Ultra Performance BEH C18 column with a 1.7 mm particle size, 2.1 x 100 mm dimension, which was coupled to a Micromass LCT Premier TOF Mass Spectrometer by Waters Corporation (Milford, MA) with electrospray ionization source. Analytes were separated using a 10-50 % acetonitrile: water (0.1 % formic acid) over 5 minutes at a flow rate of 0.5 mL min^{-1}. For MS analyses, the capillary and sample cone voltages were 3000 V and 30 V, respectively. The source and desolvation temperatures were 100 °C and 300 °C, respectively. The cone and desolvation gas flow were 60 and 800 L hr^{-1}. Exact mass measurements were made on a Bruker Daltonics APEXIV 4.7 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (FT-ICR-MS) with electrospray ionization.

HPLC separations were performed on a Lichrosorb reverse phase column (Select B, 25 cm x 4.0 mm column, 5 μm particle size) using 20-50 % acetonitrile: water (0.1 % trifluoroacetic acid) over 20 minutes and a flow rate of 1 mL min^{-1}. Preparative HPLC was performed on a Beckman Coulter System Gold equipped with a 125 solvent module,

and a 166P detector using the same gradient at a flow rate of 19 mL min^{-1}. Analytes were separated on a reverse-phase column (Grace Vydac 2.2cm x 25 cm, 10 µm particle size). ^{1}H NMR, ^{13}C NMR, COSY, and ^{1}H, ^{13}C HSQC spectra were recorded on a Bruker 400 MHz spectrometer.

4.4.2 Sequence and expression data

The *C. roseus* transcriptome and expression mapping data were generated by the NIH-GO Medicinal Plant Consortium at Michigan State University. Illumina sequencing technology was used to analyze transcripts expressed in different *C. roseus* plant tissues such as cell suspension cultures, non-induced and methyl jasmonate induced hairy roots and seedlings, mature plants, and TDC-silenced hairy roots. The data can be accessed at http://medicinalplantgenomics.msu.edu/.

4.4.3 Hierarchical clustering and (co)-expression analysis for P450s in *C. roseus*

Dr. David Liscombe subjected expression levels of annotated P450 candidates (by pfam analysis) and known alkaloid biosynthetic genes, in control and elicited seedlings, and cell suspension cultures (CSC) to hierarchical clustering analyses using CLUSTER 3.0 for Mac (http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm). The resulting dendrogram was visualized with Treeview [26].

4.4.4 Gene cloning

A T16H gene, codon-optimized for expression in *S. cerevisiae* was synthesized (GenScript, Piscataway, NJ) with BamHI and XhoI restriction sites for directional

cloning into pYeDP60. Dr. David Liscombe amplified CYP71BJ1, "Locus_41747", and

CYP81Z1 open reading frames (ORFs) from elicited seedling cDNA (6 days after

elicitation) using sense and antisense primers listed in Table 4.1 containing restriction

sites (underlined) and start codons (bold). Using directional cloning, the candidate P450

ORFs were cloned into pYeDP60, sequenced, and transformed into *S. cerevisiae* WAT11

cells [25].

The 16-hydroxytabersonine 16-*O*-methyltransferase (16OMT) ORF was amplified from

elicited seedling cDNA using the primers listed in Table 4.1. Directional cloning was

used to clone the ORF into pYES3-CT. After the construct was sequenced, lithium

acetate was used to transform the 16OMT-pYES3-CT construct into *S. cerevisiae*

WAT11 cells harboring the T16H construct [40].

Table 4.1 Primers for the amplification of full-length P450 and 16OMT genes

Primer name	Primer sequence 5' to 3'
CYP71BJ1-BamHI_fwd	AAAAAAGGATCCATGTTGTCTTCATTGAAAGAT
CYP71BJ1-EcoRI_rev	AAAAAAGAATTCTTAAAAAATGGTAACCGGAGTTG
41747-BamHI_fwd	AAAAAAGGATCCATGAACTTCTCTCTCACCTCTCCCATTTTCC
41747-EcoRI_rev	AAAAAAGAATTCTTAGTTTCCTTCAACTACAGTTGAGATGCTAGG
CYP81Z1-BamHI_fwd	AAAAAAGGATCCATGGAGGTTTCCTTTTTCTACACCTC
CYP81Z1-SmaI_rev	AAAAAACCCGGGTTATGGGTTATTTTCCAT
16OMT-BamHI_fwd	AATAAAGGATCCATGGATGTTCAATCTGAGG
16OMT-XhoI_fwd	ATTTTATTTTCTCGAGTCAAGGATAAACCTCAATG

4.4.5 Whole cell assay for the detection of P450 activity

Growth conditions and protein expression of T16H (or both T16H and 16OMT) in *S. cerevisiae* WAT11 strain were conducted according to Pompon et al [25]. *S. cerevisiae* WAT11 cell cultures expressing T16H were grown for 5 hours after induction at 28°C, at which point cell cultures were supplemented with 147 µM of tabersonine (from a 73.3 mM DMSO stock solution). Every 24 hours, the optical density at 600 nm was recorded on a UV-visible spectrophotometer. Additionally, 522 µL of yeast cell culture was centrifuged at maximum speed to remove cells and 400 µL of clarified media was extracted into 400 µL ethyl acetate. After vortexing and centrifuging the mixture to separate aqueous and organic layers, 200 µL of the ethyl acetate layer was evaporated and reconstituted in 400 µL of HPLC grade methanol. An aliquot of 5 µL of the methanol solution was diluted into 1 mL of HPLC grade methanol and 2 µL of this mixture were analyzed by LC-MS.

### 4.4.6 Yeast strain, growth, and whole cell assay for activity of candidate P450s

Yeast *S. cerevisiae* WAT11 strain harboring the desired plasmid to express candidate P450 genes was grown in 5 mL cultures and protein expression was induced as previously reported by Pompon et al. The whole cell assay discussed in section 4.4.5 was used to detect product formation in order to characterize CYP71BJ1, "Locus_41747", and CYP81Z1.

4.4.7 Isolation of CYP71BJ1 enzymatic product

Two 500 mL cultures of *S. cerevisiae* WAT11 cells harboring the CYP71BJ1 construct were grown and induced. The medium was supplemented with 147 µM of tabersonine after 5 hours of protein expression. Cells were removed 24 hours later by centrifugation at 3810 x *g* for 15 min. The medium was then extracted with 1 L of ethyl acetate three times and concentrated. The CYP71BJ1 enzymatic product was purified by column chromatography using a mobile phase of hexanes: ethyl acetate (1:1) for the first column, and dichloromethane: methanol (99.5: 0.5) was used for the second column. ^1H NMR, ^{13}C NMR, COSY, and ^1H, ^{13}C HSQC spectra were recorded for the isolated compound.

