Exploiting Alkaloid Biosynthesis in Madagascar Periwinkle to Obtain Natural Product Derivatives and New Biocatalysts
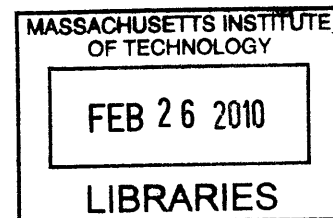
by

Peter Bernhardt

M.Sc. Biotechnology
Kungliga Tekniska Högskolan (KTH), Sweden, 2004

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

FEBRUARY 2010

Signature of Author: _____

Department of Chemistry
November 19, 2009

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 Studies

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

Chair: _____

Alexander M. Klibanov
Professor of Chemistry and Biological Engineering

Thesis Supervisor: _____

Sarah E. O'Connor
Associate Professor of Chemistry

Committee Member: _____

Catherine L. Drennan
Professor of Chemistry and Biology

2

EXPLOITING ALKALOID BIOSYNTHESIS IN MADAGASCAR PERIWINKLE TO OBTAIN NATURAL PRODUCT DERIVATIVES AND NEW BIOCATALYSTS

by

Peter Bernhardt

Submitted to the Department of Chemistry on November 19, 2009
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Chemistry

ABSTRACT

Plant alkaloid biosynthesis produces many natural products with medicinal value. For example, vinblastine and vincristine from *Catharanthus roseus* monoterpene indole alkaloid biosynthesis, and camptothecin derivatives from *Ophiorrhiza pumila* quinoline alkaloid biosynthesis, are anticancer agents currently used in the clinic. Strictosidine synthase is a key enzyme in the biosynthesis of these medicinal natural products, but its narrow substrate scope limits precursor-directed biosynthesis of alkaloid analogs in plant cell cultures. I describe two new assays to monitor strictosidine synthase activity, which enable the rapid screening of enzyme mutant libraries to identify two strictosidine synthase variants that accept new substrates. A transgenic plant cell culture that contains one of these mutants generated "unnatural" monoterpene indole alkaloids in *C. roseus*. I also describe the characterization of *O. pumila* strictosidine synthase, which has considerably broader substrate specificity than the homologous enzyme from *C. roseus*. This alternative catalyst is a candidate enzyme for construction of transgenic cell cultures, and potentially useful as a biocatalyst, since it catalyzes the asymmetric Pictet-Spengler reaction to form tetrahydro-β-carboline pharmacophores. I used computer modeling to propose a model for how strictosidine synthase achieves its high stereoselectivity; this model may be used to engineer a Pictet-Spenglerase that forms the alternative stereoisomer. Such a stereocomplementary catalyst would be useful in biocatalysis, giving the synthetic organic chemist access to both stereoisomers in high enantiomeric purity. Finally, I describe the total synthesis of stereoisomer mixtures of indole alkaloid precursors, and use these mixtures to determine the stereoselectivity of strictosidine synthase and two subsequent enzymes in monoterpene indole alkaloid biosynthesis. The combination of chemical synthesis and the recruitment of enzymes from unrelated biosynthetic pathways could generate diverse alkaloid libraries, containing different stereoisomers, for bioactivity evaluation.

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

Acknowledgments

I would like to thank a number of people in my life that have helped me get to where I am today:

First, Sarah O'Connor, my thesis advisor, for your intelligence, enthusiasm, and support. It has been an invaluable experience to be a part of your research group and I can only hope to someday emulate your vision and clarity.

Alex Klibanov and Cathy Drennan, my thesis committee members, for useful discussions, your insightful questions, and constructive feedback on my research and writing. JoAnne Stubbe, for a most inspiring enzymology class and for providing constructive critique during my second- and third-year oral exams.

The O'Connor lab members, past and present, not only for your help and advice at work, but also for the many informal discussions at the Muddy, the chess games, and the lab outings. A special thanks to Shi Chen and Carmen Galan for getting me started in molecular biology and organic synthesis, respectively; Weerawat "Ricky" Runguphan, for showing that my enzyme works; Nancy Yerkes, for a fruitful collaboration using her reductase; Lesley-Ann Giddings, for taking over some exciting projects on enzyme engineering and evolution; and undergrads Joe Gregg, Ken Loh, and Keri Garel, who I had the pleasure of working with during various parts of my research.

Summer Thyme, in the laboratory of David Baker at the University of Washington, Seattle, for inspiring discussions on homology modeling, enzyme evolution, and protein structure using Rosetta. Michael Hicks, in the laboratory of Patsy Babbit at University of California at San Francisco, for help with structure-based sequence alignments and cluster analysis of the β-propeller superfamily.

Romas Kazlauskas at the University of Minnesota and Karl Hult at Kungliga Tekniska Högskolan who provided me with the tools, opportunities, and guidance that led me to where I am today.

Thanks to all of my friends who provide so much joy. A special thanks to Pam, Scott, Amanda, and Kevin who traveled all the way to Trinidad and Tobago to attend my wedding.

Finally, but most importantly, my family. My wife, Clarisse, for your love and support; for helping me to focus on what is important, even though I sometimes forget; for keeping my spirits up when I am down; and, for your advice both in research and in life. My mother, Maarit, who has always believed in me and that has made many sacrifices to help me and my sister. This thesis is dedicated to my father, Eberhard. I wish that you were able to see me graduate.

Peter Bernhardt

Table of Contents

List of Figures

List of Tables

Conventions and Abbreviations

This thesis adopts the convention of abbreviating amino acids by a three-letter code and nuclotides by a one-letter code. NMR data are listed in the Experimental Data sections and presented by the chemical shift (δ) in parts-per-million (ppm), multiplicity (singlet, s; doublet, d; triplet, t; etc.), and J-coupling constant in Hertz (Hz); NMR spectra are referenced to the NMR-solvents' residual protium (for ^{1}H NMR) or most intense carbon isotope signal (for ^{13}C NMR signal). Compounds are numbered consecutively, starting from **1** in each chapter; only compounds that are illustrated in Figures, Schemes, or Tables (at some point) are numbered.

°C	temperature according to the Celsius scale
i-Bu	*iso*-butyl
t-Bu	*tert*-butyl
BTB	bromothymol blue
cDNA	complementary deoxyribonucleic acid
CHCl₃	chloroform
CNBr	cyanogen bromide
Cy	cyclohexyl
CrSTS	*Catharanthus roseus* strictosidine synthase
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCM	dichloromethane
DMAPP	dimethyl allyl pyrophosphate
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
ee	enantiomeric excess
ESI-MS	electrospray ionization mass spectrometry
EtOAc	ethyl acetate
G10H	geranyl 10-hydroxylase
GI	GenBank index number
Glc	glucose
HPLC	high-performance liquid chromatography
HRMS	high-resolution mass spectrometry
Hz	Hertz, s^{-1}
IPP	isopentenyl pyrophosphate
IPTG	isopropyl β-D-1-thiogalactopyranoside
K_M	Michaelis constant; *the substrate concentration at one half of the maximum reaction velocity ($V_{max}/2$)*
kDa	one thousand Dalton
LB	Luria-Bertani (media)
LC-MS	liquid chromatography coupled to mass spectrometry
MeCN	acetonitrile
MeOD	tetradeuterated (d₄) methanol

MeOH	methanol
MEP	methyl erythritol pathway
MIA	monoterpene indole alkaloid
MF	mating factor
MS	mass spectrometry
NAA	naphthyl acetic acid
NADP$^+$	oxidized nicotinamide adenine dinucleotide phosphate
NADPH	reduced nicotinamide adenine dinucleotide phosphate (hydride)
NMR	nuclear magnetic resonance
NR	neutral red
OD$_x$	optical density at x nm
OpSTS	*Ophiorrhiza pumila* strictosidine synthase
PCR	polymerase chain reaction
PDA	photodiode array (detector)
PDB	protein data bank
pNP	*p*-nitrophenol
ppm	parts per million
i-PrOH	*iso*-propanol
QM	quantum mechanics
rpm	revolutions per minute
RsSTS	*Rauvolfia serpentina* strictosidine synthase
r_t	retention time
SAR	structure-activity relationship
SCMM-U	*Saccharomyces cerevisiae* uracil dropout minimal media
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGD	strictosidine glucosidase
SLS	secologanin synthase
STS	strictosidine synthase
TDC	tryptophan decarboxylase
TEA	triethylamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin-layer chromatography
TOF	time-of-flight
TTN	total turnover number
UPLC	ultra-performance liquid chromatography
UV	ultraviolet (light)
V_{max}	maximum reaction velocity as $[S] \to \infty$

CHAPTER 1

BACKGROUND AND SIGNIFICANCE

1.1 Natural products and biosynthetic pathways

Natural products are organic molecules not directly involved in the normal growth, development or reproduction of organisms.[1] Yet, natural products are prevalent in nature and enable their producers to scavenge resources from the environment, and attract, repel, or destroy other organisms.[1] Although the mode-of-action and assembly of many natural products remain unknown, natural products have been a source and inspiration for human medicines since ancient times. The discovery of new natural products and the ability to make analogs of existing natural products are vital aspects of drug discovery and development.[2-4]

An SAR (structure-activity relationship) study assesses how structural variations of the same natural product scaffold affect biological activity, and is an essential tool to optimizing bioactivity.[5] However, to perform rigorous SAR studies one needs methods to access a large number of different natural product analogs. While total- and semi-synthesis are useful tools to generate these modified structures, metabolic engineering is emerging as a versatile and eco-friendly alternative. Unlike synthetic methods, however, metabolic engineering strategies require a detailed understanding of the processes involved in the assembly of natural products, and the development of tools to alter their production.

Biosynthetic pathways assemble natural products in an orchestrated fashion using a series of enzyme catalyzed and non-catalyzed reactions, starting from building blocks derived

wyerone (antifungal agent)
fatty acid produced by *Vicia faba*

(i)

zearalenone (toxin)
polyketide produced by *Gibberella zeae*

(i)

taxol (anti-cancer agent)
diterpene produced by *Taxus brevifolia*

(iii)

vancomycin (antibiotic)
nonribosomal peptide produced
by *Amycolatopsis orientalis*

(v)

trans-resveratrol (anti-cancer,
anti-inflammatory,
blood-sugar lowering, etc.)
phenylpropanoid produced in several plants

(ii)

(-)-hyoscine (toxin, anticholinergic)
alkaloid produced in several plants

(iv)

streptomycin (antibiotic)
carbohydrate from *Streptomyces griseus*

(vi)

Figure 1.1 Bioactive molecules representative of the six classes of natural products are shown along with their common name, biological activity, and producing organism. The diterpene unit of taxol is highlighted in blue.

from primary metabolism.[1] Natural products fall into six classes based on the primary metabolites used in their construction: (i) fatty acids and polyketides, (ii) aromatic amino acids and phenylpropanoids, (iii) terpenoids and steroids, (iv) alkaloids, (v) peptides including non-ribosomal peptides, and (vi) carbohydrates (Figure 1.1).[1] By changing the chemical transformations that construct natural products, it is, in theory, possible to form any natural product analog.

1.2 Enzyme engineering in natural product biosynthesis

The following sections summarize two biosynthetic pathway classes and describe selected recent examples of enzyme engineering, which are also published in *Curr. Opin. Chem. Biol.* **2009**, *13*, 35-42.[6]

I categorize natural product biosynthetic pathways into megasynthases or dissociated-type pathways; each category offers unique challenges and opportunities for enzyme engineering as a tool to generate natural product derivatives (Figure 1.2).[6]

1.2.1 Megasynthases

Megasynthases, large protein complexes that shuttle biosynthetic intermediates between enzyme active sites, include polyketide synthases (PKSs)[7] and non-ribosomal peptide synthetases (NRPSs).[8] The bacteria and fungi that harbor megasynthases often encode these large enzyme complexes in colinear gene clusters that, typically, contain all of the genes necessary for the assembly of a natural product. The colinearity of these clusters

A   **Natural product megasynthases**

Module 1

domain-domain interactions

domain

CO_2H

megasynthase product

$[$ *tailoring steps* → final natural product $]$

module-module interactions

CO_2H
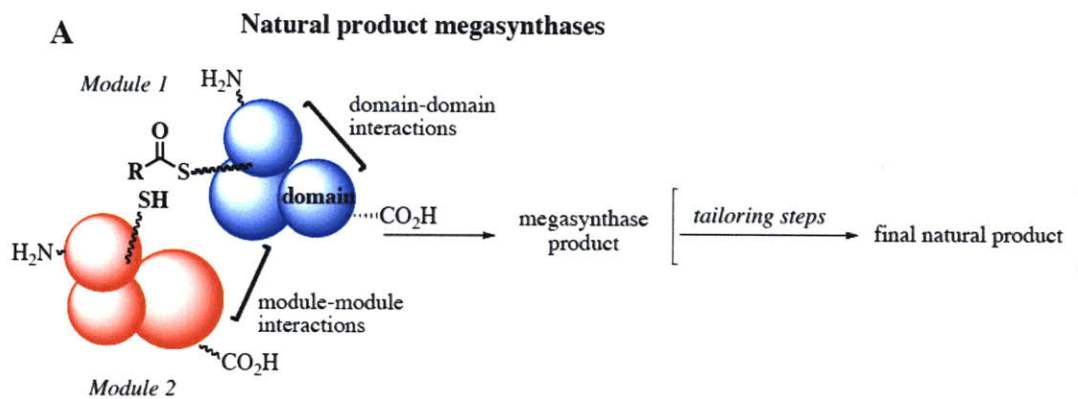
Module 2

Typical natural product classes:
- type I and II polyketides
- nonribosomal peptides

Opportunities for enzyme engineering:
- Engineer individual domains
- Optimize interactions after domain or module replacements
- Reengineering of tailoring enzymes

Challenges for enzyme engineering:
- Maintain proper protein-protein interactions
- Limited number of products from one megasynthase

B   **Dissociated biosynthetic pathways**

Enzyme-substrate complementarity

Substrate

Enzyme 1

enzyme transformations

central intermediate
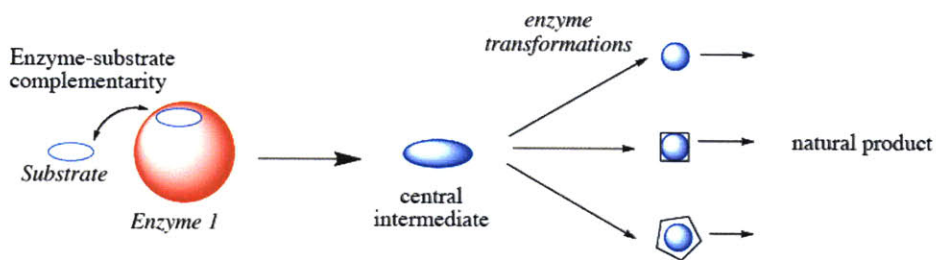
natural product

Typical natural product classes:
- alkaloids
- isoprenoids
- flavonoids
- type III polyketides

Opportunities for enzyme engineering
- Treat enzymes in isolation
- Generate larger libraries due to divergent pathways

Challenges for enzyme engineering
- Limited substrate scope
- Enzyme and intermediate compartmentalization
- Intermediate shuttling
- Often no gene clustering that facilitate cloning

Figure 1.2 Biosynthetic pathways can be grouped into two categories depending on how the core structure of a natural product is assembled. (A) Megasynthases are modular structures where the growing substrate is shuttled between different domains and protein–protein interactions are key for activity. (B) Dissociated pathways are arrays of individual enzymes that turn over a set of given substrates. Substrate specificity is often tightly regulated, since enzyme–substrate interactions determine the order of biosynthetic events.

16

and the modularity of their expression products enable the use of bioinformatics to assign putative gene functions, and allows the rational targeting of mutations to produce natural product analogs. Strategies to generate natural product analogs by engineering mega--synthases include domain and module swapping,[9,10] modification of domain-domain interfaces,[11] and mutagenesis of individual domains.[10] These engineering approaches often have negative impact on product titers and enzyme activities, since megasynthases depend on both covalent and non-covalent protein-protein interactions.[12,13]

Multimodular polyketide biosynthesis

Birch's model of the construction of polyketides (PKs),[14] inspired by early work on the multiple keten group,[15] spurred efforts to investigate the details of PK biosynthesis.[7] However, reprogramming of PK biosynthesis remained elusive until the seminal discovery that some PK synthases (PKSs) are modular; for example, 6-deoxy-erythronolide B synthase from *Saccharopolyspora erythraea,* which makes the precursor of erythromycin.[16,17]

Multimodular PKSs consist of one or more polypeptide chains (typically >100 kDa), where each polypeptide chain has protein domains that work together to catalyze one cycle of carbon-carbon bond formation (Claisen condensation) and optional steps catalyzed by ketoreductases (KRs), dehydratases (DHs) and enoyl-reductases (ERs). Numerous genetic tools are available to make PK analogs that result in the production of different PK backbones, reduction levels, ring sizes, and stereochemistry, including: in-

frame or site-directed functional deletions,[17,19] modification of module order,[20] addition

of non-natural modules,[21] and numerous examples of domain swapping.[18]

Ketoreductases (KRs) and enoyl reductases (ERs) are part of the PKS megasynthase, and

are responsible for modifying the β-ketoacyl group of the extending PK molecule. KRs

use the co-factor NADPH to reduce carbonyl groups in all PKS systems (Figure 1.3A).

KRs with different stereospecificity exist, and sequence patterns can predict which

hydroxyl group stereoisomer forms in the reduction reaction.[22] Therefore, mutagenesis

of KR domains is an appealing method to alter the substituent chirality of PK products.[23-25] After a dehydration reaction catalyzed by the dehydratase (DH) domain (Figure 1.3B),

a second group of NADPH-dependent enzymes, the enoyl reductases (ERs), generate a

fully saturated polyketide backbone (Figure 1.3B). ERs that act on an *i*-butenyl group (a

group derived from methyl malonyl-CoA) introduce a methyl substituent with either *(S)*

or *(R)* configuration (Figure 1.3B). If the ER active site contains a tyrosine residue, the

reduced product will contain a methyl group with *(S)* configuration.[26] By mutating Tyr

to Val in an ER from the erythromycin producer *S. erythraea*, the methyl group

configuration in the product becomes *(R)*.[26] Other residues may be responsible for this

change in stereospecificity, since the Val to Tyr mutation of an *(R)*-selective ER from

rapamycin biosynthesis, is not sufficient to change its stereoselectivity.[26]
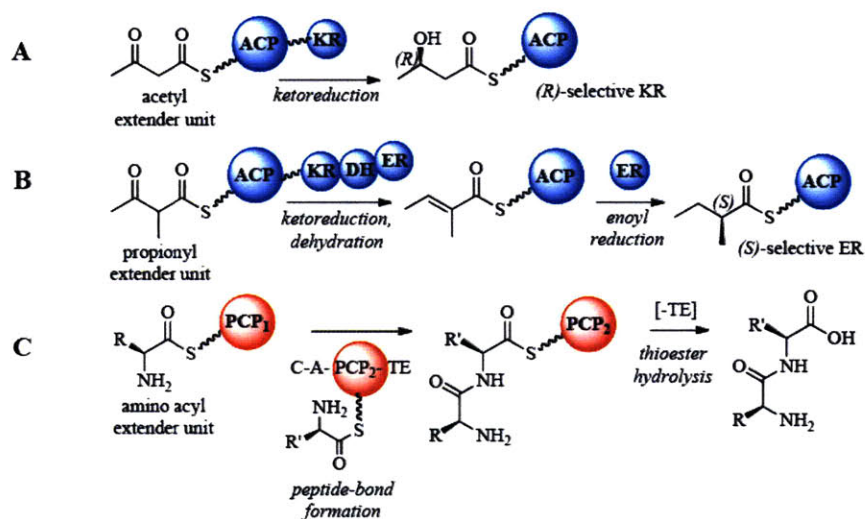
Figure 1.3 Simplified illustration of reactions catalyzed by megasynthase domains. (A) KR domains use NADPH to reduce carbonyl groups to alcohols; here, an acetyl extender unit (malonyl group). (B) The KR and DH domains convert a propionyl extender unit (methyl-malonyl group) into an *i*-butenyl moiety. The NADPH-dependent, stereoselective, ER domain reduces this olefin to form a fully saturated chain. (C) A typical NRPS module is composed of an A domain (loading an amino acid onto the PCP domain), a C domain (catalyzing the peptide bond formation), and a PCP domain (carrying the amino acyl chain). Terminal modules often have TE domains, which hydrolyzes the thioester linkage to release the natural product.

Iterative polyketide biosynthesis

Eukaryotic iterative type I PKSs are monomodular megasynthases that produce aromatic polyketides.[27] Crawford *et al.* reported the deconstruction and reassembly of selected domains in *Aspergillus parasiticus* (fungal) PksA, which makes a precursor to the environmental carcinogen aflatoxin.[28] By using high-resolution mass spectrometry (HRMS), the authors show how individual domains control PK chain-length, cyclization, and product release. A product template (PT) domain, previously identified in iterative fungal PKSs, plays a central role in controlling product formation.[28] The crystal structure reveals that the fully extended PK product of PksA binds to the PT domain correctly oriented for two regiospecific intramolecular aldol condensations, including the subsequent aromatization reactions (Townsend and Tsai, in press). Therefore, mutagenesis of the PT domain may lead to altered regiospecificity and different product outcome.

Bacterial iterative type II PKSs are non-modular and consist of individual protein domains that form intimate non-covalent associations. Ames *et al.* obtained a 1.9 Å resolution x-ray structure of an aromatase in *Streptomyces lividans* tetracenomycin synthase, Tcm.[29] Based on how the substrates dock into the active site of Tcm, the authors constructed eleven mutations of five Tcm residues and analyzed the cyclization pattern of the products. Two residues (Tyr35 and Arg69) did not change the cyclization outcome, but were responsible for anchoring the substrate to the binding site. Several other residues, however, had a marked effect on the product outcome.[29]

20

Nonribosomal peptide biosynthesis

A prototypical NRPS contains a peptidyl carrier protein (PCP), a condensation (C) domain, and an adenylation (A) domain (Figure 1.3C). In addition, the termination module usually includes a thioesterase (TE) domain that releases the natural product from the megasynthase by hydrolysis of a thioester linkage. Marahiel and co-workers recently solved the x-ray crystal structure of the intact 144-kDa NRPS termination module (C-A-PCP-TE) of surfactin synthetase, SrfA-C, to 2.6 Å resolution (PDB ID: 2VSQ).[30] The crystal structure shows the extensive contact interfaces between the different domains and suggests that large structural reorganizations take place during the catalytic cycle. The authors suggest that a promising approach to generate unnatural NRP derivatives via domain swapping, is to exchange C-A domain pairs and then optimize the PCP/C-A contact interactions using directed evolution: a cycle of mutagenesis and screening that gradually improves production rates or titers.[30]

The key challenge for using directed evolution to obtain improved catalysts is how to develop a suitable high-throughput screening or selection method. Siderophores provide a unique handle for selection, since they enable host organisms to survive on iron-limiting media. For example, enterobactin synthetase (EntB-F) from *Escherichia coli*,[33] is a two-module NRPS that allows the bacterium to grow under iron-limiting conditions by secreting the NRP siderophore, enterobactin. EntB is an aryl carrier protein (ArCP) that incorporates the 2,3-dihydroxybenzoyl group into enterobactin; the *entB⁻* knockout strain requires iron as a supplement for survival. Walsh and co-workers used this *E. coli* strain

21

to select functional NRPS mutants by growing mutant libraries on low-iron-content media. Application of this selection method on combinatorial mutagenesis libraries allowed mapping of contact interfaces between EntB and neighboring domains,[34,35] and enabled the replacement of EntB with two mutated non-cognate ArCP domains.[36] For example, three rounds of error-prone PCR on one *entB* gene homolog increased the rate of formation of enterobactin by 500-fold compared to the wild-type protein.[36] The results show that domain-domain and domain-module surface interactions impair NRPS to recognize non-cognate ArCPs.[36]

Recent solution NMR structures of NRPSs shed light on the structural changes that occur upon domain-domain interactions.[31,32] NMR structures of the PCP-TE bidomain of enterobactin synthetase module EntF reveal that other proteins that participate in the biosynthesis of enterobactin modulate the structure of the domain interface.[31] NMR structures of the type II TE in *Bacillus subtilis* surfactin synthetase, which repairs misprimed PCP-domains by hydrolyzing incorrect amino acyl thioester linkages, clarify that appropriate protein-protein interactions are key for stabilizing a certain conformation.[32] While multiple conformations are present when the type II TE domain is free in solution, a single conformation is favored when the domain is in the presence of its native substrate: a modified holo-PCP domain.[32] Hence, replacing a cognate domain with a non-cognate domain, as in the ArCP example, disrupts protein-protein interactions, which may affect megasynthase catalysis.

1.2.2 Dissociated-type pathways

Examples of dissociated-type biosynthetic pathways are type III PKSs (chalcone synthases), isoprenoid biosynthesis, and alkaloid biosynthesis. Substrate specificities of enzymes are critical for determining the order of chemical transformations in dissociated-type biosynthetic pathways, though there are several other factors to consider (Figure 1.2B). Dissociated-type biosynthetic pathways are present in eukaryotes and, thus, biosynthetic pathway enzymes are often in several different compartments; the study of dissociated-type pathways must consider enzyme location (membrane, cytosol, or organelle), enzyme environment, and metabolite transporter specificity. Nevertheless, there are no, or fewer, protein-protein interactions to consider, which may render dissociated-type pathway enzymes better suited as leads for biocatalysts.

While megasynthases typically generate a few products, most dissociated-type pathways lead to a multitude of (structurally diverse) natural products from a central intermediate, called the biogenetic compound. One strategy to generate natural product analogs in this type of pathway, or network of pathways, focuses on manipulating enzymes that catalyze the formation of the biogenetic compound. This approach assumes that a modified biogenetic compound, when placed in a biosynthetic context, forms a range of modified natural products. A key challenge is the potentially narrow substrate specificity of enzymes in the biosynthetic pathway (both reactions preceding and succeeding the biogenetic compound), which could lead to accumulation of intermediates, generation of

off-pathway products, and the absence of some or all natural products in biosynthetic pathway branches.

Isoprenoid biosynthesis

Isoprenoids constitute the largest fraction of known natural products and include sterols, carotenoids, and terpenes.[37] The carbon skeletons of isoprenoids are constructed from 3-methyl-1-butyl units and four coupling reactions are responsible for all known isoprenoid biosynthesis: chain elongation, branching, cyclopropanation and cyclobutanation (Figure 1.4A). Thulasiram *et al.* constructed eleven chimerae of chrysanthemyl diphosphate synthase (CPPase), an enzyme with cyclopropanation activity, and farnesyl diphosphate synthase (FPPase) with chain-elongation activity.[38] The authors show that hybrids of CPPase and FPPase catalyze all four coupling reactions and suggest that the four coupling reactions are evolutionarily related.[38]

Isoprenoid biosynthetic enzymes are amenable to enzyme engineering, since a relatively small number of targeted mutations can lead to significantly modulated activity and specificity. γ-Humulene synthase is sesquiterpene cyclase from *Abies grandis* that catalyzes the cyclization of FPP to numerous different products.[39] Yoshikuni *et al.* mutated nineteen residues near the active site that they believed would alter the product spectrum. Four out of the nineteen residues have >100-fold effect on cyclization selectivity, and recombination of these four residues by combinatorial mutagenesis

24
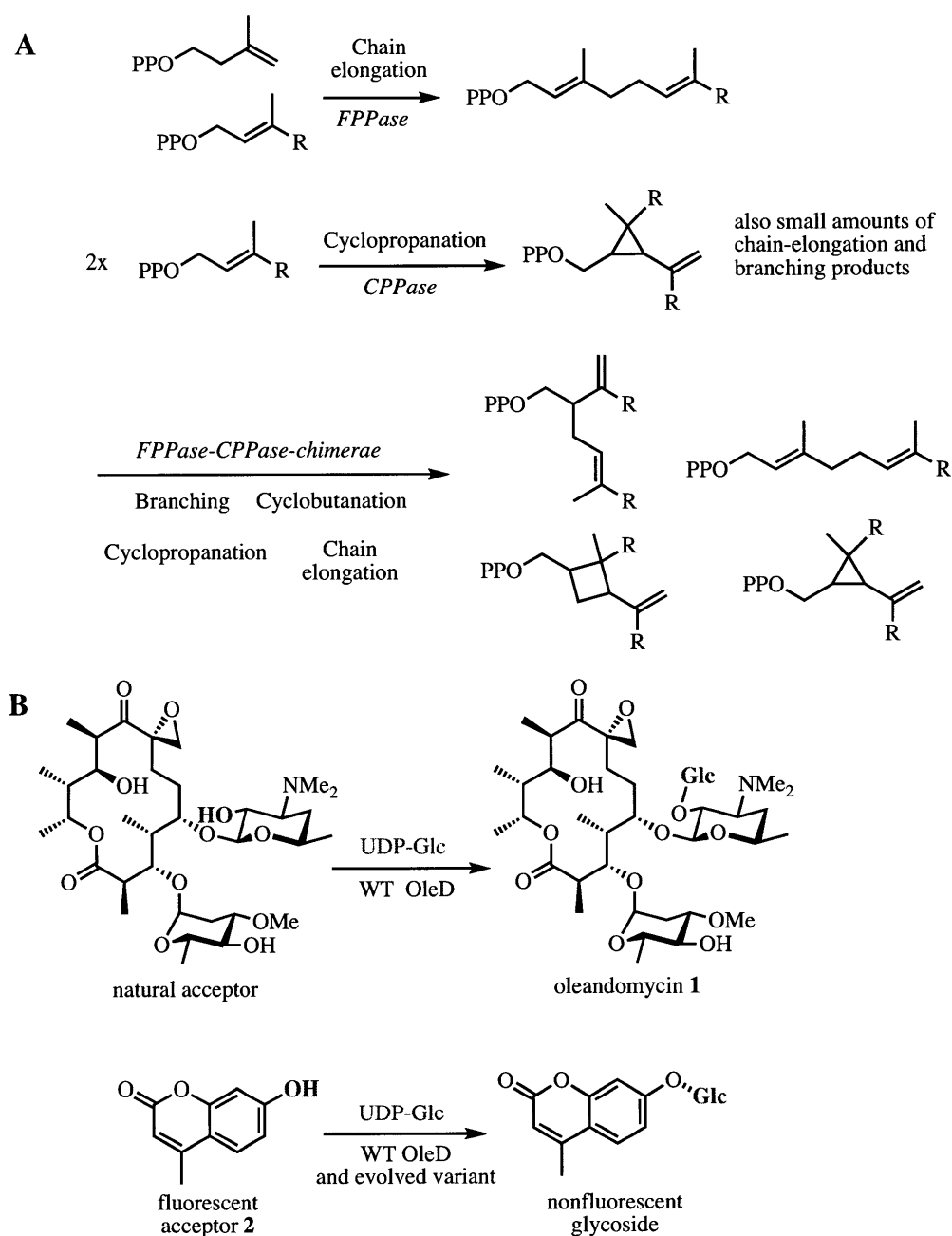
Figure 1.4 Examples of previously published enzyme engineering work on dissociated pathways. (A) Four coupling reactions assemble all isoprenoids: chain elongation, branching, cyclopropanation, and cyclobutanation. Chimerae of CPPase and FPPase catalyze all four coupling reactions. (B) OleD catalyzes the C–O bond formation between the C-2 hydroxyl group of oleandomycin **1** (acceptor) and the anomeric carbon of UDP-glucose (donor). OleD is 300-fold less efficient in transferring UDP-glucose onto 4-methylumbelliferone **2**, a fluorescent surrogate substrate.

25

results in the optimized formation of different cyclization products. The authors used the same approach to alter the specificity of seven other terpene cyclases; four of the evolved enzymes form new natural products.[39] Therefore, the considerable plasticity of the terpene cyclase active site allows several types of rearrangements and cyclizations, demonstrating the opportunity to use enzyme engineering to construct new terpenoid scaffolds. Chapter 3 describes enzyme engineering in alkaloid biosynthesis.

Tailoring enzymes

Tailoring enzymes catalyze reactions that modify completed megasynthase products. A common tailoring step, which often modulates the bioactivity of megasynthase products, is natural product glycosylation catalyzed by glycosyltransferases (GTs). Custom glycosylation of natural products could be a means to diversify bioactivity, but the GTs that transfer a sugar from a nucleotide-sugar donor to a natural product acceptor often have narrow substrate specificity. Williams *et al.* expanded the substrate scope of OleD from *Streptomyces antibioticus*, which transfers a UDP-Glc to oleandomycin **1** (Figure 1.4B).[40] They used three rounds of error-prone PCR on *oleD* and screened for activity with a fluorescent surrogate acceptor substrate **2**.[41] The result was a triple mutant of OleD that increases the activity by 60-fold using **2** as the acceptor. The mutant accepts twelve out of 22 unnatural UDP-sugar donors and also some alternate acceptors in addition to **1** and **2**. Williams *et al.* then subjected the three residues to site-saturation mutagenesis to optimize the specificity toward an unnatural acceptor.[42]

1.3 Monoterpene indole alkaloid biosynthesis

Monoterpene indole alkaloids (MIAs) are products of dissociated-type biosynthetic pathways and constitute a large family of structurally diverse plant-derived nitrogen-containing natural products.[43-46] Several MIAs, including vinblastine **3**, quinine **4**, camptothecin **5**, ajmaline **6**, and serpentine **7**, have important pharmacological activities (Figure 1.5).

All MIAs derive from tryptophan **8** and the seco-iridoid β-D-glucoside secologanin **9**.[47-49] Tryptophan decarboxylase (TDC), a pyridoxal dependent enzyme,[50-52] converts tryptophan **8** to tryptamine **10**.[53] The isoprenoid precursors of **9**, isopentenyl pyrophosphate (IPP) **11** and dimethylallyl pyrophosphate (DMAPP) **12** derive from the MEP (non-mevalonate) pathway (Figure 1.6).[54] In the first committed step of the iridoid terpene biosynthesis that leads to **9**, geraniol-10-hydroxylase (G10H) hydroxylates geraniol, which originates from isopentenyl pyrophosphate (IPP).[55,56] Feeding experiments with ^{3}H-labeled terpene intermediates suggest that 10-hydroxygeraniol, iridodial, and iridotrial are intermediates in the secologanin biosynthetic pathway,[57,58] though the details of the transformations between these intermediates are unknown. After oxidation of the iridotrial aldehyde to the carboxylic acid, followed by esterification and glucosyl transfer to give deoxyloganin, subsequent hydroxylation of deoxyloganin yields loganin. Secologanin synthase (SLS), an NADPH-dependent P450 oxidase isolated from a cDNA library of an alkaloid-producing *C. roseus* cell culture,[59] then converts loganin to secologanin **9**, presumably via a radical-mediated, oxidative cleavage mechanism.[60]
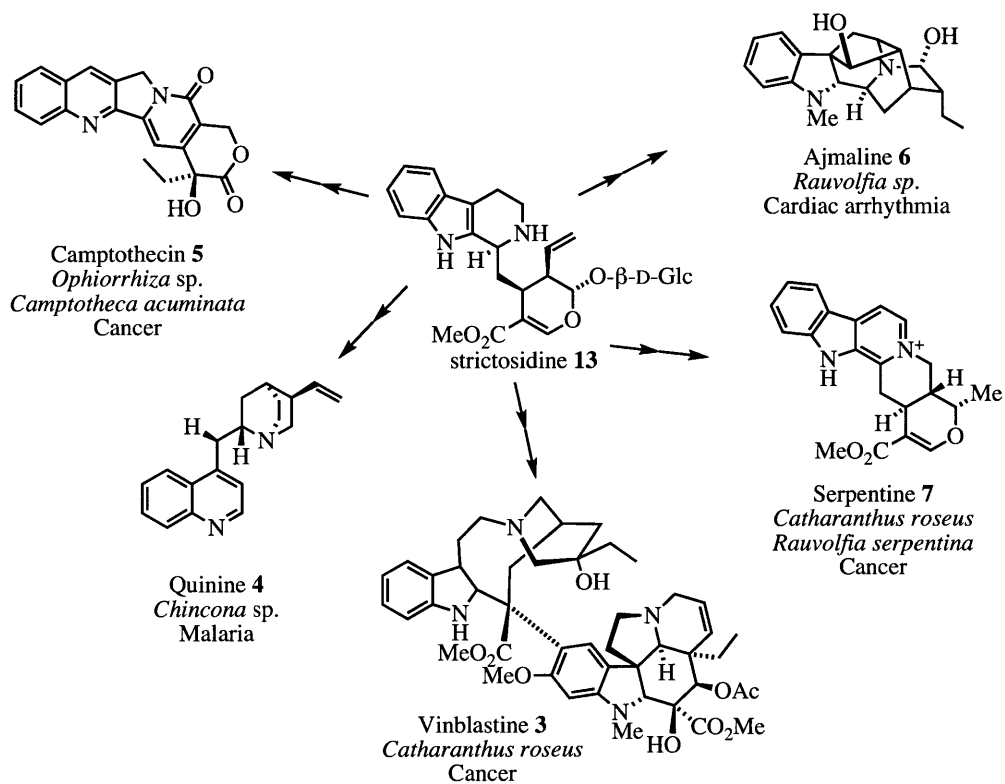
Figure 1.5 Examples of biologically active monoterpene indole alkaloids, such as, vinblastine **3**, quinine **4**, camptothecin **5**, ajmaline **6**, and serpentine **7**, as well as the source organisms and the diseases they treat.
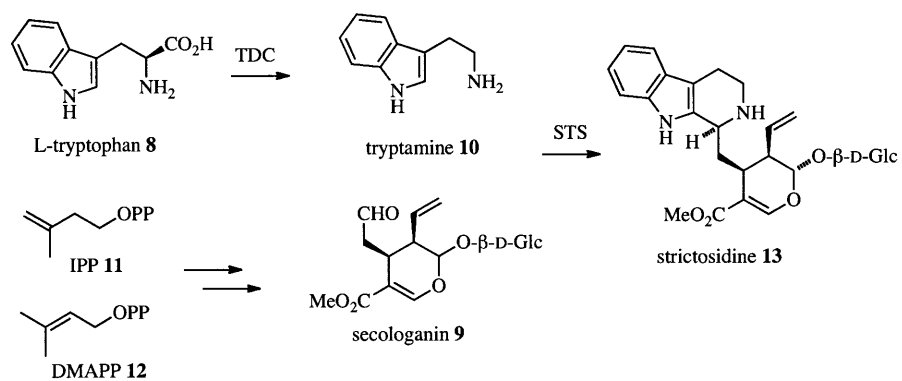
Figure 1.6 Precursors of MIA biosynthesis. The non-mevalonate pathway (MEP) forms isoprenoids that are the progenitors of secologanin **9**. Tryptamine **10** is formed by the decarboxylation of tryptophan **8**. A Pictet-Spengler reaction catalyzed by strictosidine synthase (STS) leads to strictosidine **13** from secologanin **9** and tryptamine **10**.
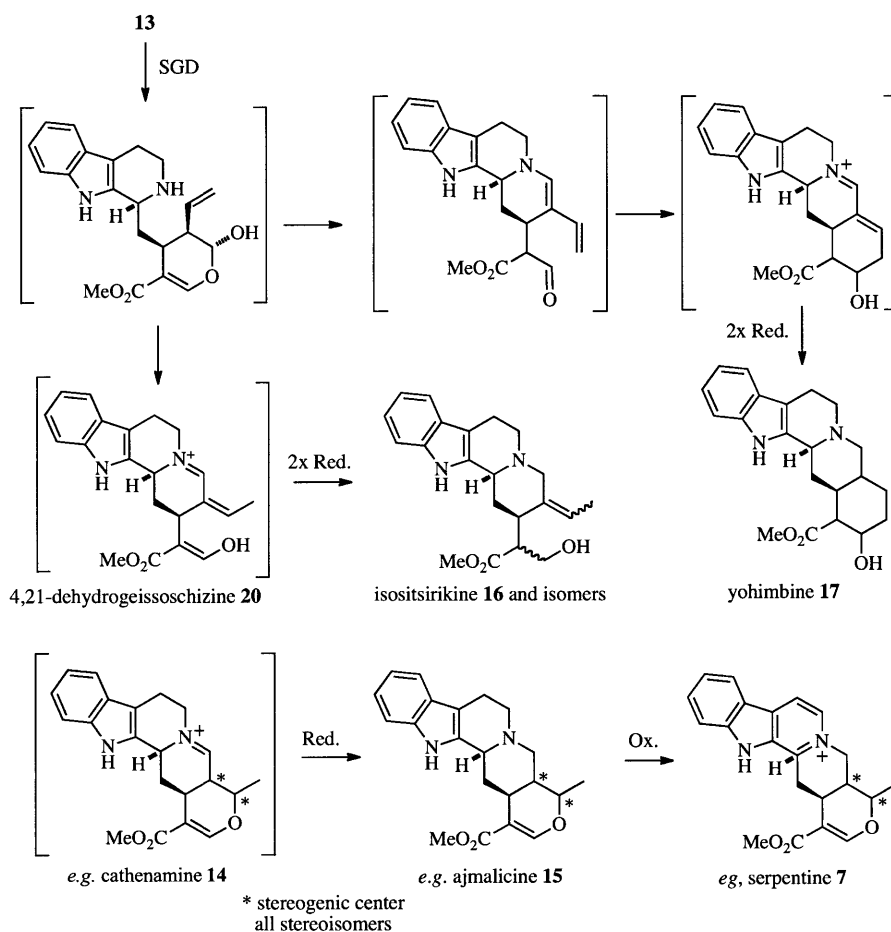
Figure 1.7 Proposed branching reactions that occur after the SGD-catalyzed deglucosylation of strictosidine **13**, result in the formation of yohimbine **17**, isositsirikine **16**, ajmalicine **15**, and serpentine **7**. The enzymatic reduction and oxidation steps are not known at the genetic or biochemical levels.
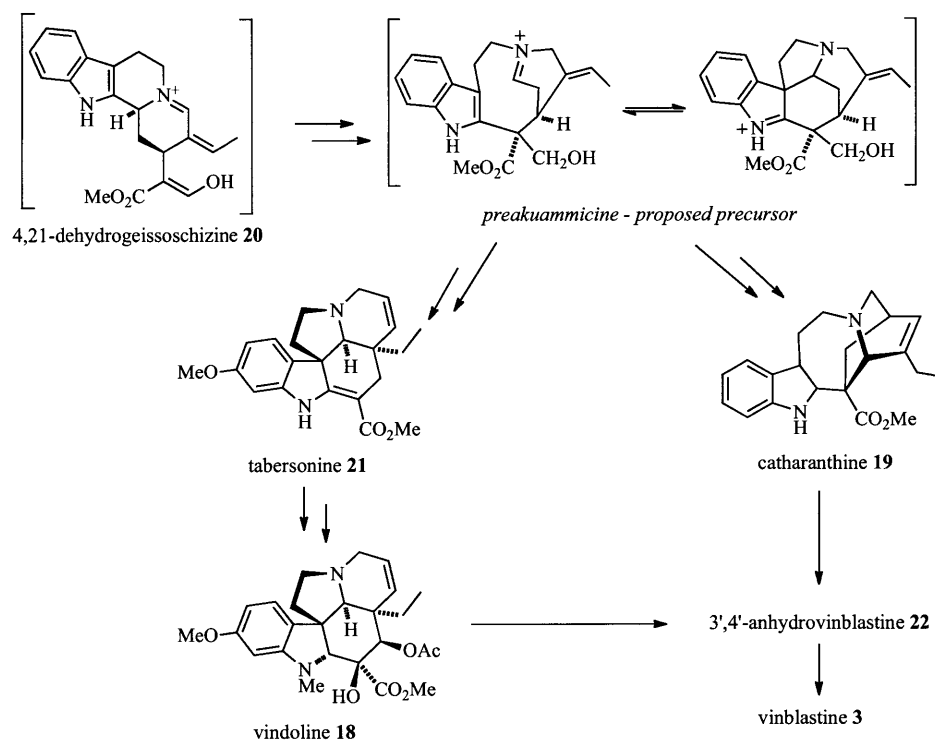
Figure 1.8 Dehydrogeissoschizine **20** is a likely precursor for the rearranged MIAs, including vindoline **18** and catharanthine **19**. Pre-akuammacine is a proposed precursor for the formation of vinblastine. The conversion of tabersonine **21** into vindoline **18** is perhaps the most well understood MIA pathway. Anhydrovinblastine **22** is formed via the oxidative dimerization between **18** and **19**. Hydration of the 3', 4' bond leads to vinblastine **3**.