(R)-19-hydroxytabersonine
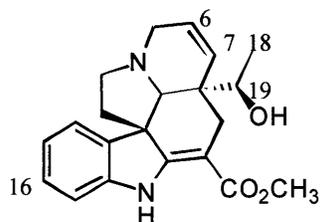
^{1}H NMR (CDCl$_3$): d 8.91 (1H, s), 7.27 (1H, d, J = 6.9), 7.15 (1H, t, J = 7.7), 6.89 (1H, t, J = 7.5), 6.82 (1H, d, J = 7.8), 5.92 (1H, dd, J = 4.9, 10.2), 5.80 (1H, d, J = 9.9), 3.79 (3H, s), 3.47 (1H, dd, J = 4.9, 15.9), 3.30 – 3.35 (1H, m), 3.23 (1H, d, J = 16.2), 3.06 (1H, t, J = 7.0), 2.89 (1H, dd, J = 15.4), 2.78 (1H, s), 2.73-2.79 (1H, m), 2.49 (1H, d, J = 15.4), 2.07-2.14 (1H, m), 1.83-1.87 (1H, m), 1.55 (1H, d, J = 3.3), 0.88 (3H, d, J = 6.4);

^{13}C NMR (CDCl$_3$): d 168.54, 166.38, 143.03, 137.71, 129.44, 127.81, 126.31, 121.54, 120.93, 109.37, 91.26, 67.00, 66.62, 55.57, 51.29, 51.02, 50.10, 46.31, 43.90, 27.57,

17.36; ESI-MS($C_{21}H_{24}N_2O_3^+$) *m/z* calculated: 353.1860 [M+H]$^+$, found: 353.1848

[M+H]$^+$. The UV absorbance maxima were observed at 229 nm, 296 nm, and 331 nm.

4.4.8 Isolation of lochnericine and hörhammericine

Alkaloids were extracted from *C. roseus* hairy roots (5 g) by grinding roots in 50 mL of

HPLC-grade methanol followed by sonication for one hour. After filtration, the methanol

extract was analyzed by HPLC at 330 nm using a 20-50 % acetonitrile: water (0.1 %

trifluoroacetic acid) gradient over 20 minutes. To isolate lochnericine and

hörhammericine, preparative HPLC was performed at 330 nm using the same gradient.

The absorbance maxima for isolated lochnericine were observed at 226 nm, 296 nm, and

330 nm [28]. Additionally, the exact mass for lochnericine was calculated to be 353.1860

[M+H]$^+$ and measured to be 353.1866 [M+H]$^+$. The absorbance maxima for isolated

hörhammericine were observed at 228, 296, 330 nm. The exact mass for hörhammericine

was calculated to be 369.1809 [M+H]$^+$ and measured to be 369.1804 [M+H]$^+$ [27, 28].

4.4.9 Preparation of microsomes, CYP71BJ1 kinetics, and substrate specificity assays

Yeast microsomes enriched with CYP71BJ1 or T16H were prepared using the high-

density procedure according to Pompon et al. [25]. The total microsomal protein content

was determined using a bichinchoninic acid assay by Pierce (Rockford, IL). CYP71BJ1

assays were prepared in a final volume of 100 μL containing 55 μg of microsomal

protein, 1 mM NADPH, 4 mM dithiothreitol, and 100 μM ajmaline internal standard in

100 mM sodium phosphate buffer (pH 7.0). Assays were initiated by the addition of

30 μM substrate unless otherwise noted. The following substrates were assayed with

microsomes: tabersonine, yohimbine, catharanthine, dihydrotabersonine, norharmane, lochnericine, 16-hydroxytabersonine (3 μM), 16-methoxytabersonine (300 nM), and vindolinine (5 μM). Dihydrotabersonine was prepared as reported by Liscombe et al [24]. Assays lacking microsomes or NADPH served as negative controls. After incubation of CYP71BJ1 with substrates at 30 °C for one hour or overnight after, 10 % of the assay volume was quenched with 1 mL of HPLC-grade methanol. Assays were clarified by centrifugation for 5 minutes in a microcentrifuge and then analyzed by LC-MS.

4.4.10 Steady-state enzyme kinetics

A DMSO stock (7.3 mM) of tabersonine was prepared and then diluted with water to a final concentration of 5.9 mM. Serial dilutions of this stock were used for enzyme assays. Enzyme assays (0.1 mL reaction volume) contained 1.2 mg of CYP71BJ1 enriched microsomes, 1 μM ajmaline (for an internal standard), 1 mM NADPH in sodium phosphate buffer (100 mM, pH 7.0), and 4 mM dithiothreitol. Reactions were initiated by the addition of tabersonine (290 nM, 550 nM, 570 nM, 590 nM, 1.2 μM, 2.3 μM, 4.8 μM, 5.3 μM, 6.2 μM, 9.2 μM) and incubated at 30 °C. Timepoints were chosen such that the rate of product formation was linear, to ensure accurate measure of initial rates. The initial rates were determined from the slope of the line made from the linear fit of four normalized data points (between 3 and 15 minutes after initiation of the reaction) for each substrate concentration. The reactions were analyzed by LC-MS and values of V_{max} and K_m were estimated using nonlinear fitting (OriginPro 7.0, Northampton, MA). The errors reported are based on the 95 % confidence of the standard deviations of three independent experiments.

4.4.11 Inhibition assays

Aqueous stocks of the P450 suicide inhibitor ABT ranging from 64 mM to 250 μM were used for inhibition assays. A concentration of 9.2 μM of tabersonine substrate in the presence of 0, 25 μM, 100 μM, 400 μM, 1.6 mM, and 6.4 mM ABT was used. The errors reported are based on the 95 % confidence of the standard deviations of three independent experiments.

4.5 References

1.   Nielsen, K.A., Møller B. L., *Cytochrome P450s in plants*, in *Cytochrome P450: Structure, mechanism, and biochemistry*, P.R. Ortiz de Montellano, Editor 2005, Kluwer Academic/Plenum Publishers: New York. p. 553-583.

2.   Nelson, D.R., *Plant cytochrome P450s from moss to poplar*. Phytochemistry Rev., 2006. **5**: p. 193-204.

3.   Mizutani, M. and D. Ohta, *Diversification of P450 genes during land plant evolution*. Annu. Rev. Plant Biol., 2010. **61**: p. 291-315.

4.   Chapple, C.L., *Molecular-genetic analysis of plant cytochrome P450-dependent monoxygenases*. Annu. Rev. Plant Physiol. Plant Mol. Biol., 1998. **49**: p. 311-43.

5.   Meunier, B., S.P. de Visser, and S. Shaik, *Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes*. Chem. Rev., 2004. **104**: p. 3947-3980.

6.   Groves, J.T., *Models and mechanisms of cytochrome P-450 action*, in *Cytochrome P450: Structure, mechanism, and biochemistry*, P.R. Ortiz de

Montellano, Editor 2005, Kluwer Academic/Plenum Publishers: New York, NY. p. 1-44.

7.   Nelson, D.R., et al., *Comparative genomics of rice and arabidopsis. Analysis of 727 cytochrome P450 genes and pseudogenes from a monocot and a dicot.* Plant Physiol., 2004. **135**(2): p. 756-772.

8.   Schuler, M.A., et al., *Arabidopsis cytochrome P450s through the looking glass: a window on plant biochemistry.* Phytochemistry Rev., 2006. **5**: p. 205-237.

9.   Jensen, K. and B.L. Møller, *Plant NADPH-cytochrome P450 oxidoreductases.* Phytochemistry, 2010. **71**(2-3): p. 132-141.

10.   O'Connor, S.E. and J. Maresh, *Chemistry and biology of monoterpene indole alkaloid biosynthesis.* Nat. Prod. Rep., 2006. **23**(4): p. 532-547.

11.   Facchini, P.J. and V. De Luca, *Opium poppy and Madagascar periwinkle: model non-model systems to investigate alkaloid biosynthesis in plants.* Plant J, 2008. **54**(4): p. 763-84.

12.   Kaltenbach, M., et al., *Flavonoid hydroxylase from Catharanthus roseus: cDNA, heterologous expression, enzyme properties and cell-type specific expression in plants.* The Plant Journal, 1999. **19**(2): p. 183-193.

13.   Hotze, M., G. Schröder, and J. Schröder, *Cinnamate 4-hydroxylase from Catharanthus roseus and a strategy for the functional expression of plant*

cytochrome P450 proteins as translational fusions with P450 reductase in *Escherichia coli.* FEBS Lett., 1995. **374**(3): p. 345-350.

14.   Collu, G., et al., *Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis.* FEBS Lett., 2001. **508**(2): p. 215-220.

15.   Irmler, S., et al., *Indole alkaloid biosynthesis in Catharanthus roseus: new enzyme activities and identification of cytochrome P450 CYP72A1 as secologanin synthase.* Plant J., 2000. **24**(6): p. 797-804.

16.   Schröder, G., et al., *Light-induced cytochrome P450-dependent enzyme in indole alkaoid biosynthesis: tabersonine 16-hydroxylase.* FEBS Lett., 1999. **458**(2): p. 97-102.

17.   Matsuno, M., et al., *Evolution of a novel phenolic pathway for pollen development.* Science, 2009. **325**(5948): p. 1688-1692.

18.   Ehlting, J., et al., *An extensive (co-)expression analysis tool for the cytochrome P450 superfamily in Arabidopsis thaliana.* BMC Plant Biology, 2008. **8**: p. 47.

19.   Laflamme, P., B. St-Pierre, and V. De Luca, *Molecular and biochemical analysis of a Madagascar periwinkle root-specific minovincinine-19-hydroxy-O-acetyltransferase.* Plant Physiol., 2001. **125**: p. 189-198.

20.   Furuya, T., et al., *Biotransformation of tabersonine in cell suspension cultures of Catharanthus roseus.* Phytochemistry, 1992. **31**(9): p. 3065-3068.

21. Shanks, J.V., et al., *Quantification of metabolites in the indole alkaloid pathways of Catharanthus roseus: Implications for metabolic engineering.* Biotechnol. Bioeng., 1998. **58**(2-3): p. 333-338.

22. Kutney, J.P., et al., *Alkaloid production in Catharanthus roseus cell cultures: Isolation and characterization of alkaloids from one cell line.* Phytochemistry, 1980. **19**(12): p. 2589-2595.

23. Morgan, J.A. and J.V. Shanks, *Inhibitor studies of tabersonine metabolism in C. roseus hairy roots.* Phytochemistry, 1999. **51**(1): p. 61-68.

24. Liscombe, D.K., A.R. Usera, and S.E. O'Connor, *Homolog of tocopherol C methyltransferases catalyzes N methylation in anticancer alkaloid biosynthesis.* Proc. Natl. Acad. Sci. USA, 2010. **107**(44): p. 18793-18798.

25. Pompon, D., et al., *Yeast expression of animal and plant P450s in optimized redox environments.* Methods in Enzymology, 1996. **272**(6): p. 51-64.

26. Saldanha, A.J., *Java treeview--extensible visualization of microarray data.* Bioinformatics, 2004. **20**(17): p. 3246-3248.

27. Abraham, D.J., et al., *Structure elucidation and chemistry of Catharanthus alkaloids. IV. Structures of hoerhammericine and hoerhammerinine.* J. Org. Chem., 1969. **34**(6): p. 1575-1576.

28. Hisiger, S. and M. Jolicoeur, *Analysis of Catharanthus roseus alkaloids by HPLC.* Phytochemistry Reviews, 2007. **6**(2-3): p. 207-234.

29.   Ortiz de Montellano, P.R. and J.M. Mathews, *Autocatalytic alkylation of the cytochrome P-450 prosthetic haem group by 1-aminobenzotriazole.*. Biochem J., 1981. **195**(3): p. 761-764.

30.   Olry, A., et al., *A medium-throughput screening assay to determine catalytic activities of oxygen-consuming enzymes: a new tool for functional characterization of cytochrome P450 and other oxygenases.* Plant J., 2007. **51**(2): p. 331-340.

31.   Kemp, C.A., J.-D. Maréchal, and M.J. Sutcliffe, *Progress in cytochrome P450 active site modeling.* Arch. Biochem. Biophys., 2005. **433**(2): p. 361-368.

32.   Mao, Y.-B., et al., *Silencing a cotton bollworm P450 monooxygenase gene by plant-mediated RNAi impairs larval tolerance of gossypol.* Nat. Biotech., 2007. **25**(11): p. 1307-1313.

33.   Negishi, M., et al., *Structural flexibility and functional versatility of cytochrome P45 and rapid evolution.* Mutation Research, 1996. **350**(1): p. 43-50.

34.   Barnes, H.J., *Maximizing expression of cytochrome P450s in Escherichia coli.* Meth. Enz., 1996. **272**: p. 3-14.

35.   Lee, C.A., T.A. Kost, and C.J. Serabjit-Sing, *Recombinant baculovirus strategy for coexpression of functional human cytochrome P450 and P450 reductase.* Meth. Enz., 1996. **272**: p. 86-95.

36.   Rodriguez, S., et al., *Jasmonate-induced epoxidation of tabersonine by a cytochrome P-450 in hairy root cultures of Catharanthus roseus.* Phytochemistry, 2003. **64**: p. 401-409.

37.   Morgan, J.A. and J.V. Shanks, *Inhibitor studies of tabersonine metabolism in C. roseus hairy roots.* Phytochemistry, 1999. **51**: p. 61-68.

38.   Langois, N. and R.Z. Andriamialiosa, *Studies on vindolinine. 6. Partial synthesis of aspidospermane-type alkaloids.* J. Org. Chem., 1979. **44**(14): p. 2468-2471.

39.   Kohl, W., B. Witte, and G. Hofle, *Alkaloids from catharanthus tissue cultures. II.* Z. Naturforsch. B, 1981. **36**: p. 1153-1162.

40.   Gietz, R.D. and R.A. Woods, *Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method.* Methods Enzymol., 2002. **350**: p. 87-96.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

5.1 Conclusions

Three unique studies described in this thesis highlight the benefits of understanding the biosynthetic machinery that plants use to produce functionally diverse natural products. The first study described in Chapter 2 utilized biochemical information derived from kinetic isotope effects, rate dependence on pH, and structural data to propose a mechanism by which strictosidine synthase (STR) catalyzes an asymmetric Pictet-Spengler reaction, a key reaction in the synthesis of alkaloids. The mechanism of the Pictet-Spengler reaction has been debated over the past 50 years [1-4] and by studying how Nature evolved an enzyme to catalyze this chemistry, we were also able to gain more insight on the mechanism of the nonenzymatic reaction.