31

Secologanin **9** and tryptamine **10** are substrates for strictosidine synthase (STS),[61] which catalyzes a Pictet-Spengler reaction to form strictosidine **13**,[62,63] the key intermediate for all MIAs (Figure 1.6), including MIAs in *Catharanthus roseus*,[64-69] *Rauvolfia serpentina*,[70-77] and *Ophiorrhiza pumila*.[78] The STS-catalyzed reaction only forms the 3*(S)* stereoisomer, whereas the Brønsted acid catalyzed reaction forms a mixture of 3*(S)*-strictosidine and its 3*(R)* epimer, vincoside.[79-84] In the second step of MIA biosynthesis, strictosidine-β-D-glucosidase (SGD) deglucosylates strictosidine **13** with retention of configuration to generate a hemiacetal.[85-88] The hemiacetal rearranges spontaneously into a mixture of products, where cathenamine **14** is likely the major product (Figure 1.7).[85,89] The enzyme-catalyzed steps that follow deglucosylation of **13** are not well understood. Pathway branches in *C. roseus* that utilize mainly reductive chemistry form several alkaloids, including ajmalicine **15**, isositsirikine **16**, and yohimbine **17**, but also serpentine **7** via oxidation of **15** (Figure 1.7).[90-95] In a different set of pathways, one or more oxidative branches form the so-called rearranged MIAs (Figure 1.8).[43] These MIAs lead to vindoline **18** and catharanthine **19**, and ultimately the anti-cancer agent vinblastine **3**. Although the details of most steps leading to the biosynthesis of these MIAs are unknown, 4,21-dehydrogeissoschizine **20** may be the precursor.[96,97] The conversion of tabersonine **21** into vindoline **18** is perhaps the most well understood biosynthetic pathway in *C. roseus* (Figure 1.8).[98,99] An oxidative dimerization reaction initiates the formation of vinblastine **3** (via anhydrovinblastine **22**) from vindoline **18** and catharanthine **19**.[100-104]

1.4 Biocatalysis

Biocatalysis is the use of enzymes to perform transformations on organic compounds, encompassing the use of isolated and whole-cell catalysts. An ideal biocatalyst should possess a broad substrate scope, a high total turnover number (TTN), and high stereoselectivity.[105,106] Industrial biotransformations frequently employ hydrolases,[106] since they possess most of the qualities listed above. Furthermore, in addition to hydrolysis, hydrolases also catalyze several related reactions: condensation,[107] alcoholysis,[108] and perhydrolysis.[109]

The multitude of biosynthetic pathways in the three kingdoms of life certainly must contain enzyme catalysts that are useful outside of the context of the biosynthetic pathways.[110] For example, biosynthetic enzymes could be useful in stereoselective organic synthesis of chemicals and pharmaceuticals[111] and for the conversion of biomass.[112,113] Useful biocatalysts derived from biosynthetic pathways are rare, likely because biosynthetic enzymes often have restrictive substrate scopes. Nevertheless, there are several examples of promising biocatalysts in biosynthetic pathways, including γ-humulene synthase (section 1.2.2), lovastatin acyltransferase (LovD),[114,115] GTs (section 1.2.3), and haloperoxidases (HPOs).[116,117] Chapter 4 describes an STS homolog from *O. pumila* with potential use as a biocatalyst for the asymmetric construction of tetrahydro-β-carbolines.[122]

1.5 Research goals and thesis overview

The following goals motivate the research summarized in this thesis:

A: Increase natural product diversity in monoterpene indole alkaloid biosynthesis. Precursor-directed biosynthesis of unnatural derivatives of monoterpene indole alkaloids (MIAs) in *C. roseus* shows that only minor structural modifications of secologanin **9** and tryptamine **10** are tolerated for conversion into higher alkaloids by the plant biosynthetic pathways.[118] These restrictions are primarily due to the narrow substrate specificity of strictosidine synthase (STS). Chapter 2 describes the development of two assays to detect STS activity.[119,120] These assays are instrumental for the rapid screening of STS mutant libraries for activity with tryptamine substrate analogs that wild-type STS does not accept.[119] Chapter 3 describes the identification of two STS variants (Val214Met and Phe232Leu) that form novel strictosidine derivatives.[119] This work was one of the first examples of enzyme engineering in medicinal plant biosynthesis, which enabled the production of "unnatural" natural products in plant cells.[121]

B: Use of strictosidine synthase as a general biocatalyst for the formation of chiral tetrahydro-β-carbolines. Chiral tetrahydro-β-carboline derivatives are important pharmacophores and a general biocatalytic method to produce these molecules would be useful. *C. roseus* STS, however, essentially only accepts secologanin **9**, which makes it unsuitable as a general biocatalyst. Chapter 4 describes the use of a STS homolog from *O. pumila* that has considerably broader aldehyde substrate specificity than *C. roseus*

34

STS.[122] A series of new chiral tetrahydro-β-carbolines form as a result of *O. pumila* STS catalysis from achiral aldehydes and tryptamine or secologanin and tryptamine analogs.

<u>C: Synthetic derivatives of secologanin to investigate the stereochemical restrictions in monoterpene indole alkaloid biosynthesis.</u> Different stereoisomers of pharmaceuticals often possess different bioactivities. Natural products typically exist as single stereoisomers, but not all natural products are bioactive or have undesired bioactivity. It is reasonable to propose that an altered stereoconfiguration can render an otherwise undesired or inactive natural product medicinally useful. Chapters 4-5 describe the first steps on the way to construct stereoisomer libraries of MIAs. Chapter 4 describes the total synthesis and biochemical evaluation of des-vinyl secologanin aglycone stereoisomers with *C. roseus* STS.[123] I observed that STS accepts alternate stereoisomers when the glucose moiety is removed. Chapter 5 extends this work by describing the synthesis and evaluation of *glucosylated* des-vinyl secologanin to identify stereochemical restrictions in reconstituted heteroyohimbine alkaloid biosynthesis.[124] Chemo-enzymatic synthesis may in the future lead to the generation of stereoisomer libraries for bioactivity evaluation.

1.6 References

1. Dewick, P.M. *Medicinal natural products: a biosynthetic approach*, 2nd ed. Wiley **2003**.

2. Koehn, F.E.; Carter, G.T. "The evolving role of natural products in drug discovery", *Nat. Rev. Drug. Discov.* **2005**, *4*, 206-220.

3. Wilkinson, B.; Micklefield, J. "Mining and engineering natural product biosynthetic pathways", *Nat. Chem. Biol.* **2007**, *3*, 379-386.

4. Baker, D.D.; Chu, M.; Oza, U.; Rajgarhia, V. "The value of natural products to future pharmaceutical discovery", *Nat. Prod. Rep.* **2007**, *24*, 1225-1244.

5. Patani, G.A.; LaVoie, E.J. "Bioisosterism: a rational approach in drug design", *Chem. Rev.* **1996**, *96*, 3147-3176.

6. Bernhardt, P.; O'Connor S.E. "Opportunities for enzyme engineering in natural product biosynthesis", *Curr. Opin. Chem. Biol.* **2009**, *13*, 35-42.

7. Hopwood, D.A.; Sherman, D.H. "Molecular genetics of polyketides and its comparison to fatty acid biosynthesis", *Annu. Rev. Genet.* **1990**, *24*, 37-66.

8. Marahiel, M.A.; Stachelhaus, T.; Mootz, H.D. "Modular peptide synthetases involved in nonribosomal peptide synthesis", *Chem. Rev.* **1997**, *97*, 2651-2673; and references therein.

9. Oliynyk, M.; Brown, M.J.B.; Cortes, J.; Staunton, J.; Leadlay, P.F. "A hybrid modular polyketide synthase obtained by domain swapping", *Chem. Biol.* **1996**, *3*, 833-839.

10. Del Vecchio, F.; Petkovic, H.; Kendrew, S.G.; Low, L.; Wilkinson, B.; Cortes, J.;

Rudd, B.A.M.; Staunton, J.; Leadlay, P.F. "Active-site residue, domain and module swaps in modular polyketide synthases" *J. Ind. Microbiol. Biotechnol.* **2003**, *30*, 489-494.

11. Fischbach, M.A.; Lai, J.R.; Roche, E.D.; Walsh, C.T.; Liu, D.R. "Directed evolution can rapidly improve the activity of chimeric assembly-line enzymes", *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11951-11956

12. Weissman, K.J. "The structural basis for docking in modular polyketide biosynthesis", *Chembiochem* **2006**, *7*, 485-494.

13. Weissman, K.J. "Single amino acid substitutions alter the efficiency of docking in modular polyketide biosynthesis", *Chembiochem* **2006**, *7*, 1334-1342.

14. Birch, A.J. "Biosynthesis of polyketides and related compounds", *Science* **1967**, *156*, 202-206.

15. Collie, J.N. "Derivatives of the multiple keten group", *J. Chem. Soc.* **1907**, *51*, 1806-1813.

16. Corter, J.; Haydock, S.F.; Robert, G.A.; Bevitt, D.J.; Leadlay, P.F. "An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*", *Nature* **1990**, *348*, 176-178.

17. Danadio, S.; Staver, M.J.; McAlpine, J.B.; Swanson, S.J.; Katz, L. "Modular organization of genes for complex polyketide biosynthesis", *Science* **1991**, *252*, 675-679.

18. Staunton, J.; Weissman, K.J. "Polyketide biosynthesis: a millenium review", *Nat. Prod. Rep.* **2001**, *4*, 380-416.

19. Donadio, S.; McAlpine, J.B.; Sheldon, P.J.; Jackson, M.; Katz, L. "An erythromycin analog produced by reprogramming of polyketide synthesis", *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7119-7123.

20. McDaniel, R.; Ebert-Khosla, S.; Hopwood, D.A.; Khosla, C. "Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic products", *Nature* **1995**, *375*, 549-554.

21. McDaniel, R.; Kao, C.M.; Fu, H.; Hevezi, P.; Gustafsson, C.; Betlach, M.; Ashley, G.; Cane, D.E.; Khosla, C. "Gain-of-function mutagenesis of a modular polyketide synthase", *J. Am. Chem. Soc.* **1997**, *119*, 4309-4310.

22. Keatinge-Clay, A.T. "A tylosin ketoreductase reveals how chirality is determined in polyketides", *Chem. Biol.* **2007**, *14*, 898-908.

23. Marsden, A.F.A.; Wilkinson, B.; Cortes, J.; Dunster, N.J.; Staunton, J; Leadlay, P.F. "Engineering broader specificity into an antibioticproducing polyketide synthase", *Science* **1998**, *279*, 199-202.

24. Baerga-Ortiz, A.; Popovic, B.; Siskos, A.P.; O'Hare, H.M.; Spiteller, D.; Williams, M.G.; Campillo, N.; Spencer, J.B.; Leadlay, P.F. "Directed mutagenesis alters the stereochemistry of catalysis by isolated ketoreductase domains from the erythromycin polyketide synthase", *Chem. Biol.* **2006**, *13*, 277-285.

25. O'Hare, H.M.; Baerga-Ortiz, A.; Popovic, B.; Spencer, J.B.; Leadlay, P.F. "High-throughput mutagenesis to evaluate models of stereochemical control in ketoreductase domains from the erythromycin polyketide synthase", *Chem. Biol.* **2006**, *13*, 287-296.

26. Kwan, D.H.; Sun, Y.; Schulz, F.; Hong, H.; Popovic, B.; Sim-Stark, J.C.; Haydock, S.F.; Leadlay, P.F. "Prediction and manipulation of the stereochemistry of enoylreduction in modular polyketide synthases", *Chem. Biol.* **2008**, *15*, 1231-1240.

27. Schneider, G. "Enzymes in the biosynthesis of aromatic polyketide antibiotics", *Curr. Opin. Struct. Biol.* **2005**, *15*, 629-636.

28. Crawford, J.M.; Thomas, P.M.; Scheerer, J.R.; Vagstad, A.L.; Kelleher, N.L.; Townsend, C.A. "Deconstruction of iterative multidomain polyketide synthase function", *Science* **2008**, *320*, 243-246.

29. Ames, B.D.; Korman, T.P.; Zhang, W.; Smith, P.; Vu, T.; Tang, Y.; Tsai, S.C. "Crystal structure and functional analysis of tetracenomycin ARO/CYC: implications for cyclization specificity of aromatic polyketides", *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 5349-5354.

30. Tanovic, A.; Samel, S.A.; Essen, L.O.; Marahiel, M.A. "Crystal structure of the termination module of a nonribosomal peptide synthetase", *Science* **2008**, *321*, 659-663.

31. Frueh, D.P.; Arthanari, H.; Koglin, A.; Vosburg, D.A.; Bennett, A.E.; Walsh, C.T.; Wagner, G. "Dynamic thiolation-thioesterase structure of a non-ribosomal peptide synthetase", *Nature* **2008**, *454*, 903-906.

32. Koglin, A.; Lohr, F.; Bernhard, F.; Rogov, V.V.; Frueh, D.P.; Strieter, E.R.; Mofid, M.R.; Guntert, P.; Wagner, G.; Walsh, C.T. "Structural basis for the selectivity of the external thioesterase of the surfactin synthetase", *Nature* **2008**, *454*, 907-911.

33. Roche, E.D.; Walsh, C.T. "Dissection of the EntF condensation domain boundary and active site residues in nonribosomal peptide synthesis", *Biochemistry* **2003**, *42*, 1334-1344.

34. Lai, J.R.; Fischbach, M.A.; Liu, D.R.; Walsh, C.T. "Localized protein interaction surfaces on the EntB carrier protein revealed by combinatorial mutagenesis and selection", *J. Am. Chem. Soc.* **2006**, *128*, 11002-11003.

35. Zhou, Z.; Lai, J.R.; Walsh, C.T. "Interdomain communication between the thiolation and thioesterase domains of EntF explored by combinatorial mutagenesis and selection", *Chem. Biol.* **2006**, *13*, 869-879.

36. Zhou, Z.; Lai, J.R.; Walsh, C.T. "Directed evolution of aryl carrier proteins in the enterobactin synthetase", *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11621-11626.

37. Buckingham J. Dictionary of Natural Products on CD-ROM (Chapman & Hall, London) **1998**, version 6.1.

38. Thulasiram, H.V.; Erickson, H.K.; Poulter, C.D. "Chimeras of two isoprenoid synthases catalyze all four coupling reactions in isoprenoid biosynthesis", *Science* **2007**, *316*, 73-76.

39. Yoshikuni, Y.; Ferrin, T.E.; Keasling, J.D. "Designed divergent evolution of enzyme function", *Nature* **2006**, *440*, 1078-1082.

40. Williams, G.J.; Zhang, C.; Thorson, J.S. "Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution", *Nat. Chem. Biol.* **2007**, *3*, 657-662.

41. Williams, G.J.; Thorson, J.S. "A high-throughput fluorescence based glycosyltransferase screen and its application in directed evolution", *Nat. Protoc.* **2008**, *3*, 357-362.

42. Williams, G.J.; Goff, R.D.; Zhang, C.; Thorson, J.S. "Optimizing glycosyltransferase specificity via 'hot spot' saturation mutagenesis presents a catalyst for novobiocin glycorandomization", *Chem. Biol.* **2008**, *15*, 393-401.

43. O'Connor, S.E.; Maresh, J.J. "Chemistry and biology of monoterpene indole alkaloid biosynthesis", *Nat. Prod. Rep.* **2006**, *23*, 532-547.

44. Wenkert, E. "Biosynthesis of indole alkaloids. The aspidosperma and iboga bases", *J. Am. Chem. Soc.* **1962**, *84*, 98-102.

45. Thomas, R. "A possible biosynthetic relationship between the cyclopentanoid monoterpenes and the indole alkaloids", *Tetrahedron Lett.* **1961**, *16*, 544-553.

46. Money, T.; Wright, I.G.; McCapra, F.; Scott, A.I. "Biosynthesis of the indole alkaloids", *Proc. Natl. Acad. Sci. U.S.A.* **1965**, *53*, 901-903.

47. Battersby, A.R.; Burnett, A.R.; Parsons, P.G. "Preparation of secologanin: its conversion into ipecoside and its role in indole alkaloid biosynthesis", *Chem. Comm. (London)* **1968**, 1280-1281.

48. Battersby, A.R.; Burnett, A.R.; Parsons, P.G. "Alkaloid biosynthesis. Part XIV. Secologanin: its conversion into ipecoside and its role as biological precursor of the indole alkaloids", *J. Chem. Soc. C.* **1969**, 1187-1192.

49. Battersby, A.R.; Byrne, J.C.; Kapil, R.S.; Martin, J.A.; Payne, T.G.; Arigoni, D.; Loew, P. "The mechanism of indole alkaloid biosynthesis", *Chem. Comm. (London)* **1968**, 951-953.

50. De Luca, V.; Marineau, C.; Brisson, N. "Molecular cloning and analysis of cDNA encoding a plant tryptophan decarboxylase: comparison with animal dopa decarboxylases", *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 2582-2586.

51. Facchini, P.J.; Huber-Allanach, K.L.; Tari, L.W. "Plant aromatic L-amino acid decarboxylases: evolution, biochemistry, regulation, and metabolic engineering applications", *Phytochemistry* **2000**, *54*, 121-138.

52. Hong, S.-B.; Peebles, C.A.M.; Shanks, J.V.; San, K.-Y.; Gibson, S.I. "Expression of the *Arabidopsis* feedback-insensitive anthranilate synthase holoenzyme and tryptophan decarboxylase genes in *Catharanthus roseus* hairy roots", *J. Biotechnol.* **2006**, *122*, 28-38.

53. Leete, E. "Biogenesis of the *Rauwolfia* alkaloids alkaloids—II : The incorporation of tryptophan into serpentine and reserpine", *Tetrahedron* **1961**, *14*, 35-41.

54. Contin, A.; van der Heijden, R.; Lefeber, A.W.M.; Verpoorte, R. "The iridoid glucoside secologanin is derived from the novel triose phosphate/pyruvate pathway in a *Catharanthus roseus* cell culture", *FEBS Lett.* **1998**, *434*, 413-416.

55. Collu, G.; Unver, N.; Peltenburg-Looman, M.G.; van der Heijden, R.; Verpoorte, R.; Memelink, J. "Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis", *FEBS Lett.* **2001**, *508*, 215-220.

56. Collu, G.; Garcia, R.; van der Heijden, R.; Verpoorte, R. "Activity of the cytochrome P450 enzyme geraniol 10-hydroxylase and alkaloid production in plant cell cultures", *Plant Sci.* **2002**, *162*, 165-172.

57. Uesato, S.K.; Iida, A.; Inouye, H.; Zenk, M.H. "Mechanism for iridane skeleton formation in the biosynthesis of secologanin and indole alkaloids in *Lonicera tatarica, Catharanthus roseus* and suspension cultures of *Rauwolfia serpentina*", *Phytochemistry* **1986**, *25*, 839-842.

58. Uesato, S.; Matsuda, S.; Inouye, H. "Mechanism for iridane skeleton formation from monoterpenes in the biosynthesis of secologanin and vindoline in *Catharanthus roseus* and *Lonicera morrowii*", *Chem. Pharm. Bull.* **1984**, *32*, 1671-1674.

59. Irmler, S.; Schroder, G.; St. Pierre, B.; Crouch, N.P.; Hotze, M.; Schmidt, J.; Strack, D.; Matern, U.; Schroder, J. "Indole alkaloid biosynthesis in *Catharanthus roseus*: new enzyme activities and identification of cytochrome P450 CYP72A1 as secologanin synthase", *Plant J.* **2000**, *24*, 797-804.

60. Yamamoto, H.; Katano, N.; Ooi, A.; Inoue, K. "Secologanin synthase which catalyzes the oxidative cleavage of loganin into secologanin is a cytochrome P450", *Phytochemistry* **2000**, *53*, 7-12.

61. Kutchan, T.M. "Strictosidine: From alkaloid to enzyme to gene", *Phytochemistry* **1993**, *32*, 493-506.

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

63. Cox, E.D.; Cook, J.M. "The Pictet-Spengler condensation: a new direction for an old reaction", *Chem. Rev.* **1995**, *95*, 1797-1842.

64. Mizukami, H.; Nordlov, H.; Lee, S.-L.; Scott, A.I. "Purification and properties of strictosidine synthase (an enzyme condensing tryptamine and secologanin) from *Catharanthus roseus* cultured cells", *Biochemistry* **1979**, *18*, 3760-3763.

65. Pfitzner, U.; Zenk, M.H. "Homogeneous strictosidine synthase isoenzymes from cell suspension cultures of *Catharanthus roseus*", *Planta Med.* **1989**, *55*, 525-530.

66. McKnight, T.D.; Roessner, C.A.; Devagupta, R.; Scott, A.I.; Nessler, C. "Nucleotide sequence of a cDNA encoding the vacuolar protein strictosidine synthase from *Catharanthus roseus*", *Nucl. Acids Res.* **1990**, *18*, 4939.

67. De Waal, A.; Meijer, A.H.; Verpoorte, R. "Strictosidine synthase from *Catharanthus roseus*: purification and characterization of multiple forms", *Biochem. J.* **1995**, *306*, 571-580.

68. Roessner, C.R.; Devagupta, R.; Hasan, M.; Williams, H.J.; Scott, I.A. "Purification of an indole alkaloid biosynthetic enzyme, strictosidine synthase, from a recombinant strain of *Escherichia coli*", *Protein Expr. Purif.* **1992**, *3*, 295-300.

69. Geerlings, A.; Redondo, F.J.; Contin, A.; Memelink, J.; van der Heijden, R.; Verpoorte, R. "Biotransformation of tryptamine and secologanin into plant terpenoid indole alkaloids by transgenic yeast", *Appl. Microbiol. Biotechnol.* **2001**, *56*, 420-424.

70. Treimer, J.F.; Zenk, M.H. "Strictosidine synthase from cell cultures of apocynaceae plants", *FEBS Lett.* **1979**, *97*, 159–162.

71. Treimer, J.F.; Zenk, M.H. "Purification and properties of strictosidine synthase, the key enzyme in indole alkaloid formation", *Eur. J. Biochem.* **1979**, *101*, 225–233.

72. Pfitzner, U.; Zenk, M.H. "Isolation and immobilization of strictosidine synthase", *Methods Enzymol.* **1987**, *136*, 342–350.

73. Kutchan, T.M.; Hampp, N.; Lottspeich, F.; Beyreuther, K.; Zenk, M.H. "The cDNA clone for strictosidine synthase from *Rauvolfia serpentina* DNA sequence determination and expression in *Escherichia coli*", *FEBS Lett.* **1988**, *237*, 40–44.

74. Kutchan, T.M. "Expression of enzymatically active cloned strictosidine synthase from the higher plant *Rauvolfia serpentina* in *Escherichia coli*", *FEBS Lett.* **1989**, *257*, 127–130.

75. Pennings, E.J.; van den Bosch, A.J.; van der Heijden, R.; Stevens, L.H.; Duine, J.A.; Verpoorte, R. "Assay of strictosidine synthase from plant cell cultures by high-performance liquid chromatography", *Anal. Biochem.* **1989**, *176*, 412–415.

76. Bracher, D.; Kutchan, T.M. "Strictosidine synthase from *Rauvolfia serpentina*: analysis of a gene involved in indole alkaloid biosynthesis", *Arch. Biochem. Biophys.* **1992**, *294*, 717–723.

77. Kutchan, T.M.; Bock, A.; Dittrich, H. "Heterologous expression of the plant proteins strictosidine synthase and berberine bridge enzyme in insect cell culture", *Phytochemistry* **1994**, *35*, 353–360.

78. Yamazaki, Y.; Sudo, H.; Yamazaki, M.; Aimi, N.; Saito, K. "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*, 395-403.

79. Scott, A.I.; Lee, S.L.; de Capite, P.; Culver, M.G.; Hutchinson, C.R. "The role of isovincoside (strictosidine) in the biosynthesis of the indole alkaloids", *Heterocycles* **1977**, *7*, 979–984.

80. Battersby, A.R.; Lewis N.G.; Tippett, J.M. "The basic glucosides related to the biosynthesis of indole and ipecac alkaloids", *Tetrahedron Lett.* **1978**, *19*, 4849–4852.

81. Heckendorf, A.H.; Hutchinson, C.R. "Biosynthesis of camptothecin. II. Confirmation that isovincoside, not vincoside, is the penultimate biosynthetic precursor of indole alkaloids", *Tetrahedron Lett.* **1977**, *18*, 4153–4154.

82. Nagakura, N.; Ruffer, M.; Zenk, M.H. "The biosynthesis of monoterpenoid indole alkaloids from strictosidine", *J. Chem. Soc., Perkin Trans. 1* **1979**, 2308–2312.

83. Stöckigt, J.; Zenk, M.H. "Strictosidine (isovincoside): the key intermediate in the biosynthesis of monoterpenoid indole alkaloids", *J. Chem. Soc., Chem. Comm.* **1977**, 646–648.

84. Brown, R.T.; Leonard, J.; Sleigh, S.K. "The role of strictosidine in monoterpenoid indole alkaloid biosynthesis", *Phytochemistry* **1978**, *17*, 899–900.

85. Geerlings, A.; Ibanez, M.M.-L.; Memelink, J.; van der Heijden, R.; Verpoorte, R. "Molecular cloning and analysis of strictosidine β-D-glucosidase, an enzyme in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*", *J. Biol. Chem.* **2000**, *275*, 3051–3056.

86. Hemscheidt, T.; Zenk, M.H. "Glucosidases involved in indole alkaloid biosynthesis of *Catharanthus* cell cultures", *FEBS Lett.* **1980**, *110*, 187–191.

87. Scott, A.I.; Lee, S.L.; Wan, W. "Indole alkaloid biosynthesis: partial purification of "ajmalicine synthetase" from *Catharanthus roseus*", *Biochem. Biophys. Res. Commun.* **1977**, *75*, 1004–1009.

88. Luijendijk, T.J.C.; Stevens, L.H.; Verpoorte, R. "Purification and characterisation of strictosidine β-D-glucosidase from *Catharanthus roseus* cell suspension cultures", *Plant Physiol. Biochem.* **1998**, *36*, 419–425.

89. Gerasimenko, I.; Sheludko, Y.; Ma, X.; Stöckigt, J. "Heterologous expression of a *Rauvolfia* cDNA encoding strictosidine glucosidase, a biosynthetic key to over 2000 monoterpenoid indole alkaloids", *Eur. J. Biochem.* **2002**, *269*, 2204–2213

90. Kan-Fan, C.; Husson, H.P. "Stereochemical control in the biomimetic conversion of heteroyohimbine alkaloid precursors. Isolation of a novel key intermediate", *J. Chem. Soc., Chem. Commun.* **1978**, 618–619.

91. Husson, H.P.; Kan-Fan, C.; Sevenet, T.; Vidal, J.-P. "Structure de la cathénamine intermédiaire clé de la biosynthèse des alcaloides indoliques", *Tetrahedron Lett.* **1977**, *18*, 1889–1892.

92. Brown, R.T.; Leonard, J.; Sleigh, S.K. "One-pot biomimetic synthesis of 19-heteroyohimbine alkaloids", *J. Chem. Soc., Chem. Commun.* **1977**, 636–638.

93. Brown, R.T.; Leonard, J. "Biomimetic synthesis of cathenamine and 19-epicathenamine, key intermediates to heteroyohimbine alkaloids", *J. Chem. Soc., Chem. Commun.* **1979**, 877–879.

94. Brown, R.T.; Leonard, J. "Reversible trapping of labile 21-dehydroheteroyohimbines as 21-cyano adducts", *Tetrahedron Lett.* **1977**, *18*, 4251–4254.

95. Stöckigt, J.; Husson, H.P.; Kan-Fan, C.; Zenk, M.H. "Cathenamine, a central intermediate in the cell free biosynthesis of ajmalicine and related indole alkaloids", *J. Chem. Soc., Chem. Commun.* **1977**, 164–166.

96. Scott, A.I.; Qureshi, A.A. "Biogenetic-type chemistry of the indole alkaloids", *J. Am. Chem. Soc.* **1969**, *91*, 5874–5876.

97. Wenkert, E.; Wickberg, B. "General methods of synthesis of indole alkaloids. IV. A synthesis of DL-eburnamonine", *J. Am. Chem. Soc.* **1965**, *87*, 1580–1589.

98. Hashimoto, T.; Yamada, Y. "New genes in alkaloid metabolism and transport", *Curr. Opin. Biotechnol.* **2003**, *14*, 163–168.

99. De Luca, V. "Biochemistry and molecular biology of indole alkaloid biosynthesis: the implication of recent discoveries", *Recent Adv. Phytochem.* **2003**, *37*, 181–202.

100. Endo, T.; Goodbody, A.; Vukovic, J.; Misawa, M. "Enzymes from *Catharanthus roseus* cell suspension cultures that couple vindoline and catharanthine to form 3',4'-anhydrovinblastine", *Phytochemistry* **1988**, *27*, 2147–2149.

101. Smith, J.I..; Amouzou, E.; Yamagushi, A.; McLean, S.; DiCosmo, F. "Peroxidase from bioreactor cultivated *Catharanthus roseus* cell cultures mediates biosynthesis of α-3'4'-anhydrovinblastine", *Biotechnol. Appl. Biochem.* **1988**, *10*, 568–574.

102. Hillou, F.; Costa, M.; Almeida, I.; Lopes Cardoso, I.; Leech, M.; Ros Barcelo, A.; Sottomayor, M. "Cloning of a peroxidase enzyme involved in the biosynthesis of pharmaceutically active terpenoid indole alkaloids in *Catharanthus roseus*" *Proc. VI*

Int. Plant Peroxidase Symp., ed. Acosta, M.; Rodriguez-Lopez, J.N.; Pedreno, M.A. Servicio de Publicaciones, Universidad de Murcia, Spain, **2002**, 152–158.

103. Sottomayor, M.; Lopez-Serrano, M.; DiCosmo, F.; Ros Barcelo, A. "Purification and characterization of α-3′,4′-anhydrovinblastine synthase (peroxidase-like) from *Catharanthus roseus* (L.) G. Don", *FEBS Lett.* **1998**, *428*, 299–303.

104. Sottomayor, M.; DiCosmo, F.; Ros Barcelo, A. "On the fate of catharanthine and vindoline during the peroxidase-mediated enzymatic synthesis of α-3′,4′-anhydrovinblastine", *Enz. Microb. Technol.* **1997**, *21*, 543–549.

105. Faber, K. *Biotransformations in organic synthesis*, 5th ed. **2008**, Springer.

106. Bornscheuer, U.T.; Kazlauskas, R.J. *Hydrolases in Organic Synthesis*, 2nd ed. **2006**, Wiley~VCH; and references therein.

107. Isowa, Y.; Ohmori, M.; Ichikawa, T.; Mori, K.; Nonaka, Y.; Kihara, K.; Oyama, K.; Saroh, H.; Nishimura, S. "The thermolysin-catalyzes condensation reactions of *N*-substituted aspartic acid and glutamic acids with phenylalannine alkyl esters", *Tetrahedron Lett.* **1979**, 2611-2612.

108. Morgan, B.; Dodds, D.R.; Zaks, A.; Andrews, D.R.; Klesse, R. "Enzymatic desymmetrization of prochiral 2-substituted-1,3-propanediols: a practical chemoenzymatic synthesis of a key precursor of SCH5108, a broad spectrum orally active antifungal agent", *J. Org. Chem.* **1997**, *62*, 7736-7743.

109. Bernhardt, P.; Hult, K.; Kazlauskas, R.J. "Molecular basis of perhydrolase activity in serine hydrolases", *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 2742-2746.

110. Lorenz, P.; Liebeton, K.; Niehaus, F.; Eck, J. "Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space", *Curr. Opin. Biotechnol.* **2002**, *13*, 572-577.

111. Zhou, H.; Xie, X.; Tang, Y. "Engineering natural products using combinatorial biosynthesis and biocatalysis", *Curr. Opin. Biotechnol.* **2008**, *19*, 590-596.

112. Liu, Z.; Weis, R.; Glieder, A. "Enzymes from higher eukaryotes for industrial biocatalysis", *Food Technol. Biotechnol.* **2004**, *42*, 237-249.

113. Bruggink, A.; Straathof, A.J.J.; Wielen, L.A.M. "A 'fine' chemical industry for life science products: green solutions to chemical challenges", *Adv. Biochem. Eng.* **2003**, *80*, 69-113.

114. Xie, X.; Watanabe, K.; Wojcicki, W.A.; Wangs, C.C.C.; Tang, Y. "Biosynthesis of lovastatin analogs with a broadly specific acyltransferase", *Chem. Biol.* **2006**, *13*, 1161-1169.

115. Xie, X.; Tang, Y. "Efficient synthesis of simvastatin by use of whole-cell biocatalysis", *Appl. Environ. Microb.* **2007**, *73*, 2054-2060.

116. Winter, J.M.; Moore, B.S. "Exploring the chemistry and biology of vanadium-dependent haloperoxidase", *J. Biol. Chem.* **2009**, *284*, 18577-18581.

117. Neumann, C.S.; Fujimori, D.G.; Walsh, C.T. "Halogenation strategies in natural product biosynthesis", *Chem. Biol.* **2008**, *15*, 99-109.

118. McCoy, E.; O'Connor, S.E. "Directed biosynthesis of alkaloid analogs in the medicinal plant periwinkle", *J. Am. Chem. Soc.* **2006**, *128*, 14276-14277.

119. Bernhardt, P.; McCoy, E.; O'Connor, S.E. "Rapid identification of enzyme variants for reengineered alkaloid biosynthesis in periwinkle", *Chem. Biol.* **2007**, *14*, 888-897.

120. Bernhardt, P.; Giddings, L.-A.; Loh, K.; O'Connor, S.E. "pH-indicator based assay for strictosidine synthase activity and its application in enzyme engineering", *manuscript in preparation.*

121. Runguphan, W.; O'Connor, S.E. "Metabolic reprogramming of periwinkle plant cell culture", *Nat. Chem. Biol.* **2009**, *5*, 151-153.

122. Bernhardt, P.; Usera, A.R.; O'Connor, S.E. "Strictosidine synthase from *Ophiorrhiza pumila* is a stereoselective and promiscuous biocatalyst for the Pictet-Spengler reaction", *submitted* **2009**.

123. Bernhardt, P.; O'Connor, S.E. "Synthesis and biochemical evaluation of des-vinyl secologanin *O*-analogs with alternate stereochemistry", *Tetrahedron Lett.* **2009**, *50*, 7118-7120.

124. Bernhardt, P.; Yerkes, N.; O'Connor, S.E. "Bypassing stereoselectivity in the early steps of alkaloid biosynthesis", *Org. Biomol. Chem.* **2009**, *7*, 4166-4168.