Isotope effects were used to study the enzymatic mechanism, and the rate-controlling step of the reaction, surprisingly, was the final deprotonation step (Fig 2.1 A). Furthermore, the enzymatic rate dependence on pH was measured with secologanin and tryptamine/[2-^{2}H]-tryptamine, and the apparent pKa values of two ionizable residues involved in catalysis were estimated to be 4.6 and 8.28. The pKa value of 4.6 implicated the Glu309 residue, which structural characterization has shown to be positioned closely to the amine and aldehyde moieties of tryptamine and secologanin, respectively. The crystal structure of STR and mutagenesis data did not reveal any other ionizable amino acid residues that could be involved in catalysis with a pKa value of 8.27. However, this value could represent the involvement of the iminium intermediate or protonated tryptamine substrate in the reaction mechanism. Based on these results, we proposed a mechanism for the enzymatic Pictet-Spengler reaction, which is shown in Chapter 2, Fig. 2.8.

To identify whether the spiroindolenine intermediate (Fig. 2.1 A, step 4a) is formed in the STR-catalyzed reaction, *ab initio* calculations were used to determine the ground and transitions states from the iminium intermediate formed in a model Pictet-Spengler reaction. The computational results indicated that forming the spiroindolenine intermediate is a slightly higher energy process than directly forming the 6-membered ring intermediate from the iminium. Furthermore, no transition state was found between the spiroindolenine intermediate and the 6-membered ring intermediate, suggesting that the proposed 1,2-shift from the spiroindolenine intermediate to form the 6-membered ring intermediate does not occur; therefore, even if the spiroindolenine intermediate is formed during the reaction, it is nonproductive.

Since the time of this study, several other mechanistic aspects of enzymatic and nonenzymatic alkaloid synthesis have been explored using isotope effects to determine if deprotonation is rate-controlling. Examples can be seen in the recently published work on the mechanisms of norcoclaurine synthase [5], epimerization of cis to trans trisubstituted 1,2,3,4-tetrahydro-β-carbolines formed from the condensation of tryptophan derivatives and aldehydes [6], and prenyl transferases [7]. These studies collectively demonstrate the significant mechanistic insights that can be gained from studying the chemistry evolved by Nature.

Chapter 3 highlights the catalytic plasticity of the β-propeller fold that, based on the few proteins that have been structurally characterized, seems to be conserved within the lactonohydrolase/paraoxonase superfamily. Global computational comparison of the

sequences of proteins within this superfamily revealed that a large subset of proteins have conserved metal coordinating residues and may catalyze metal-dependent hydrolysis [8]. The functionally characterized proteins that belong to this superfamily appear mainly to catalyze ester hydrolysis whereas STR catalyzes the highly selective Pictet-Spengler reaction [8, 9]. To experimentally determine whether the STR β-propeller fold could be used as a scaffold to introduce hydrolase activity, rational mutagenesis was used to create three STR hybrid proteins, P1, P2, and P3. These hybrids contained metal-binding residues and the conserved residues amongst closely related STR homologs that have been identified as putative hydrolases.

Mutations made to STR were based on sequence alignments of closely related STR homologs from *Arabidopsis thaliana* and *Vitis vinifera*, homology models of STR homologs from *V. vinifera*, and the crystal structures of STR and paraoxonase (PON1). Gratifyingly, one of the STR homologs from *V. vinifera* and the three STR hybrid proteins displayed hydrolase activity with *p*-nitrophenyl acetate at rates that were above levels of the spontaneous hydrolysis of *p*-nitrophenyl acetate that occurs in solution. Additionally, the P3 *Cr*STR hybrid became inactive in the presence of the metal chelator EDTA and calcium inhibitor terbium chloride, indicating that the hydrolase activity observed is calcium-dependent. However, upon mutation of four of the five putative calcium-binding residues in P3 to alanine, hydrolase activity was not eliminated, suggesting that the mechanism of hydrolysis for P3 and the other hybrids is distinct from that of PON1. The lack of two histidine residues in the putative active sites of *Cr*STR hybrids that have been implicated in the catalytic mechanism of PON1 further

178

emphasizes the differences in catalytic mechanisms between STR hybrids and PON1. Moreover, when three of the four metal-binding residues in the *V. vinifera* hydrolase were mutated to alanine, the triple mutants were still active, suggesting possible mechanistic similarities between the *V. vinifera* hydrolase and STR hybrids. Additional mutations that eliminate hydrolase activity need to be made to validate this hypothesis, and other substrates also need to be assayed with P3 and the *V. vinifera* hydrolase to gain more insight on the native substrates for these proteins.

Typically, in the evolution of proteins within a superfamily, aspects of a catalytic mechanism is conserved [10]. For example, in the mechanistically diverse enolase superfamily, the base abstraction of protons alpha to the carboxylate moieties to form enolates is conserved [11, 12]. Yet mechanistic similarities of hydrolysis or Pictet-Spengler condensation between the enzymes in the lactonohydrolase/paraoxonase family are not completely obvious as the substrates and chemical mechanisms vary. Both reactions do appear to involve a nucleophilic attack of an electrophile. Once the mechanisms of the *V. vinifera* hydrolase and STR hybrids are elucidated, more insight can be gained on how the Pictet-Spengler reaction evolved from hydrolytic reactions. Structural characterization of the *V. vinifera* hydrolase and STR hybrids may also help elucidate potential mechanistic differences between PON1 and other characterized members of the lactonohydrolase/paraoxonase superfamily.

The final example in this thesis that emphasizes the importance of studying plant biosynthetic enzymes is described in Chapter 4. Tabersonine is an alkaloid that can be

transformed into several oxygenated species in addition to vindoline biosynthetic intermediates such as 16-hydroxytabersonine. Morgan and coworkers suggested that the oxygenation of tabersonine to produce other alkaloids, such as lochnericine and hörhammericine, are P450-dependent reactions [13]. This chapter describes the identification and biochemical characterization of a unique P450 enzyme that hydroxylates the tabersonine at the 19-position to form 19-hydroxytabersonine.

Using recently sequenced *C. roseus* transcriptome data, hierarchical clustering and co-expression analyses were used to identify orphan P450 transcripts that have a similar expression profile to genes known to be involved in alkaloid biosynthesis. Three full-length putative P450 transcripts were found to cluster with minovincinine 19-hydroxy-*O*-acetyltransferase (MAT), an enzyme known to catalyze *O*-acetylation of the 19-hydroxyl group of hörhammericine and minovincinine, which are both derived from tabersonine. A whole cell assay was developed whereby the media of yeast cell-culture expressing tabersonine-16-hydrolase or candidate P450s was supplemented with tabersonine. Remarkably, oxygenated tabersonine products could be detected in culture media. Using this assay, the candidate P450, CYP71BJ1, was found to produce a product with a mass consisted with epoxidation or hydroxylation of tabersonine. NMR structural characterization revealed the enzymatic product to be (*R*)-19-hydroxytabersonine.

Substrate specificity studies indicated that lochnericine was the only other substrate accepted by CYP71BJ1. The physiological substrate for CYP71BJ1 is not clear as both lochnericine and tabersonine substrates are accepted by this enzyme. A qualitative

competition assay of CYP71BJ1 incubated with equivalent amounts of lochnericine and tabersonine revealed that lochnericine is the preferred substrate. However, gene silencing of CYP71BJ1 using RNA interference will hopefully provide greater insight on the physiological substrate of this enzyme. Additionally, with the sequences of the *C. roseus* transcriptome now available, the cDNAs that encode other enzymes that are able convert tabersonine into vindoline as well as other oxygenated alkaloids such as lochnericine and hörhammercine will be rapidly identified and characterized. Once these genes are identified, gene silencing may reveal how to redirect tabersonine to upregulate vindoline production, making it easier to access large quantities of vindoline for the semi-synthesis of bisindole alkaloids.