CHAPTER 2

DEVELOPING ASSAYS TO MONITOR STRICTOSIDINE SYNTHASE ACTIVITY

2.1 Introduction

Strictosidine synthase (STS) catalyzes a Pictet-Spengler reaction between tryptamine **1** and secologanin **2** to form strictosidine **3**.[1] Strictosidine **3** is the precursor for all monoterpene indole alkaloids (MIAs) and numerous quinoline alkaloids found in medicinal plants, such as *Catharanthus roseus*, *Rauvolfia serpentina*, *Ophiorrhiza pumila*, and *Chinchona officinalis* (Chapter 1).[1]

Precursor directed biosynthesis, which uses cell cultures to convert starting materials into natural products, is a useful strategy to generate alkaloid analogs in *C. roseus* (Chapter 3).[2] However, the narrow substrate scope of STS limits the application of precursor-directed biosynthesis in plant cell cultures from readily available tryptamine analogs.[3,4] This limitation can be circumvented by introducing STS variants with broad substrate specificities into plant cells (Chapter 3).[4,5]

Enzyme engineering, as a method to broaden the substrate scope of STS, requires efficient screening or selection methods to detect STS activity in mutant libraries. Since the only available analytical technique to detect STS activity was an HPLC-assay,[6] I sought higher-throughput methods. This chapter describes two visual detection methods of Pictet-Spenglerase activity. The first assay uses a pigmentation reaction to detect STS activity in the supernatant of yeast cell cultures that export STS.[4] The second assay depends on a dye that increases the assay sensitivity and enables activity screening directly on agar plates.[7]

2.2. Results and discussion

2.2.1 Development of a yeast expression system for STS activity

Saccharomyces cerevisiae is a useful microorganism for protein engineering, since a high *in vivo* homologous recombination frequency allows the ligation-independent library construction via co-transfomation of a linearized expression vector and a pool of mutated genes. *S. cerevisiae* also facilitates activity screening because a protein of interest can be exported by appending an N-terminal Mating Factor (MF)-Alpha signal sequence.[8]

The *sts* gene from *C. roseus* encodes a putative signal peptide (residues 1-26) for localization to the vacuole. Deletion of these 26 residues and insertion of the resulting *sts* gene in-frame with an encoded MF-Alpha signal sequence enabled the export of STS to the culture media after expression in *S. cerevisiae* (Figure 2.1, Lane 2). A C-terminal FLAG-peptide allowed purification by anti-FLAG affinity chromatography (Figure 2.1, Lanes 3-4).[9]

Characterization of wild-type STS purified from the yeast culture supernatant by a quantitative HPLC assay afforded kinetic constants for tryptamine **1** (Table 2.1). The kinetic parameters of this yeast construct are similar to parameters reported for STS isolated from *C. roseus* cells[10] and heterologously expressed STS from *R. serpentina* (88% sequence similarity to *C. roseus* STS).[11]
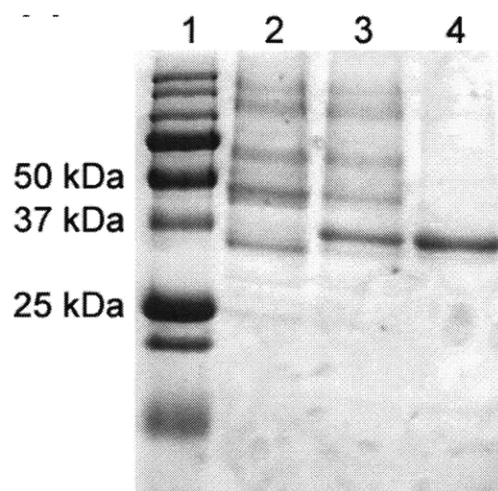
Figure 2.1 SDS-PAGE of STS expressed in *S. cerevisiae* BJ5465. STS has a calculated molecular weight of 37 kDa. Lane 1: Relative mass marker. Lane 2: STS found in the supernatant after expression (no FLAG-tag). Lane 3: STS found in the supernatant after expression (with FLAG-tag). Lane 4: STS purified from the supernatant by anti-FLAG antibody affinity chromatography. FLAG-tag: Asp-Tyr-Lys-Asp$_4$-Lys.

Mutagenesis of *R. serpentina* STS suggested that Glu309 is important for catalysis (corresponding to Glu315 in *C. roseus*).[11] The >3,000,000 reduced catalytic efficiency of the analogous mutant of *C. roseus* STS confirms that Glu315 is key for catalysis also in *C. roseus* STS.[4]

Table 2.1. Kinetic constants for STS with tryptamine **1**.[a]

Enzyme	V_{max} [U mg^{-1}] [b]	K_M [μM]	V_{max}/K_M [U mg^{-1} mM^{-1}] [b]
C. roseus STS heterologous expression	1.6 ± 0.1	1.5 ± 0.4	1000 ± 600
R. serpentina STS heterologous expression [c]	2.1	6.0	350
C. roseus STS cell culture [d]	10	7.0	1400
Glu315Ala *C. roseus* STS	0.00023 ± 0.00001	440 ± 50	0.0005 ± 0.0001
Glu309Ala *R. serpentina* STS [e]	0.0024	0.0054	0.45

[a] Kinetic assays were performed in sodium phosphate buffer (0.1 M, pH 7.0) using an internal standard (NAA, 0.075 mM). Samples were analyzed by HPLC for formation of product and the product was quantitated by peak integration; [b] 1 U = 1 μmol product formed per minute at pH 7.0 and 30 °C; [c] Reference 11; [d] Reference 10 – average of several isoforms; [e] Reference 11.

2.2.2 Development of a pigment-forming assay to detect STS activity

A yellow pigment forms when loganin is added to yeast cells that express secologanin synthase (SLS), STS, and strictosidine glucosidase (SGD).[12,13] I observed that a pigment also forms in assays containing tryptamine **1**, secologanin **2**, STS, and SGD, and hypothesized that this yellow pigment is related to a previously described amine adduct that appear when deglycosylated **3** is incubated in the presence of ammonium sulfate.[14]

The SGD-catalyzed deglucosylation reaction of strictosidine **3** yields an unstable hemi-acetal **6** that rearranges into an equilibrium mixture of isomers, including 4,21-dehydrogeissoschizine **4** and cathenamine **5** (Figure 2.2).[1,15] LC-MS analysis, however,

showed that compounds with *m/z* 493, 491, and 489 form in the pigmentation reaction at the expense of **4** and **5** (*m/z* 351, Figure 2.3). The additional 142 a.m.u. may derive from the Schiff-base condensation between the deglucosylated intermediate and tryptamine **1**. Thus, instead of ammonium sulfate,[14] tryptamine is a suitable amine donor for pigment formation. I also examined the products in the pigment reaction using *5(R)*-hydroxymethyl strictosidine **3b** (*m/z* 561), which adds 60 a.m.u. to the mass of strictosidine **3**, and found that compounds with *m/z* 553 and 551 form in the presence of 2'*(R)*-tryptophanol **1b**, consistent with the formation of a tryptamine adduct (Figure 2.4).

The gradual appearance of *m/z* 491 and 489 at the expense of *m/z* 493 (not shown) suggests that the tryptamine adducts oxidize. Amine adducts with deglucosylated strictosidine **3** can rearrange to form a dihydropyridine ring in the presence of ammonium sulfate; the dihydropyridine easily oxidizes to give a pyridinium ring.[14] Partial purification of the compounds corresponding to *m/z* 491 and 489 by preparative HPLC verified that these compounds are yellow-colored; the UV-VIS spectrum of the compound corresponding to *m/z* 489 shows absorption maxima at 314 nm and 427 nm and *m/z* 491 has one absorption maximum at 470 nm. ^{1}H NMR signals characteristic of indole and pyridinium protons support the assignment of the chromophore (Figure 2.2 and section 2.4.4).

Any screening method that employs this pigmentation assay relies on the substrate scope of SGD and that a pigment forms with alternative amines. I therefore synthesized and
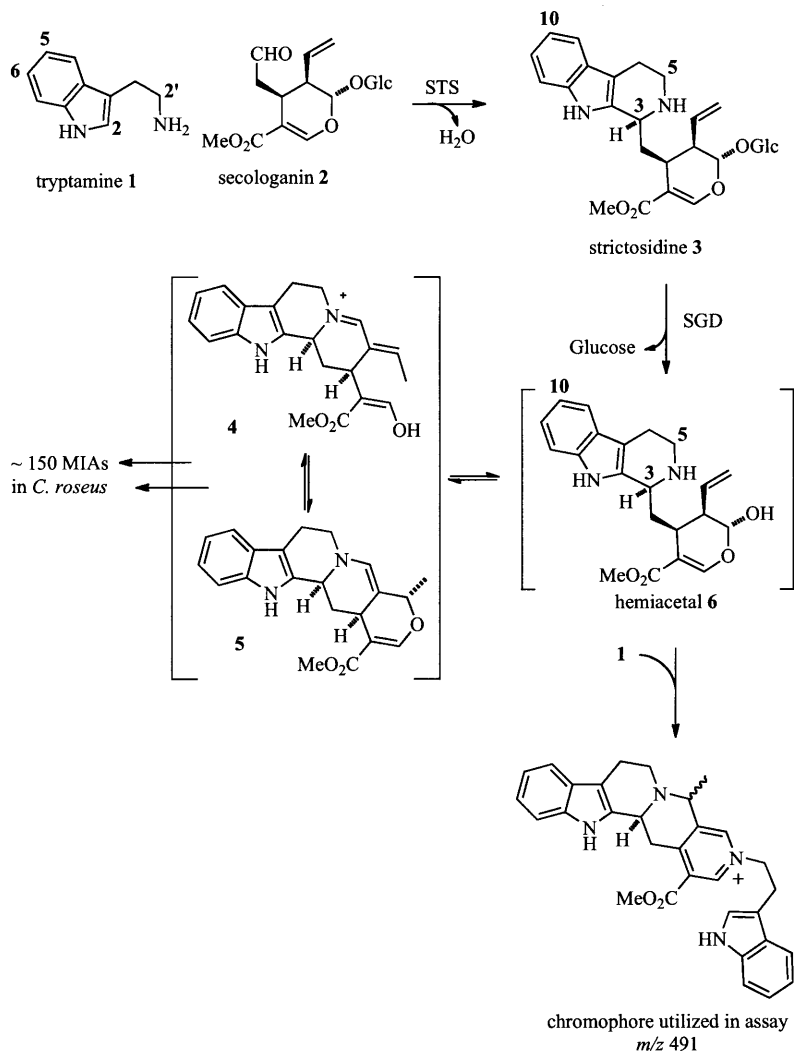
Figure 2.2 STS catalyzes the Pictet-Spengler reaction between tryptamine **1** and secologanin **2** to form strictosidine **3**. SGD catalyzes the deglucosylation of **3** and generates a mixture of products (the hemi-acetal that initially forms rearranges in different ways), including 4,21-dehydrogeissoschitzine **4** and cathenamine **5** (enamine form of **5** is shown). *In vitro*, the aglucone intermediate can react with excess amine substrate to form an adduct; for example, the chromophore utilized in the pigmentation assay. *In vivo*, the deglucosylated product mixture is channeled into the various branches of monoterpene indole alkaloid (MIA) biosynthesis.

58

m/z = 493.2
Putative tryptamine adduct

Assay

No strictosidine synthase

No strictosidine glucosidase

Tim

2.00   3.00   4.00   5.00   6.00   7.00   8.00   9.00   10.00   11.00   12.00
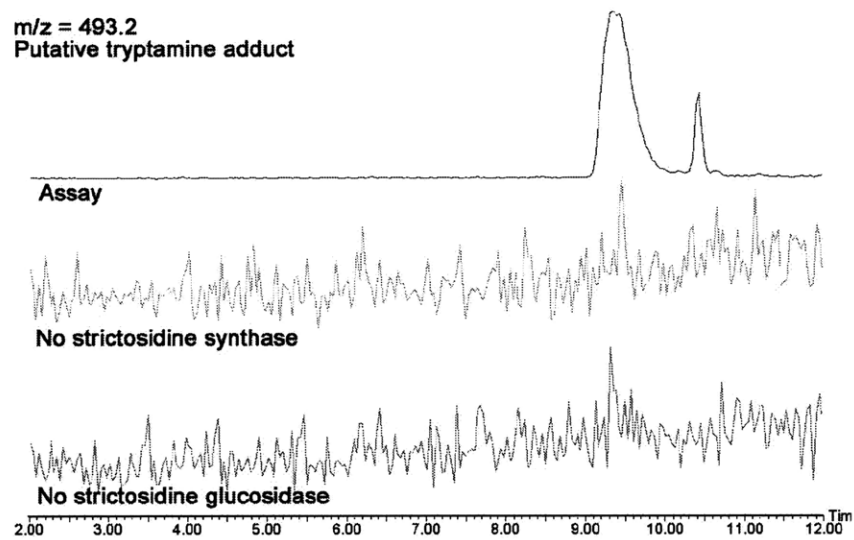
Figure 2.3 LC-MS traces. Formation of a tryptamine adduct (*m/z* 493) is strictly dependent on the presence of both strictosidine synthase (STS) and strictosidine glucosidase (SGD) (top trace). The adduct is not observed when either STS (middle trace) or SGD (bottom trace) is absent. The second, smaller, peak is the 493-isotope of *m/z* 491 that is also formed in the pigmentation assay.
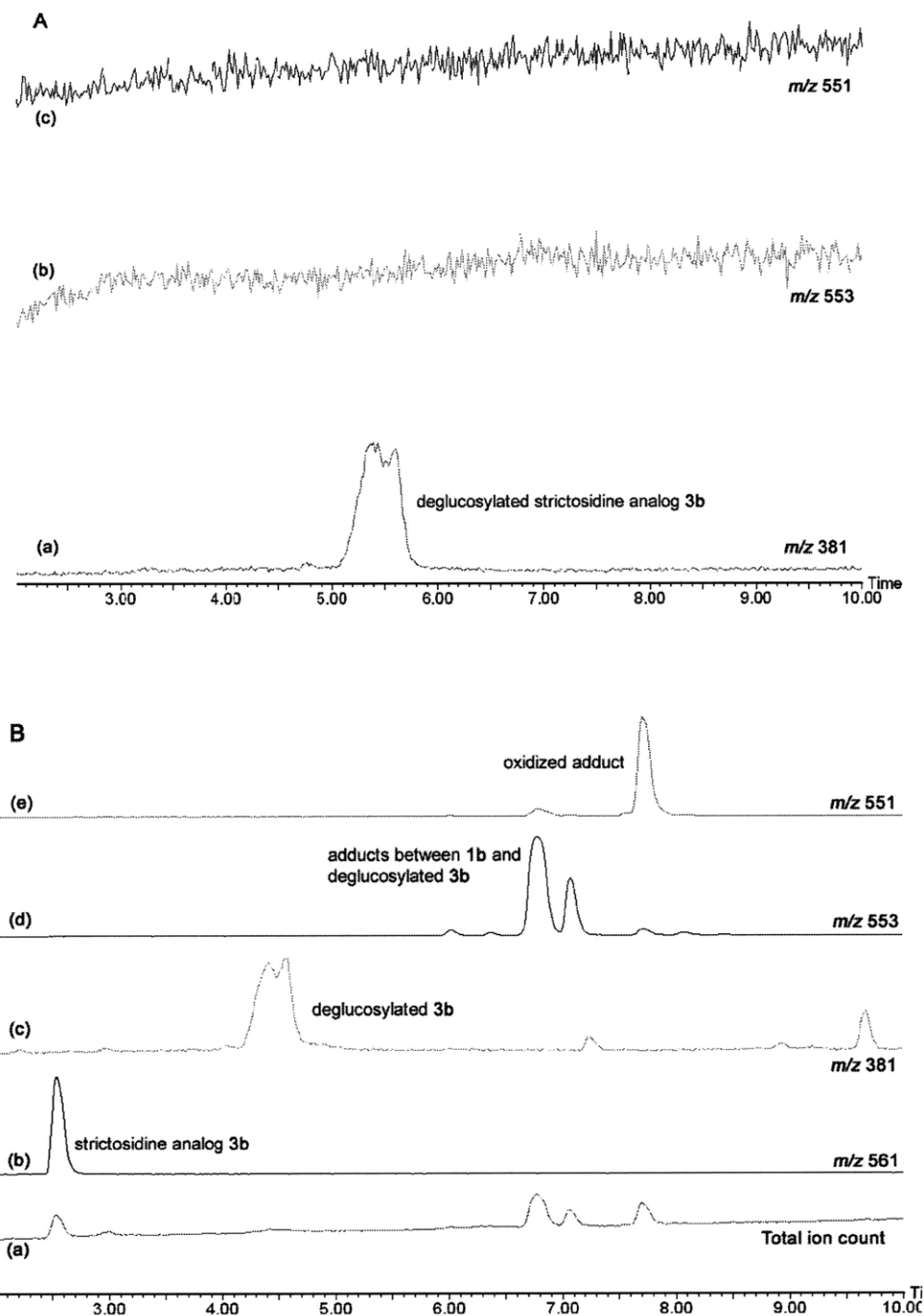
Figure 2.4 LC-MS traces. (A) Purified *5(R)*-hydroxymethyl strictosidine **3b** (*m/z* 561) and SGD were reacted for 3 h in the absence of 2'*(R)*-tryptophanol **1b**. Only the corresponding deglucosylated strictosidine analog is present (*m/z* 381) (trace a) and no adducts with 2'*(R)*-tryptophanol **1b** (*m/z* 553 or 551) are observed (traces b and c). (B) After addition of **1b** to a reaction containing deglucosylated **3b** (*m/z* 381), two new masses (*m/z* 553 and 551) are formed at the expense of the initial deglucosylated product (*m/z* 381).

60

purified diastereomeric mixtures of a variety of strictosidine analogs, including those derived from 5-methyl-, 5-methoxy-, and 2'(R)-tryptophanol **1b**. Pigments formed when I incubated the synthetic strictosidine analogs with SGD and an excess of the tryptamine analog (2 eq), suggesting that the assay is useful for screening enzyme libraries with tryptamine analogs. I replaced wild-type STS with Glu315Ala STS in the assay, and found that the weakly active mutant lead to the formation of a faint pigment (not shown), suggesting that this screen is sufficiently sensitive to screen for mutants with low activity.

2.2.3 Validation of the pigment-forming assay

Saturation mutagenesis introduces a degenerate codon into a gene, which after expression results in a protein library that contains all twenty naturally occurring amino acids at one specific position.[16] To validate the expression system and screening method, I constructed a saturation mutagenesis library of a tyrosine residue distant from the active site (Tyr167Ala reduces STS catalytic activity by 40-fold compared to wild-type STS, unpublished data) (Figure 2.5), expecting a pigmentation pattern where a different color intensity corresponds to a different catalytic activity. I constructed the gene library by *in vivo* homologous recombination in *S. cerevisiae* BJ5465 by co-transforming the linearized expression vector, pGAL-MF, and the mutated *sts* gene (Figure 2.5 and section 2.4.6). Inoculation of 88 colonies – an oversampling that ensures the library is >94% complete[17] – in expression media in 96-well deep-well plates and induction of protein expression, gave an STS mutant library. After centrifugation to remove the cells, I
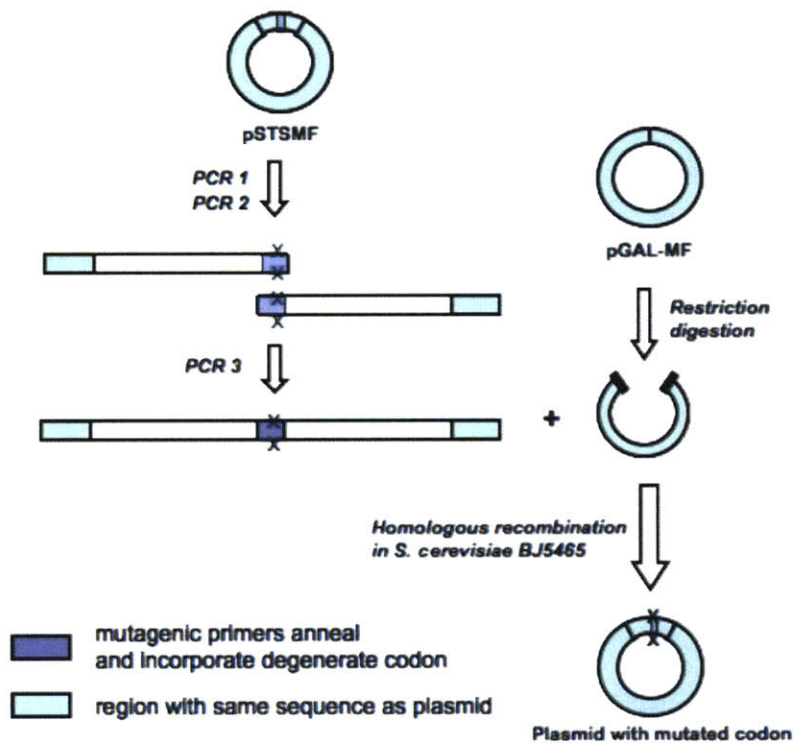
Figure 2.5 Saturation mutagenesis approach to construct STS mutant libraries. Three PCR reactions were used to assemble the mutated gene (using the NNK degenerate codon) with regions that flank the plasmid. The plasmid library was constructed by homologous recombination of the linearized experssion vector, pGAL-MF, and the mutated gene in *S. cerevisiae* BJ5465.
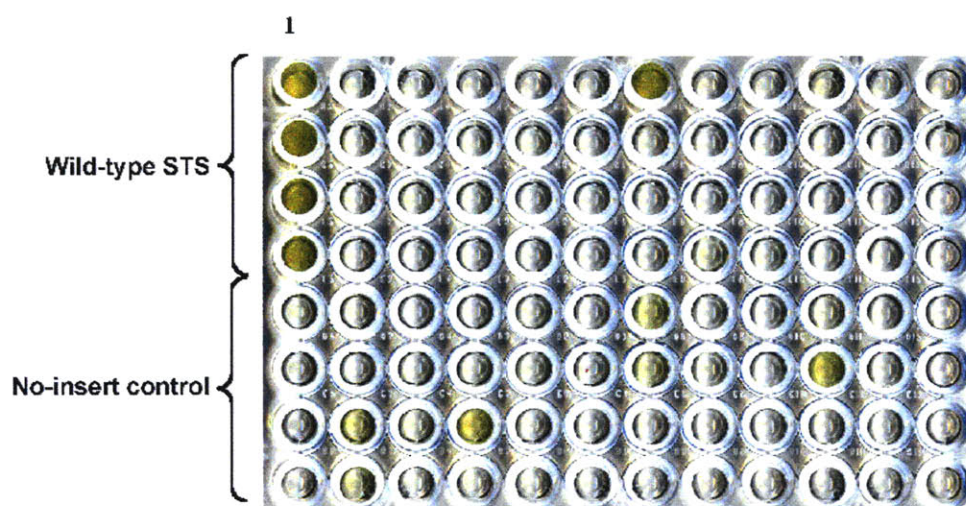
Figure 2.6 Tyr167Xxx STS library screened using the pigment-forming assay. Column 1 contains positive (wild-type STS) and negative (no-insert) controls. Several wells are yellow colored and contain either wild-type STS or other active STS variants. See Chapter 3 for an application of this screening method to expand the substrate scope of STS.

assayed the supernatant, which contains the exported enzyme, with tryptamine **1**, secologanin **2**, and SGD (Figure 2.6, section 2.4.6). DNA sequencing showed that wells that produce a strong yellow pigment in the assay contain yeast cells that export wild-type STS. Conversely, wells that have weaker or no pigmentation contain yeast cells that export STS variants with different mutations at tyrosine 167. This test-assay, which revealed different color intensities for different levels of activity with the natural substrates, indicated that the library construction–screening approach could be used to search for STS variants with altered substrate specificities (Chapter 3).

2.2.4 Use of bromothymol blue to monitor STS activity

pH-indicators can monitor enzymatic reactions that produce or consume protons.[18,19] Selected examples include glycosyl transferases,[20] hydrolases,[21,22] haloalkane dehalogenases,[23] and kinases.[24] The STS-catalyzed Pictet-Spengler reaction, and the subsequent SGD-catalyzed deglucosylation reaction, are proton neutral and it is not clear whether a pH-indicator can detect STS or STS-SGD-coupled activity. Nevertheless, when I incubated STS and SGD in a buffer at pH 7.1 with tryptamine **1** (1 mM) and secologanin **2** (1 mM) in the presence of the pH-indicator bromothymol blue **7** (BTB, pK_a 7.1; visual transition interval: pH 6.0-7.6), I observed a rapid color change (within a few minutes) from blue to yellow (relative to the pigmentation assay; see Figure 2.7). Control experiments showed that the color change is dependent on the deglucosylation of strictosidine **3** (Figure 2.7).
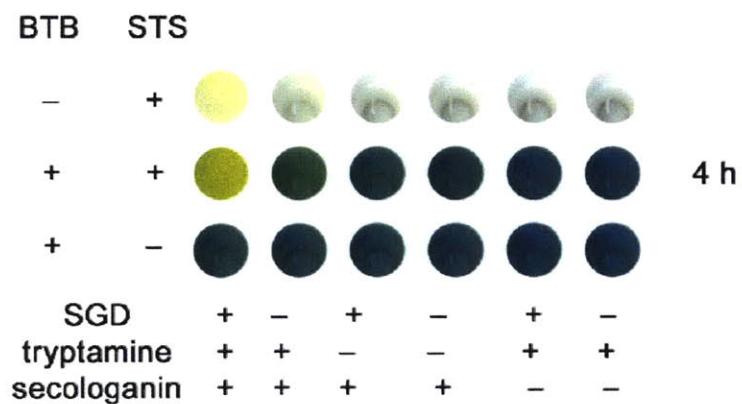
Figure 2.7 Comparison of the pigmentation assay (top row) and the indicator assay (middle row). The response time is significantly faster in the presence of the indicator BTB **7**. Omission of STS results in no color change; both substrates and SGD are necessary for the color change.
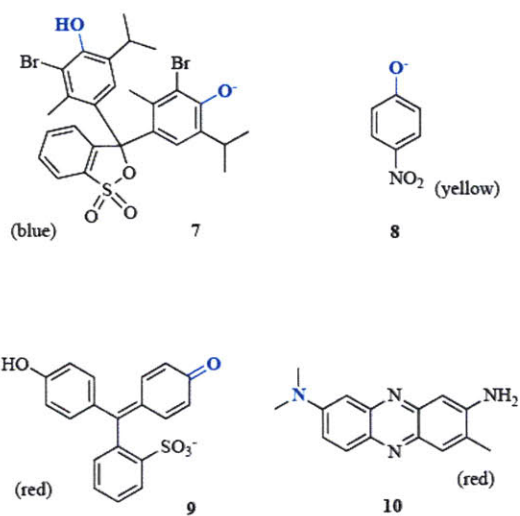
Figure 2.8 Structures of the deprotonated form of indicators mentioned in the text. The atom that becomes protonated is highlighted in blue; the color of the deprotonated species is indicated in parenthesis. Bromothymol blue (BTB, **7**); *p*-nitrophenolate (pNP, **8**); phenol red (**9**); and neutral red (NR, **10**).

During reactions where protons are released, the proton can add to the indicator or to the basic form of the buffering agent. Protonation of the indicator is what results in a color switch; thus, a low buffer strength is required for good sensitivity. Unlike known pH-indicator assays, however, the color change I measured at 616 nm is buffer-strength *independent* because buffer concentrations of >0.2 M still allowed a rapid color change. Although the oxidation reactions that result in the formation of the yellow-colored pigment (section 2.2.2) generate protons, they occur after several hours, and are therefore not on the time-scale of the blue-to-yellow color change, which is complete after a few minutes with sufficient amounts of substrates.

Two other pH-indicators have pK_a-values of approximately 7: *p*-nitrophenol (**8**, pNP, pK_a 7.1; visual transition interval: pH 5.0-7.0) and phenol red (**9**, pK_a 7.7; visual transition interval: pH 6.8-8.2) (Figure 2.8). pNP **8** has an absorption maximum at 404 nm and is unsuitable for use in the STS-SGD coupled assay, since the yellow-colored pigment, which also forms in the assay, absorbs at 404 nm (section 2.2.2). Phenol red **9** has a maximum absorbance at 563 nm and is therefore an alternative to BTB **7**. Titration of **9** by hydrochloric acid (1 mM) at pH 7.1 resulted in the expected, linear, proton-dependent blue shift. However, there was no corresponding color change in the enzymatic reaction at the same pH and buffer and indicator concentrations. Therefore, the color switch is likely a BTB-specific response.

The blue-to-yellow color switch corresponds to a negative to netural charge change of BTB **7**. Conversely, the red-to-yellow color change of phenol red **9** occurs when the indicator becomes a positively charged ion (Figure 2.8). Previous reports indicated that BTB **7** can quantify the amount of positively charged active components in pharmaceutical preparations by forming an ion pair.[25] Therefore, the color change observed in the STS-SGD coupled assay may be the result of an ion pair formation between the BTB indicator and a positively charged deglucosylated product.

2.2.5 Detection of STS activity on agar plates

Qualitative detection of a color change directly on agar plates is useful for high-throughput screening. After considerable experimentation using contrasting dyes, I found a combination of neutral red (**10**, NR) and BTB **7** that results in a blue-green to pink-brown color transition and that thereby can detect STS activity on BTB-dyed agar plates (Figure 2.9 and section 2.4.7). Yeast colonies that export the wild-type enzyme to the surrounding agar-medium, produce a zone of pink-brown color. Conversely, yeast colonies that export inactive enzyme form no color zones (Figure 2.9). Since the screened yeast colonies are viable, this method does not rely on replica plating.

To demonstrate the application of the indicator assay, I screened a triple-site saturation mutagenesis library of the *sts* gene, leading to Val214Xxx-Pro215Xxx-Gly216Xxx *C roseus* STS, using tryptamine **1** and secologanin **2** (section 2.4.8). DNA sequencing of halo-forming colonies identified that Val214Met STS is active, which is the same mutant

Wild-type STS
(active)

Glu315Ala STS
(inactive)
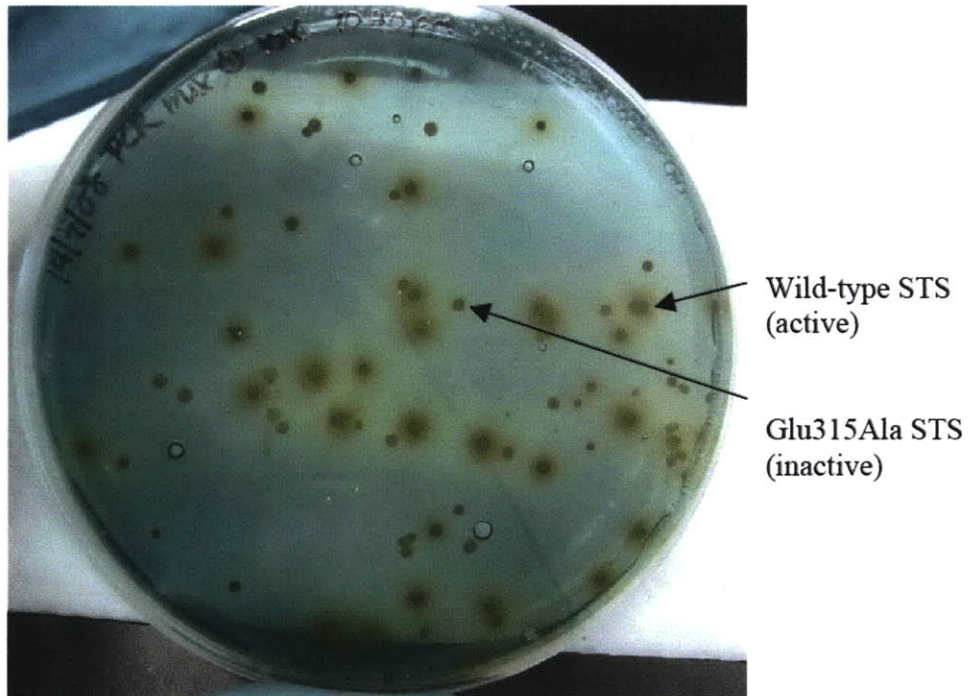
Figure 2.9 *S. cerevisiae* BJ5465, containing plasmids that encode either inactive (Glu315Ala) or active (wild-type) STS, were spread as a 1:1 mixture on yeast minimal media containing agar (1%), potassium phosphate buffer (0.15 M, pH 7.1), galactose (2%), glucose (0.8%), and BTB **7**. During a 72 h incubation at 30 °C, yeast colonies appear that continuously export STS to the surrounding media. The activity of the exported enzymes is assayed by pouring a molten solution (~3 mL) of agarose (0.5%), BTB **7** (60 mg L^{-1}), NR **10** (20 mg L^{-1}), tryptamine **1** (2 mM), secologanin **2** (1 mM), and SGD (~10 μg) in potassium phosphate buffer (20 mM, pH 7.1). After the agarose top layer had solidified, the plate was incubated for 1 h at 30 °C. DNA sequencing of plasmids isolated from yeast cells showed that colonies that express wild-type STS result in a colored zones on the screening plate, whereas yeast colonies that express Glu315Ala STS result in no colored zones.

identified in the single-site saturation mutagenesis library, Val214Xxx STS (Chapter 3). The screening also identified double mutants (Val214Gly/Pro215Gly,Ala) that retain the hydrophobic character of the tryptamine binding site (section 2.4.8). This shows that a triple-NNK approach can identify variants not identified using single-NNK library construction (Chapters 3). Chapter 7 describes several future directions that apply the indicator screening method for enzyme engineering.

2.3. Conclusions

This chapter describes the development of two new methods to detect STS activity in the culture supernatant after expression in *S. cerevisiae*. The pigment assay, which relies on the formation of an adduct between tryptamine **1** and deglucosylated strictosidine **3**, is useful for the 96-well plate screening of STS activity in saturation mutagenesis libraries. This assay was instrumental for changing the substrate scope of *C. roseus* STS (Chapter 3).

Addition of a pH-indicator, bromothymol blue (BTB), to the STS-SGD-coupled activity, dramatically increases the screening throughput and sensitivity, compared to the pigment assay. Although the pigment eventually reinforces the blue-to-yellow color change of the indicator, the time-scale of the indicator color change is very different from that of the pigment, suggesting an independent mechanism of color change. Although this mechanism remains unclear, the response is BTB-specific and may involve the formation of an ion pair between the negatively charged BTB species and a positively charged

deglucosylated product. I applied the agar plate adapted indicator method to screen a triple-site-saturation mutagenesis library; future experiments that use this assay are in progress.

2.4 Experimental methods

2.4.1 General methods and analytical techniques

Secologanin **2** was obtained by MeOH-extraction of *Lonicera tatarica* as described previously.[26] Tryptamine **1** and tryptamine analogs were purchased in highest available purity and used without further purification.

LC-MS analysis was performed on a Waters LC-MS system. Compounds were separated by an Acquity UPLC BEH C18-column (1.7 μm, 2.1 x 100 mm) equipped in tandem with an MS technologies Micromass LCT Premier with an electrospray ionization (ESI) source and time-of-flight (TOF) detector. The chromatography solvent system consisted of MeCN in water with 0.1% formic acid. The capillary and sample cone voltages were 2000 V and 30 V, respectively. The desolvation and source temperature were 350 °C and 100 °C, respectively. The cone and desolvation gas flow-rates were 20 L h^{-1} and 700 L/h, respectively. HRMS spectra were obtained in accurate mass mode (W-mode) with reserpine (*m/z* 609.2812, as 1 nM solution, 40 μL min^{-1}) as reference. The HPLC-based assay was carried out on a Beckman-Coulter System Gold system equipped with a 125 solvent module, 125 autosampler and a 168 detector. Data was processed (integrated) using 32Karat (v 7). A solvent system composed of MeCN in water with 0.1% TFA was

used with a reverse-phase column (HiBar RT 250-4, RP-select B, 5 μm). NMR spectra were acquired on Varian Inova 500 MHz or Bruker 600 MHz spectrometers.

2.4.2 Heterologous expression of *C. roseus* STS in *S. cerevisiae*

pGAL-MF (Dualsystems Biotech AG) is an expression vector for *S. cerevisiae* that encodes an MF-alpha signal-peptide directly upstream of the multiple cloning site. This yeast signal sequence enables secretion of heterologously expressed protein into the culture media during expression. pGAL-MF also contains a galactose-inducible promoter and confers resistance to ampicillin (bacteria) or growth in uracil dropout media (yeast, deficient in the biosynthesis of uracil). *S. cerevisiae* cells without pGAL-MF are maintained in yeast-peptone-dextrose media (YPD, composition per liter: 10 g yeast extract, 20 g peptone and 2% (w/v) glucose); cells with pGAL-MF are maintained on a minimal media depleted of uracil (SCMM-U, *S. cerevisiae* minimal media, composition per liter: 6.7 g yeast extract without amino acids, 1.92 g yeast supplemental media without uracil, and 2% (w/v) glucose). To induce expression the SCMM-U media is also supplemented with 2% (w/v) galactose (glucose may be replaced with raffinose for optimized protein expression).

A synthetic gene encoding *C. roseus* STS, codon-optimized for expression in yeast was amplified from the lab-stock-plasmid pMCR (Kan^R) with PCR primers that introduce recognition sites for *Xho* I and *Hind* III (underlined) and the FLAG-tag epitope (italics):

5'-CCG<u>AAG CTT</u>TCACCAATCTTGAAGAAG-3'

5'-TCCTACGTTTCCTCC*GGATCTGACTACAAGGATGACGACAACAAGTAG*

<u>CTCGAG</u> CGG-3'

These primers were designed to eliminate the first 26 residues corresponding to a putative native *N*-terminal signal sequence for vacuolar localization. The PCR product was purified using a PCR cleanup kit (Aldrich) and subcloned into the pGEM-T Easy Vector (Promega, Madison, WI) using the manufacturer's instructions. After DNA sequencing, which verified the integrity of the TA-region, the gene was cloned from the shuttle vector into pGAL-MF (Dualsystems Biotech AG) to obtain pSTSMF-FLAG, which was transformed into electro-competent TOP10 *E. coli* cells. After DNA sequencing, the correct construct was transformed into the expression host, *S. cerevisiae* BJ5465.

Transformed *S. cerevisiae* BJ5465 harboring pSTSMF-FLAG was grown overnight in SCMM-U supplemented with 2% glucose. The cells were recovered from an overnight culture (OD_{600} ~ 10) and washed in SCMM-U without any carbon source to remove the glucose. The overnight culture was then diluted with SCMM-U media containing 2% (w/v) raffinose and 2% (w/v) galactose to yield a culture with an OD_{600} = 1. The culture was incubated at 25 °C for 72 h at 225 rpm. The cells were then removed by centrifugation (5 min, 3,000*g*, 4 °C) and the supernatant was concentrated and buffer exchanged to give an enzyme preparation in sodium phosphate buffer (0.05 M, pH 7.0).

The enzyme preparation was then subjected to purification using anti-FLAG M2 resin (Sigma) according to the manufacturer's instructions. STS was eluted with FLAG peptide, buffer exchanged to sodium phosphate buffer (0.05 M, pH 7.0, 20% glycerol) and stored in 0.02 mL aliquotes at -20 °C.

2.4.3 Site-directed mutagenesis

The Glu315Ala mutant of STS was constructed using the complementary primer approach applying Stratagene's site-directed mutagenesis kit according to the manufacturer's instructions. The following primer-pair was used:

5'-GAAGGTGAACACTTC**GCT**CAAATTCAAGAGCAC-3'

5'-GTGCTCTTGAATTTG**AGC**GAAGTGTTCACCTTC-3'

2.4.4 Expression and purification of SGD

C. roseus SGD was expressed as a 6x-His construct in *E. coli* as described previously,[27] except that elution from the column was accomplished by a step-wise increase in imidazole concentration.

2.4.5 HRMS and NMR of tryptamine adducts

The tryptamine adducts were analyzed by HRMS: ESI-MS($C_{31}H_{33}N_4O_2^+$) *m/z* calculated: 493.2604 [M+H]$^+$, found: 493.2579 [M+H]$^+$; ESI-MS($C_{31}H_{31}N_4O_2^+$) *m/z* calculated: 491.2447 [M]$^+$, found: 491.2458 [M]$^+$; ESI-MS($C_{31}H_{29}N_4O_2^+$) *m/z* calculated: 489.2291

74

[M]$^+$, found: 489.2290 [M]$^+$. A pyridinium protons exist in *m/z* 491; NMR (600 MHz,

CH$_3$CN): δ 9.50 (s, indole-NH), 9.26 (s, indole-NH), 7.78 (s, 1H, pyridinium-H), 7.57 (d,

1H, *J* = 7.7, Ar-H, cosy: indole a), 7.43-7.40 (m, 3H, Ar-H, cosy: indole a,b), 7.20 (dt,

1H, Ar-H, *J* = 0.9, 7.7 cosy: indole a), 7.13 (t, 1H, *J* = 0.8, 7.8, Ar-H, cosy: indole a),

7.07 (t, 1H, *J* = 0.8, 7.9, Ar-H, cosy: indole b), 7.06 (d, *J* = 1.5, pyrrole-H), 6.94 (t, 1H, *J*

= 0.8, 7.9, Ar-H, cosy: indole b), 6.82 (s, 1H, pyridinium-H). A pyridinium ring also

exists in *m/z* 489: NMR (600 MHz, CD$_3$Cl): δ 9.86 (s, indole-NH), 9.14 (s, indole-NH),

8.44 (s, 1H, pyridinium-H), 7.82 (s, 1H, pyridinium-H), 7.61 (d, 1H, Ar-H), 7.45 (t, 1H,

Ar-H), 7.37 (t, 1H, Ar-NH), 7.30-7.16 (m, 4H, Ar-H, pyrrole-H), 6.91 (t, 1H, Ar-H), 6.74

(t, 1H, Ar-H).

2.4.6 Construction of the saturation mutagenesis library

Two flanking primers (pGALMF_fwd, pGALMF_rev, Table 2.2) and complementary

mutagenic primers for the Tyr167Xxx saturation mutagenesis library were designed to

incorporate the NNK (MNN, reverse sense, **N** = A, C, G, T; **K** = G, T; **M** = A, C). PCR

reactions were carried out using Expand High-Fidelity PCR Kit (Roche Biosciences, Inc.)

following the manufacturer's instructions, except that 1.0 μl DMSO was added to each

50 μl PCR reaction. The thermocycling program started with denaturation (94 °C, 3 min),

followed by 29 cycles of denaturation (94 °C, 30 s), annealing (55 °C, 30 s) and extension

(72 °C, 1 min) and ended with an extension (72 °C, 7 min). Increased product yield was

sometimes achieved by lowering the annealing temperature to 52 °C.

Table 2.2. PCR primers for saturation mutagenesis.

Primer name	Primer sequence (5'->3')
pGALMF_fwd	G G G G T A T C T T T G G A T A A A A G A G A G G C T G A A G C T G A A T T C G A T A T C
pGALMF_rev	G G G A G G G C G T G A A T G T A A G C G T G A C A T A A C T A A T T A C A T G A C
Tyr167Xxx_fwd	G G T A G A T T G A T G A A G <u>N N K</u> G A C C C A T C C A C T A A G
Tyr167Xxx_rev	C T T A G T G G A T G G G T C <u>M N N</u> C T T C A T C A A T C T A C C

PCR products were separated by agarose gel (1%) electrophoresis and the excised gel bands were extracted and purified using GenElute Gel Extraction Kit (Sigma-Aldrich, St. Louis, MO). In the first PCR reaction, the forward flanking primer (pGALMF_fwd) was added in equimolar concentration to the reverse mutagenic primer. The reverse flanking primer (pGALMF_rev) was likewise added with the forward mutagenic primer in a second PCR reaction. The two resulting PCR fragments both contained the degenerate codon (Figure 2.5). The PCR products were purified by spin-column (after digestion by *Dpn* I to remove parental plasmid DNA) and then added together in a splicing-overlap extension (SOE) PCR, to produce the full-length mutated gene. The mutated gene was separated using agarose (1%) gel electrophoresis and the 1.1-kb DNA band was excised and purified. The pGAL-MF plasmid was digested with *Hind* III and *Xho* I restriction enzymes in the recommended buffer and purified by agarose gel electrophoresis.

The purified and digested vector and the mutated gene were added in ratios ranging from 1:8–1:2 to electro-competent *S. cerevisiae* BJ5465 cells, subjected to electroporation, diluted in 1 mL, 1 M ice-cold sorbitol and then plated onto two Petri dishes (15-cm in diameter) with SCMM-U agar media supplemented with 2% (w/v) glucose. The plates were incubated at 30 °C and colonies appeared after 48-72 hours. The 42-45 base pair

overlap between the vector and the gene allowed homologous recombination in *S. cerevisiae*, which produced a plasmid with the mutated inserted gene (Figure 2.5). The recombined plasmid allowed growth on the uracil-deficient (selection) media.

2.4.7 Expression and screening of saturation mutagenesis libraries

Individual colonies from the library of *S. cerevisiae*, containing pGAL-MF with mutated *sts* (section 2.4.6), were inoculated in 96-well flat-bottom deep-well plates with 0.2 mL SCMM-U broth supplemented with 2% (w/v) glucose and incubated at 30 °C for 24 h at 225 rpm in a 96-well deep-well plate. To this seed culture, 0.8 mL SCMM-U broth containing 1% (w/v) glucose and 2% (w/v) galactose was added and the plate was incubated at 20 °C for an additional 72 h. The minimal amount of sodium phosphate buffer (1 M, pH 7.0) was added using a multi-channel pipette to adjust to neutral pH and the 96-well expression plate was then centrifuged (4 °C, 3,000*g*, 10 min) to pellet the cells.

The buffered yeast culture supernatant (50 µl) was added to a two-fold concentrated assay solution (relative the final concentrations) (50 µl) in 96-well microtiter plates (0.25-mL well volume) containing the following: tryptamine **1** (4 mM), secologanin **2** (2 mM), and SGD in sodium phosphate buffer (50 mM, pH 7.0). The plates were incubated at 30 °C for up to 48 hours during which time active STS variants produced a yellow color (Figure 2.6). Wells on the expression plate corresponding to color formation on the screening plate were recultured in SCMM-U media supplemented with 2% (w/v) glucose.

Plasmid was extracted from the yeast cells using a Zymoprep Plasmid Extraction Kit according to the manufacturer's instructions (Zymo Research). Plasmid was transformed into and isolated from *E. coli* TOP10 (this results in cleaner DNA isolation) and then submitted for sequencing analysis using the following sequencing primers:

5'-GCTAAAGAAGAAGGGGTATCT-3'

5'-GGAGGGCGTGAATGTAAGCG-3'

2.4.8 Media recipes and agar plate screening-procedure using bromothymol blue

Agar (7.5 g), yeast supplemental media (1.44 g), yeast nitrogen base (5.03 g), and dd-water (0.53 L) was added to a 1 L Erlenmeyer flask. After autoclaving, the agar solution was supplemented with potassium phosphate buffer (0.115 L, 1 M, pH 7.0), galactose (0.075 L, 20% filter sterilized), glucose (0.03 L, 20% filter sterilized), and BTB **7** (to a final concentration of 60 mg L^{-1}). The solution was suspended in petri dishes (0.75 L makes around 30 standard plates) giving agar plates with the following final concentrations: agar (1%), phosphate buffer (0.15 M), 2% (w/v) galactose, 0.8% (w/v) glucose, and BTB **7** (60 mg L^{-1}).

S. cerevisiae cells, *e.g.* electroporated with DNA, were spread on the agar-indicator plates. After 72 h at 30 °C the yeast colonies had exported sufficient amounts of STS in their surroundings to be screened. An agarose solution was first prepared in a 25-mL autoclaved Erlenmeyer flask, containing agarose (0.5 g), BTB **7** (24 mL of 0.25 g L^{-1}

autoclaved stock solution), neutral red **10** (2 mL of 1 g L^{-1} autoclaved stock solution), potassium phosphate buffer (50 mL, 0.05 M, pH 7.0), and dd-water (24 mL). The solution was heated to a boil until the agarose had *completely* dissolved. The molten agarose solution was equilibrated to 55 °C in a water bath. About 8 mL of agarose solution was required to screen two standard 8.4-cm in diameter plates with yeast colonies. Tryptamine **1** (0.15 mL, 0.05 M filter sterilized stock), secologanin **2** (0.15 mL, 0.05 M filter-sterilized stock), and SGD (~10 µg) were added to 7.7 mL of the indicator-buffer-agarose solution. This mix was then *immediately* added on top of the agar plates with the yeast colonies. After the agarose had solidified, the plate was incubated at 30 °C; colored zones appeared after about 1 h (Figure 2.7).

2.4.9 Construction of triple-NNK focused libraries

A triple-NNK library of the *sts* gene was constructed using Stratagene's site-directed mutagenesis kit according to the manufacturers instructions using the following two primers that encode STS Val214Xxx, Pro215Xxx, Gly216Xxx:

5'-GAAGGAGTTGCAC<u>NNKNNKNNK</u>GGTGCTGAGATTTC-3'

5'-GAAATCTCAGCACC<u>MNNMNNMNN</u>GTGCAACTCCTTC-3'

The wild-type template was digested using *Dpn* I and the mutated plasmid was transformed into the XL-1 Blue cells. Plasmid isolated from the bacterial cells was transformed into *S. cerevisiae* BJ5465 and sorbitol (1 M, 1 mL) was added. After

incubation at 30 °C for 48 h, the cells were pelleted and diluted 1000-fold in sterile water. This dilution was spread on 8.4-cm in diameter agar-indicator plates (section 2.4.8) and the plates were incubated at 30 °C for 72 h (colonies appear after 48 h). The agarose assay solution described in section 2.4.8 was aliquoted (4 mL) onto the colony-containing agar plates and the plates were incubated at 30 °C. Halo-formation was observed and noted up to 24 h after incubation. Active colonies were picked from underneath the agarose top-layer and restreaked on SCMM-U agar plates supplemented with 2% (w/v) glucose. Colonies were inoculated in SCMM-U broth with 2% (w/v) glucose and the resulting cells were harvested from the yeast culture (1 mL) using a yeast miniprep kit (Zymoprep, Zymo Research). The plasmid obtained from yeast was transformed (3 µL) into electrocompetent *E. coli* TOP10 cells. After selection of plasmid-harboring *E. coli* cells, plasmid was isolated and submitted for sequencing. The sequencing data for ten representative colonies are given below. In summary, the sampled active colonies consisted of 50% unique active variants, 70% non-wild-type active variants, and 30% wild-type sequences.

Sample	DNA sequence	Peptide sequence
WT	GTT CCA GGT	V-P-G
1	GTT CCA GGT	V-P-G
2	GTT CCA GGT	V-P-G
3	GGT GTG CAT	G-V-H
4	GTT CCA GGT	V-P-G

5	GGT GGG GGG	G-G-G
6	GGG GCG GGT	G-A-G
7	GGT GGG GGG	G-G-G
8	GGG GCG GGG	G-A-G
9	GGG GGG GCT	G-G-A
10	ATG CCA GGT	M-P-G

2.5 References

1. O'Connor, S.E.; Maresh, J.J. "Chemistry and biology of monoterpene indole alkaloid biosynthesis", *Nat. Prod. Rep.* **2006**, *23*, 532-547.

2. McCoy, E.; O'Connor, S.E. "Directed biosynthesis of alkaloid analogs in the medicinal plant periwinkle", *J. Am. Chem. Soc.* **2006**, *128*, 14276-14277.

3. McCoy, E.; Galan, M.C.; O'Connor, S.E. "Substrate specificity of strictosidine synthase", *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2475-2478.

4. Bernhardt, P.; McCoy, E.; O'Connor, S.E. "Rapid identification of enzyme variants for reengineered alkaloid biosynthesis in periwinkle", *Chem. Biol.* **2007**, *14*, 888-897.

5. Runguphan, W.; O'Connor, S.E. "Metabolic reprogramming of periwinkle plant cell culture", *Nat. Chem. Biol.* **2009**, *5*, 151-153.

6. Pennings, E.J.; van den Bosch, A.J.; van der Heijden, R.; Stevens, L.H.; Duine, J.A.; Verpoorte, R. "Assay of strictosidine synthase from plant cell cultures by high-performance liquid chromatography", *Anal. Biochem.* **1989**, *176*, 412–415.

7. Bernhardt, P.; Giddings, L.-A.; Loh, K.; O'Connor, S.E. "pH-indicator based assay for strictosidine synthase activity and its application in enzyme engineering", *manuscript in preparation.*

8. Brake, A.J.; Merryweather, J.P.; Coit, D.G.; Heberlein, U.A.; Masiarz, F.R.; Mullenbach, G.T.; Urdea, M.S.; Valenzuela, P.; and Barr, P.J. "Alpha-factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*", *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 4642-4646.

9. Knappik, A.; Pluckthun, A. "An improved affinity tag based on the FLAG peptide for the detection and purification of recombinant antibody fragments", *Biotechniques* **1994,** *17,* 754-761.

10. de Waal, A.; Meijer, A.H.; Verpoorte, R. "Strictosidine synthase from *Catharanthus roseus*: Purification and characterization of multiple forms", *Biochem. J.* **1995,** *306,* 571-580.

11. Ma, X.; Panjikar, S.; Koepke, J.; Loris, E.; Stockigt, J. "The structure of *Rauvolfia serpentina* strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins", *Plant Cell* **2006,** *18,* 907-920.

12. Geerlings, A.; Redondo, F.J.; Memelink, J.; Contin, A.; van der Heijden, R.; Verpoorte, R. "Screening method for cDNAs encoding putative enzymes converting loganin into secologanin by a transgenic yeast culture", *Biotech. Tech.* **1999,** *13,* 605-608.

13. Luijendijk, T.J.C.; Stevens, L.H.; Verpoorte, R. "Reaction for the localization of strictosidine glucosidase activity on polyacrylamide gels", *Phytochem. Anal.* **1996,** *7,* 16-19.

14. Heinstein, P.; Höfle, G.; and Stöckigt, J. "Involvement of cathenamine in the formation of N-analogues of indole akaloids", *Planta Med.* **1979,** *37,* 349-357.

15. Brown, R.T.; Leonard, J. "Biomimetic synthesis of cathenamine and 19-epicathenamine, key intermediates to heteroyohimbine alkaloids", *J. Chem. Soc. Chem. Comm.* **1979,** 877-879.

16. Derbyshire, K.M.; Salvo, J.J.; Grindley, N.D. "A simple and efficient procedure for saturation mutagenesis using mixed oligodeoxynucleotides", *Gene* **1986**, *46*, 145-152.

17. Patrick, W.M.; Firth, A.E.; and Blackburn, J.M. "User friendly algorithms for estimating completeness and diversity in randomized protein encoding libraries", *Protein Eng.* **2003**, *16*, 451-457.

18. Wahler, D.; Reymond, J.-L. "High-throughput screening for biocatalysts", *Curr. Opin. Biotechnol.* **2001**, *12*, 535-544.

19. Reymond, J.-L. ed. "Enzyme assays: high-throughput screening, genetic selection, and fingerprinting", Wiley~VCH **2006**.

20. Persson, M.; Palcic, M.M. "A high-throughput pH indicator assay for screening glycosyltransferase saturation mutagenesis libraries", *Anal. Biochem.* **2008**, *378*, 1-7.

21. Moris-Varas, F.; Shah, A.; Aikens, J.; Nadkarni, N.P.; Rozzell, J.D.; Demirjian, D.C. "Visualization of enzyme-catalyzed reactions using pH indicators: rapid screening of hydrolase libraries and estimation of the enantioselectivity", *Bioorg. Med. Chem.* **1999**, *7*, 2183-2188.

22. Janes, L.E.; Löwendahl, A.C.; Kazlauskas, R.J. "Rapid quantitative screening of hydrolases using pH indicators: finding enantioselective hydrolases", *Chem. Eur. J.* **1998**, *4*, 2324-2331.

23. Holloway, P.; Trevors, J.T.; Lee, H. "A colorimetric assay for detecting haloalkane dehalogenase activity", *J. Microb. Met.* **1998**, *32*, 31-36.

24. Chapman, E.; Wong, C.-H. "A pH sensitive colorimetric assay for the high-throughput screening of enzyme inhibitors and substrates: a case study using kinases", *Bioorg. Med. Chem.* **2002**, *10*, 551-555.

25. For example: Rahman, N.; Hejaz-Azmi, S.N. "Extractive spectrophotometric methods for determination of diltiazem HCl in pharmaceutical formulations using bromothymol blue, bromophenol blue and bromocresol green", *J. Pharm. Biomed. Anal.* **2000**, *24*, 33-41.

26. Galan, M.C.; O'Connor, S.E. "Semi-synthesis of secologanin analogues", *Tetrahedron Lett.* **2006**, *47*, 1563-1565.

27. Yerkes, Y.; Wu, J.X.; McCoy, E.; Galan, M.C.; Chen, S.; O'Connor, S.E. "Substrate specificity and diastereoselectivity of strictosidine glucosidase", *Bioorg. Med. Chem. Lett.* **2008**, *15*, 3095-3098.

CHAPTER 3

REENGINEERING STRICTOSIDINE SYNTHASE

3.1 Introduction

Monoterpene indole alkaloids (MIAs) from *Catharanthus roseus* (Madagascar periwinkle) are structurally complex natural products.[1] Several MIAs have important biological activities,[2] including the anti-hypertensive compound ajmalicine **1**, the topoisomerase type-II inhibitor serpentine **2**, and the anticancer agents vinblastine **3** and vincristine **4** (Figure 3.1).[2-4] Derivatives of vinblastine **3**, such as Vinflunine **5** (4'-deoxy-20',20'-difluoro-C'-norvincaleukoblastine), show improved anti-cancer activity compared to vinblastine **3**, indicating that small changes to an alkaloid scaffold can improve or modulate bioactivity (Figure 3.1).[5-7] The current supply of MIA analogs, however, depends on low-yielding semi-syntheses from isolated pathway intermediates. Metabolic engineering of *C. roseus* to obtain natural product analogs is an environmentally friendly and potentially efficient alternative to semi-synthesis and total-synthesis.

Precursor-directed biosynthesis with *C. roseus* hairy-root cultures and seedlings utilizes the plant biosynthetic pathways to convert precursor derivatives into MIA analogs.[8] However, the narrow substrate scope of strictosidine synthase (STS), the enzyme at the entry point in the pathway, limits precursor-directed biosynthesis efforts.[9] STS catalyzes a stereoselective Pictet-Spengler reaction between tryptamine **6a** and secologanin **7** to yield the tetrahydro-β-carboline 3*(S)*-strictosidine **8a** as a single stereoisomer (Figure 3.2).[10] Expanding the substrate scope of STS may be instrumental in bestowing the plant – or a microbe with a reconstituted plant pathway – the ability to make unnatural MIA

ajmalicine **1**   serpentine **2**

vinblastine, **3**: R = Me
vincristine, **4**: R = CHO
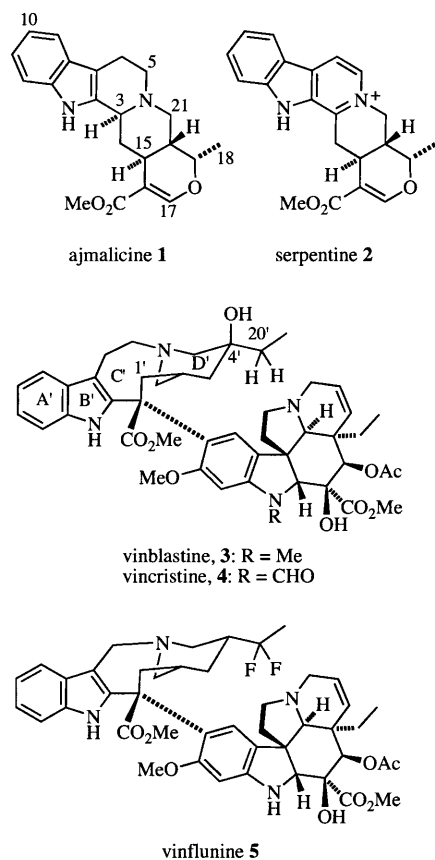
vinflunine **5**

Figure 3.1 Structures of selected medicinal MIAs produced by, or derived from, *C. roseus*. Ajmalicine **1** acts as an antihypertensive agent, serpentine **2** is a topoisomerase type-II inhibitor, vinblastine **3** and vincristine **4** are anticancer agents, and vinflunine **5** is a vinblastine analog with improved anti-cancer activity compared to **3** and **4**.
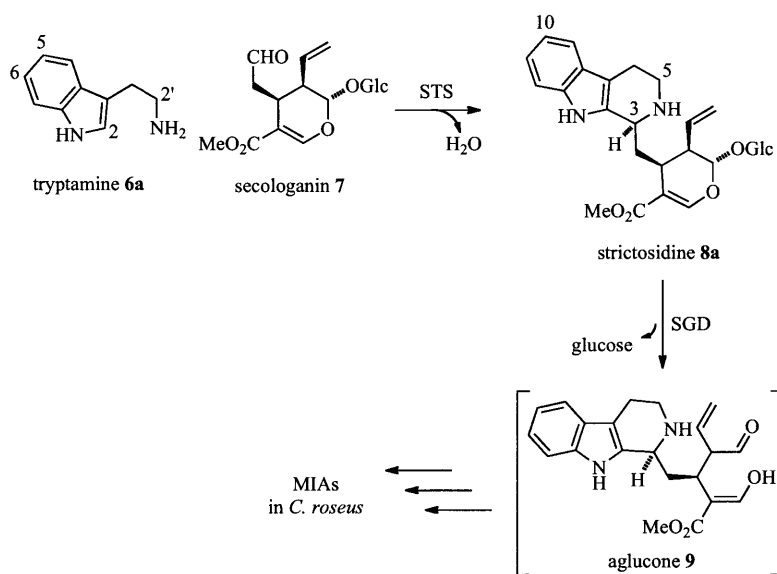
Figure 3.2 Strictosidine synthase (STS) catalyzes the Pictet-Spengler reaction between tryptamine **6a** and secologanin **7** to form strictosidine **8a**. Strictosidine glucosidase (SGD) catalyzes the deglucosylation of **8a** to form aglucone **9**, which is channeled into different pathway branches to form the monoterpene indole alkaloids (MIAs).

derivatives. Deglucosylation of strictosidine **8a** by the second enzyme of the pathway, strictosidine glucosidase (SGD), leads to aglucone **9** that is converted into the structurally diverse group of MIAs produced in *C. roseus* (Figure 3.2).[1,2]

Semi-rational design of enzyme catalysts combines the benefits of rational design and random mutagenesis and relies in part on structure and sequence information.[13] Semi-rational design is more efficient than random mutagenesis in improving properties such as substrate- and stereoselectivity.[14] Site-directed mutagenesis in the secologanin binding site of STS can alter the aldehyde specificity;[11] for example, Chen *et al.* reported an STS mutant (Asp177Ala) that selectively turns over a secologanin analog, which the wild-type enzyme already accepted.[11] The authors rationalized the change in specificity based on the x-ray crystal structure of *Rauvolfia serpentina* STS in complex with secologanin (PDB ID: 2FPC).[11,12] The crystal structure of *R. serpentina* STS in complex with tryptamine (PDB ID: 2FPB) is also available (Figure 3.3A),[12] though the structure does not fully explain the amine substrate specificity of STS. For example, it is not clear why STS accepts 5-hydroxytryptamine **6f** but not 5-methyltryptamine **6l** (Figure 3.3A).[9] To obtain a STS-catalyst with broader tryptamine substrate specificity, I used a semi-rational strategy, saturation mutagenesis,[15] to probe all twenty naturally occurring amino acids at four amino acid positions in the tryptamine binding site.[16] This chapter describes the identification and characterization of two STS mutants and the products of their new catalytic abilities.[16]

3.2 Results and discussion

3.2.1 Selection of substrates for screening

I selected amine substrates for screening based on the known substrate specificity of STS (Figure 3.3B).[9] *C. roseus* STS accepts tryptamine analogs with a fluorine substitution (**6b, 6e, 6g,** and **6i**) and analogs with a methyl substituent at the 4- or 7-position (**6c** and **6j**).[9] STS, however, imposes more restrictions to substitutions at the 5- and 6-positions of tryptamine, accepting 5-hydroxytryptamine **6f** and 6-methoxytryptamine **6h**, but not 5-methyltryptamine **6l**. Therefore, I selected 5-methyltryptamine **6l** to screen STS mutant libraries for increased tolerance of indole-5-substitutions.

STS also discriminates against substitutions at the 2'-position of the tryptamine side-chain; only a methyl substitution is accepted.[9] I assayed wild-type STS with several tryptamine analogs and found that STS does not turn over 2'*(R)*-trytophanol **6m**, 2'*(S)*-tryptophanol, 2'*(R)*-tryptophan, 2'*(S)*-tryptophan, and 2'*(R)*-tryptophan methyl ester. 2'*(R)*-tryptophanol **6m** has the smallest substituent size and I selected **6m** to screen STS libraries for increased acceptance of 2'-substitutions.

3.2.2 Selection of residues for saturation mutagenesis

The B-factor of tryptamine **6a** in the crystal structure of *R. serpentina* STS is approximately three-fold higher than the remainder of the protein, suggesting that the orientation of **6a** in the binding site is flexible.[12] With the exception of Glu315, which is crucial for catalysis of the Pictet-Spengler reaction,[12,17] the tryptamine binding site

consists primarily of hydrophobic and aromatic residues, including Trp153, Tyr155, Val171, Val182, Val214, and Phe232 (Figure 3.3A).[12] Hydrophobic active site residues tend to favor non-specific enzyme–substrate interactions, and therefore, the tryptamine binding pocket may foster flexible binding of tryptamine **6a** and then accommodate the structural changes that occur during iminium formation and cyclization to form the tetrahydro-β-carboline product.[17]

STS may prevent substitutions at the 5-position of tryptamine by the proximity of Val171 and Val214; the closest distances between these residues and the 5-position are 4.4 Å and 3.2 Å, respectively. Similarly, STS may prevent large substitutions at the 2'-position of tryptamine, since the 2'pro*(S)* hydrogen atom is 3.3 Å from the carboxylate oxygen of Glu315 and the 2'pro*(R)* hydrogen atom is 3.8 Å from the phenyl ring of Phe232. I hypothesized that saturation mutagenesis of residues Val171 or Val214 allows turnover of 5-methyltryptamine **6l** and that mutagenesis of Phe232 or Glu315 would allow turnover of 2'*(R)*-tryptophanol **6m** (Figure 3.3.A).

3.2.3 Screening of saturation mutagenesis libraries

I constructed saturation mutagenesis libraries of STS, randomized at Val171, Val214, Phe232, and Glu315 using methods described in Chapter 2. After protein expression, the media containing the enzyme was transferred to microtiter plates for screening using the pigment assay solution (section 2.4.7). The Val171Xxx and Glu315Xxx STS libraries did not produce positive hits when screened with 5-methyltryptamine **6l** and 2'*(R)*-

A

B

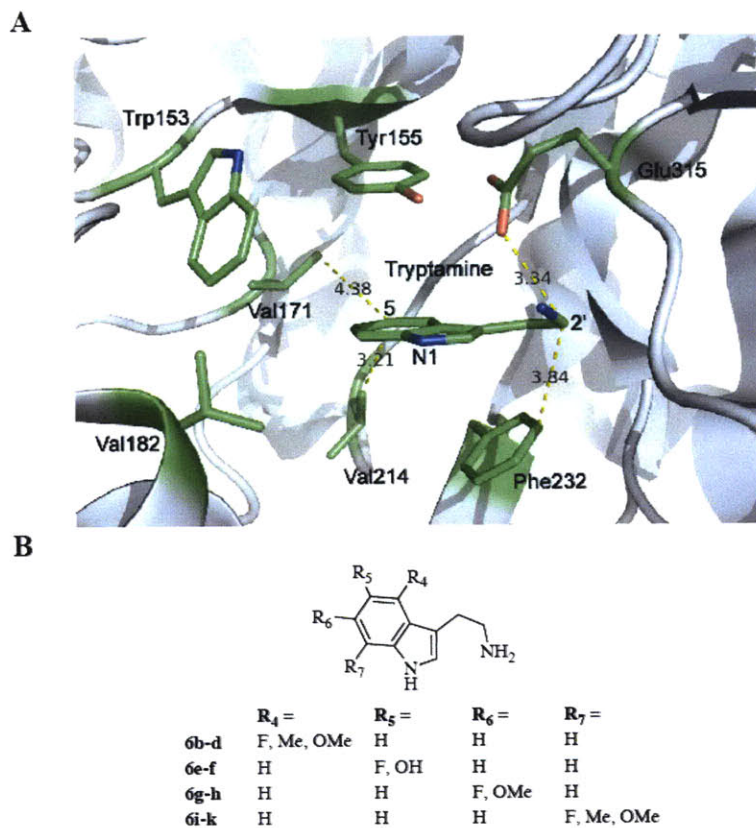|	**R$_4$ =**	**R$_5$ =**	**R$_6$ =**	**R$_7$ =**
6b-d	F, Me, OMe	H	H	H
6e-f	H	F, OH	H	H
6g-h	H	H	F, OMe	H
6i-k	H	H	H	F, Me, OMe

Figure 3.3 (A) Tryptamine binding site of *R. serpentina* STS (PDB ID: 2FPB; residue numbering according to the *C. roseus* amino acid sequence). Side-chains interacting with **6a** are shown in sticks and the main-chain is represented in cartoon form. (B) Analogs of **6a** with indole ring substitutions that wild-type STS accepts.
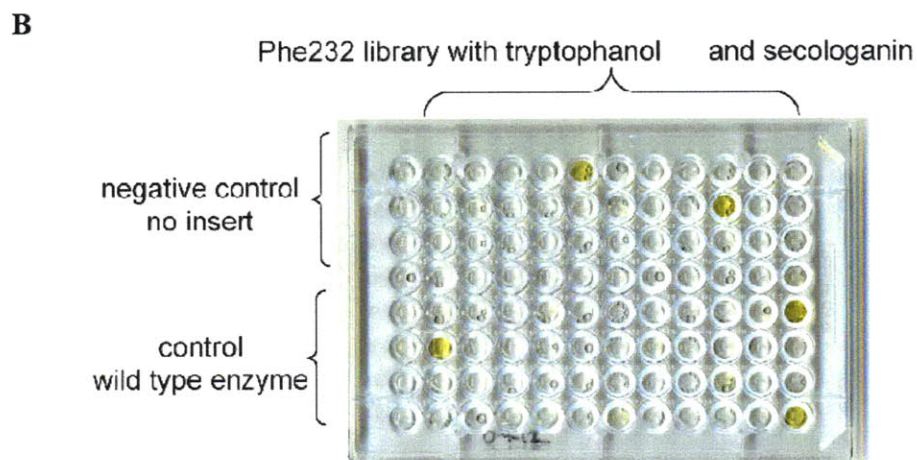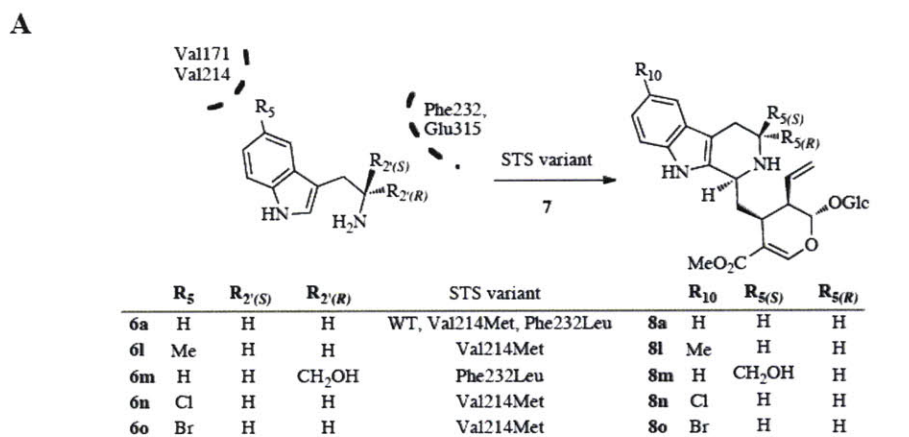
A

	R$_5$	R$_{2'(S)}$	R$_{2'(R)}$	STS variant		R$_{10}$	R$_{5(S)}$	R$_{5(R)}$
6a	H	H	H	WT, Val214Met, Phe232Leu	**8a**	H	H	H
6l	Me	H	H	Val214Met	**8l**	Me	H	H
6m	H	H	CH$_2$OH	Phe232Leu	**8m**	H	CH$_2$OH	H
6n	Cl	H	H	Val214Met	**8n**	Cl	H	H
6o	Br	H	H	Val214Met	**8o**	Br	H	H

B

Phe232 library with tryptophanol   and secologanin

Figure 3.4 (A) Formation of strictosidine analogs from tryptamine analogs with the STS mutants identified by screening of saturation mutagenesis libraries. (B) Screening of Phe232Xxx library using the pigment assay described in Chapter 2. Each well contains supernatant from one STS mutant expression culture, 2'(R)-tryptophanol **6m**, secologanin **7**, and SGD. Negative and positive controls are in the first column (rows 1-4: no-insert vector, rows 5-8: wild-type STS).

tryptophanol **6m**. Nevertheless, several members of the Val214Xxx STS library showed activity with **6l**. Sequencing of three active mutants revealed a Val214Met mutation; the Met substitution removes the β-branching of Val, which could hinder reaction between **6l** and **7** in wild-type STS. Since no mutants containing smaller amino acid residues at position 214 were isolated from the screen, I speculate that a longer side-chain is key for orienting the tryptamine analog optimally for catalysis. Stöckigt and co-workers independently identified that the Val208Ala mutant of *R. serpentina* STS (corresponding to Val214Ala in *C. roseus* STS) accepts **6l** and 5-methoxytryptamine.[20] The Val214Ala mutant of *C. roseus* STS does not endow STS with a broadened substrate scope. Further evaluation of the Val214Met STS mutant showed that it also accepts 5-chlorotryptamine **6n** and 5-bromotryptamine **6o**. Halogenated metabolic precursors are useful because aryl chlorides and bromides can be derivatized by cross-coupling reactions;[18] also, there is evidence that bromination of synthetic vinblastine analogs at this position increases anti-tumor activity.[19]

The pigment assay indicated that members of the Phe232 library accept **6m** (Figure 3.4B). Sequencing of six active clones showed a substitution at residue 232 to leucine (four of the six sequenced genes) or methionine (two of the six sequenced genes). The 2'pro*(R)* hydrogen atom projects toward Phe232 (closest distance: 3.4 Å) and mutagenesis of Phe232 to the smaller leucine or methionine residue may help accommodate the hydroxymethylene substituent. I selected the Phe232Leu mutant of STS for further characterization, and found that it is not active toward tryptamine analogs

with a 2'pro*(S)* hydrogen substitution, likely since this substituent points toward the catalytic Glu315 residue. Turnover of L-tryptophan by STS could impact the cellular supply of tryptophan and the enzyme may have evolved to discriminate against tryptophan in the plant cell.

3.2.4 Kinetic analysis of Val214Met and Phe232Leu STS

I expressed the Val214Met and Phe232Leu STS variants in *S. cerevisiae* BJ5465, concentrated the supernatant, and purified the FLAG-tagged protein by anti-FLAG affinity chromatography (section 3.4.3). Pseudo-steady-state kinetics of Val214Met and Phe232Leu *C. roseus* STS, shows that both mutations significantly impact the activity of STS (Table 3.1). Both decreased V_{max}-values (5-100-fold) and increased K_M-values (200-1000-fold) contribute to the decreased catalytic efficiencies (7,000-150,000-fold) of the mutant enzymes with tryptamine (Table 3.1).

Table 3.1. Kinetic constants for STS.[a]

Variant, substrate	V_{max} $[U\ mg^{-1}]$ [b]	K_M $[\mu M]$	V_{max}/K_M $[U\ mg^{-1}\ M^{-1}]$ [b]
Wild-type, **6a**	1.6 ± 0.1	1.5 ± 0.4	1,000,000 ± 600,000
Val214Met, **6a**	0.32 ± 0.03	600 ± 200	500 ± 100
Val214Met, **6l** [c]	0.0067 ± 0.0003	800 ± 100	8 ± 2
Phe232Leu, **6a**	0.015 ± 0.001	900 ± 200	17 ± 6
Phe232Leu, **6m** [c]	0.0057 ± 0.0003	1800 ± 300	3.2 ± 0.8

[a] Kinetic assays were performed in sodium phosphate buffer (0.1 M, pH 7.0) using an internal standard (NAA, 0.075 mM) and analyzed by HPLC for formation of product; [b] 1 U = formation of 1 μmol product (**6a** or analog of **6a**) per minute at pH 7.0 and 30 °C; [c] Compounds **6l** and **6m** are not accepted by wild-type STS.

Phe232 may help orient the substrate correctly for catalysis and stabilize high-energy intermediates in the reaction pathway; therefore, a leucine in its position can dramatically

reduce the catalytic activity of STS. The 60- and 5-fold lower V_{max}/K_M-values for Val214Met and Phe232Leu, respectively, show that both mutants prefer tryptamine **6a** to the unnatural analogs (Table 3.1).

3.2.5 Preparative enzymatic synthesis of strictosidine analogs **8l-n**

To show that the enzyme mutants form true tetrahydro-β-carboline products, I used Val214Met and Phe232Leu STS in the mmol-scale synthesis of 10-methylstrictosidine **8l**, 5(R)-hydroxymethyl strictosidine **8m**, and 10-chlorostrictosidine **8n**. I cloned the *sts* gene from *C. roseus* into the *Escherichia coli* expression vector pET-28 and introduced both mutations into this bacterial construct using site-directed mutagenesis (section 3.4.5). Expression of the STS mutants in *E. coli*, followed by biocatalytic transformations in cleared *E. coli* cell lysates, afforded milligram quantities of each strictosidine analog after preparative HPLC purification. ^{1}H NMR confirmed the structure of the enzymatic products (section 3.4.5).

3.2.6 Feeding strictosidine analogs **8m** and **8o** to *C. roseus* hairy-root cultures

C. roseus plant cell cultures can take up the relatively large strictosidine biosynthetic intermediate,[21] and the MIA biosynthetic pathways can convert tryptamine and secologanin analogs to higher alkaloid derivatives.[8,11] Prior to constructing transgenic plants with Val214Met and Phe232Leu STS, however, it is important to validate whether subsequent MIA biosynthetic enzymes process the new strictosidine analogs. Therefore, I

studied which MIA analogs formed after feeding 5(R)-hydroxymethyl strictosidine **8m** and 10-bromomstrictosidine **8o** to two-week-old *C. roseus* hairy-root cultures.

I used established protocols to extract the hairy-root cultures one week after feeding,[22,23] and used the available information on naturally occurring MIA metabolites to analyze the culture extracts by LC-MS.[2,17] The extract of the control culture, supplemented with tryptamine **6a**, contained alkaloids such as ajmalicine **1** (*m/z* 353), serpentine **2** (*m/z* 349), yohimbine **10** (*m/z* 355), and akuammacine **11** (*m/z* 323) (Figures 1.5, 1.6, 1.8, and 3.1).[8] The culture extracts of the **8m**- and **8o**-supplemented hairy-root cultures, however, contained compounds with masses corresponding to MIA analogs (expected masses of analogs: **8m**: M+30; **8o**, M+78/80) not present in the control culture (Appendix A).

MIA analogs derived from halogenated strictosidine precursors have characteristic isotope patterns, *e.g.* the equal abundance of ^{79}Br and ^{81}Br, and based on this information, I found that *C. roseus* converted **8o** into analogs of known MIA: ajmalicine **1**, serpentine **2**, yohimbine **10**, and akuammacine **11** (Figure 3.5A and Appendix A).[16] The recent construction of a transgenic *C. roseus* hairy-root line that expresses Val214Met STS, validated that the Val214Met enzyme mutant functions *in vivo* to produce methylated, chlorinated, and brominated MIA analogs that the plant cell normally cannot produce.[24-26]
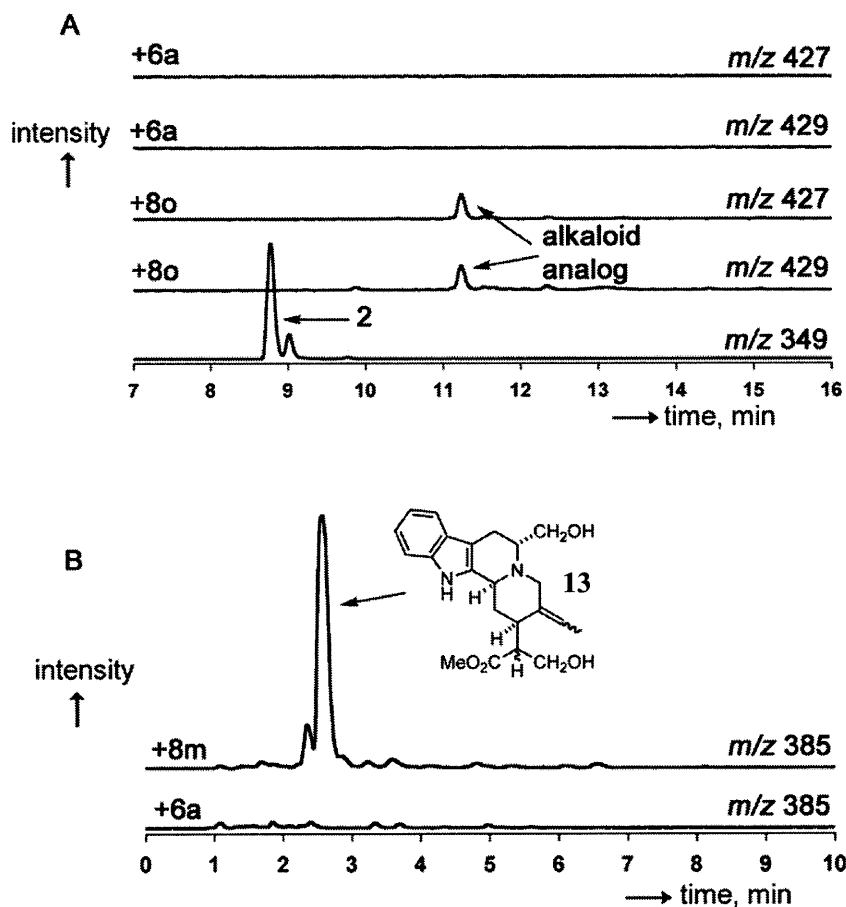
Figure 3.5 LC-traces of selected alkaloid analogs formed by feeding strictosidine analogs to two-week-old *C. roseus* hairy root cultures. (A) Formation of a putative serpentine or serpentine isomer analog upon feeding of 10-bromostrictosidine **8o** to periwinkle. The two top traces show that no compounds corresponding to *m/z* 427 and 429 exist in the culture supplemented with tryptamine **6a**. The bottom trace shows the elution of an authentic standard of serpentine **2**. The remaining two traces show the formation of an alkaloid analog displaying the expected isotopic pattern that could be a brominated MIA analog. (B) Formation of isositsirikine analog **13** in a feeding study of 5*(R)*-hydroxymethyl strictosidine **8m**. The top trace shows the expected mass of **13**; the bottom trace shows that no naturally occurring alkaloid of *m/z* 385 co-elutes with **13**.