5.2 Future Directions

5.2.1 Reconstituting vindoline biosynthesis

The whole cell assay used to identify P450 candidates in Chapter 4 was initially developed to reconstitute vindoline biosynthesis in yeast (*S. cerevisiae*). Since the tabersonine substrate is commercially available, and five of the six enzymatic steps that convert tabersonine to vindoline are fully characterized, we can begin to test strategies to reconstitute vindoline biosynthesis from the tabersonine starting substrate in microbial hosts. Moreover, preliminary data indicate that the vindoline biosynthetic enzyme T16H (Fig. 4.3), which hydroxylates tabersonine, can also hydroxylate the CYP71BJ1 product, 19-hydroxytabersonine (Fig. 5.1). This result suggests that novel products can be fermented by co-expressing the biosynthetic enzymes 16OMT, T16H, and CYP71BJ1 together in yeast. Examples of the potential products from this synthetic biology

approach are shown in Fig. 5.2. The possible products that could be produced are limited since the substrate scope of 16OMT is highly specific [14]. However, additional efforts to reengineer this enzyme to accept unnatural substrates may also increase the functional diversity of the products formed. Together these experiments will provide a platform for exploring the challenges of reconstitution in yeast such as the efficiency of substrate uptake into the yeast cultures, product export into the media, and the expression ratios of the three enzymes that maximize product yield.
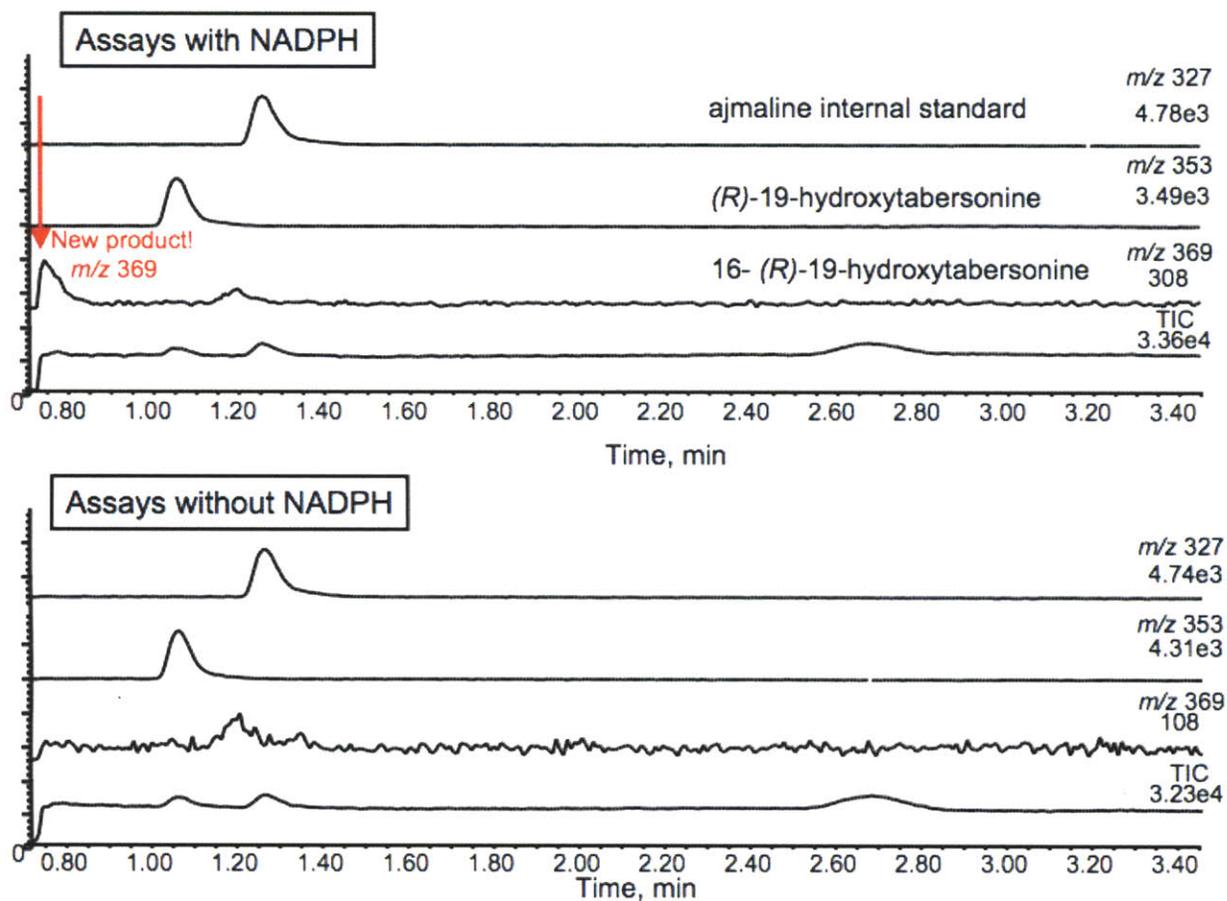
Figure 5.1 Selected ion LC-MS chromatograms of assays of T16H incubated with 19-hydroxytabersonine, 1 mM NADPH, and 4 mM dithiothreitol with and without NADPH for 4.5 hours at 30 °C.
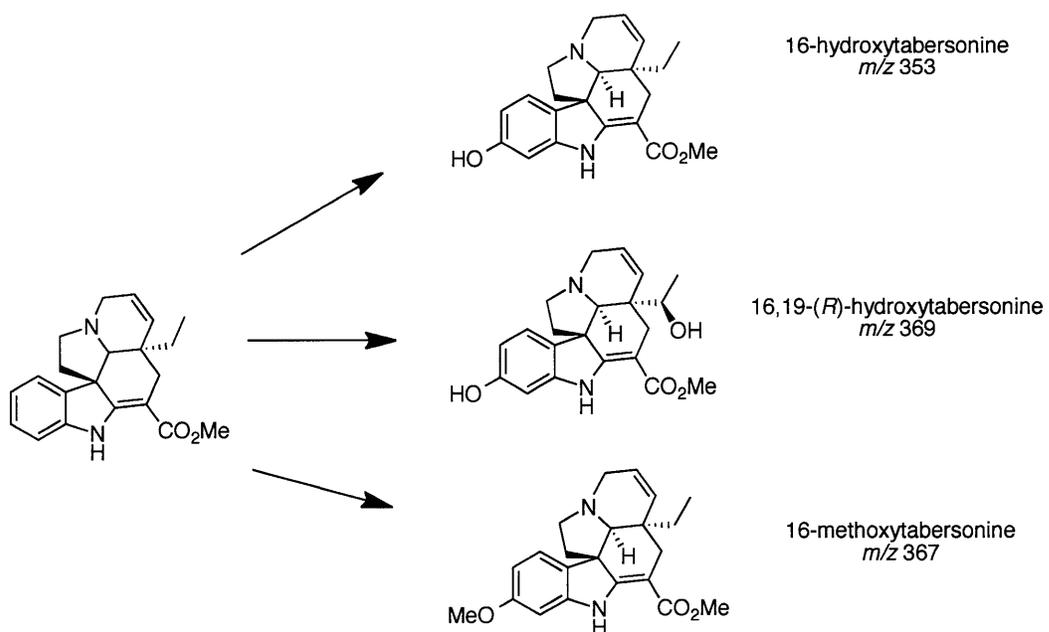
Figure 5.2 Predicted products and corresponding masses produced in the media of yeast cultures co-expressing of T16H, 16OMT, and CYP71BJ1 that have been supplemented with tabersonine.