LC-MS analysis of hairy-root cultures supplemented with 5*(R)*-hydroxymethyl strictosidine **8m** showed three new peaks, not found in the control culture supplemented with tryptamine **6a**, containing compounds with *m/z* 383 (**12**), 385 (**13**), and 326 (**14**). Compound **12** may be an ajmalicine analog (ajmalicine: *m/z* 353), but the low production titer of **12** precluded isolation and structural characterization. I isolated compounds **13** (2 mg, 3% isolated yield) and **14** (1 mg, 1%) from 0.5 L hairy-root cultures by preparative HPLC. Compound **14** (*m/z* 426) consisted of several stereoisomers; ^{1}H-NMR signals for the indole ring strongly suggest the assignment as an MIA derivative (section 3.4.6). Compound **13** is an isositsirikine analog with a hydroxy-methylene substitutent (Figure 3.5B),[27] since it contained an unsubstituted indole (7.5-7.1 ppm, 4H), an intact hydroxymethylene group (4.0 and 3.8 ppm, 2H; compared to 4.0 and 3.8 ppm for **8m**), and a methyl ester (s, 3.8 ppm, 3H), as well as an ethylidene group containing a trisubstituted alkene, since one olefinic hydrogen atom appears as a quartet (6.0 ppm, *J* = 6.8) coupled to three hydrogen atoms that appear as a doublet (1.9 ppm, *J* = 6.8).[27]

The hairy-root culture supplemented with **8m** does not form a serpentine analog, suggesting that the oxidation reaction of ajmalicine **1** to form serpentine **2** requires a 5'*pro(R)*-hydrogen atom (Figures 1.7 and 3.4A). Furthermore, the culture does not form analogs of akuammicine **11**, suggesting that the unknown enzyme(s) forming **11** does not accept **8m** as a precursor. McCoy and O'Connor observed that MIA biosynthetic pathways in hairy-roots and seedlings often channel supplemented precursor analogs – particularly those with large substituents that cannot be converted into "normal" MIAs

such as serpentine **2** and akuammacine **11** – into the reduction steps that form isositsirikine.[28] I observed that MIA derivatives **13** and **14** are present in both the culture media and the cells, while natural alkaloids are not found in the culture media; I speculate that the plant cell cultures have a mechanism to export accumulated or unwanted compounds into the culture media.

3.3 Conclusions

This chapter describes the first example of enzyme reengineering in plant alkaloid biosynthesis. I expanded the substrate specificity of *C. roseus* STS – a critical first step toward pathway engineering – using saturation mutagenesis. Two STS mutants (Val214Met and Phe232Leu), identified using the expression system and pigment assay described in Chapter 2, catalyze the formation of strictosidine analogs that are not catalyzed by wild-type STS.

The Val214Met mutant of STS catalyzes the formation of methylated and halogenated strictosidine derivatives; *C. roseus* hairy-root cultures supplemented with 10-bromostrictosdine **8o** incorporate the bromine atom into higher MIAs. The Phe232Leu mutant of STS mutant triggered the formation of a hydroxylated strictosidine analog. Feeding studies of 5*(R)*-hydroxymethyl strictosidine **8m** show that the addition of a hydroxyl group causes major perturbations of the alkaloid profile of *C. roseus*.

Runguphan *et al.* recently demonstrated that the Val214Met mutant of STS functions in a cellular context, producing natural product analogs that the plant cannot normally produce.[25] Runguphan *et al.* also developed a transgenic *C. roseus* cell culture without alkaloid production (obtained by RNAi-silencing of tryptophan decarboxylase transcripts); MIA biosynthesis commences when tryptamine substrates are supplemented to this transgenic cell culture.[26] Therefore, the silenced strain, in conjunction with reengineered STS variants, will likely lead to the efficient transgenic production of MIA derivatives.

3.4 Experimental methods

3.4.1 General methods and analytical techniques

Tryptamine **6a** and analogs of **6a** were purchased in highest available purity and used without further purification. Secologanin **7** was obtained by MeOH-extraction of *Lonicera tatarica* as described previously.[29]

LC-MS analysis was performed on a Waters LC-MS system. Compounds were separated by an Acquity UPLC BEH C18-column (1.7 μm, 2.1 x 100 mm) equipped in tandem with MS technologies Micromass LCT Premier with an electrospray ionization (ESI) source and time-of-flight (TOF) detector. The chromatography solvent system consisted of MeCN in water with 0.1% formic acid. The capillary and sample cone voltages were 2000 V and 30 V, respectively. The desolvation and source temperature were 350 °C and 100 °C, respectively. The cone and desolvation gas flow-rates were 20 L h^{-1} and 700 L/h,

respectively. HRMS were obtained in accurate mass mode (W-mode) with reserpine (m/z 609.2812, as 1 nM solution, 40 μL min^{-1}) as reference. The HPLC-based assay was carried out on a Beckman-Coulter System Gold system equipped with a 125 solvent module, 125 autosampler and a 168 detector. Data was processed (integrated) using 32Karat (v 7). A solvent system composed of MeCN in water with 0.1% TFA was used with a reverse-phase column (HiBar RT 250-4, RP-select B, 5 μm). NMR spectra were recorded on Varian 500 MHz or Bruker 600 MHz spectrometers.

3.4.2 Construction, expression, and screening of saturation mutagenesis libraries

Saturation mutagenesis libraries were constructed according to the procedure described in section 2.4.6, using the primers listed in Table 3.2. Expression and assays of saturation mutagenesis libraries were conducted in 96-well plates as described in section 2.4.6.

Table 3.2. PCR primers for saturation mutagenesis.

Primer name	Primer sequence (5'->3')
pGALMF_fwd	GGGGTATCTTTGGATAAAAGAGAGGCTGAAGCTGAATTCGATATC
pGALMF_rev	GGGAGGGCGTGAATGTAAGCGTGACATAACTAATTACATGAC
Val171Xxx_fwd	CTTCACTGACNNKTCCTCTATTC
Val171Xxx_rev	GAATAGAGGAMNNGTCAGTGAAG
Val214Xxx_fwd	GAAGGAGTTGCACNNKCCAGGTGGTGCTG
Val214Xxx_rev	CAGCACCACCTGGMNNGTGCAACTCCTTC
Phe232Xxx_fwd	GTTGTTGCTGAGNNKTTGTCCAACAGAATC
Phe232Xxx_rev	GATTCTGTTGGACAAMNNCTCAGCAACAAC
Glu315Xxx_fwd	GAAGGTGAACACTTCNNKCAAATTCAAGAGC
Glu315Xxx_rev	GCTCTTGAATTTGMNNGAAGTGTTCACCTTC

3.4.3 Expression and purification of FLAG-tagged STS mutants

To ensure that no other mutations were responsible for the activity observed in the screening, the Val214Met- and Phe232Leu-encoding mutations were introduced into

103

pSTSMF-FLAG using Stratagene's site-directed mutagenesis kit (Table 3.3). A mutation

of the *sts* gene that results in Val214Ala STS was also introduced to test if this mutant

enzyme accepts new substrates (Table 3.3).

Table 3.3. PCR primers for site-directed mutagenesis.

Primer name	Primer sequence (5'->3')
Val214Ala_fwd	G A A G G A G T T G C A C G C C C C A G G T G G T G C T G
Val214Ala_rev	C A G C A C C A C C T G G G G C G T G C A A C T C C T T C
Val214Met_fwd	G A A G G A G T T G C A C A T G C C A G G T G G T G C T G
Val214Met_rev	C A G C A C C A C C T G G C A T G T G C A A C T C C T T C
Phe232Leu_fwd	G T T G T T G T T G C T G A G C T G T T G T C C A A C A G A A T C
Phe232Leu_rev	G A T T C T G T T G G A C A A C A G C T C A G C A A C A A C A A C

S. cerevisiae BJ5465 harboring pSTSMF-FLAG with the indicated mutations were grown

overnight in SCMM-U supplemented with 2% (w/v) glucose. The cells were recovered

from an overnight culture (OD_{600} ~ 10) and washed in SCMM-U without any carbon

source to remove the glucose. The overnight culture was then diluted with SCMM-U

containing 2% (w/v) raffinose and 2% (w/v) galactose to yield a culture with an OD_{600} =

1. The culture was incubated at 25 °C for 72 h shaking at 225 rpm. The cells were then

removed by centrifugation (5 min, 3,000g, 4 °C) and the supernatant was concentrated

and buffer exchanged to afford enzyme preparations in sodium phosphate buffer (0.05 M,

pH 7.0) The concentrated STS mutant preparations were then purified using anti-FLAG

M2 agarose beads (Sigma Aldrich, St. Louis, MO) according to the manufacturer's

instructions (the column-bound protein was eluted with FLAG-peptide). The eluted

enzyme was buffer exchanged to sodium phosphate buffer (0.05 M, pH 7.0) and stored in
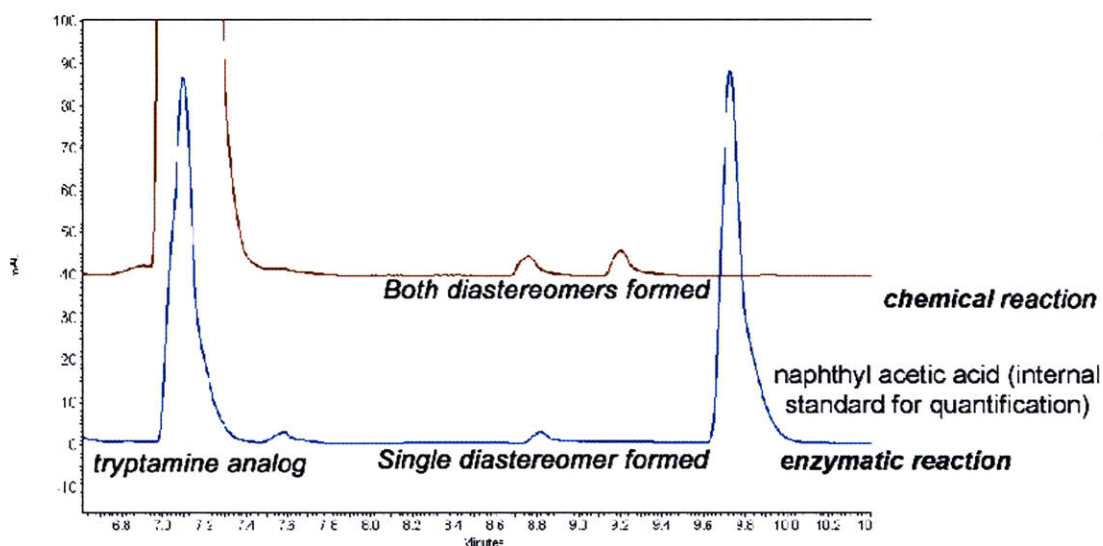
20% glycerol at -20 °C.

104

Figure 3.6 HPLC trace (280 nm) of a representative spectrum used in analysis of 5-methyltryptamine **61** kinetics using the Val214Met mutant of STS. The top trace (brown) represents the chemical reaction control, which shows formation of two peaks corresponding to the two diastereomers, 10-methylstrictosidine **81** and 10-methylvincoside, respectively. The blue trace represents the enzymatic reaction with Val214Met STS and 5-methyltryptamine **61** in which only one diastereomer is formed. NAA is added as an internal standard. Retention times: **61** (r_t = 7.1 min), secologanin **7** (r_t = 7.6 min, low absorbance at 280 nm), 3*(S)*-epimer of **81** (r_t = 8.8 min), 3*(R)*-epimer of **81** (r_t = 9.2 min), and NAA (r_t = 9.8 min).

105

3.4.4 HPLC-kinetics of Val214Met and Phe232Leu STS

All enzyme assays (0.100 mL) contained secologanin **7** (4.0 mM) in sodium phosphate buffer (0.1 M, pH 7.0) with internal standard (0.075 mM naphthylacetic acid, NAA) and one of several tryptamine or tryptamine analog concentrations (added from a 10-fold concentrated stock solution). Product formation was monitored by reverse-phase HPLC at 280 nm (section 3.4.1); the detection-limit precluded the use of substrate concentrations <1 μM. Wild-type STS reactions were quenched by two equivalents of sodium hydroxide (relative to the buffer concentration); a quench that resulted in the quantitative formation of strictosamide [ESI-MS($C_{26}H_{31}N_2O_8^+$) *m/z* calculated: 499.2091 [M+H]$^+$, found: 499.2080 [M+H]$^+$].[30] The chromatogram peaks were integrated, normalized by the internal standard (NAA), and converted to concentrations using a response factor for strictosamide. Corrected experimental data were fit to the Michaelis-Menten or Northrop equations by non-linear regression analysis using Origin (v 7.0552) (OriginLab Corp., Northampton, MA) to directly get catalytic parameters V_{max}, V_{max}/K_M, and K_M. All kinetic assays were repeated three times and only conversions below 15% were considered. A representative HPLC trace using the Val214Met mutant and 5-methyltryptamine **6l** is shown in Figure 3.6.

## 3.4.5 Heterologous expression of *C. roseus* STS in *E. coli* and enzymatic syntheses

The following PCR primers were used to amplify the *sts* gene in pSTSMF-FLAG (section 2.4.2) and introduce sites for *Nco* I and *Xho* I (underlined):

5'-CAT G<u>CC ATG GG</u>C TCA CCA ATC TTG AAG AAG ATC-3'

5'-CCG <u>CTC GAG </u>GGA GGA AAC GTA GGA GTT TCC C-3'

The PCR product was cloned into pET-28a(+) in-frame with nucleotides in the vector that encode for a hexa-histidine tag, resulting in construct pSTS-His. Mutations Val214Met and Phe232Leu were introduced by Stratagene's site-directed mutagenesis kit using pSTS-His as a template (Table 3.3). pSTS-His was transformed into *E. coli* BL21(DE3) for protein expression. A single colony of plasmid-containing bacteria encoding either mutant enzyme was inoculated in LB-media (30 μg mL^{-1} kanamycin) and incubated overnight at 37 °C and 200 rpm. The seed culture was then inoculated (1% v/v) in fresh LB-media (4 L in 2.8-L Fernbach flasks, 30 μg mL^{-1} kanamycin) and incubated until OD$_{600 \, nm}$ = 1 (approximately 2.5 h at 37 °C and 200 rpm). The culture was chilled to 4 °C, IPTG was added (1 mM final concentration) to induce protein expression, and the culture was incubated at 18 °C for 24 h at 200 rpm. The cell material was recovered by centrifugation (15 min, 3000g, 4 °C) to an approximate wet-weight of 7 g L^{-1} culture. The cells were resuspended in five volumes sodium phosphate buffer (0.1 M, pH 7.0) and lysed by passing the suspension twice through a French (hydraulic) press. The DNA was sheared using a 22-gauge syringe needle and the crude cell extract was recovered by centrifugation (120 min, 10000g, 4 °C). The cell extract was used directly for synthesis as described below.

10-methylstrictosidine 8l
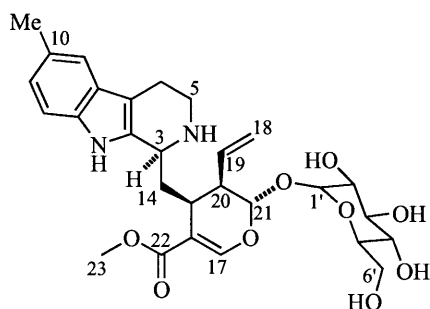
5-methyltryptamine **6l** hydrochloride (105 mg, 0.5 mmol, 1 eq) and secologanin **7** (194 mg, 0.5 mmol, 1 eq) were dissolved in 225 mL sodium phosphate buffer (25 mM, pH 7.0). The cell-free extract containing Val214Met STS, prepared as described above, was then added (25 mL, 10% v/v) and the reaction was stirred for 72 h at 30 °C. MeOH was added (4 eq) to precipitate the protein, the resulting heterogenous mixture was filtered to remove the particulates, and most of the solvent was evaporated *in vacuo*. Preparative HPLC using a reverse-phase column, 1 mL injections, and a gradient of 20-90% MeCN in water with 0.1% TFA over 20 min afforded 10-methylstrictosidine **8l** (*d.e.* >99.5%). ^{1}H NMR (MeOD, 500 MHz): δ 7.81 (s, 1H, H-17), 7.26 (s, 1H, H-9), 7.19 (d, 1H, *J* = 8.3, H-12), 6.98 (dd, 1H, *J* = 1.3, 8.4, H-11), 5.85 (ddd, 1H, *J* = 7.6, 10.6, 17.5, H-19), 5.84 (d, 1H, *J* = 9.1, H-21), 5.35 (dd, 1H, *J* = 1.3, 17.4, H-18*trans*), 5.28 (dd, 1H, *J* = 1.1, 10.6, H-18*cis*), 4.80 (d, 1H, *J* = 7.9, GlcH-1'), 4.64 (br-d, 1H, *J* = 10.6, H-3), 3.98 (dd, 1H, *J* = 2.1, 11.8, GlcH-6'a), 3.80 (s, 3H, H-23), 3.73 (ddt, 1H, *J* = 5.0, 12.5, GlcH-6'b), 3.49-3.42 (m, 1H, H-5a), 3.40 (t$_{app}$, 1H, *J* = 9.1, GlcH-3'), 3.39-3.35 (m, 1H, H-5b), 3.25-3.19 (m, 2H, GlcH-2', GlcH-4'), 3.10-2.97 (m, 3H, H-6a, H-6b, H-15), 2.75 (ddd, 1H, *J* = 5.1, 7.8, 8.7, H-20), 2.40 (s, 3H, H-23), 3.36-2.16 (m, 2H, H-14pro*S*, H14pro*R*); ^{13}C

NMR (MeOD, 125 MHz): δ 171.43, 157.05, 136.67, 135.62, 130.30, 130.03, 127.82, 125.24, 119.90, 118.91, 112.10, 109.08, 106.79, 100.56, 97.43, 78.96, 78.135, 74.80, 71.88, 63.15, 53.19, 52.74, 45.52, 42.80, 34.92, 32.75, 21.71, 19.71; ESI-MS($C_{28}H_{37}N_2O_9^+$) *m/z* calculated: 545.2499 [M+H]$^+$, found: 545.2523 [M+H]$^+$.

5(R)-hydroxymethyl strictosidine **8m**
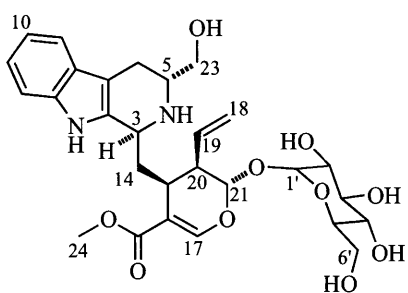
2'(R)-tryptophanol **6m** (190 mg, 1 mmol, 1 eq) was added as a solution in MeCN (25 mL, 5% v/v final concentration of organic solvent) and secologanin **7** was added as a powder (388 mg, 1 mmol, 1 eq) to 450 mL stirred sodium phosphate buffer (25 mM, pH 7.0). Cell-extract containing Phe232Leu STS, prepared as described above, was then added (25 mL, 5% v/v) and the reaction was stirred for 46 h at 30 °C. Some of the water was co-evaporated *in vacuo* with toluene, after which MeOH was added (4 eq) to precipitate the protein. After the resulting heterogenous mixture was filtered to remove the particulates, most of the solvent was evaporated *in vacuo*. Preparative HPLC using a reverse-phase column, 1 mL injections, and a gradient of 20-60% MeCN in water with 0.1% TFA over 20 min afforded *5(R)*-hydroxymethyl strictosidine **8m** (190 mg, 0.34 mmol, 43%, *d.e.* >99.5%). ^{1}H NMR (MeOD, 500 MHz): δ 7.80 (s, 1H, H-17), 7.46 (td, 1H, *J* = 1.0, 7.8, H-9), 7.33 (td, 1H, *J* = 0.8, 8.2, H-12), 7.15 (ddd, 1H, *J* = 1.1, 7.1, 8.2,

H-11), 7.05 (ddd, 1H, J = 0.9, 7.1, 8.2, H-10), 5.84 (ddd, 1H, J = 7.6, 10.6, 17.4, H-19),

5.80 (d, 1H, J = 8.8, H-21), 5.37 (ddd, 1H, J = 1.0, 1.3, 17.4, H-18*trans*), 5.28 (ddd, 1H, J

= 1.0, 1.3, 10.6, H-18*cis*), 4.78 (d, 1H, J = 7.9, GlcH-1'), 4.67 (dd, 1H, J = 4.0, 10.9, H-

3), 4.03 (dd, 1H, J = 3.4, 10.7, H-23), 3.94 (dd, 1H, J = 2.1, 11.8, GlcH-6'a), 3.84-3.75

(m, 2H, H-5a, H-23), 3.78 (s, 3H, H-24), 3.62 (dd, 1H, J 6.9, 11.9, GlcH-6'b), 3.40 (t_{app},

1H, J = 9.1, GlcH-3'), 3.36-3.32 (m, 1H, GlcH-5'), 3.25-3.19 (m, 2H, GlcH-2',4'), 3.12

(dd, 1H, J = 4.8, 16.3, H-6a), 3.04 (td, 1H, J = 4.6, 11.2, H-15), 2.89 (dd, 1H, J = 8.9,

16.4, H-6b), 2.77 (ddd, 1H, J = 5.0, 7.8, 8.6, H-20), 2.32 (ddd, 1H, J = 4.4, 11.1, 15.2, H-

14pro*R*), 2.24 (ddd, 1H, J = 4.2, 11.4, 15.2, H-14pro*S*); ^{13}C NMR (MeOD, 125 MHz): δ

171.28, 156.93, 138.45, 135.51, 130.46, 127.50, 123.79, 120.76, 120.08, 119.24, 112.48,

108.84, 106.25, 100.56, 97.54, 78.82, 78.08, 74.78, 71.82, 63.02, 62.03, 54.10, 52.76,

51.16, 45.34, 35.38, 32.63, 21.89; ESI-MS($C_{28}H_{37}N_2O_{10}{}^+$) *m/z* calculated: 561.2448

[M+H]$^+$, found: 561.2448 [M+H]$^+$.

10-chlorostrictosidine 8n
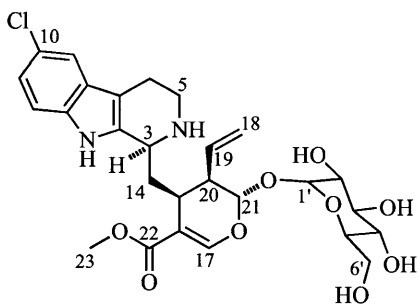
5-chlorotryptamine **6n** hydrochloride salt (230 mg, 1 mmol, 1 eq) and secologanin **7** (388

mg, 1 mmol, 1 eq) were dissolved in 450 mL stirred sodium phosphate buffer (25 mM,

pH 7.0). The cell-free extract containing Val214Met STS extract, prepared as described above, was then added (25 mL, 5% v/v) and the reaction was stirred for 72 h at 30 °C. Some of the water was co-evaporated *in vacuo* with toluene, after which MeOH was added (4 eq) to precipitate the protein. After the resulting heterogenous mixture was filtered to remove the particulates, most of the solvent was evaporated *in vacuo*. Preparative HPLC using a reverse-phase column, 1 mL injections, and a gradient of 20-90% MeCN in water with 0.1% TFA over 15 min afforded 10-chlorostrictosidine **8n** (*d.e.* >99.5%). ^{1}H NMR (MeOD, 500 MHz): δ 7.81 (s, 1H), 7.47 (dd, 1H, J = 0.6, 2.1, H-9), 7.29 (dd, 1H, J = 0.6, 8.7, H-12), 7.11 (dd, 1H, J = 2.0, 8.7, H-11), 5.85 (ddd, 1H, J = 7.7, 10.7, 17.4, H-19), 5.84 (d, 1H, J = 9.1, H-21), 5.35 (dt$_{app}$, 1H, J = 1.4, 17.4, H-18*trans*), 5.28 (dt$_{app}$, 1H, J = 1.3, 10.7, H-18*cis*), 4.80 (d, 1H, J = 7.9, H-1'), 4.68 (br-d, 1H, J = 10.4, H-3), 3.98 (dd, 1H, J = 2.0, 11.8, H-6'a), 3.80 (s, 3H, H-23), 3.74 (td, 1H, J = 5.2, 12.5, H-5eq), 3.64 (dd, 1H, J = 7.1, 11.8, H-6'b), 3.47 (ddd, 1H, J = 5.6, 8.5, 12.5, H-5ax), 3.40 (t$_{app}$, 1H, J = 9.1, H-3'), 3.39-3.34 (m, 1H, H-5'), 3.23 (dd, 1H, J = 8.9, 9.9, H-4'), 3.22 (dd, 1H, J = 7.9, 9.3, H-2'), 3.13-3.05 (m, 2H, H-6ax, H-15), 3.02 (td$_{app}$, 1H, J = 4.5, 16.3, H-6eq), 2.75 (dddd, 1H, J = 0.5, 5.0, 7.8, 8.8, H-20), 2.32 (ddd, 1H, J = 3.0, 12.1, 15.0, H-14pro*R*), 2.23 (ddd, 1H, J = 3.7, 11.8, 15.2, H-14pro*S*); ^{13}C NMR (MeOD, 125 MHz): δ 171.40, 157.06, 136.70, 135.54, 132.19, 128.65, 126.51, 123.80, 119.97, 118.84, 113.67, 109.03, 107.19, 100.54, 97.41, 78.95, 78.12, 74.79, 71.87, 63.15, 53.05, 52.74, 45.50, 42.63, 34.82, 32.68, 19.52; ESI-MS($C_{27}H_{34}N_2O_9{}^{35}Cl^+$) *m/z* calculated: 565.1953 [M+H]$^+$, found: 565.1959 [M+H]$^+$; ESI-MS($C_{27}H_{34}N_2O_9{}^{37}Cl^+$) *m/z* calculated: 567.1923 [M+H]$^+$, found: 567.1945 [M+H]$^+$.

3.4.6 Feeding study and alkaloid isolation

5*(R)*-hydroxymethyl strictosidine **8m** was synthesized enzymatically as described (see section 3.4.5). *C. roseus* hairy-root cultures were inoculated in half-strength Gamborg media (0.4 L), grown for two weeks (32 °C, 60 rpm), and then supplemented with **8m** (0.5 mM, 110 mg as filter-sterilized solution in water). After one week, the media was separated from the hairy roots and the roots were washed in water several times. Hairy roots were ground with a mortar and pestle in MeOH (3 x 30 mL) and after filtration, the solvent was evaporated *in vacuo* to give a yellow residue. The residue was dissolved in aqueous hydrochloric acid (10 mM, pH 2) and the acidified natural product mixture was extracted with hexanes. The aqueous layer was basified (pH 9) with ammonium hydroxide and then extracted with DCM. The organic solvent was evaporated to give a yellow oil, which was dissolved in water and analyzed by LC-MS. Compounds **13** and **14** were isolated by preparative HPLC using the following gradients: (1) 20-40% MeCN in water with 0.1% TFA for 15 minutes; then, (2) 30-40% MeCN in water with 0.1% TFA for 10 minutes. This isolated isositsirikine analog **13** in 2.0 mg isolated yield (5 μmol, 3%): ^{1}H NMR (600 MHz, MeOD): δ 7.51 (d, 1H, *J* = 7.7, H-9), 7.38 (d, 1H, *J* = 7.6, H-12), 7.26 (t, 1H, *J* = 7.6, H-10), 7.09 (t, 1H, *J* = 7.4, H-11), 5.96 (q, 1H, *J* = 6.8, H-19), 4.94-4.87 (m, 1H, H-3), 4.20 (d, 1H, *J* = 14, H-21a), 4.13 (d, 1H, *J* = 14, H-21b), 4.04 (dd, 1H, *J* = 4.0, 12, H-17a), 3.82 (dd, 1H, *J* = 4.7, 12, H-17b), 3.99-3.85 (m, 2H, H-23a, H-23b), 3.76 (s, 3H, H-22), 3.67-3.58 (m, 1H, H-5), 3.20 (dd, 1H, *J* = 7.0, 14, H-6a), 3.17-3.09 (m, 1H, H-16), 2.98 (dd, 1H, *J* = 7.0, 15, H-6b), 2.87-2.82 (m, 1H, H-15), 2.59-2.42 (m, 2H, H-14), 1.87 (d, 3H, *J* = 6.8, H-18); ESI-MS($C_{22}H_{29}N_2O_4^+$) *m/z* calculated:

385.2127 [M+H]$^+$, found: 385.2134 [M+H]$^+$. HPLC-isolation of **14** and partial NMR assignments show that this compound is also an indole-containing compounds: ^{1}H NMR (600 MHz, MeOD): δ 7.41 (d, 1H, J = 7.7), 7.34 (d, 1H, J = 8.2), 7.14 (t, 1H, J = 7.6), 7.04 (t, 1H, J = 7.5). 10-bromostrictosidine **8o** was chemically synthesized together with its vincoside diastereomer (Dr. E. McCoy). Briefly, 5-bromotryptamine **6o** hydrochloride was added to secologanin **7** in maleic acid buffer (10 mM, pH 2.0) and reacted overnight to afford the product [ESI-MS($C_{27}H_{34}N_2O_9{}^{79}Br^+$) *m/z* calculated: 609.1448 [M+H]$^+$, found: 609.1456 [M+H]$^+$; ESI-MS($C_{27}H_{34}N_2O_9{}^{81}Br^+$) *m/z* calculated: 611.1427 [M+H]$^+$, found: 611.1453 [M+H]$^+$]. The reaction was filter-sterilized and added to hairy root cultures as described above (final volume 20 mL half-strength Gamborg media, compound concentration 0.5 mM). After seven days the hairy roots were extracted into MeOH, filtered and injected on the LC-MS. Assignments as brominated compounds are based on isotopic signals of ^{79}Br and ^{81}Br incorporated in the alkaloid analog. Bromine-substituted analogs are expected to have later elution times (more hydrophobic) than the natural MIA alkaloids. See Appendix A for LC-MS traces.

3.5 References

1. O'Connor, S.E.; Maresh, J.J. "Chemistry and biology of monoterpene indole alkaloid biosynthesis", *Nat. Prod. Rep.* **2006**, *23*, 532-547.

2. van der Heijden, R.; Jacobs, D.I.; Snoeijer, W.; Didier, H.; Verpoorte, R. "The *Catharanthus* alkaloids: pharmacognosy and biotechnology", *Curr. Med. Chem.* **2004**, *11*, 607-628.

3. Johnson, I.S.; Wright, H.F.; Svoboda, G.H. "Experimental basis for clinical evaluation of antitumor principles from *Vinca rosea* Linn", *J. Lab. Clin. Med.* **1959**, *54*, 830-838.

4. Svoboda, G.H. "Alkaloids of *Vinca rosea* Linn. (*Catharanthus roseus*). 1X: extraction and characterization of leurosidine and leurocristine", *Llyodia* **1961**, *24*, 173-178.

5. Fahy, J. "Modifications in the "upper" or velbenamine part of the vinca alkaloids have major implications for tubulin interacting activities", *Curr. Pharm. Des.* **2001**, *7*, 1181-1197.

6. Fahy, J.; Duflos, A.; Ribet, J.-P.; Jacquesy, J.-C.; Berrier, C.; Jouannetaud, M.-P.; and Zunino, F. "Vinca alkaloids in superacidic media: a method for creating a new family of antitumor derivatives", *J. Am. Chem. Soc.* **1997**, *119*, 8576-8577.

7. Kruczynski, A.; Hill, B.T. "Vinflunine, the latest Vinca alkaloid in clinical development: a review of its preclinical anticancer properties", *Oncology and Hematology* **2001**, *40*, 159-173.

8. McCoy, E.; O'Connor, S.E. "Directed biosynthesis of alkaloid analogs in the medicinal plant periwinkle", *J. Am. Chem. Soc.* **2006**, *128*, 14276-14277.

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

10. Kutchan, T.M. "Strictosidine: from alkaloid to enzyme to gene", *Phytochemistry* **1993**, *32*, 493-506.

11. Chen, S.; Galan, M.C.; Coltharp, C.; and O'Connor, S.E. "Redesign of a central enzyme in alkaloid biosynthesis", *Chem. Biol.* **2006**, *13*, 1137-1141.

12. Ma, X.; Panjikar, S.; Koepke, J.; Loris, E.; and Stöckigt, J. "The structure of *Rauvolfia serpentina* strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins", *Plant Cell* **2006**, *18*, 907-920.

13. Chica, R.A.; Doucet, N.; Pelletier, J.N. "Semi-rational approaches to engineering enzyme activity: combining benefits of directed evolution and rational design", *Curr. Opin. Biotechnol.* **2005**, *16*, 378-384.

14. Morley, K.L.; Kazlauskas, R.J. "Improving enzyme properties: when are closer mutations better?", *Trends Biotechnol.* **2005**, *23*, 231-237.

15. Derbyshire, K.M.; Salvo, J.J.; Grindley, N.D. "A simple and efficient procedure for saturation mutagenesis using mixed oligodeoxynucleotides", *Gene* **1986**, *46*, 145-152.

16. Bernhardt, P.; McCoy, E.; O'Connor, S.E. "Rapid identification of enzyme variants for reengineered alkaloid biosynthesis in periwinkle", *Chem. Biol.* **2007**, *14*, 888-897.

17. Maresh, J.J.; Giddings, L.A.; Friedrich, A.; Loris, E.A.; Panjikar, S.; Trout, B.L.; Stöckigt, J.; Peters, B.; O'Connor, S.E. "Strictosidine synthase: mechanism of a Pictet-Spengler catalyzing enzyme", *J. Am. Chem. Soc.* **2008**, *130*, 710-723.

18. Miyaura, N.; Suzuki, A. "Palladium-catalyzed cross-coupling reactions of organoboron-compounds", *Chem. Rev.* **1995**, *95*, 2457-2483.

19. Trouet, A.B.L.; Hannart, J.A.A.; and Rao, K.S.B. "Vinblastin-23-oyl amino acid derivatives", US Pat. 4,639,456. **1987**; *Chem. Abs.* 107, 40181.

20. Loris, E.; Panjikar, S.; Ruppert, M.; Barleben, L.; Unger, M.; Schübel, H.; Stöckigt, J. "Structure-based engineering of strictosidine synthase: auxiliary for alkaloid libraries", *Chem. Biol.* **2007**, *14*, 979-985.

21. Hutchinson, C.R.; Heckendorf, A.H.; Straughn, J.L.; Daddona, P.E.; and Cane, D.E. "Biosynthesis of camptothecin. 3. Definition of strictosamide as the penultimate biosynthetic precursor assisted by carbon-13 and deuterium NMR spectroscopy", *J. Am. Chem. Soc.* **1979**, *101*, 3358-3369.

22. Kohl, W.; Witte, B.; and Höfle, G. "Alkaloide aus *Catharanthus roseus* Zellkulturen, III", *Z. Naturforsch. B., Anorg. Chem. Org. Chem.* **1982**, *37b*, 1346-1351.

23. Kohl, W.; Witte, B.; and Höfle, G. "Alkaloide aus *Catharanthus roseus* Zellkulturen, II", *Z. Naturforsch. B., Anorg. Chem. Org. Chem.* **1981**, *36b*, 1153-1162.

24. Hughes, E.H.; Hong, S.-B.; Shanks, J.V.; San, K.-Y.; Gibson, S.I. "Characterization of an inducible promoter system in *Catharanthus roseus* hairy roots", *Biotechnol. Prog.* **2002**, *18*, 1183-1186.

25. Runguphan, W.; O'Connor, S.E. "Metabolic reprogramming of periwinkle plant culture", *Nat. Chem. Biol.* **2009**, *5*, 151-153.

26. Runguphan, W.; Maresh, J.J. O'Connor, S.E. "Silencing of tryptamine biosynthesis for production of nonnatural alkaloids in plant culture", *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13673-13678.

27. Kutney, J.P.; and Brown, R.T. "The structural elucidation of sitsirikine, dihydrositsirikine and isositsirikine", *Tetrahedron* **1966**, *22*, 321-336.

28. McCoy, E.A.; O'Connor, S.E.; *manuscript in preparation* **2009**.

29. Galan, M.C.; O'Connor, S.E. "Semi-synthesis of secologanin analogues", *Tetrahedron Lett.* **2006**, *47*, 1563-1565.

30. Hutchinson, C.R.; Heckendorf, A.H.; Straughn, J.L.; Daddona, P.E.; and Cane, D.E. "Biosynthesis of camptothecin. 3. Definition of strictosamide as the penultimate biosynthetic precursor assisted by carbon-13 and deuterium NMR spectroscopy", *J. Am. Chem. Soc.* **1979**, *101*, 3358-3369.

CHAPTER 4

SUBSTRATE SPECIFICITY OF *OPHIORRIZA PUMILA* STRICTOSIDINE SYNTHASE

AND A STEREOCHEMICAL MODEL FOR STRICTOSIDINE SYNTHASE

4.1 Introduction

The Pictet-Spengler reaction, a special case of the Mannich reaction, is where a β-arylethylamine, such as tryptamine **1** or dopamine **4**, undergoes ring closure after condensation with an aldehyde.[1] This reaction generates one carbon-carbon bond and one stereogenic center. There are several Pictet-Spengler reactions in natural product biosynthetic pathways, including reactions between tryptamine **1** and secologanin **2** to form strictosidine **3**; between dopamine **4** and secologanin **2** to form deacetylisoipecoside **5**; between dopamine **4** and 4-hydroxyphenyl acetaldehyde **6** to form norcoclaurine **7**; and, between dopamine **4** and 4-hydroxydihydrocinnamaldehyde **8** to form autumnaline **9** (Figure 4.1).[2]

The tetrahydro-β-carboline moiety of strictosidine **3** is a heterocycle found in a wide range of natural products and synthetic pharmaceuticals. Catalytic methodology exists to stereoselectively generate the carbon-carbon bond formed in this Pictet-Spengler reaction,[3-9] although these chemical methods often suffer from low stereoselectivity, involve the use of organic solvents, and require activated amine substrates or harsh reaction conditions. Since enzymes can be useful for preparative stereoselective synthetic transformations,[10,11] and since natural product biosynthetic pathways are potentially rich sources for new biocatalysts, I sought a biocatalytic alternative to forming tetrahydro-β-carbolines.

119
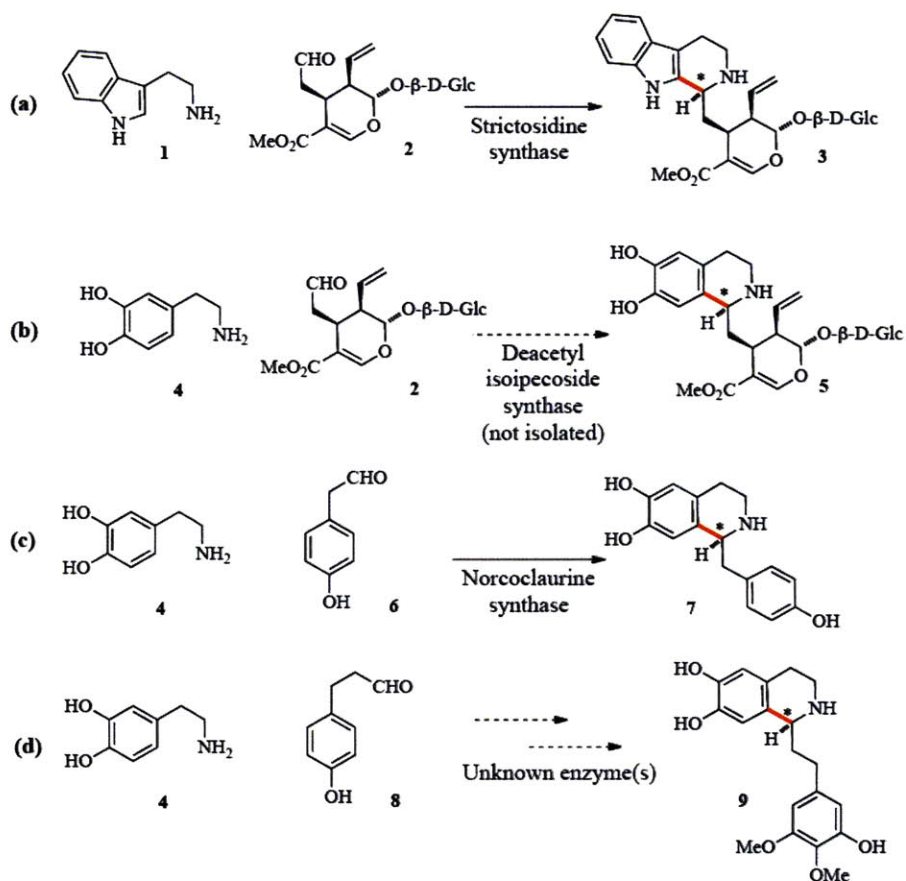
Figure 4.1 Pictet-Spengler reactions in natural product biosynthesis. The new carbon-carbon bond is shown in red-bold and the new stereogenic center (C3) is marked with an asterisk. (a) Formation of strictosidine **3** by strictosidine synthase; (b) formation of deacetylisoipecoside **5**; (c) formation of norcoclaurine **7**; and (d) formation of autumnaline **9** by a set of unknown enzymes (in addition to the Pictet-Spenglerase, two hydroxylases and two *O*-methyl transferases may be involved).