5.2.2. RNAi-induced gene silencing of CYP71BJ1

There are currently no practical synthetic procedures available to obtain the "blockbuster" antimitotic drugs vinblastine and vincristine. Total synthesis of vinblastine requires a 67-step synthesis that is impractical for obtaining gram-quantities of material [15]. Currently, 46.6 μg g^{-1} dry weight of vinblastine and trace amounts of vincristine are isolated from the aerial parts of *Catharanthus roseus* for clinical use [14]. Alternatively, vinblastine can be generated semi-synthetically from isolated vindoline and catharanthine alkaloids, which can be dimerized in the presence of a peroxidase [16] or iron (III) chloride [17]. While catharanthine can be fermented in large-scale cell suspension cultures, limited success has been made with vindoline due to the enzymes under developmental regulation that are involved in vindoline biosynthesis [18].

One way to obtain vindoline is to isolate the alkaloid from the leaves of *C. roseus*. Since *C. roseus* plants can be regenerated from hairy root cultures that have been infected with *Agrobacterium rhizogenes* [19], we can use RNA-mediated suppression to generate seedlings in which genes, other than T16H, that are involved in metabolizing tabersonine, the precursor to vindoline, are silenced. We hypothesize that silencing the genes involved in the tabersonine branchpoint, thereby closing the biosynthetic "valves" to shunt pathways, will increase levels of tabersonine that can be turned over by downstream enzymes involved in vindoline biosynthesis. The suppression of CYP71BJ1, a tabersonine 19-hydrolase described in Chapter 4, is a good place to start "tuning" vindoline biosynthesis to produce more vindoline that can be isolated from *C. roseus* leaves.

In order to generate CYP71BJ1 RNAi-silencing constructs, 400-500 bp fragments of the CYP71BJ1 gene (Table 5.1) will be cloned into the pHELLSGATE vector using Gateway® technology followed by transformation into *A. rhizogenes*. The pCAMBIA1300 empty vector will also be transformed into *A. rhizogenes* to serve as a control for normal levels of alkaloid production. Following the methods of Runguphan et al, hairy roots with suppressed levels of CYP71BJ1 gene expression can be generated, and LC-MS analysis will demonstrate the effects of silencing CYP71BJ1 on the hairy root metabolome [20]. Additionally, we intend to regenerate the whole plants from CYP71BJ1-suppressed hairy roots using the methods of Choi and coworkers to hopefully observe increased levels of vindoline and/or vindoline intermediates in *C. roseus* [19].

Table 5.1 Primers for amplification of CYP71BJ1 gene fragments for RNA-mediated suppression

Primer name	Primer sequence 5' to 3'
BeginCYP71BJ1_fwd	ATGTTGTCTTCATTGAAAGATTTCTTCGTT
BeginCYP71BJ1_rev	TTGATTTATTTTTTCCAACATTGCCCTTAT
MiddleCYP71BJ1_fwd	ATGTTGGAAAAAATAAATCAAGCTAGTAATAATTCAAGT
MiddleCYP71BJ1_rev	CTCTGTGATTACCCAATGCAAAAGCA
EndCYP71BJ1_fwd	ATGACTAAATTACAAAAAGAGGTGAGAGAAATAGTCG
EndCYP71BJ1_rev	AAAAATGGTAACCGGAGTTGCCA

5.2.3. Heterologous expression and functional characterization of STR human homolog, C20orf3

Chapter 3 described an STR homolog (20 % homology to *C. roseus* STR) found in humans, C20orf3 [21]. Our laboratory has verified that this enzyme does not catalyze the Pictet-Spengler reaction [22]. Recent global sequence comparison of the sequences, structure, and functions of the lactonohydrolase/paraoxonase superfamily revealed conserved metal-binding residues that have been found in proteins that do not display Pictet-Spenglerase activity but are very similar to strictosidine synthase [8]. Based on homology modeling and global computational analyses of the amino acid sequences of members of the lactonohydrolase/paraoxonase family, Babbitt and Hicks speculated that C20orf3 was a putative hydrolase with a conserved metal-binding site comprised of Glu49, Asn147, Asn206, and Asp252 residues [23]. To test this hypothesis, we cloned C20orf3 into pYES2-CT with a C-terminal hexahistidine tag for heterologous expression in yeast. The cells were lysed by sonication and protein was solubilized with 0.1 % tergitol detergent. Ni-NTA affinity chromatography yielded protein that was buffer exchanged into 50 mM HEPES pH 8 buffer with 1 mM $CaCl_2$, 0.2 M NaCl, and 10 % glycerol. Using phenyl acetate as a substrate, the protein was assayed for hydrolase activity using an HPLC-based assay (Fig. 5.3). Notably, we observed hydrolase activity greater than background levels of the spontaneous hydrolysis of phenyl acetate that occurs at pH 8. Furthermore, when this protein was exchanged into buffer containing 10 mM EDTA, the activity diminished to background levels, suggesting that C20orf3 is a metal-dependent hydrolase (Fig. 5.3).

In 2008 Ilhan and coworkers expressed C20orf3 in insect and *E. coli* cells to raise antibodies against different C20orf3 epitopes [24]. These antibodies were used to identify and isolate C20orf3 expressed in human liver and parts of the kidney. The isolated C20orf3 protein was assayed with β-napthyl acetate and phenyl acetate substrates and hydrolysis was observed. Although these were qualitative assays, and experiments to control for background hydrolysis were not performed, these results corroborate the results we obtained with protein heterologously expressed in yeast.

To continue further functional characterization of C20orf3, a faster, higher yielding *E. coli* protein expression system is needed to obtain the quantities of proteins required for rigorous biochemical assays. More substrates can then be tested with this heterologous protein. The substrate preferences of this enzyme may provide clues about its physiological role in human biochemistry. Additionally, mutational studies to identify residues important for hydrolysis will also shed light on the mechanism of C20orf3 hydrolysis, and provide insights on the evolution of hydrolase activity in the lactonohydrolase/paraoxonase superfamily.
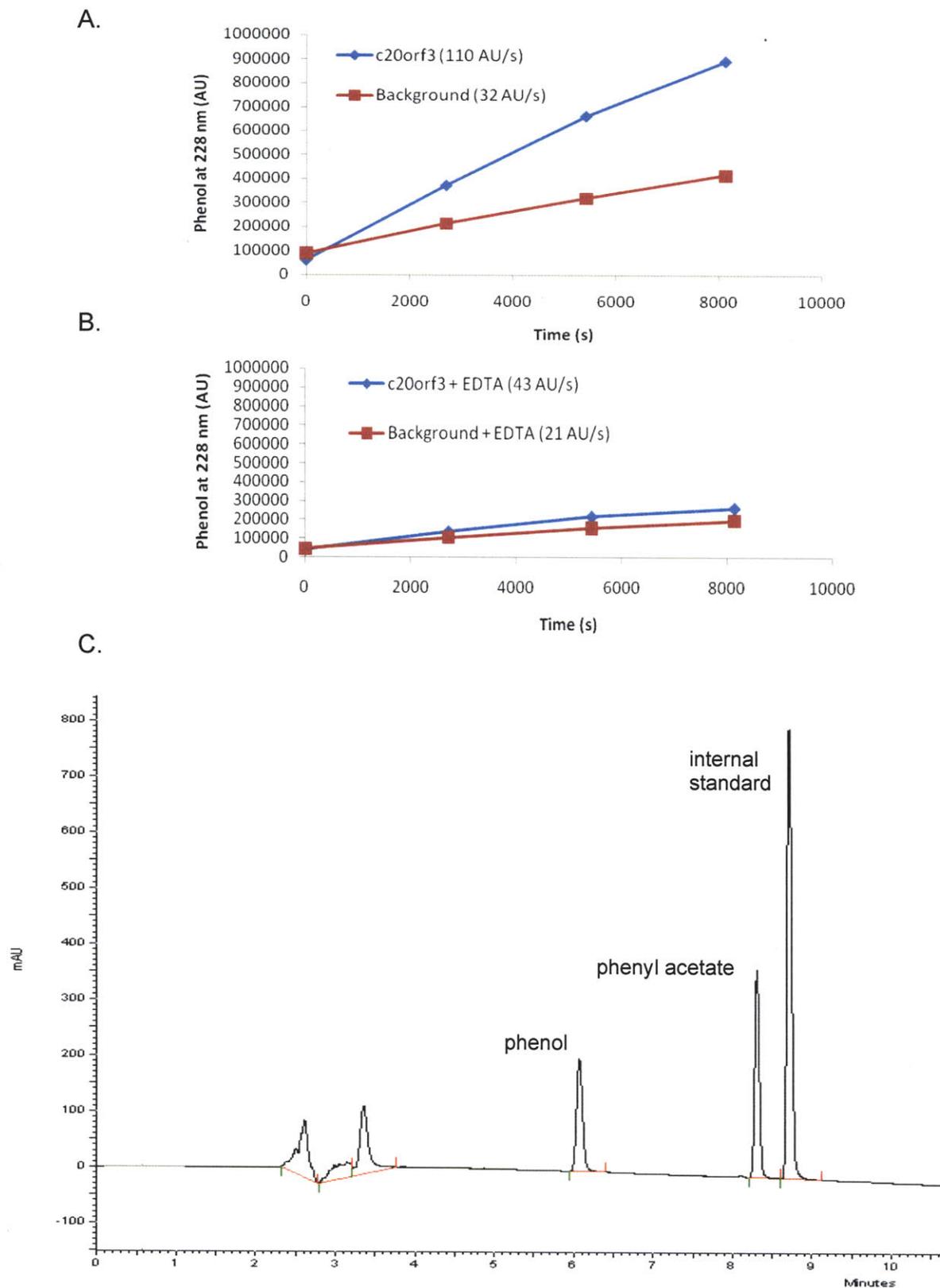
Figure 5.3 Phenol produced from C20orf3-catalyzed hydrolysis of A) phenyl acetate B) phenyl acetate in the presence EDTA. Spontaneous (background) hydrolysis of phenyl acetate has been included in both A and B. C) A representative HPLC trace of the assay of C20orf3 incubated with phenyl acetate at 228 nm.

5.2.4 Evolving *V. vinifera* hydrolase

There are major challenges faced when working with proteins from the lactonohydrolase/paraoxonase family. Poor protein expression, solubility, and stability are characteristics of many of these proteins. To overcome these difficulties, directed evolution has been used to generate stable and active proteins within this family; notably, a high-expressing, stable paraoxonase variant has been evolved [25, 26]. Since the mechanism of the *V. vinifera* hydrolase is not entirely clear, it is critical to generate abundant amounts of this proteins for more detailed studies. As such, directed evolution using DNA shuffling and consensus/ancestor mutagenesis of the closely related STR homologs from *V. vinifera* (CAO67974.1, CAN77945.1, CAO67963.1, CAO66499.1) and STR hybrids is an effective strategy for the production of abundant stable and active protein. DNA shuffling and consensus/ancestor mutagenesis involves iterative rounds of digesting 50-200 bp fragments of different members of a gene family and oligomers of ancestral mutations followed by self-reassembly using PCR, selection, and screening to create a library of mutants. This method of directed evolution was used to evolve PON3, a member of the lactonohydrolase/paraoxonase superfamily, and we plan to follow a similar procedure [27]. Moreover, the high-throughput methods that were successful in evolving paraoxonases described by Aharoni et al will be employed to generate more stable and abundant hydrolases [26].

5.2.5 Functional characterization of closely-related STR homologs

Chapter 3 described the use of five closely related STR homologs that were used as target sequences to convert CrSTR into a hydrolase. We were able to functionally characterize one of the homologs from V. vinifera (CAN77945.1) and we can utilize the same strategy to express the other four homologs from V. vinifera and A. thaliana (CAO67974.1, CAO67963.1, CAO66499.1, NP_177542.1) to determine if they also exhibit hydrolase activity with p-nitrophenyl acetate. Based on the global computational sequence comparison of Hicks and coworkers, we expect that these enzymes will also be metal-dependent [8]. As such, we also expect to observe reduced hydrolase activity in the presence of metal chelators such as EDTA. Having a subset of functionally characterized STR homologs that do not display Pictet-Spenglerase activity will enable us to better understand the sequence and mechanistic requirements of hydrolase activity within the lactonohydrolase/paraoxonase superfamily.

5.3 References

1. Bailey, P.D., *Direct proof of the involvement of a spiro intermediate in the Pictet-Spengler reaction.* J. Chem. Res., Synopses, 1987: p. 202-203.

2. Kowalski, P. and J.L. Mokrosz, *Structure and spectral properties of β-carbolines. Part 9. New arguments against direct rearrangement of the spiroindolenine intermediate into β-carboline system in the Pictet-Spengler cyclization. An MNDO approach.* Bulletin des Societes Chimiques Belges, 1997. **106**(3): p. 147-149.

3.   Ungemach, F. and J.M. Cook, *The spiroindolenine intermediate. A review.* Heterocycles, 1978. **9**: p. 1089-1119.

4.   Ibaceta-Lizana, J.S., et al., *Electrophilic substitution in indoles. 13. The synthesis and rearrangement of 2-deuteriospiro[cyclopentane-3'-indolenine].* J. Chem. Soc. Perkin Trans. II, 1987(9): p. 1221-1226.

5.   Luk, L.Y.P., et al., *Mechanistic studies on norcoclaurine synthase of benzylisoquinoline alkaloid biosynthesis: An enzymatic Pictet-Spengler reaction.* Biochemistry, 2007. **46**(35): p. 10153-10161.

6.   Van Linn, M.L. and J.M. Cook, *Mechanistic studies on the cis to trans epimerization of trisubstituted 1,2,3,4-tetrahydro-β-carbolines.* J. Org. Chem., 2010. **75**(11): p. 3587-3599.