Strictosidine synthase (STS) catalyzes the formation of strictosidine **3** in the biosynthesis of monoterpene indole alkaloids (MIAs) and some quinoline alkaloids.[12] Two STS homologs from the medicinal plants *Rauvolfia serpentina* (RsSTS)[13] and *Catharanthus roseus* (CrSTS),[14] with 82% amino acid sequence identity and near-identical active sites, have been studied in detail. RsSTS and CrSTS have narrow substrate scopes that preclude their application as general biocatalysts, although mutagenesis can expand the substrate scope somewhat.[15-19] In this chapter I describe the characterization of *Ophiorrhiza pumila* STS (OpSTS), which has a considerably broader substrate specificity compared to RsSTS and CrSTS or any mutant of these two enzymes.[20] I also propose a model for how STS achieves its high stereoselectivity; this model may be used to construct a Pictet-Spenglerase with the alternative stereoselectivity.

4.2 Results and discussion

4.2.1 Sequence and homology model of *O. pumila* STS

OpSTS, like CrSTS and RsSTS, also catalyzes the Pictet-Spengler reaction between tryptamine **1** and secologanin **2** to form strictosidine **3**, but has a lower sequence identity to both CrSTS (54%) and RsSTS (60%), compared to the high sequence identity between CrSTS and RsSTS (82%).[21] Specifically, sequence alignments showed that OpSTS has a four-amino acid deletion that corresponds to $\Delta[\text{GluLeuAspGly}]^{277\text{-}280}$ in CrSTS and $\Delta[\text{GluLeuAspGly}]^{271\text{-}274}$ in RsSTS, and several other differences in this sequence region (Figure 4.2). The crystal structure of RsSTS in complex with secologanin **2** (PDB: 2FPC)[22] suggests that the deletion in OpSTS is near the secologanin binding site.

```
OpSTS  -------MHSSEAMVVSILCALFLSSLSLVSSSPEFFEFIEAPSYGPNAYAFDSDGE-LY  52
CrSTS  MANFSESKSMMAVFFMFFLLLLSSSSSSSSSSSPILKKIFIESPSYAPNAFTFDSTDKGFY  60
RsSTS  ----MAKLSDSQTMALFTVFLLFLSSSLALSSPILKEILIEAPSYAPNSFTFDSTNKGFY  56

OpSTS  ASVEDGRIIKYDKPSNKFLTHAVASPIWNNALCENNTNQDLKPLCGRVYDFGFEYETQRL  112
CrSTS  TSVQDGRVIKYEGPNSGFTDFAYASPFWNKAFCENSTDPEKRPLCGRTYDISYDYKNSQM  120
RsSTS  TSVQDGRVIKYEGPNSGFVDFAYASPYWNKAFCENSTDAEKRPLCGRTYDISYNLQNNQL  116

OpSTS  YIADCYFGLGFVGPDGGHAIQLATSGDGVEFKWLYALAIDQQAGFVYVTDVSTKYDD--R  170
CrSTS  YIVDGHYHLCVVGKEGGYATQLATSVQGVPFKWLYAVTVDQRTGIVYFTDVSSIHDDSPE  180
RsSTS  YIVDCYYHLSVVGSEGGHATQLATSVDGVPFKWLYAVTVDQRTGIVYFTDVSTLYDD--R  174

OpSTS  GVQDIIRINDTTGRLIKYDPSTEEVTVLMKGLNIPGGTEVSKDGSFVLVGEFASHRILKY  230
CrSTS  GVEEIMNTSDRTGRLMKYDPSTKETTLLLKELHVPGGAEISADGSFVVVAEFLSNRIVKY  240
RsSTS  GVQQIMDTSDKTGRLIKYDPSTKETTLLLKELHVPGGAEVSADSSFVLVAEFLSHQIVKY  234

OpSTS  WLKGPKANTSEFLLKVRGPGNIKRTKDGDFWVASSD----NNGITVTPRGIRFDEFGNIL  286
CrSTS  WLEGPKKGSAEFLVTIPNPGNIKRNSDGHFWVSSSEELDGGQHGRVVSRGIKFDGFGNIL  300
RsSTS  WLEGPKKGTAEVLVKIPNPGNIKRNADGHFWVSSSEELDGNMHGRVDPKGIKFDEFGNIL  294

OpSTS  EVVAIPLPYKGEHIEQVQEHDGALFVGSLFHEFVGILHNYKSSVDHHQEKNSGGLNASFK  346
CrSTS  QVIPLPPPYEGEHFEQIQEHDGLLYIGSLFHSSVGILVYDDHDNKGNSYVSS--------  352
RsSTS  EVIPLPPPFAGEHFEQIQEHDGLLYIGTLFHGSVGILVYDKKGNSFVSSH----------  344

OpSTS  EFSSF 351
CrSTS  -----
RsSTS  -----
```

Figure 4.2 Sequence alignments of OpSTS, CrSTS, and RsSTS. The four-amino acid deletion and the Gly that replaces His283 in CrSTS are highlighted in yellow; the corresponding residues in CrSTS and RsSTS are highlighted in bold.
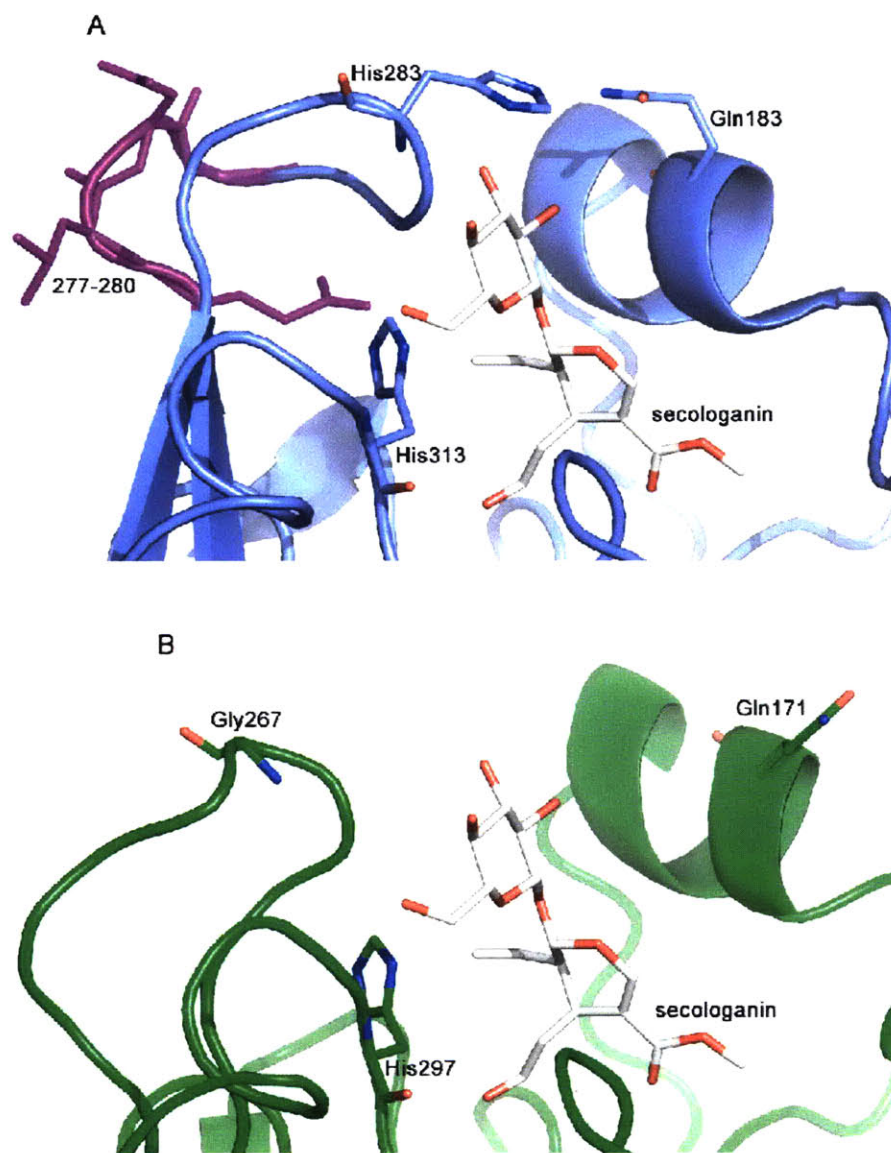
Figure 4.3 (A) Crystal structure of RsSTS with the CrSTS amino acid sequence numbering. The glucose moiety of secologanin **2** binds in a pocket defined by a loop-loop-helix motif, which contains two histidine residues (His283 and His313). (B) Homology model of OpSTS based on the crystal structure of RsSTS. The glucose moiety of secologanin **2** likely has fewer interactions with this loop structure due to a four-amino acid deletion and histidine to glycine substitution.

Moreover, His283 (His277 in RsSTS), which is substituted for a glycine in OpSTS (Figure 4.2) and on the same loop as the deletion, is within hydrogen bonding distance to the glucose moiety of **2**.

To obtain further insight into the differences between OpSTS an RsSTS, I constructed a homology model of OpSTS based on the crystal structure of RsSTS and the amino acid sequence of OpSTS, using PHYRE (Figure 4.3B).[23,24] Since the amino acid sequence and the amine and aldehyde substrate scopes of CrSTS and RsSTS are nearly identical, I presumed that the structure of RsSTS is representative of CrSTS. The homology model of OpSTS suggests that the region in the secologanin binding site is different in OpSTS compared to RsSTS and these differences prompted me to assess whether OpSTS has a broader, and therefore potentially more useful, aldehyde substrate scope.

4.2.2 Expression, purification, and kinetics of OpSTS

I codon optimized the gene encoding OpSTS for expression in *Escherichia coli*, removed the nucleotides that encode the first 25 amino acids in the published peptide sequence (GI: 13928598) to remove a putative signal peptide,[25,26] and then cloned the optimized *opsts* gene into the *E. coli* expression vector pET-28 (section 4.4.3). After overexpression of OpSTS and a similar construct of CrSTS in *E. coli* (Chapter 3), I purified both enzymes using affinity chromatography (section 4.4.3).[17] Pseudo steady-state kinetics for secologanin **2** with OpSTS and CrSTS verified that heterologously expressed OpSTS and CrSTS have similar catalytic efficiencies (Table 4.1).[20]

Table 4.1. Kinetic constants.[a]

Enzyme, substrate	V_{max} [U mg^{-1}] [b]	K_M [μM]	V_{max}/K_M [U mg^{-1} mM^{-1}] [b]
CrSTS, secologanin **2**	7.4 ± 0.6	40 ± 20	200 ± 150
OpSTS, secologanin **2**	1.7 ± 0.1	21 ± 5	80 ± 30
OpSTS, 5-methoxytryptamine **24**	0.0043 ± 0.0003	40 ± 10	0.11 ± 0.04
OpSTS, 2'*(R)*-tryptophanol **25**	0.0359 ± 0.0003	18 ± 1	2.0 ± 0.2

[a] Kinetic assays were performed in sodium phosphate buffer (0.05 M, pH 7.0) using an internal standard (NAA, 0.075 mM) and analyzed by HPLC for formation of product; [b] 1 U = formation of 1 μmol product per minute at pH 7.0 and 30 °C.

4.2.3 Aldehyde substrate scope of OpSTS

To examine the aldehyde substrate scope of OpSTS, I selected a series of aldehydes **10-20** with the following properties: aliphatic, aromatic, linear, branched, hydrogen bond-accepting, and functionalized (Table 4.2). I then incubated aldehydes **10-20** (1 mM) and tryptamine **1** (1 mM) with OpSTS (0.2 mol% catalyst, 2 μM) or CrSTS (1.0 mol% catalyst, 10 μM) in aqueous buffer (pH 7.0). Analysis of organic extracts of the reaction mixtures showed that OpSTS is capable of catalyzing the conversion of aldehydes **10-16** and **19-20** into the corresponding tetrahydro-β-carboline products, as evidenced HRMS.[20]

I prepared authentic standards of tetrahydro-β-carbolines resulting from reaction of aldehydes **10-16** and tryptamine **1** using conditions where the product precipitates during the reaction or on standing, presumably by forming a hydrochloride salt. This method is compatible with green chemistry, as no organic solvent is required for either reaction or workup. I characterized the racemic mixtures by chiral HPLC, HRMS, and NMR (section 4.4.2 and Appendix B), and found that the tetrahydro-β-carboline standards co-elute with the products resulting from OpSTS catalysis.

Table 4.2 Aldehyde scope for OpSTS and CrSTS

Starting material R =	k_{obs} [min^{-1}][a] OpSTS	CrSTS	Starting material R =	k_{obs} [min^{-1}][a] OpSTS	CrSTS
(structure) 10	0.048	0.0002	(structure) 16	<0.0001[c]	not obsvd.[b]
(structure) 11	0.037	not obsvd.[b]	(structure) 17	not obsvd.[b]	not obsvd.[b]
(structure) 12	0.032	not obsvd.[b]	(structure) 18	not obsvd.[b]	not obsvd.[b]
(structure) 13	0.022	not obsvd.[b]	(structure) 19	n.d.[d]	not obsvd.[b]
(structure) 14	0.013	not obsvd.[b]	(structure) 20	n.d.[d]	not obsvd.[b]
(structure) 15	<0.0001[c]	not obsvd.[b]			

[a] k_{obs} measured by monitoring product formation with **1** at 280 nm using reverse-phase HPLC with a PDA detector; [b] not observed, detection limit ~10^{-5} min^{-1}; [c] quantification limit ~10^{-4} min^{-1}. k_{obs} of **10** is 230-fold higher for OpSTS than CrSTS; k_{obs} of **11-14** is > 3700-fold higher with OpSTS than CrSTS, given a detection limit of 10^{-5} min^{-1}; [d] not determined.

Poor solubility of the aldehyde analogs precluded a full kinetic analysis (estimated K_M-values for the aldehydes are ≥10 mM). However, I derived k_{obs}-values by HPLC kinetics at 1 mM substrate concentration (section 4.4.7). Although OpSTS favors the natural substrate, secologanin **2**, by at least 1000-fold, it catalyzes the conversion of a relatively broad range of aldehyde substrates, including hydrogen bond-accepting (**10**), aliphatic (**11-12** and **14**), and aromatic (**13**) aldehydes with similar rates (k_{obs} = 0.013-0.048 min^{-1}). The kinetics also demonstrated that OpSTS clearly prefers aldehydes **10-14** to aldehydes **15-16**; particularly aldehydes unsubstituted at the alpha position (compare **10-14** with **16-18**). CrSTS accepts substrate **10**, although k_{obs} is 230-fold higher for OpSTS than CrSTS, but not aldehydes **11-14** since k_{obs} is >3700-fold higher for OpSTS compared to CrSTS (Table 4.2).

The enantiomeric excess (*ee*) of the tetrahydro-β-carboline products derived via OpSTS catalysis is >98% as evidenced by chiral HPLC (section 4.4.5 and Appendix B). To determine the absolute configuration of the enzymatic products, I immobilized OpSTS onto solid support (CNBr-activated Sepharose, section 4.4.4)[27] and used this immobilized catalyst to catalyze the 1-mmol scale reaction between tryptamine **1** and *n*-hexanal **14**. Chiral HPLC analysis of the enzymatic product and an authentic 3*(S)*-standard obtained by Jacobsen's organocatalytic method (section 4.4.6),[3] showed that the absolute configuration is 3*(S)*, which is the configuration observed with the natural substrate, secologanin **2**. The absolute configuration is assumed 3*(S)* for all enzymatic products.

4.2.4 Tryptamine substrate scope of OpSTS

I hypothesized that specific interactions between tryptamine **1** and the enzyme active site entrance partly determine the amine substrate specificity. Hence, if the secologanin binding site, which is closer to the surface of the enzyme (PDB ID: 2FPC), is different in OpSTS compared to CrSTS it is possible that the tryptamine substrate scope is altered.

I assayed a series of tryptamine analogs **21-28** – which previous studies showed that CrSTS or RsSTS do not accept – with secologanin **2** and OpSTS (Table 4.3). LC-MS analysis showed that OpSTS has a different tryptamine substrate scope than CrSTS or any CrSTS mutant. While 5-methyl- **21**, 5-chloro- **22**, and 5-bromotryptamine **23** are not substrates for OpSTS, 5-methoxytryptamine **24** *is* a substrate (Table 4.3); conversely,

Val214Met CrSTS accepted **21-23**, but not **24** (Chapter 3).[17] In addition, 2'*(R)*-tryptophanol **25**, which is converted by the Phe232Leu mutant of CrSTS but not wild-type CrSTS,[17] is a substrate for OpSTS (Table 4.3).

Table 4.3 Tryptamine scope of OpSTS[a]
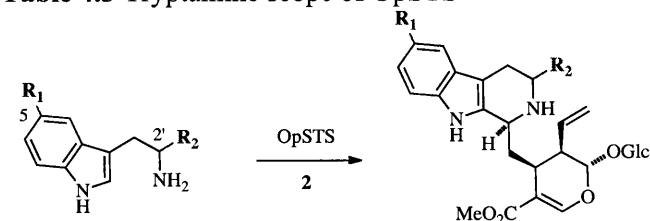
Starting material R$_1$ =	R$_2$ =	Outcome	Starting material R$_1$ =	R$_2$ =	Outcome
21 Me	H	–	25 H	(R)-CH$_2$OH	+
22 Cl	H	–	26 H	(S)-CH$_2$OH	–
23 Br	H	–	27 H	(R)-CO$_2$H	–
24 OMe	**H**	**+**	28 H	(S)-CO$_2$H	–

[a] The reaction outcome was determined by monitoring product appearance by LC-MS; CrSTS does not catalyze the Pictet-Spengler reaction with any of the assayed tryptamine analogs **21-28**.

Pseudo steady-state kinetics with OpSTS using **24** and **25** showed that both substrates follow the Michealis-Menten model (Table 4.1 and section 4.4.7).[20] The catalytic constants suggest that OpSTS is a significantly better catalyst than Phe232Leu STS with 2'*(R)*-tryptophanol **25** (Chapter 3).[17]

The factors that contribute to the dramatic changes in tryptamine substrate scope are at present unknown, though I speculate that the electronic properties of the substrate *and* the enzyme are important in the binding of tryptamine substrates to OpSTS. Therefore, monitoring the effects on the substrate specificity of CrSTS, as a result of mutations of residues at the entryway to the active site, may lead to insights into the mechanism of

tryptamine substrate specificity. Nevertheless, the broader tryptamine substrate specificity of OpSTS renders this enzyme a candidate for metabolic engineering of *C. roseus* cell cultures to produce new MIA.[28]

4.2.5 Towards elucidating the molecular basis for the different substrate specificities

In an attempt to elucidate the molecular basis for the broadened specificity in OpSTS, I modified amino acid residues in the secologanin binding site of CrSTS that are different in OpSTS (Figure 4.2 and section 4.4.8). I substituted His283, which hydrogen bonds to the glucose moiety of secologanin, with Gly, Leu, and Phe, and deleted the four amino acids present in CrSTS but absent in OpSTS (section 4.2.1). These changes, however, were not sufficient to expand the aldehyde scope of CrSTS; future experiments may shed light on the different substrate specificities of these two enzymes (Chapter 7).

4.2.6 Model for stereoselectivity in STS

The stereoselectivity observed in the reaction between **1** and **2** with RsSTS, CrSTS, and OpSTS is preserved with achiral aldehydes (section 2.4.3), indicating that stereoselectivity does not rely on the chiral substrate **2**. The configuration-determining step may be the 6-endo-trig cyclization[29] of iminium **29** (derived from condensation of **1** and **2**), which gives a high-energy intermediate **30**, containing two stereogenic centers: C2 and C3; H2 is deprotonated to form **3** (Figure 4.4).[30] I hypothesized that formation of **30** is the configuration determining step, and docked QM/MM models of **30** into OpSTS

using a molecular mechanics force field to gain insight into the molecular basis of stereoselectivity (section 4.4.9).

The x-ray structure of RsSTS in complex with secologanin **2** (PDB ID: 2FPC)[22] served as starting structure for docking of the four six-membered ring structures comprising **30** (section 4.4.9), which are the products of cyclization from *cis* and *trans* iminium precursors **29** from the top or bottom face of the indole plane, respectively. The four structures of **30** were docked into RsSTS while rotating the C3-C14 and C14-C15 bonds of the ligand and the following bonds in the tryptamine binding site: Trp149 (β-γ), Val176 (α-β), Val208 (α-β) and Glu309 (α-β, β-γ, γ-δ).

Glu309 (Glu315 in CrSTS) is essential in the catalytic cycle, where it participates as a general acid/base catalyst. I define a productive docked structure (one that leads to the formation of product) as one that contains two hydrogen bonds (Tyr151-Glu309 and Glu309-N_4H), a π-stacking interaction between the indole moiety and Phe226, and proximity of Glu309 to H2. I define a non-productive docked structure as one that differ in one or more of these criteria. Based on these criteria, only one intermediate, **30-2*(R)*3*(S)*,** shows all productive interactions, including the proximity of H2 to Glu309, the Tyr151-Glu309 (2.8 Å) and Glu309-N_4H (2.8 Å) hydrogen bonds, and a π-stacking similar to that observed in the x-ray structures of RsSTS in complex with tryptamine or an inhibitor (Figure 4.5A).[22,30] Its H-2 epimer, **30-2*(S)*3*(S)*,** which also results in
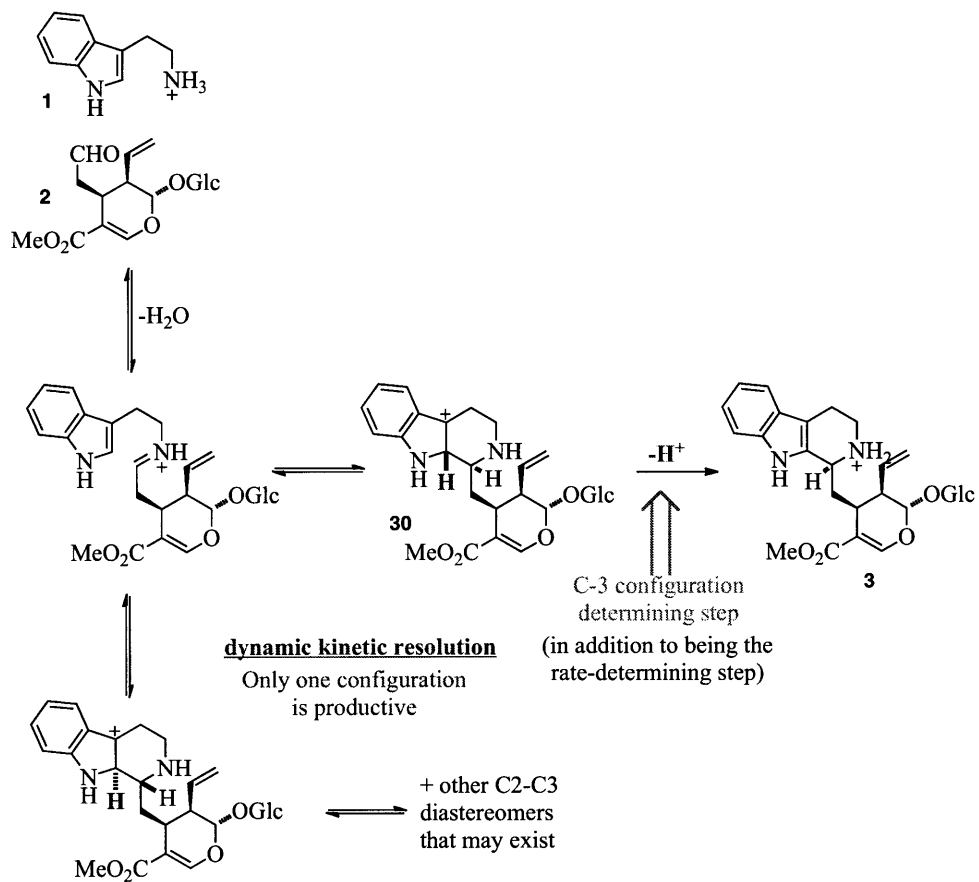
Figure 4.4 The Pictet-Spengler reaction between tryptamine **1** and secologanin **2** leads to strictosidine **3** via iminium **29** and the proposed "productive" high-energy intermediate, **30**-*2(R)3(S)*.

A

Glu309

Tyr151

30-2*(R)*3*(S)*
(to strictosidine **3**)

2.8 Å

H-2   3.3 Å

3.5 Å   2.8 Å

H-3

Interactions:
Glu309-Oε... HO-Tyr151
Tyr151 π-stacking
Glu309-Oε... H2
Glu309-Oε... H-N4

Absent interactions:
N/A

Productive intermediate **30**:
H-2 pointing towards Glu309
hydrogen bonds and π-stacking present

Phe226

B

Glu309

Tyr151

30-2*(S)*3*(S)*
(to strictosidine **3**)

2.9 Å

4.4 Å

H-2   H-3

Interactions:
Glu309-Oε... HO-Tyr151

Absent interactions:
π-stacking
Glu309-Oε... H2
Glu309-Oε... H-N4

Nonproductive intermediate **30**:
H-2 pointing away from Glu309
missing hydrogen bonds and π-stacking

Phe226

Glu309

Tyr151

30-2(R)3(R)
(to vincoside)

2.7 Å

2.7 Å

3.7 Å
H-2

H-3

Interactions:
Glu309-Oε... HO-Tyr151
Glu309-Oε... H-N4

Absent interactions:
π-stacking
Glu309-Oε... H2

Glu309 changed position

*Nonproductive intermediate **30**:*
H-2 pointing towards Glu309, though Glu309 is reoriented
contains key hydrogen bonds but lacks π-stacking

Phe226

C

Glu309

Tyr151

30-2(S)3(R)
(to vincoside)

2.8 Å

2.8 Å
H-3

Interactions:
Glu309-Oε... HO-Tyr151
Phe226 π-stacking
Glu309-Oε... H-N4

H-2

Absent interactions:
Glu309-Oε... H2

3.8 Å

*Nonproductive intermediate **30**:*
H-2 pointing away from Glu309
hydrogen bonds and π-stacking present

Phe226

D

133

E

F

Figure 4.5 (A-D) High-energy reaction intermediates **30** docked into the active site of RsSTS (PDB ID: 2FPC) as described in the text and in section 4.4.9. Each panel shows interactions that are present or absent; an intermediate is defined as non-productive if it lacks one or several of these interactions. Intermediates with 3*(S)* configuration, leading to strictosidine **3**, are shown in a white-tone CPK color scheme; intermediates with 3*(R)* configuration, leading to vincoside, are depicted with a blue-tone CPK color scheme. (E) X-ray crystal structure of RsSTS in complex with tryptamine (PDB ID: 2FPB). (F) X-ray crystal structure of RsSTS in complex with strictosidine (PDB ID: 2V91).

134
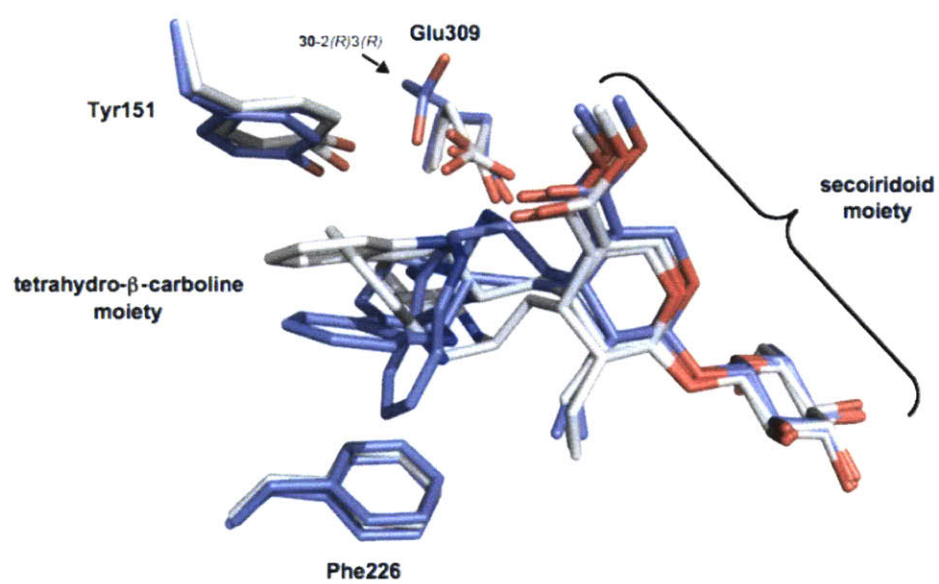
Figure 4.6 Overlay of the docked high-energy intermediates **30**. One intermediate, **30**-2*(R)*3*(R)*, which leads to vincoside, docks in a way that reorients Glu309.

strictosidine formation only preserves the Glu309-Tyr151 (2.9 Å) interaction (Figure 4.5B). The diastereomer of **30-2*(S)*3*(S)***, **30-2*(R)*3*(R)***, twists the Glu309 side-chain orientation and lacks the π-stacking interaction (Figures 4.5C and 4.6). The H-2 epimer of **30-2*(S)*3*(S)***, **30-2*(S)*3*(R)***, leading to vincoside formation has all key interactions (Figure 4.5D), but displays the H2 proton on the opposite face of the indole plane compared to Glu309. Therefore, I hypothesize that the general base, Glu309, can deprotonate **30-2*(R)*3*(S)*** (closest distance is 3.4 Å), while Glu309 cannot deprotonate **30-2*(S)*3*(R)***. Since kinetic data suggests that the iminium-cyclization step is reversible, the "wrong" stereoisomer may reform the iminium and then have another chance to cyclize.

In summary, based on the docked structures, facial deprotonation is likely key for determining the stereochemical outcome. This facial deprotonation model is consistent with the reversible nature of the Pictet-Spengler reaction, but should be interpreted cautiously, since it has not been experimentally validated. However, this model serves as a basis for understanding and engineering the stereoselectivity of Pictet-Spenglerases (Chapter 7).

4.3 Conclusions

This chapter describes the first report of a Pictet-Spenglerase with broad substrate specificity. The broader *aldehyde* substrate specificity is potentially useful for biocatalytic applications. Future experiments involve increasing the thermostability of OpSTS that would allow reaction at higher temperatures and, presumably, lead to higher

product yields (Chapter 7). The broader *tryptamine* substrate specificity makes OpSTS a candidate for constructing transgenic plants to produce new monoterpene indole alkaloid analogs.

The evolutionary distance between the apocynaceae plant family, from which CrSTS and RsSTS are derived, and the rubiaceae family, from which OpSTS is derived, may explain the different substrate specificities. Rubiaceae plants contain chiral tetrahydro-β-carbolines different from those found in the apocynaceae family,[31] and although the enzymes that catalyze the formation of the precursor for these rubiaceae natural products are unknown, I speculate that they contain a Pictet-Spenglerase that resembles OpSTS. Screening of rubiaceae plants for new STS homologs and assay of the predicted substrates can experimentally test this possibility.

OpSTS catalyzes the formation of the 3*(S)* stereogenic center even with smaller aldehydes; I presented a stereochemical model that accounts for how STS maintains high stereoselectivity irrespective of the size and chirality of the aldehyde substrate. This model may facilitate mutagenesis to alter the stereoselectivity of OpSTS.

4.4 Experimental methods

4.4.1 General methods and analytical techniques

All purchased chemicals were used as received from the supplier. Achiral aldehydes were obtained from VWR distributors or Sigma-Aldrich, except THPCH$_2$CHO **10**, which was

obtained from Biofine International (Blaine, WA), and CyCH$_2$CHO **11**, which was obtained from BetaPharma (New Haven, CT); secologanin was obtained by extraction of *Lonicera tatarica* as previously described.[32]

UPLC and MS analysis were performed in tandem using an Acquity Ultra Performance BEH C18 column with a 1.7 μm particle size, 2.1 x 100 mm dimension, and a flow rate of 0.6 mL min^{-1}. MS analysis was carried out using a Micromass LCT Premier TOF Mass Spectrometer with an ESI source. Both modules are from Waters Corporation (Milford, MA). The capillary and sample cone voltages were 3000 V and 30 V, respectively. The desolvation and source temperatures were 300 and 100 °C, respectively. The cone and desolvation gas flow rates were 60 and 800 L h^{-1}. HPLC analysis was performed on a Beckman Coulter Gold System equipped with a 125 solvent module, 125 autosampler and a 168 detector. Data were processed (integrated) using 32Karat (v 7). Analytes were separated on a reverse-phase column (HiBar RT 250-4, RP-select B, 5 μm). Chiral HPLC analysis was performed on an Agilent 1100 series LC (Agilent Technologies, Inc., Santa Clara, CA). Data were processed using ChemStation for LC (Rev. A 10.01 [1635]). Analytes were separated on a normal-phase chiral column (Chiralcel OD-H 0.46 cm x 25 cm) from Daicel Chemical Industries, LTD; Chiral Technologies, Inc., West Chester, PA). ^{1}H NMR and ^{13}C NMR spectra were recorded on Varian 300 MHz, Varian 500 MHz, or Brucker 400 MHz spectrometers.
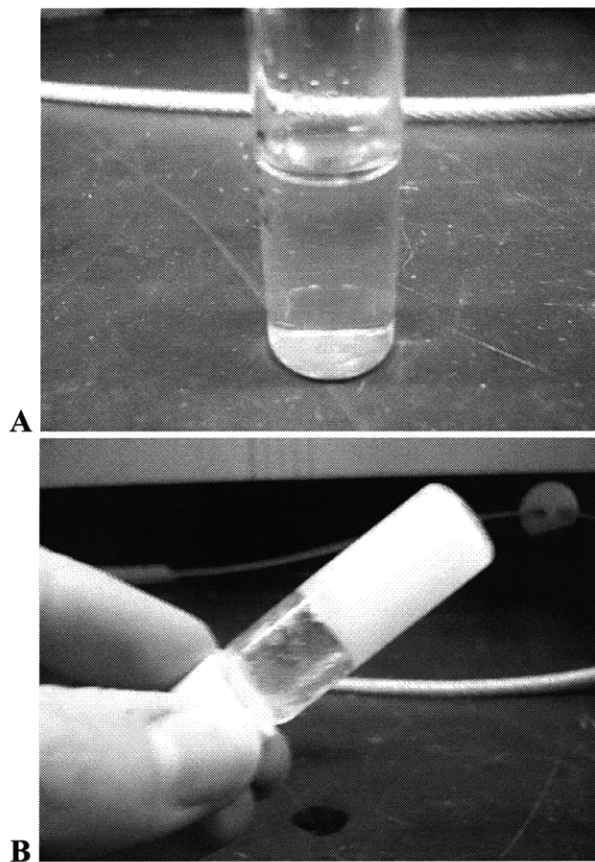
Figure 4.7 Precipitation of tetrahydro-β-carbolines in the synthesis of racemic standards. (A) Tryptamine **1** and hexanal **14** were added in equimolar ratio to maleic acid (pH 2) buffer. (B) The reaction solidifies after 20 min at 60 °C. Water was then added and the precipitate was filtered and washed in water to obtain the analytically pure tetrahydro-β-carboline product.

4.4.2 Chemical synthesis of authentic standards

Authentic standards were synthesized using a chemical Pictet-Spengler reaction in aqueous maleic acid buffer (1 M substrate concentration, 10 mM, pH 2.0). During incubation at 60 °C or on standing, the tetrahydro-β-carboline product precipitated (Figure 4.7). After filtration, and several washes with water, the dried precipitate contained >99% pure product.

All authentic standards were verified by ^{1}H and ^{13}C NMR (Appendix B) and used as authentic standards for *ee* and absolute configuration determination (sections 4.4.5-4.4.6), to confirm enzyme catalyzed product formation, and to obtain response factors for HPLC kinetics (section 4.4.7).

*1-((tetrahydro-2H-pyran-4-yl)methyl)-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (from 1 and 10)*
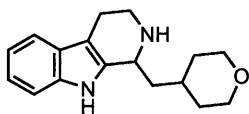
Tryptamine hydrochloride (1 M) was added to the tetrahydropyran aldehyde **10** (1 M) in aqueous maleic acid buffer (10 mM, pH 2.0) and the mixture was incubated at 60 °C for 18 h. The precipitate was rec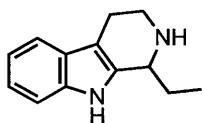overed, washed with water, and dried: ^{1}H NMR (MeOD): δ 7.48 (1H, dt, *J* = 0.9, 7.9), 7.36 (1H, dt, *J* = 0.9, 8.2), 7.16 (1H, td, *J* = 1.1, 8.2), 7.06 (1H, td, *J* = 0.9, 8.0), 4.86-4.82 (1H, m), 4.03-3.96 (2H, m), 3.75 (1H, dt, *J* = 1.3, 4.1, 12.4), 3.55-3.44 (3H, m), 3.12-3.07 (2H, m), 2.21-2.15 (1H, m), 1.95-1.83 (3H, m), 1.68 (1H,

td, J = 1.9, 13.1), 1.50-1.37 (2H, m); ^{13}C NMR (MeOD): δ 137.61, 129.58, 126.73, 122.80, 119.91, 118.34, 111.63, 106.54, 68.03, 67.87, 51.08, 42.39, 39.94, 33.93, 32.40, 31.09, 18.80; UV-max: 228, 279 nm; ESI-MS($C_{17}H_{23}N_2O^+$) *m/z* calculated: 271.1810 [M+H]$^+$, found: 271.1805 [M+H]$^+$.

1-(cyclohexylmethyl)-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (from 1 and 11)
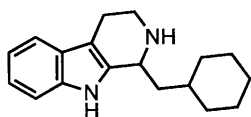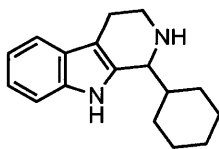

Tryptamine hydrochloride was added to the cyclohexyl aldehyde **11** in aqueous maleic acid buffer (10 mM, pH 2.0) and the mixture was incubated at 60 °C for 20 h. The precipitate was recovered, washed with water, and dried: ^{1}H NMR (MeOD): δ 7.47-7.44 (1H, m), 7.36 (1H, td, J = 0.8, 8.2), 7.14 (1H, ddd, J = 1.1, 7.1, 8.2), 7.04 (1H, ddd, J = 1.0, 7.1, 8.0), 4.81-4.74 (1H, m), 3.71 (1H, ddd, J = 3.5, 5.5, 12.5), 3.41 (1H, ddd, J = 5.5, 9.9, 12.6), 3.09 (1H, dddd, J = 1.9, 5.5, 9.8, 15.4), 3.05-2.99 (1H, m), 2.12 (1H, ddd, J = 3.7, 10.2, 14.3), 2.08-2.00 (1H, m), 1.87-1.60 (6H, m), 1.49-1.34 (2H, m), 1.33-1.22 (1H, m), 1.19-1.02 (2H, m); ^{13}C NMR (MeOD): δ 137.59, 129.95, 126.75, 122.69, 119.84, 118.31, 111.64, 106.39, 51.57, 42.46, 40.53, 34.56, 33.59, 32.41, 26.74, 26.54, 26.16, 18.81; ESI-MS($C_{18}H_{25}N_2^+$) *m/z* calculated: 269.2012 [M+H]$^+$, found: 269.2014 [M+H]$^+$.