7.   Luk, L.Y.P. and M.E. Tanner, *Mechanism of dimethylallyltryptophan synthase: Evidence for a dimethylallyl cation intermediate in an aromatic prenyltransferase reaction.* J. Am. Chem. Soc., 2009. **131**(39): p. 13932-13933.

8.   Hicks, M.A., et al., *The Evolution of New Function in Strictosidine Synthase-like Proteins*, 2011: San Francisco. p. 1-28.

9.   Kobayashi, M., et al., *Lactone-ring-cleaving enzyme: genetic analysis, novel RNA editing, and evolutionary implication.* Proc. Natl. Acad. Sci. USA, 1998. **95**(22): p. 12787-12792.

10.   Toscano, M.D., K.J. Woycechowsky, and D. Hilvert, *Minimalist active-site redesign: teaching old enzymes new tricks.* Angew. Chem., 2007. **46**(18): p. 3212-3236.

11.   Gerlt, J.A. and P.C. Babbitt, *Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprafamilies.* Annu. Rev. Biochem., 2001. **70**: p. 209-246.

12.   Glasner, M.E., J.A. Gerlt, and P.C. Babbitt, *Evolution of enzyme superfamilies.* Current Opinion in Chemical Biology, 2006. **10**(5): p. 492-497.

13.   Morgan, J.A. and J.V. Shanks, *Inhibitor studies of tabersonine metabolism in C. roseus hairy roots.* Phytochemistry, 1999. **51**(1): p. 61-68.

14.   Levac, D., et al., *Application of carborundum abrasion for investigating the leaf epidermis: molecular cloning of Catharanthus roseus 16-hydroxytabersonine-16-O-methyl transferase.* Plant J., 2008. **53**(2): p. 225-236.

15.   Yokoshima, S., et al., *Stereocontrolled total synthesis of (+)-vinblastine.* J. Am. Chem. Soc., 2002. **124**(10): p. 2137-2139.

16.   Sottomayor, M., et al., *Peroxidase and the biosynthesis of terpenoid indole alkaloids in the medicinal plant Catharanthus roseus (L.) G. Don.* Phytochemistry Rev., 2004. **3**(1-2): p. 159-171.

17.   Ishikawa, H., et al., *Total synthesis of vinblastine, vincristine, related natural products, and key structural analogues.* J. Am. Chem. Soc., 2009. **131**(13): p. 4904-4916.

18.   Moreno, P.R.H., R. Van der Heijden, and R. Verpoorte, *Cell and tissue cultures of Catharanthus roseus: a literature survey II.* Plant Cell Tiss. Org. Cult., 1995. **42**(1): p. 1-25.

19. Choi, P.S., et al., *Plant regeneration from hairy-root cultures transformed by infection with Agrobacterium rhizogenes in Catharanthus roseus.* Plant Cell Rep., 2004. **22**(11): p. 828-831.

20. Runguphan, W., J. Maresh, and S.E. O'Connor, *Silencing of tryptamine biosynthesis for production of nonnatural alkaloids in plant culture.* Proc. Natl. Acad. Sci. USA, 2009. **106**(33): p. 13673-13678.

21. Morita, M., et al., *Genomic construct and mapping of gene for CMAP and identification of a proximal novel gene, BSCv.* Genomics, 2000. **67**(1): p. 87-91.

22. Hatzios, S., *Human alkaloid biosynthesis: Chemical inducers of Parkinson's disease*, in *Chemistry* 2005, Massachusetts Institute of Technology: Cambridge.

23. Babbitt, P.C., *Human protein in strictosidine synthase network*, S.E. O'Connor, Editor 2008: San Francisco.

24. Ilhan, A., et al., *Localization and characterization of the novel protein encoded by C20orf3.* Biochem. J., 2008. **414**(3): p. 485-495.

25. Aharoni, A., et al., *Directed evolution of mammalian paraoxonases PON1 and PON3 for bacterial expression and catalytic specialization.* Proc. Natl. Acad. Sci. USA, 2003. **101**(2): p. 482-487.

26. Aharoni, A., et al., *Directed evolution of mammalian paraoxonases PON1 and PON3 for bacterial expression and catalytic specialization.* Proc. Natl. Acad. Sci. USA, 2004. **101**(2): p. 482-487.

27. Khersonsky, O., et al., *Directed evolution of serum paraoxonase PON3 by family shuffling and ancestor/consensus mutagenesis, and its biochemical characterization.* Biochemistry, 2009. **48**(28): p. 6644-6654.

CV- Lesley-Ann Giddings

Education
2005-2011   **Ph.D.** Chemistry, MIT
2001-2005   **B.A.** Chemistry, Smith College

Research Experience
2011-   Postdoctoral fellow, National Institutes of Health
2005-2011   Research Assistant, MIT
2002-2005   Undergraduate Researcher, Smith College
2003-2005   Summer student, National Institutes of Health

Publications
1. Justin Maresh, **Lesley-Ann Giddings**, Anne Friedrich, Elke A. Loris, Santosh Panjikar, Bernhardt L.Trout, Joachim Stöckigt, Baron Peters, and Sarah E O'Connor (2008). Strictosidine Synthase: Mechanism of a Pictet-Spengler Catalyzing Enzyme. *Journal of the American Chemical Society.* 130, 710-723.
2. John W. Daly, H. Martin Garraffo, Thomas F. Spande, **Lesley-Ann Giddings**, Ralph A. Saporito, David R. Vieites, and Miguel Vences (2008). Individual and Geographic Variation of Skin Alkaloids in Three Species of Madagascan Poison Frogs (Mantella). *Journal of Chemical Ecology.* 34, 252-279.
3. **Lesley-Ann Giddings***, David K. Liscombe*, and Sarah E. O'Connor. A Stereoselective Hydroxylation Step of Alkaloid Biosynthesis by a Unique Cytochrome P450 in *Catharanthus roseus* (2011). *Journal of Biological Chemistry.* 286, 16751-16757.

<u>Submitted manuscript</u>
Michael A. Hicks, Alan E. Barber II, **Lesley-Ann Giddings**, Jenna Caldwell, Sarah E. O'Connor, Patricia C. Babbitt. The Evolution of New Function in Strictosidine Synthase-Like Proteins (2011). *Proteins: Structure, Function, and Bioinformatics.*

<u>Manuscripts in preparation</u>
Lesley-Ann Giddings*, Jenna Caldwell*, Summer Thyme, Peter Bernhardt, Michael Hicks, Patsy Babbit, David Baker, and Sarah E. O'Connor. Reengineering Strictosidine Synthase: Conversion of a Pictet-Spenglerase into a Hydrolase.
Nirina R. Andriamaharavo, H. Martin Garraffo, Thomas F. Spande, **Lesley-Ann Giddings**, Ralph A. Saporito, David R. Vieites, and Miguel Vences. Individual and Geographic Variation of Skin Alkaloids in Three Swamp-Forest Species of Madagascan Poison Frogs (*Mantella*).

* these authors contributed equally.

Selected Honors and Awards
MIT Graduate Fellow, 2005-2006.

National Science Foundation Predoctoral Fellow, 2007-2010.

MIT Masamune Summer Fellow, Summer 2010.

Ford Foundation Predoctoral Fellow, 2006-2007 and 2010-2011.