141

1-isobutyl-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (from 1 and 12)
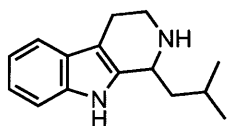
Tryptamine hydrochloride (1 M) was added to isovaleryl aldehyde **12** (1 M) in aqueous maleic acid buffer (10 mM, pH 2.0) and the mixture was incubated at 60 °C for 18 h. The precipitate was recovered, washed with water, and dried. The white shiny powder had poor solubility, but could be dissolved in DMSO-d_6 after sonication and brief heating. ^{1}H NMR (10:1, MeOD/DMSO-d_6) δ 7.50 (1H, td, J = 1.0, 7.8), 7.39 (1H, td, J = 0.9, 8.2), 7.17 (1H, ddd, J = 1.2, 7.1, 8.2), 7.08 (1H, ddd, J = 1.0, 7.1, 8.0), 4.82-4.79 (1H, m), 3.74 (1H, ddd, J = 3.9, 5.3, 12.6), 3.47 (1H, ddd, J = 5.9, 9.3, 12.6), 3.16-3.02 (2H, m), 2.08 (1H, ddd, J = 3.9, 10.2, 14.2), 1.99 (1H, dqd, J = 3.9, 6.5, 19.5), 1.85 (1H, ddd, J = 3.9, 10.3, 14.2), 1.15 (3H, d, J = 6.4), 1.10 (3H, d, J = 6.5); ^{13}C NMR (DMSO-d_6): δ 137.22, 131.64, 126.98, 122.74, 120.07, 119.09, 112.46, 106.81, 51.51, 41.94, 24.67, 24.53, 22.51, 19.16; ESI-MS($C_{15}H_{21}N_2^+$) m/z calculated: 229.1699 [M+H]$^+$, found: 229.1697 [M+H]$^+$.

1-benzyl-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (from 1 and 13)
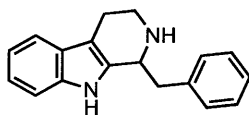
Tryptamine hydrochloride (1 M) was added to benzyl aldehyde **13** (1 M) in aqueous maleic acid buffer (10 mM, pH 2.0) and the mixture was incubated at 60 °C for 20 h. The precipitate was recovered, washed with water, and dried: ^{1}H NMR (MeOD): δ 7.55-7.37

(m, 7H), 7.21 (1H, ddd, J = 1.1, 7.1, 8.2,), 7.11 (1H, dd, J = 1.0, 7.1, 8.0), 5.06-5.01 (1H,

m), 3.79-3.73 (1H, m), 3.62 (1H, ddd, J = 3.6, 5.3, 12.5), 3.36 (1H, ddd, J = 5.7, 9.6,

12.6), 3.21-3.01 (3H, m); ^{13}C NMR (MeOD, MHz): δ 136.68, 135.82, 130.10, 129.87,

127.71, 126.26, 122.36, 119.54, 118.59, 111.90, 106.72, 54.20, 41.90, 37.76, 18.45; ESI-

MS($C_{18}H_{19}N_2^+$) m/z calculated: 263.1543 [M+H]$^+$, found: 263.1540 [M+H]$^+$.

1-pentyl-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (from 1 and 14)
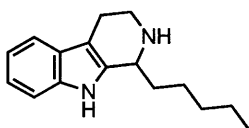
Tryptamine hydrochloride (1 M) was added to *n*-hexanal **14** (1 M) in aqueous maleic acid

buffer (10 mM, pH 2.0) and the mixture was incubated at 60 °C for 45 min. The

precipitate was recovered, washed with water, and dried: ^{1}H NMR (MeOD): δ 7.48 (1H,

dt, J = 0.9, 7.9), 7.36 (1H, dt, J = 0.9, 8.2), 7.15 (1H, td, J = 1.1, 8.2), 7.06 (1H, td, J =

0.9, 8.0), 4.69 (1H, dd, J = 4.1, 8.9) 3.73 (1H, ddd, J = 3.8, 5.5, 12.5), 3.44 (1H, ddd, J =

5.6, 9.6, 12.6), 3.11 (1H, dddd, J = 1.9, 5.6, 9.5, 15.1), 3.05 (1H, dddd, J = 1.3, 3.8, 5.3,

16.3), 2.31-2.23 (1H, m), 1.97-1.89 (1H, m), 1.64-1.54 (2H, m), 1.53-1.37 (4H, m), 0.97

(3H, t, J = 7.1); ^{13}C NMR (MeOD): δ 138.36, 130.32, 127.46, 123.48, 120.59, 119.10,

112.39, 107.18, 55.17, 43.19, 33.38, 32.79, 25.92, 23.48, 19.54, 14.37; ESI-

MS($C_{16}H_{23}N_2^+$) m/z calculated: 243.1856 [M+H]$^+$, found: 243.1851 [M+H]$^+$.

1-ethyl-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (from 1 and 15)

Tryptamine hydrochloride (1 M) was added to propanal **15** (1 M) in aqueous maleic acid buffer (10 mM, pH 2.0) and the mixture was incubated at 60 °C for 20 h. The precipitate formed upon standing (approximately 48 h), was recovered, washed with water, and dried: ^{1}H NMR (MeOD): δ 7.47 (1H, td, J = 0.9, 7.9), 7.37 (1H, td, J = 0.9, 8.2), 7.15 (1H, ddd, J = 1.1, 7.1, 8.2), 7.06 (1H, ddd, J = 1.0, 7.1, 8.0), 4.63-4.57 (1H, m), 3.71 (1H, ddd, J = 3.7, 5.6, 12.5), 3.39 (1H, ddd, J = 5.5, 9.7, 12.6), 3.10 (1H, dddd, J = 2.0, 5.6, 9.7, 15.3), 3.01 (1H, dddd, J = 1.4, 3.7, 5.3, 16.3), 2.33 (1H, dqd, J = 4.2, 7.6, 15.1), 2.05-1.93 (1H, m), 1.19 (3H, t, J = 7.5); ^{13}C NMR (MeOD): δ 138.30, 130.14, 127.42, 123.46, 120.57, 119.13, 112.38, 107.25, 56.19, 43.15, 26.33, 19.52, 9.93; ESI-MS($C_{13}H_{17}N_2^+$) *m/z* calculated: 201.1386 [M+H]$^+$, found: [M+H]$^+$ 201.1385.

1-cyclohexyl-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (from 1 and 16)

Tryptamine hydrochloride (1 M) was added to cyclohexyl aldehyde **16** (1 M) in aqueous maleic acid buffer (10 mM, pH 2.0) and the mixture was incubated at 60 °C for 20 h. The precipitate was recovered, washed with water, and dried: ^{1}H NMR (MeOD): δ 7.45 (1H, d, J = 7.9), 7.38 (1H, d, J = 8.2), 7.14 (1H, ddd, J = 1.1, 7.2, 8.2), 7.04 (1H, ddd, J = 1.0,

144

7.2, 8.0), 4.60-4.57 (1H, m), 3.71 (1H, ddd, J = 3.2, 5.5, 12.4), 3.38 (1H, ddd, J = 5.4,

10.1, 12.5), 3.10 (1H, dddd, J = 2.0, 5.6, 10.1, 15.8), 3.04-2.95 (1H, m), 2.35-2.22 (1H,

m), 1.98-1.68 (4H, m), 1.57-1.10 (6H, m); ^{13}C NMR (MeOD): δ 137.64, 128.27, 126.71,

122.76, 119.84, 118.31, 111.65, 107.40, 59.19, 43.15, 40.32, 30.12, 27.07, 26.82, 26.43,

26.41, 18.70; ESI-MS($C_{17}H_{23}N_2^+$) m/z calculated: 255.1856 [M+H]$^+$, found: 255.1858

[M+H]$^+$.

4.4.3 Cloning of *opsts* into pET-28a(+) and expression and purification of OpSTS and CrSTS

The *opsts* gene (GI: 13928597) was codon optimized for expression in *E. coli*. The

nucleotides encoding the first 25 amino acids in the published peptide sequence (GI:

13928598) were omitted to remove a putative signal peptide (SignalP 3.0).[25,26] Upstream

Nco I and downstream *Xho* I restriction sites (underlined), a sequence that encodes for a

hexa-histidine tag (yellow), and a stop codon (red) were added to codon optimized *opsts*

to give the following nucleotide sequence:

CCATGGGCTCTCCTGAGTTTTTCGAATTTATTGAAGCACCGTCTTATGGTCCAAATGCGTATGCGTTCGACAGCGACGGCGAGTTGTATGC
GAGCGTGGAAGACGGTCGTATTATCAAGTACGACAAGCCTTCTAACAAATTCCTGACTCATGCTGTTGCCAGCCCGATCTGGAACAATGC
CCTGTGTGAGAATAATACCAACCAAGACCTGAAGCCGCTGTGCGGTCGCGTCTACGACTTTGGTTTTCATTATGAAACGCAGCGCCTGTA
CATTGCAGATTGCTACTTCGGCTTGGGCTTTGTTGGTCCGGACGGCGGTCACGCGATTCAACTGGCAACCTCCGGTGATGGCGTTGAGTTC
AAGTGGCTGTACGCGTTGGCGATCGACCAACAGGCAGGCTTCGTCTACGTGACGGACGTTTCTACTAAGTACGATGATCGTGGTGTTCAG
GACATTATTCGCATTAATGATACCACGGGTCGCCTGATTAAGTATGACCCTTCGACCGAAGAGGTGACCGTGCTGATGAAAGGCCTGAAT
ATTCCGGGCGGTACCGAGGTTAGCAAAGACGGTAGCTTTGTGCTGGTTGGTGAGTTCGCGTCGCATCGTATCCTGAAGTACTGGCTGAAG
GGTCCGAAGGCCAATACCAGCGAGTTTCTGCTGAAGGTGCGCGGTCCAGGTAATATCAAACGTACCAAAGATGGTGATTTCTGGGTTGCG
TCCAGCGATAACAACGGCATCACGGTGACGCCACGTGGTATCCGCTTCGATGAGTTTGGCAACATTCTGGAGGTCGTTGCTATTCCGCTG
CCGTATAAAGGTGAACATATCGAGCAGGTCCAAGAACACGACGGCGCCCTGTTCGTGGGTAGCCTGTTTCATGAGTTCGTCGGCATCCTG

CATAACTATAAGAGCAGCGTTGACCATCATCAGGAAAAGAACTCGGGTGGTCTGAACGCGAGCTTCAAGGAGTTCTCTTCGTTTGGATCT
CATCATCACCATCACCACTAGCTCGAG

The designed nucleotide sequence was synthesized and supplied as a sequenced construct in pJ201:12909 by DNA2.0 (Menlo Park, CA). Plasmid pJ201 (containing *opsts*) was subjected to a double restriction endonuclease reaction using *Nco* I and *Xho* I and the digestion products were ligated into pET-28a(+) (Novagen, EMD Biosciences, Merck KgaG, Darmstadt, Germany) using T4 DNA ligase. DNA sequencing was used to confirm the integrity of the resulting construct, pOpSTS. pOpSTS was transformed into *E. coli* BL21(DE3) for expression. pCrSTS was also available in *E. coli* BL21(DE3) (Chapter 3).[17] Purification of OpSTS and CrSTS were carried out as described previously.[17] The enzymes were buffer exchanged into sodium phosphate buffer (0.05 M, pH 7.0, 20% glycerol), analyzed by SDS-PAGE (Figure 4.8) and stored in 0.02 mL aliquots at -20 °C. The enzyme preparations tolerated repeated freeze-thaw cycles and are stable for at least one year without noticeable loss of activity. Crude extracts of *E. coli* BL21(DE3) that lacked plasmid and boiled OpSTS or CrSTS preparations did not catalyze the Pictet-Spengler reaction. Moreover, no chemical background reaction was detectable after two days of incubation of 2 mM tryptamine **1** and 2 mM secologanin **2** or aldehydes **10-20** in pH 7.0-buffer.
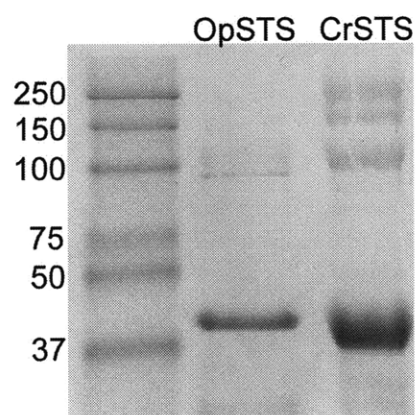
Figure 4.8 Purified, concentrated, and buffer exchanged OpSTS (1.2 mg mL^{-1}) and CrSTS (4.1 mg mL^{-1}) were loaded onto SDS-PAGE gels, separated, and stained using coomassie brilliant blue. The expected molecular weight of OpSTS and CrSTS is 37 kDa; the approximate loading was: OpSTS (9 μg) and CrSTS (30 μg).

4.4.4 Immobilization of OpSTS on CNBr-activated Sepharose 4B and enzymatic synthesis of tetrahydro-β-carbolines

CrSTS from partially purified plant tissue extract was previously immobilized on Sepharose 4B by the group of M. Zenk and used for the preparative synthesis of gram-quantities of strictosidine **3**.[27] OpSTS (2 U total activity, where 1 U = 1 μmol strictosidine formed per minute) was immobilized on cyanogen bromide (CNBr)-activated Sepharose 4B (1.4 mL) according to the manufacturer's instructions (71-7086-00 AD, GE Healthcare, Uppsala, Sweden). Immobilized OpSTS was active, where 1 U total activity was immobilized on the resin (51% recovered yield). Immobilized OpSTS could be reused at least three times with no noticeable loss of activity, similar to Zenk's previous observations with CrSTS.[27]

Enzyme-catalyzed reactions (total volume: 4 mL) contained sodium phosphate buffer (0.025 M, pH 7.0), tryptamine **1** (1 mM), aldehyde analog (5 mM), and immobilized OpSTS (0.02 U, based on strictosidine-forming activity). After incubation for four days at 100 rpm and 30 °C, the Sepharose beads were settled and the supernatant was extracted three times with DCM and once with EtOAc. The combined organic layers were evaporated *in vacuo* and the residue was dissolved in 2 mL of 1:1 MeCN-water. The tetrahydro-β-carboline product was isolated by semi-preparative HPLC using a solvent gradient of 30-100% MeCN in water with 0.1% TFA over 20 min. The collected fractions were concentrated *in vacuo* and the residue was separated by chiral HPLC (section 4.4.5).

n-Pentyl-tetrahydro-β-carboline, from substrates **1** and **14**, was isolated from a scaled up reaction (mg-scale) using immobilized OpSTS.

4.4.5 Determination of enantiomeric excess (*ee*)

Purified enzymatic products, obtained using enzymatic synthesis with immobilized OpSTS, were dissolved in isopropanol and injected on a chiral HPLC column. For all tetrahydro-β-carbolines except the CH_2THP analog (derived from **10**), a solvent composition of 10% *i*-PrOH in hexanes for 10 min, followed by a gradient of 10-30% *i*-PrOH in hexanes over 40 min was used. For the product of tryptamine **1** and aldehyde **10**, a solvent composition of 20% *i*-PrOH in hexanes for 10 min followed by a gradient of 20-25% *i*-PrOH in hexanes over 40 min was used. The authentic standards prepared as described in section 4.4.2 were run in parallel to enzymatic reactions. Peaks were detected by UV light at 230 nm, close to the maximal absorption of the tetrahydro-β-carbolines products. The products of compounds **1** and **10-11**, **13-14** separated on Chiracel OD-H (Chiral Technologies, Inc., PA); 2.6 cm x 25 cm; with the following retention times: Chemical standards (tetrahydro-β-carboline C3-substituent): CH_2THP (E1: 17 min, 18 min, E2: 24 min, 26 min); CH_2Cy (E1: 19 min, E2: 23 min); benzyl (E1: 21 min, E2: 27 min); *n*-pentyl (E1: 17 min, E2: 23 min). Enzymatic products (tetrahydro-β-carboline C3-substituent): CH_2THP (E1: 16 min, 17 min; E2: not found), *ee* >98%; CH_2Cy (E1: not found, E2: 23 min), *ee* >98%; CH_2Cy (E1: not found, E2: 23 min), *ee* >98%; *n*-pentyl (E1: not found, E2: 22 min), *ee* >98%. Despite considerable efforts using different chiral HPLC columns and mobile phases, the enantiomers of the

N-Ac-3-n-pentyl
(from 1 and 14)

i *OpSTS-derived*

ii chemical standard (*ref. 3*)
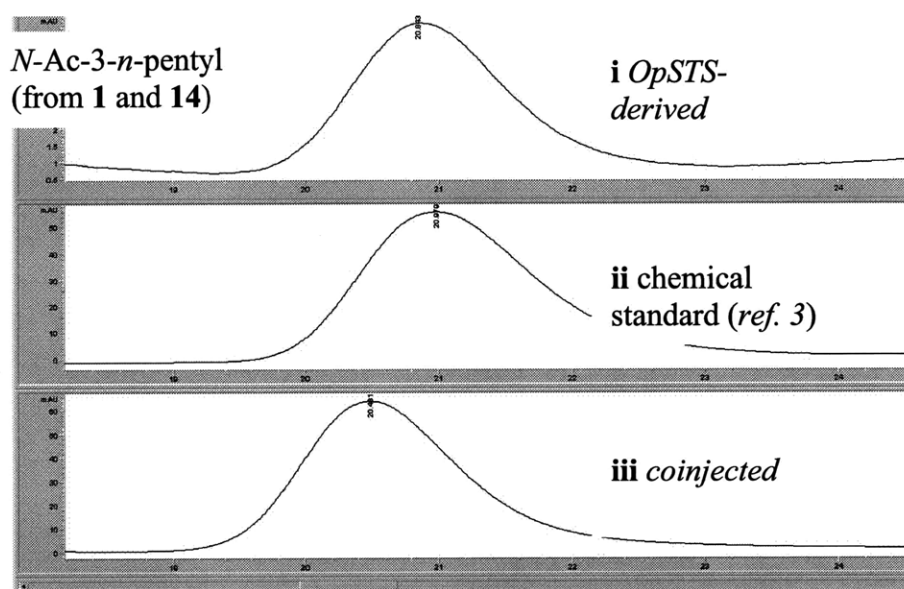
iii *coinjected*

Figure 4.9 Chiral HPLC trace for determination of the absolute configuration of the enzymatic products. *i.* *N*-acetyl derivatized OpSTS product from **1** and **14**. *ii.* Chiral authentic standard prepared according to reference 4. *iii.* Co-injection of *i* and *ii*.

secologanin **2**

how **2** is connected to QM-
calculated 6-exo-trig products

Docked intermediates
30
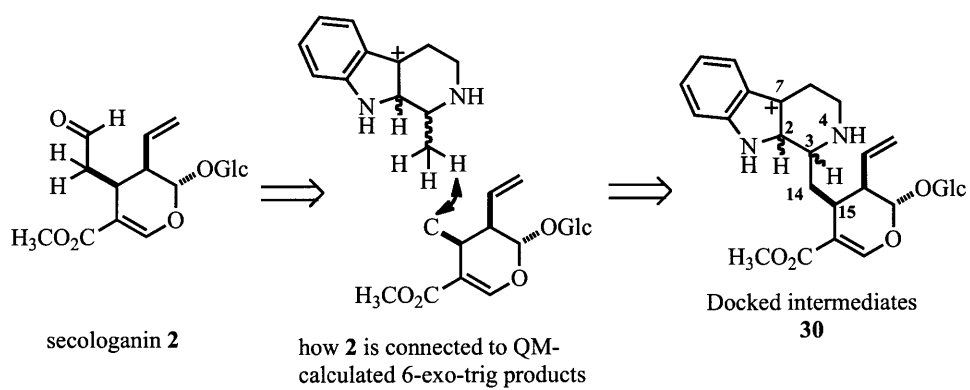
Figure 4.10 Models for **30** were generated based on the orientation of secologanin **2** in the active site of RsSTS and quantum mechanically derived structures of 6-exo-trig products.

tetarhydro-β-carboline product resulting tryptamine **1** and isobutanal **12** could not be separated. The chiral HPLC traces can be found in Appendix B.

4.4.6 Determination of absolute configuration

n-Pentyl tetrahydro-β-carboline (derived from **14**), prepared from an enzymatic reaction using immobilized-OpSTS (section 4.4.4), was dissolved in DCM and *N*-acetylated with acetic anhydride (1.0 eq) in the presence of TEA (1.0 eq) at -80 °C. The reaction was monitored by LC-MS, and upon completion, the reaction was quenched by MeOH and concentrated *in vacuo*. The residue was dissolved in 2:1 hexanes/EtOAc and purified by flash column chromatography (SiO$_2$, 2:1, hexanes/EtOAc) yielding a white solid. The purified *N*-acetylated enzymatic product was compared to an authentic standard of the 3*(S)* enantiomer of *N*-acetylated *n*-pentyl-tetrahydro-β-carboline (*ee* = 95%), prepared as described in ref. 3 (Figure 4.9). Since only one enantiomer formed in the enzymatic reaction, and since this enantiomer coeluted with the authentic 3*(S)* standard, a 3*(S)* configuration was assigned to the enzymatic product. It is assumed that the absolute configuration of the other enzymatic products is also 3*(S)*.

4.4.7 Enzyme assays and kinetics

Aldehydes 10-20

A stock (1 M) in DMSO of each aldehyde analog was carefully prepared and then diluted with water (final concentration of 10 mM) to minimize DMSO in the enzyme assay

(0.1% v/v solvent in each assay). Enzyme assays (0.2 mL reaction volume) contained tryptamine **1** (1 mM) and aldehyde analog **10-20** (1 mM) in sodium phosphate buffer (0.05 M, pH 7.0) and reactions were initiated by the addition of OpSTS (0.2 mol% relative to the substrate concentration) or CrSTS (1.0 mol%). The reactions were quenched by injection onto the HPLC column (27 µL). The area under the product peak was converted to concentration via a standard curve obtained by injecting known amounts of authentic standard (section 4.4.2) onto the HPLC. The reaction rates were normalized by the enzyme concentration in the reaction and reported as k_{obs} [min^{-1}].

Secologanin 2

Pseudo steady-state kinetics for OpSTS and CrSTS with secologanin **2** was carried out by fixing the concentration of tryptamine **1** (1 mM) while varying the concentration of **2**. Each reaction (0.2 mL) contained sodium phosphate buffer (0.025 M, pH 7.0), internal standard (NAA, 0.075 mM), tryptamine **1** (1 mM), and secologanin **2** (10, 20, 40, 100, 250 or 1000 µM), equilibrated to 30 °C. The reaction was initiated by adding either OpSTS (1.6 nM) or CrSTS (0.90 nM). At time points (10, 20, 30, and 40 min), an aliquot (0.05 mL) was removed from the reaction and quenched in sodium hydroxide (4 N, 1 µL, pH >12). This quench forms strictosamide quantitatively within minutes.[33] The quenched samples were stored at 10 °C in the HPLC autosampler tray until injected (0.027 mL) on the HPLC column. The chromatogram was integrated and the area under the peaks corresponding to strictosamide and NAA were noted. The relative area was compared to a standard curve where known amounts of strictosamide had been injected

under conditions identical to the kinetic assays. The kinetic data were collected three times for each concentration of secologanin **2** and the average initial rate (<15% conversion) was corrected for enzyme concentration and plotted against the concentration of **2**. The data were fit by non-linear regression analysis according to the Michaelis-Menten equation to obtain V_{max} (U mg^{-1}), K_M (µM), and V_{max}/K_M (U mg^{-1} mM^{-1}).

Tryptamine analogs

Pseudo steady-state kinetic analysis was performed by varying the concentration of tryptamine analog and keeping the secologanin **2** concentration at 1 mM ($K_{M,2}$ = 0.021 mM). Each reaction (0.1 mL) contained sodium phosphate buffer (0.125 M, pH 7.0), internal standard (NAA, 0.075 mM), secologanin **2** (1 mM), and tryptamine analog (0.025, 0.050, 0.10, 0.25, and 0.50 mM for 5-methoxytryptamine **24**; 0.010, 0.025, 0.050, 0.10, and 0.50 mM for 2'*(R)*-tryptophanol **25**), equilibrated to 30 °C. The reaction was initiated by adding OpSTS (1.4 µM). At time points (15, 30, 45, and 60 min), an aliquot (0.025 mL) was removed from the reaction and added to 0.1 mL MeOH to precipitate salts and protein. The quenched samples were stored at 10 °C in the HPLC autosampler tray until injected (0.027 mL) on the HPLC column. The chromatogram was integrated and the area under the peaks corresponding to the strictosidine analog and NAA were noted. The relative area was compared to a standard curve where known amounts of strictosidine **3** had been injected under conditions identical to the kinetic assays. The kinetic data were collected three times for each concentration of tryptamine analog **24** or **25**, and the average initial rate (<15% conversion) was corrected for enzyme

154

concentration and plotted against the concentration of **24** or **25**. The data were fit by non-linear regression analysis according to the Michaelis-Menten equation to obtain V_{max} (U mg^{-1}), K_M (μM), and V_{max}/K_M (U mg^{-1} mM^{-1}).

4.4.8 Site-directed mutagenesis

Site-directed mutagenesis was carried out using the complementary primer method using Stratagene's Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA). The primers listed in Table 4.4 were used.

Table 4.4 Primers used for site-directed mutagenesis.

Primer name	Primer sequence (5'->3')
del-fwd[a]	G T T T C C T C A T C T G A A G G T C A A C A C G G T C G T G
del-rev[a]	C A C G A C C G T G T T G A C C T T C A G A T G A G G A A A C
His283Gly-fwd	G G A C G G T G G T C A A G G C G G T C G T G T T G T T T C
His283Gly-rev	G A A A C A A C A C G A C C G C C T T G A C C A C C G T C C
His283Leu-fwd	G G A C G G T G G T C A A G T G G G T C G T G T T G T T T C
His283Leu-rev	G G A C G G T G G T C A A C A C G G T C G T G T T G T T T C
His283Phe-fwd	G G A C G G T G G T C A A T T C G G T C G T G T T G T T T C
His283Phe-rev	G G A C G G T G G T C A A G A A G G T C G T G T T G T T T C

[a] del = deletion of residues 277-280 in the *C. roseus* sequence.

The following CrSTS sequence was used as template to construct site-directed mutants:

TCACCAATCTTGAAGAAGATCTTCATCGAGTCTCCATCTTACGCTCCAAACGCTTTCACTTTCGACTCTAC

TGACAAGGGATTCTACACTTCCGTTCAGGACGGTAGAGTTATCAAGTACGAGGGTCCAAACTCTGGTTTC

ACTGACTTCGCTTACGCTTCTCCATTTTGGAACAAGGCTTTCTGTGAGAACTCTACTGACCCAGAGAAGA

GACCATTGTGTGGTAGAACTTACGACATCTCCTACGACTACAAGAACTCCCAGATGTACATTGTTGACGG

TCACTACCACTTGTGTGTTGTTGGTAAGGAAGGAGGTTACGCTACTCAATTGGCTACATCCGTTCAGGGT

GTTCCATTCAAGTGGTTGTACGCTGTTACTGTTGACCAGAGAACTGGTATCGTTTACTTCACTGACGTTTC

CTCTATTCACGACGATTCTCCAGAAGGTGTTGAGGAGATCATGAACACTTCCGACAGAACTGGTAGATTG

ATGAAGTACGACCCATCCACTAAGGAGACTACTTTGTTGTTGAAGGAGTTGCACGTTCCAGGTGGTGCTG

AGATTTCTGCTGACGGTTCCTTCGTTGTTGTTGCTGAGTTCTTGTCCAACAGAATCGTTAAGTACTGGTTG

155

GAGGGTCCAAAGAAGGGTTCTGCTGAGTTCTTGGTTACTATCCCAAACCCAGGTAACATCAAGAGAAACT

CCGATGGTCACTTCTGGGTTTCCTCATCTGAAGAATTGGACGGTGGTCAACACGGTCGTGTTGTTTCCAGA

GGTATCAAGTTCGACGGTTTCGGTAACATCTTGCAGGTTATCCCATTGCCACCACCATACGAAGGTGAAC

ACTTCGAGCAAATTCAAGAGCACGACGGATTGTTGTACATTGGTTCCTTGTTTCACTCCTCCGTTGGTATC

TTGGTTTACGACGACCACGATAACAAGGGAAACTCCTACGTTTCCTCC

4.4.9 Molecular modeling

Molecular mechanics was performed using Sybyl Molecular Expert Modeling Software (Version 7.3 including the FlexiDock application, Tripos, Inc.). Energy minimizations were carried out until the energy of the system remained constant, using a dielectric constant scaling factor of 4, a non-bonded cutoff at 8 Å, and a 1-4 scaling factor of 0.5. The steepest descent algorithm was used in the beginning of each energy minimization, followed by the Powell algorithm. The x-ray structure of RsSTS in complex with secologanin 2 (PDB ID: 2FPC)[22] was first edited to remove the B monomer. The termini were capped, hydrogens were added in random orientations and His307 was defined as protonated. Kollman All Atom types were assigned with Pullman charges for the biopolymer and ligands. Geometry minimization of only the enzyme hydrogens and water molecules was followed by minimization including the side-chains and lastly, the main-chain. The free carbon atom on C14 (obtained after deletion of acetaldehyde and two protons on secologanin 2 in the minimized crystal structure) was joined to one of the methyl hydrogen atoms of the small-molecule QM-derived structures (Figure 4.10), which had been loaded into the molecular modeling software. The atoms were defined with Kollman types. N4 and C7 were assigned a formal positive charge of unity that was distributed as partial charges according to Pullman.

4.5 References

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

2.   Dewick, P.M. *Medicinal natural products: a biosynthetic approach*, 2nd ed. Wiley **2003**.

3.   Taylor, M.S.; Jacobsen, E.N. "Highly enantioselective catalytic acyl-Pictet–Spengler reactions", *J. Am. Chem. Soc.* **2004**, *126*, 10558-10559.

4.   Seayad, J.; Seayad, A.M.; List, B. "Catalytic asymmetric Pictet-Spengler reaction", *J. Am. Chem. Soc.* **2006**, *128*, 1086-1087.

5.   Raheem, I.T.; Thiara, P.S.; Peterson, E.A.; Jacobsen, E.N. "Enantioselective Pictet–Spengler-type cyclizations of hydroxylactams: H-bond donor catalysis by anion binding", *J. Am. Chem. Soc.* **2007**, *129*, 13404-13405.

6.   Wanner, M.J.; van der Haas, R.N.; de Cuba, K.R.; van Maarseveen, J.H.; Hiemstra, H. "Catalytic asymmetric Pictet-Spengler reactions via sulfenyliminium ions", *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 7485-7487.

7.   Sewgobind, N.V.; Wanner, M.J.; Ingemann, S.; de Gelder, R.; van Maarseveen, J.H.; Hiemstra, H. "Enantioselective BINOL-phosphoric acid catalyzed Pictet–Spengler reactions of N-benzyltryptamine", *J. Org. Chem.* **2008**, *73*, 6405-6408.

8. Klausen, R.S.; Jacobsen, E.N. "Weak Brønsted acid–thiourea co-catalysis: enantioselective, catalytic protio-Pictet–Spengler reactions", *Org. Lett.* **2009**, *11*, 887-890.

9. Bou-Hamdan, F.R.; Leighton, J.L. "Highly enantioselective Pictet-Spengler reactions with α-ketoamide-derived ketimines: access to an unusual class of quaternary α-amino amides", *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 2403-2406.

10. Faber, K. *Biotransformations in Organic Synthesis*, 5th ed. **2008**, Springer.

11. Bornscheuer, U.T.; Kazlauskas, R.J. *Hydrolases in Organic Synthesis*, 2nd ed. **2006**, Wiley~VCH.

12. O'Connor, S.E.; Maresh, J.J. "Chemistry and biology of monoterpene indole alkaloid biosynthesis", *Nat. Prod. Rep.* **2006**, *23*, 532-547.

13. Kutchan, T.M.; Hampp, N.; Lottspeich, F.; Beyreuther, K.; Zenk, M.H. "The cDNA clone for strictosidine synthase from *Rauvolfia serpentina* DNA sequence determination and expression in *Escherichia coli*", *FEBS Lett.* **1988**, *237*, 40-44.

14. McKnight, T.D.; Roessner, C.A.; Devagupta, R.; Scott, A.I.; Nessler, C.L. "Nucleotide sequence of a cDNA encoding the vacuolar protein strictosidine synthase from *Catharanthus roseus*", *Nucl. Acids Res.* **1990**, *18*, 4939.

15. McCoy, E.A.; Galan, M.C.; O'Connor, S.E. "Substrate specificity of strictosidine synthase", *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2475-2478.

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

17. Bernhardt, P.; McCoy, E.A.; O'Connor, S.E. "Rapid identification of enzyme variants for reengineered alkaloid biosynthesis in periwinkle", *Chem. Biol.* **2007**, *14*, 888-897.

18. Loris, E.A.; Panjikar, S.; Ruppert, M.; Barleben, L.; Unger, M.; Schübel, H.; Stöckigt, J. "Structure-based engineering of strictosidine synthase: auxiliary for alkaloid libraries", *Chem. Biol.* **2007** *14*, 979-85.

19. Chen, S.; Galan, M.C.; Coltharp, C.; O'Connor, S.E. "Redesign of a central enzyme in alkaloid biosynthesis", *Chem. Biol.* **2006**, *13*, 1137-1141.

20. Bernhardt, P.; Usera, A.R.; O'Connor, S.E. "Strictosidine synthase from *Ophiorrhiza pumila* is a stereoselective and promiscuous biocatalyst for the Pictet-Spengler reaction", *submitted* **2009**.

21. Yamazaki, Y.; Sudo, H.; Yamazaki, M.; Aimi, N.; Saito, K. "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*, 395-403.

22. Ma, X,; Panjikar, S.; Koepke, J.; Loris, E.; Stöckigt, J. "The structure of *Rauvolfia serpentina* strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins", *Plant Cell* **2006**, *23*, 532-547.

23. Protein Homology/analogY Recognition Enginer, www.sbg.bio.ic.ac.uk/phyre

24. Kelley, L.A.; Sternberg, M.J.E. "Protein structure prediction on the web: a case study using the Phyre server", *Nat. Protocols* **2009**, *4*, 363 - 371.

25. Bendtsen, J.D.; Nielsen, H.; von Heijne, G.; Brunak, S. "Improved prediction of signal peptides: SignalP 3.0", *J. Mol. Biol.* **2004**, *340*, 783-795

26. Nielsen, H.; Engelbrecht, J.; Brunak, S.; von Heijne, G. "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites", *Protein Eng.* **1997**, *10*, 1-6.

27. Pfitzner, U.; Zenk, M.H. "Isolation and immobilization of strictosidine synthase", *Methods Enzymology,* **1987**, *136*, 342-350.

28. Runguphan, W.; O'Connor, S.E. "Metabolic reprogramming of periwinkle plant culture", *Nat. Chem. Biol.* **2009**, *5*, 151-153.

29. Baldwin, J.E. "Rules for ring closure. Stereoelectronic control in the endocyclic alkylation of ketone enolates", *J. Chem. Soc., Chem. Commun.* **1977**, 233-235.

30. Maresh, J.J.; Giddings, L.A.; Friedrich, A.; Loris, E.A.; Panjikar, S.; Trout, B.L.; Stöckigt, J.; Peters, B.; O'Connor, S.E. "Strictosidine synthase: mechanism of a Pictet-Spengler catalyzing enzyme", *J. Am. Chem. Soc.* **2008**, *130*, 710-723.

31. Lopes, S.; von Poser, G.L.; Kerber, V.A.; Farias, F.M.; Konrath, E.L.; Moreno, P.; Sobral, M.E.; Zuanazzi, J.A.S.; Henriques, A.T. "Taxonomic significance of alkaloids and iridoid glucosides in the tribe Psychotrieae (rubiaceae)", *Biochem. Syst. Ecol.* **2004**, *32*, 1187-1195.

32. Galan, M.C.; O'Connor, S.E. "Semisynthesis of secologanin analogues", *Tetrahedron Lett.* **2006**, *47*, 1563-1565.

33. Hutchinson, C.R.; Heckendorf, A.H.; Straughn, J.L.; Daddona, P.E.; and Cane, D.E. "Biosynthesis of camptothecin. 3. Definition of strictosamide as the penultimate

biosynthetic precursor assisted by carbon-13 and deuterium NMR spectroscopy", *J.*

Am. Chem. Soc. **1979**, *101*, 3358-3369.

CHAPTER 5

SYNTHESIS AND BIOCHEMICAL EVALUATION OF DES-VINYL SECOLOGANIN

AGLYCONES WITH ALTERNATIVE STEREOCONFIGURATION

5.1 Introduction

Biogenetic compounds are polyfunctional molecules that occupy a central position in biosynthetic pathways and are, therefore, precursors for a multitude of natural products.[1] For example, the seco-iridoid β-D-glucoside, secologanin **1**, is the precursor for all monoterpene indole alkaloids (MIAs) in several medicinal plants, including *Catharanthus roseus*, *Ophiorrhiza pumila*, and *Rauvolfia serpentina*.[2] In these plants **1** reacts with tryptamine **2** to form strictosidine **3**; a reaction catalyzed by strictosidine synthase (STS).[2] Secologanin **1** is a densely functionalized molecule that contains three stereogenic centers on its dihydropyran core [2*(S)*3*(R)*4*(S)*], and only this stereoisomer of **1** occurs naturally (Figure 5.1).

Galan *et al.* modified secologanin **1** by semi-synthesis – the ester group by transesterification reactions and the vinyl group by olefin metathesis reactions.[3] STS does not accept secologanin analogs with olefin substitutions, but *C. roseus* cell cultures that contain STS convert secologanin analogs with alternate ester groups into MIA analogs.[4] Access to aglycone analogs and alternative stereosiomers of **1**, however, is not possible by semi-synthesis (Dr. C. Galan, unpublished). These analogs can be synthesized using a tandem Knoevenagel-hetero-Diels-Alder (KHDA) reaction where an electron-rich vinyl ether reacts with an *in situ*-formed diene.[1,5,6] This chapter describes the application of the KHDA domino reaction in the synthesis of a series of des-vinyl secologanin aglycone stereoisomers.[7] I show that STS accepts more than one stereoisomer,[7] and describe a synthetic route to des-vinyl strictosidine analogs with
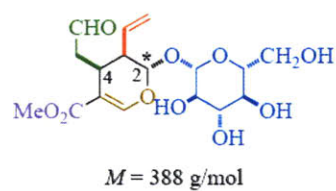
Functionality

alkene
aldehyde
ester
enol (masked aldehyde)
*acetal (masked aldehyde)
β-glucose

M = 388 g/mol

Stereochemistry

Three stereogenic centers on DHP:
    2-*(S)*, 3-*(R)*, 4-*(S)*

β-anomer

Figure 5.1 Secologanin **1** is a densely functionalized molecule with three stereogenic centers on its dihydropyran (DPH) core. The glycoside is linked in a β-anomer configuration.
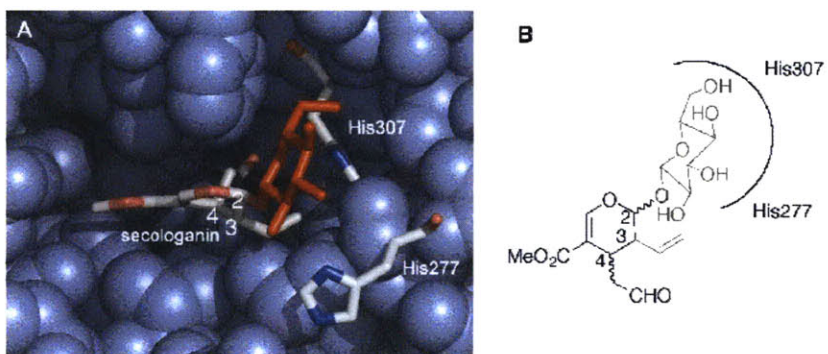
Figure 5.2 (A) *Rauvolfia serpentina* STS in complex with secologanin (PDB ID: 2FPC). The glucose moiety (red) hydrogen bonds with two histidine residues (His277 and His307). (B) Without the glucose group (gray) the secologanin binding pocket could potentially accommodate substrates with different geometries.

photo-labile acetal protective groups. These orthogonally protected molecules, which are stable to deglucosylation by strictosidine glucosidase (SGD), are potentially useful in the temporal control of alkaloid biosynthesis.

5.2 Results and discussion

The x-ray crystal structure of *R. serpentina* STS in complex with secologanin **1** (PDB ID: 2FPC) suggests that the glucose moiety forms hydrogen bonds with His277 and His307 (Figure 5.2).[8] I hypothesized that the glucose moiety plays a key role in correctly orienting secologanin and, therefore, impedes catalysis using alternative secologanin stereoisomers; removing or replacing the glucose moiety would leave sufficient space in the secologanin binding site to accommodate the alternative stereoisomers.

5.2.1 Synthesis of KHDA substrates

Tietze's tandem Knoevenagel-hetero-Diels-Alder reaction (KHDA) gives 2,4-disubstituted dihydropyrans from three accessible substrates: an enol, an aldehyde, and a vinyl ether.[1,5,6] Tietze *et al.* describe the synthesis of enol **5** in two steps (Figure 5.3, reactions a-b) (section 5.4.3).[9] The first step is a conjugate addition reaction between ethyl vinyl ether **10** and trichloroacetyl chloride **4**. The second step is the reaction of the α,β−unsaturated carbonyl intermediate with formic acid to give **4**. ^{1}H NMR revealed that **4** exists as the *cis*-enol isomer (enol/aldehyde, 96:4, only *cis*).

The aldehyde substrate of the KHDA reaction is a monoprotected malondialdehyde, since the prochiral aldehyde atom of secologanin **1** must be protected until the final step in the synthesis. Previous reports use a cyclic acetal protective group, which cannot be removed selectively in the presence of another acid-labile acetal (C2). I envisioned that an acyclic acetal, which permits the use of milder deprotection conditions, would leave the C2-acetal intact. This turned out to be true and KHDA reactions that proceed at room temperature employ diethoxyacetal **8**. Acetal **8**, however, decomposes upon heating and KDHA reactions that require heating therefore use the more stable dioxane aldehyde **9**, although the dioxane protective group is difficult to remove, leading to low yields in the final deprotection step (section 5.2.3).

Several protocols, based on non-selective deprotection-reprotection sequences, exist for the synthesis of monoprotected malondialdehydes. These protocols require two arduous purification steps resulting in low yields (typically, 30-50%). I therefore sought a more convenient source of dialdehyde equivalent. Dess-Martin periodinane (DMP)[11-15] triggered the formation of aldehydes **8** and **9** via oxidation of the commercially available alcohols: 3,3-diethoxy-1-propanol **6** and 2-(5,5,-dimethyl-1,3-dioxan-2-yl)ethanol **7**, respectively, in high yields (>90%) (Figure 5.3, reactions c and d).

Vinyl ethers with ethyl, *i*-butyl (*i*-Bu), *t*-butyl (*t*-Bu), and cyclohexyl (Cy) substituents **10-13** (R group in Figure 5.3) are commercially available; *o*-nitrobenzyl vinyl ether **16**
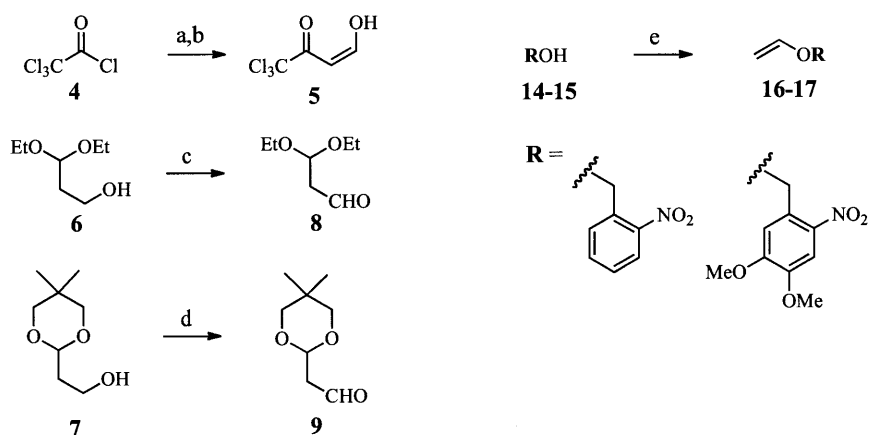
Figure 5.3 Synthesis of starting materials for the KHDA reaction: (a) ethyl vinyl ether **10** (1.5 eq), pyridine (1 eq); (b) formic acid, 88% over 2 steps; (c) DMP (1.1 eq), DCM, 90%; (d) DMP (1.2 eq), DCM, 97%; (e) vinyl acetate (2 eq), [IrCl(cod)]₂ (2 mol-%), NaCO₃ (0.6 eq) or NaOAc (1.2 eq), toluene, 56-80%.

and 4,5-dimethoxy-2-nitrobenzyl vinyl ether **17** are synthesized in the dark from the corresponding alcohols **14-15** and vinyl acetate in a vinyl transfer reaction catalyzed by the air-stable dimeric Ir(I)-complex, [Ir(cod)Cl]$_2$ (Figure 5.3, reaction e) (section 5.4.3).[16]

5.2.2 Synthesis of des-vinyl secologanin analogs

The KHDA reaction between chloroacetyl **5**, aldehyde **8**, and vinyl ethers **10-13**, in the presence of potassium fluoride, resulted in the formation of a mixture of cyclodduct stereoisomers (Figure 5.4A, section 5.4.3). After a brief silica gel chromatography step that partially purified the desired regioisomers,[5] methanolysis in the presence of 1,8-diazabicycloundec-7-ene (DBU) lead to the formation of methyl esters **18-21**. Finally, mild hydrolysis of the acyclic acetal of **18-21** in aqueous oxalic acid and silica gel[17] afforded aldehydes **22-25**. The two sets of enantiomers separate during silica gel chromatography; the 2,4-*trans* and 2,4-*cis* configurations are distinguished by the splitting of the acetal proton signal ~ 5 ppm: *cis*: (t, J = 2-3), *trans*: (dd, J = 2-3, 7-8).[5]

O-nitrobenzyl ethers can be deprotected with UV-light (350 nm) to generate nitrosobenzaldehyde along with the deprotected alcohol.[18] Thus, substitution of the glucose moiety of secologanin **1** for an *o*-nitrobenzyl group generates an unstable hemi-acetal upon irradiation at 350 nm. This photodeprotection reaction is analogous, but orthogonal, to the SGD-catalyzed deglucosylation reaction, since only the synthetic strictosidine analog contains a photocleavable group.

A

Compound	R =	
1	β-D-Glc	3-vinyl
10, 18, 22	Et	
11, 19, 23	*i*-Bu	des-vinyl
12, 20, 24	*t*-Bu	
13, 21, 25	Cy	

B
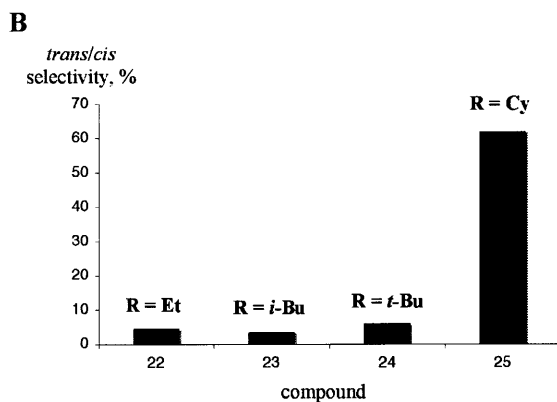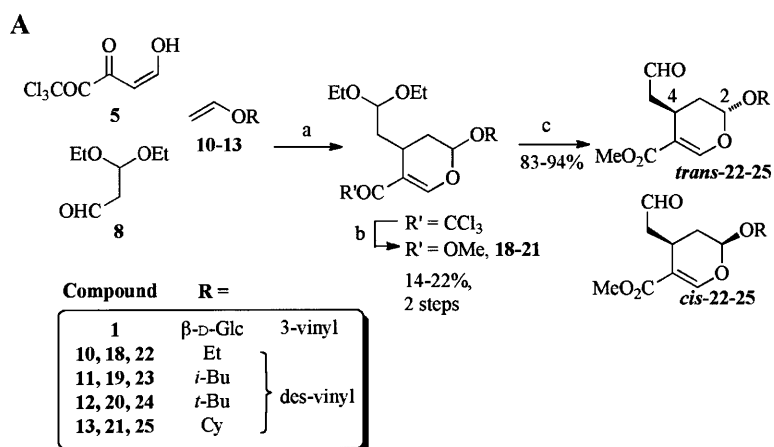
Figure 5.4 (A) Synthesis of des-vinyl secologanin *O*-analogs. (a) postassium fluoride (50 mg mmol⁻¹ **10-13**), DCM; (b) DBU (0.01 eq), MeOH, 14-22% over 2 steps; (c) oxalic acid and silica gel (each 10% w/v), THF/water (4:1), 83-94%. (B) Diastereoselectivity of *C. roseus* STS for **22-25**.
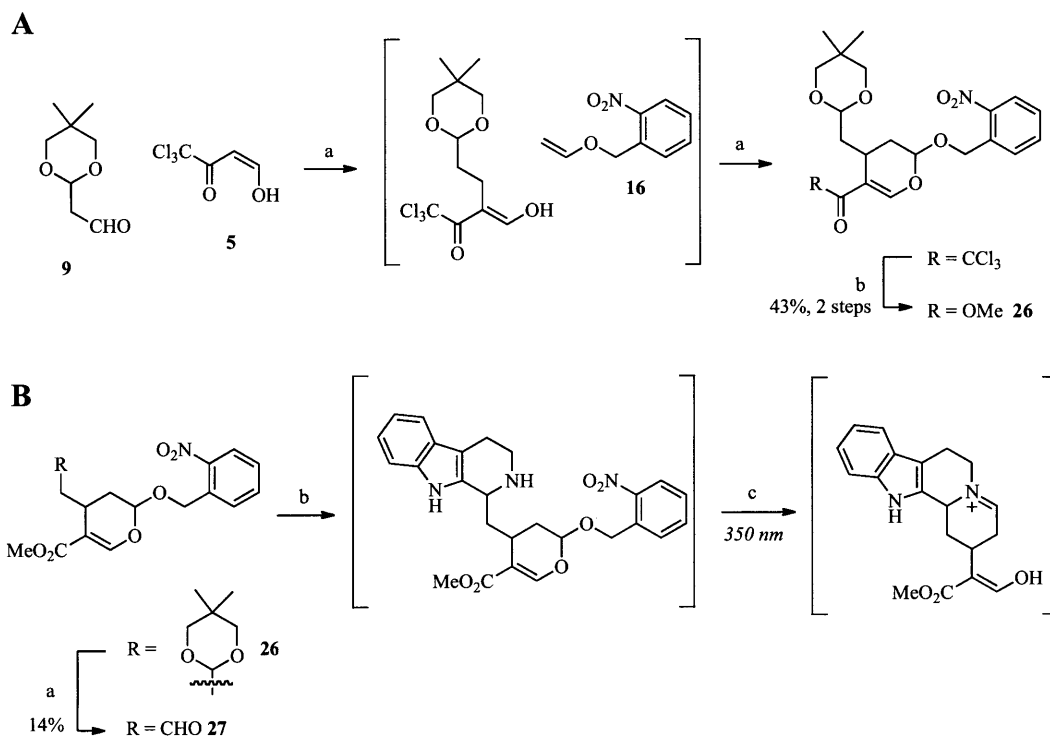
Figure 5.5 Synthesis and application of *o*-nitrobenzyl protected des-vinyl secologanin *O*-analog 27. (A) Synthesis of dioxane protected methyl ester 26. (a) postassium fluoride (50 mg mmol^{-1} 16), toluene, Δ; (b) DBU (0.01 eq), MeOH, 43% over 2 steps. (B) Synthesis and deprotection of a tetrahydro-β-carboline derived from 27. (a) aq. HCl in THF/water (4:1), Δ, 14%; (b) tryptamine 2, TFA (cat.), DCM; (c) irradiation at 350 nm (Appendix C).

The heated KHDA reaction (performed in the dark, to prevent unwanted photodecomposition) between chloroacetyl 5, aldehyde 9, and o-nitrobenzyl vinyl ether 16, followed by methanolysis, prepared dioxane-protected methyl ester 26 in two steps with 43% overall yield (Figure 5.5A). After considerable optimization, I found deprotection conditions that produced aldehyde 27 in low yield (14%), albeit with good recovery of the starting material (Figure 5.5A and section 5.4.3).

5.2.3 Biochemical characterization of des-vinyl secologanin analogs

There were no previous reports on the characterization of des-vinyl secologanin analogs as substrates for STS. I therefore determined the activity of STS with each diastereomer of aldehydes 22-25 by monitoring the appearance of tetrahydro-β-carboline products 30-33 in the presence of tryptamine 2 using LC-MS (Figure 5.4B, section 5.4.2). STS catalyzed the Pictet-Spengler reaction with all aldehyde substrates: the 2,4-*trans* configuration of secologanin 1 and the unnatural 2,4-*cis* configuration (Appendix C). For smaller aldehydes 22-24, the *trans/cis* selectivity is low (*de* = 5-10%), whereas for 25 the selectivity for the *trans* configuration is significantly higher (*de* >60%) (Figure 5.4B). Consequently, increased acetal substituent size correlates with a preference by STS for the *trans* configuration – the relative stereochemistry of the natural substrate, secologanin 1.

Enzyme assays with photolabile *o*-nitrobenzyl ether analogs revealed that they are not accepted by STS, since no product forms when **27** is incubated when tryptamine **2** in the presence of the enzyme. Therefore, the secologanin binding site is likely not compatible with the larger aromatic *O*-substituent. Nevertheless, I used a chemical Pictet-Spengler reaction between tryptamine **2** and **27** in the presence of TFA in DCM (Figure 5.5B). Under these conditions the tetrahydro-β-carboline product formed with good conversion (Appendix C). Subsequent irradiation by UV-light at 350 nm led to the appearance of the expected deprotected compound (Figure 5.5B and Appendix C), though one of the next enzymes in the MIA pathway – a partially purified reductase (N. Yerkes, unpublished) – showed no activity under the deprotection conditions.

5.3 Conclusions

This chapter describes the total synthesis and biochemical evaluation of a series of des-vinyl secologanin analogs (**22-25** and **27**). STS accepts unnatural stereoisomers and the stereochemical flexibility of STS is higher when the acetal substituent is smaller. There is, thus, a trade-off between higher activity and relaxed stereoselectivity. For example, STS from the camptothecin-producing plant, *O. pumila*, catalyzes the Pictet-Spengler reaction between tryptamine **2** and small achiral aldehydes, but with rates ~1000-fold lower than when using secologanin **1**.[19] Des-vinyl secologanin derivatives are all better substrates for STS than the achiral aldehydes, though **22-25** are, qualitatively, still less accepted than **1**. Because STS prefers secologanin **1**, the secologanin binding site of STS requires larger aldehydes for optimal activity, preferably aldehydes that contain the
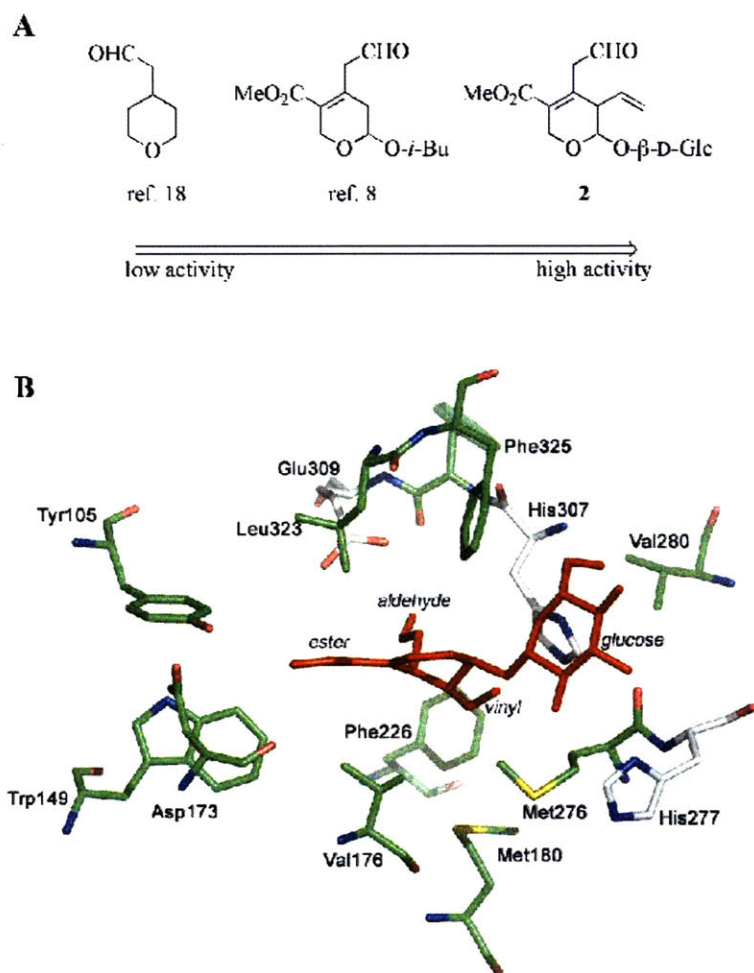
Figure 5.6 (A) Aldehydes of increasing size and hydrophilicity are preferred by STS. (B) The glucose and seco-iridoid moieties form several interactions with the secologanin binding site. By decreasing size of the aldehyde substrate, the binding site accommodates alternate aldehyde stereoisomers, albeit at the cost of reduced activity.

174
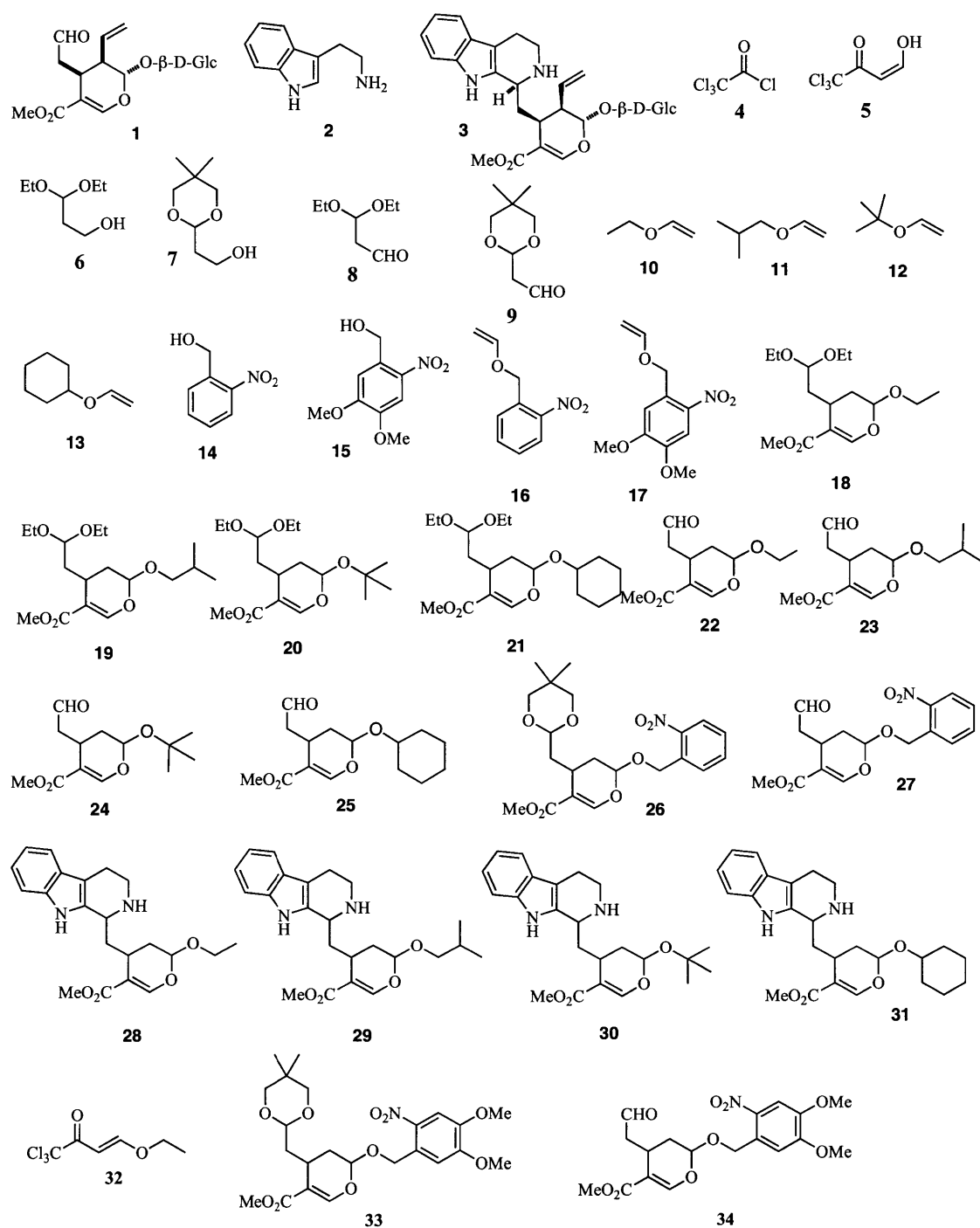
Figure 5.7 Structures of compounds mentioned in this chapter.

glucose moiety (Figure 5.6A).[20] The glucose moiety, however, restricts the ability of STS to form unnatural stereoisomers (Chapter 6) and removing the glucose moiety allows the turnover of alternative stereoisomers (Figure 5.6B). Mutations targeting the substrate binding sites of STS modulate substrate specificity,[21,22] and mutations that decrease the volume or change the electrostatic properties of the secologanin binding site could potentially increase the activity of STS with smaller, more hydrophobic, analogs of **1**.

The introduction of an orthogonal acetal protective group to strictosidine **3** avoids the deglucosylation process and enables the selective cleavage of the unnatural *o*-nitrobenzyl ether linkage using UV light. This strategy may be useful in the temporal control of pathway induction. However, installing the vinyl group is likely key for practical use, since that would allow access to the full profile of monoterpene indole alkaloids.

5.4 Experimental methods

5.4.1 General methods and analytical techniques

All chemicals were used as received from the supplier. All glassware were oven dried, evacuated, and filled with argon. STS was expressed in *E. coli* BL21(DE3) and purified by affinity chromatography using a hexa-His tag as previously described.[22]

UPLC analysis was performed using an Acquity Ultra Performance BEH C18 column with a 1.7 μm particle size, 2.1 x 100 mm dimension, and a flow rate of 0.6 mL min^{-1}. The column elution was coupled to MS analysis carried out using a Micromass LCT

Premier TOF Mass Spectrometer with an ESI source. Both modules are from Waters Corporation (Milford, MA). The capillary and sample cone voltages were 3000 V and 30 V, respectively. The desolvation and source temperatures were 300 and 100 °C, respectively. The cone and desolvation gas flow rates were 60 and 800 L h^{-1}. Analysis was performed with MassLynx 4.1 and integrations were carried out using QuantLynx. ^{1}H NMR and ^{13}C NMR spectra were recorded on Varian 300 MHz, Varian 500 MHz, or Brucker 400 MHz spectrometers.

5.4.2 Diastereoselectivity of STS

Secologanin aglycone analogs **22-25** (1 mM) were incubated with tryptamine **2** (0.5 mM) in the presence of *C. roseus* STS (0.5 μM) in sodium phosphate buffer (0.025 M, pH 7.0) at 30 °C. Control reactions containing enzyme storage buffer were treated the same way and were used for comparison. The reaction mixtures were quenched by addition of HPLC-grade MeOH containing yohimbine (internal standard) and then analyzed by UPLC-MS using the following method: 25% 5 min, 25-90% 3 min, 90-25% 1 min, 25% 1 min; flow-rate: 0.6 mL min^{-1}; MS-method: *m/z* 200-700 detected as centroided masses in W-mode using a reference compound (reserpine). Exact masses: compound **28**: (*O*-ethyl), ESI-MS($C_{21}H_{27}N_2O_4^+$) *m/z* calculated: 371.1971 [M+H]$^+$, found: 371.1976 [M+H]$^+$; Compound **29**: (*O-i*-Bu), ESI-MS($C_{23}H_{31}N_2O_4^+$) *m/z* calculated: 399.2284 [M+H]$^+$, found: 399.2366 [M+H]$^+$; Compound **30**: (*O-t*-Bu), ESI-MS($C_{23}H_{31}N_2O_4^+$) *m/z* calculated: 399.2284 [M+H]$^+$, found: 399.2379 [M+H]$^+$; Compound **31**: (*O*-Cy), ESI-MS($C_{25}H_{33}N_2O_4^+$) *m/z* calculated: 425.2440 [M+H]$^+$, found: 425.2488 [M+H]$^+$. The

initial rate was monitored by integrating the product peaks and normalizing the areas to the internal standard peak (yohimbine). Relative rates representing the diastereoselectivity of STS are reported as the average of three experiments.

5.4.3 Chemical synthesis of des-vinyl secologanin analogs

1,1,1-trichloro-4-ethoxybut-3-en-2-one **32**
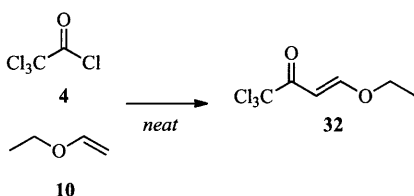
Ethyl vinyl ether **10** (33 mL, 75 mmol, 1.5 eq) was added to trichloroacetyl chloride **4** (28 mL, 50 mmol, 1.0 eq) under argon at 0 °C. The reaction was allowed to warm to room temperature and was stirred for 17 h. Pyridine (20 mL, 55 mmol, 1.1 eq) was then added dropwise under vigorous stirring at 0 °C resulting in a color change from yellow to dark red, accompanied by gas evolution and the formation of a precipitate. The reaction was allowed to warm to room temperature until TLC (hex/ether, 5:1) indicated complete conversion (~1 h). The reaction was first diluted with 10 volumes of water and then extracted with ether. Filtration of emulsified organics through a combined celite – silica gel plug facilitated removal of salts. Evaporation of the solvent left a red oil that was purified by flash column chromatography (SiO$_2$, hex/EtOAc, 15:1) to afford **32** as a yellow oil. r$_f$ = 0.56 (hex/ether, 5:1); ^{1}H NMR (300 MHz, CDCl$_3$): δ 7.86 (dt, 1H, J = 0.34, 12.1), 6.15 (d, 1H, J = 12.1), 4.10 (dq, 2H, J = 0.34, 7.1), 1.40 (t, 3H, J = 7.1); ^{13}C NMR (125 MHz, CDCl$_3$): δ 181.38, 167.79, 97.04, 96.42, 68.98, 14.72.

1,1,1-trichloro-4-hydroxybut-3-en-2-one **5**
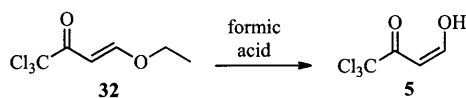
Formic acid (50 mL) was added to 1,1,1-trichloro-4-ethoxybut-3-en-2-one **32** (10 mmol) and the mixture was stirred at room temperature until TLC (hex/ether, 5:1) indicated complete conversion (24 h). The reaction was extracted with ether (5 x 60 mL) and the combined organic fractions were carefully evaporated under reduced pressure (20 °C), to afford a volatile yellow oil in analytical purity (8.4 g, 44 mmol, 88% over 2 steps). Compound **4** is not stable under silica gel chromatography conditions but is stable for several months when stored at -80 °C. ^{1}H NMR (300 MHz, CDCl$_3$): δ 12.44 (s, 1H), 7.61 (d, 1H, J = 5.7), 6.18 (d, 1H, J = 5.7); ^{13}C NMR (75 MHz, CDCl$_3$): δ 186.9, 171.6, 95.7.

3,3-diethoxypropanal **8**
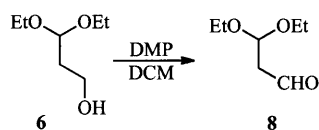
Dess-Martin periodinane (DMP, 4.8 g, 11 mmol, 1.1 eq) was added to a solution of 3,3-diethoxy-1-propanol **6** (1.5 g, 10 mmol, 1.0 eq) in DCM (100 mL) at 0 °C. The reaction was stirred at 0 °C for 5 min, warmed to room temperature and then stirred for an additional 45 min. The reaction was diluted with DCM (100 mL) and quenched by the successive addition of a saturated aqueous solution of sodium bicarbonate (50 mL) and a concentrated aqueous solution of sodium thiosulfate (50 mL) until the white precipitate dissolved. The layers were separated and the aqueous layer was extracted with DCM (3 x

20 mL). The combined organic layers were washed with a (1:1) mixture of sodium bicarbonate/sodium thiosulfate, water, and brine, and dried over anhydrous sodium sulfate. After evaporation of the solvent *in vacuo*, the residue was purified by column chromatography (SiO$_2$, hex/ether, 3:1) to afford 3,3-diethoxy-1-propanal **8** (1.3 g, 9.0 mmol, 90%). r$_f$ = 0.56 (hex/ether 1:3); ^{1}H NMR (300 MHz, CDCl$_3$): δ 9.72 (t, 1H, *J* = 2.3), 4.93 (t, 1H, *J* = 5.5), 3.65 (qd, 1H, *J* = 7.1, 9.3), 3.52 (qd, 1H, *J* = 7.1, 9.3), 2.70 (dd, 1H, *J* = 2.3, 5.5), 1.78 (t, 1H, *J* = 7.1); ^{13}C NMR (75 MHz, CDCl$_3$): δ 200.17, 98.85, 62.30, 48.22, 15.40.

2-(5,5-dimethyl-1,3-dioxan-2-yl)ethanal **9**
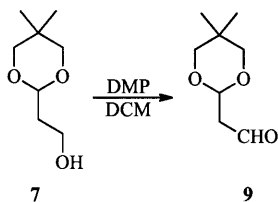
A reaction between DMP (5.1 g, 12 mmol, 1.2 eq) and 2-(5,5-dimethyl-1,3-dioxan-2-yl)ethanol **7** (1.6 g, 10 mmol, 1.0 eq) was treated as described for the synthesis of 3,3-diethoxy-1-propanal **8**. Purification by flash column chromtography (SiO$_2$, hex/ether 3:1) afforded aldehyde **9** (1.5 g, 9.7 mmol, 97%) as a clear oil. r$_f$ = 0.69 (hex/ether 1:3); ^{1}H NMR (300 MHz, CDCl$_3$): δ 9.81 (1H, t, *J* = 2.4), 4.86 (1H, t, *J* = 4.5), 3.61 (2H, d, *J* = 10.0), 3.45 (2H, d, *J* = 10.5), 2.68 (1H, dd, *J* = 2.4, 4.6), 1.17 (3H, s), 0.72 (3H, s); ESI-MS(C$_8$H$_{14}$O$_3$$^+$) *m/z* calculated: 159.1016 [M+H]$^+$, found: 159.1008 [M+H]$^+$

o-nitrobenzyl vinyl ether **16**

In a Schlenk flask under a stream of argon in the dark, [IrCl(cod)]₂ (0.067 g, 0.1 mmol, 0.01 eq) was added to a toluene solution (10 mL) of *o*-nitrobenzyl alcohol **14** (1.5 g, 10 mmol, 1.0 eq), sodium carbonate (0.64 g, 6.0 mmol, 0.6 eq) and vinyl acetate (1.7 g, 20 mmol, 2.0 eq). The mixture was heated to 100 °C and stirred at this temperature for 7 h. The reaction was then diluted with wet ether (20 mL) and filtered through SiO_2, eluting with ether. After evaporation of the solvents, the residue was purified by flash column chromatography (SiO_2, hex/DCM, 2:1) to afford vinyl ether **16** (1.0 g, 5.6 mmol, 56%). r_f = 0.57 (hex/DCM, 1:2); ^1H NMR (300 MHz, $CDCl_3$): δ 8.10 (dd, 1H, J = 1.3, 8.2), 7.69 (td, 1H, J = 0.4, 7.0), 7.64 (dt, 1H, J = 0.3, 7.9), 7.44 (dt, 1H, J = 0.7, 7.3), 5.86 (dd, 1H, J = 6.5, 14.0), 5.15 (s, 2H), 4.33 (dd, 1H, J = 2.5, 14.3), 4.13 (dd, 1H, J = 2.4, 6.8); ^{13}C NMR (75 MHz, $CDCl_3$): δ 151.19, 146.95, 134.05, 133.87, 128.60, 128.35, 124.992, 88.56, 66.93; ESI-MS($C_9H_9NO_3^+$) *m/z* calculated: 202.0475 [M+Na]⁺, found: 202.0482 [M+Na]⁺.

4,5-dimethoxy-2-nitrobenzyl vinyl ether **17**

In a Schlenk flask under a stream of argon in the dark, [IrCl(cod)]$_2$ (0.027 g, 0.04 mmol, 0.02 eq) was added to a toluene solution (5 mL) of 4,5-dimethoxy-2-nitrobenzyl alcohol 15 (0.42 g, 2.0 mmol, 1.0 eq), sodium carbonate (0.17 g, 1.6 mmol, 0.6 eq), and vinyl acetate (0.34 g, 4.0 mmol, 2.0 eq). The mixture was heated to 110 °C and stirred at this temperature for 3 h. The reaction was then diluted with wet ether (20 mL) and filtered through SiO$_2$ eluting with ether. After evaporation of the solvents, the residue was purified by flash column chromatography (SiO$_2$, hex/EtOAc, 5:1 to 2:1) to afford vinyl ether 17 (0.30 g, 1.2 mmol, 62%). ^{1}H NMR (300 MHz, CDCl$_3$): δ 7.73 (s, 1H), 7.22 (s, 1H), 6.57 (dd, 1H, J = 6.8, 14.3), 5.18 (s, 2H), 4.39 (dd, 1H, J = 2.4, 14.3), 4.16 (dd, 1H, J = 2.4, 6.8), 3.98 (s, 3H), 3.94 (s, 3H); ESI-MS(C$_{11}$H$_{14}$NO$_5^+$) m/z calculated: 240.0866 [M+H]$^+$, found: 240.0869 [M+H]$^+$.

*methyl 4-(2,2-diethoxyethyl)-2-ethoxy-3,4-dihydro-2H-pyran-5-carboxylate* 18

Potassium fluoride (100 mg) was added to a solution of trichloroacetyl 5 (1.0 g, 5.2 mmol, 2.0 eq) and 3,3,-diethoxy-1-propanal 8 (0.58 g, 5.2 mmol, 2.0 eq) in anhydrous DCM (10 mL) under an argon atmosphere, accompanied by a color change from yellow to orange-red. Ethyl vinyl ether 10 (0.19 g, 0.25 mL, 2.6 mmol, 1.0 eq) was then added dropwise and the reaction was allowed to stir at 22 °C for 16 h. The reaction was diluted with ether and filtered through water-deactivated aluminum oxide (eluted with ether).

After evaporation of the solvents, the residue was partially purified by flash column chromatography (SiO$_2$, hex/MTBE, 5:1). The partially purified cycloadducts (0.52 mg) were dissolved in anhydrous methanol (20 mL) and DBU was added (0.1 mL). When TLC indicated a completed reaction (10 min), the solvent was evaporated under reduced pressure and the residue was purified by flash column chromatography (hex/ether, 5:1) to afford ester **18** as a clear oil (0.15 g, 0.49 mmol, 19% over 2 steps, 63:37 *cis* vs. *trans* cycloadduct). ***cis-18***: ^{1}H NMR (600 MHz, CDCl$_3$): δ ppm 7.43 (s, 1H), 5.10-5.07 (m, 1H), 4.64 (dd, 1H, J = 5.1, 7.0), 3.93 (dq, 1H, J = 7.1, 14.3), 3.70-3.40 (m, 5H), 3.66 (s, 3H), 2.66-2.59 (m, 1H), 2.10-2.03 (m, 1H), 2.03-1.89 (m, 2H), 1.80 (ddd, 1H, J = 2.9), 1.21 (t, 3H, J = 7.1), 1.19-1.13 (m, 6H); ^{13}C NMR (200 MHz, CDCl$_3$): δ ppm 168.01b, 167.85a, 153.44b, 152.25a, 110.63a, 109.80b, 101.70b, 101.52a, 98.44a,b, 65.30b, 64.74a, 62.41b, 61.78a, 59.85b, 59.25a, 51.32a,b, 51.30a,b, 38.53b, 36.15a, 31.46b, 29.60a, 26.60b, 24.54a, 15.52-15.29; ***trans-18***: ^{1}H NMR (600 MHz, CDCl$_3$): δ ppm 7.43 (s, 1H), 4.97 (dd, 1H, J = 1.7, 9.1), 4.58 (t, 1H, J = 5.7), 3.79 (dq, 1H, J = 7.1, 14.2), 3.70-3.40 (m, 5H), 3.66 (s, 3H), 2.75-2.70 (m, 1H), 2.10-2.03 (m, 1H), 2.03-1.89 (m, 2H), 1.72 (ddd, 1H, J = 5.9, 9.1, 14.3), 1.47 (ddd, 1H, J = 5.4, 9.8, 14.5, 1.21), 1.19-1.13 (m, 9H, b); ESI-MS(C$_{15}$H$_{26}$O$_6$) *m/z* calculated: 325.1622 [M+Na]$^+$, found: 325.1623 [M+Na]$^+$.

methyl 2-ethoxy-4-(2-oxoethyl)-3,4-dihydro-2H-pyran-5-carboxylate **22**

183

A slurry (1 mL water with 100 mg dissolved oxalic acid and 100 mg silica gel) was added to a solution of acetal **18** (0.15 g, 0.50 mmol) in THF (4 mL) and the mixture was stirred at 75 °C for 3 h. One water volume-equivalent of toluene was added to the reaction and the solvents were evaporated under reduced pressure. The residue was dissolved in ether with 1% TEA and filtered through silica gel. The filtrate was evaporated under reduced pressure and the residue was purified by flash column chromatography (SiO$_2$, hex/MTBE/TEA, 5:1:0.01) to afford separated diastereomers of **22** (d.e. >95%): *cis*-adduct (76 mg, 0.33 mmol, 66%); *trans*-adduct (24 mg, 0.10 mmol, 21 %). **cis-22**: ^{1}H NMR (CDCl$_3$, 400 MHz): δ 9.74 (t, 1H, J = 1.2), 7.50 (d, 1H, J = 0.5), 5.14 (t, 1H, J = 2.5), 3.78 (dq, 1H, J = 7.1, 9.5), 3.68 (s, 3H), 3.50 (dq, 1H, J = 7.1, 9.5), 3.11-3.05 (m, 1H), 2.99 (ddd, 1H, J = 1.3, 9.6, 17.7), 2.75 (ddd, 1H, J = 1.0, 3.4, 17.4), 2.00 (td, 1H, J = 2.2, 14.4), 1.89 (ddd, 1H, J = 2.8, 6.7, 14.4), 1.16 (t, 3H, J = 7.1; ^{13}C NMR (CDCl$_3$, 100 MHz): δ 202.43, 167.67, 153.10, 108.81, 97.96, 64.87, 51.46, 47.87, 30.06, 22.34, 15.29; **trans-22**: ^{1}H NMR (CDCl$_3$, 500 MHz): δ 9.74 (t, 1H, J = 1.2), 7.49 (d, 1H, J = 1.1), 4.97 (dd, 1H, J = 2.4, 7.5), 3.90 (dq, 1H, J = 7.1, 9.5), 3.68 (s, 3H), 3.58 (dq, 1H, J = 7.1, 9.5), 3.24-3.18 (m, 1H), 2.87 (ddd, 1H, J = 1.2, 4.2, 16.9), 2.38 (ddd, 1H, J = 2.2, 9.0, 16.9), 1.93 (ddd, 1H, J = 6.2, 7.3, 13.7), 1.77 (ddd, 1H, J = 2.4, 5.2, 13.9), 1.21 (t, 3H, J = 7.1); ESI-MS(C$_{11}$H$_{16}$O$_5{}^+$) *m/z* calculated: 251.0890 [M+Na]$^+$, found: 251.0884 [M+Na]$^+$.

Methyl 4-(2,2-diethoxyethyl)-2-isobutoxy-3,4-dihydro-2H-pyran-5-carboxylate **19**

Potassium fluoride (0.2 g) was added to a solution of 3,3-diethoxy-1-propanal **8** (0.48 g, 3.3 mmol, 1.5 eq) and 1,1,1-trichloro-4-hydroxybut-3-en-4-ol **5** (0.63 g, 3.3 mmol, 1.5 eq) in anhydrous DCM (10 mL) under an argon atmosphere. The mixture quickly changed color from yellow to dark orange. Isobutyl vinyl ether **11** was then added drop-wise (0.22 g, 2.2 mmol, 1 eq) and the reaction was continued for 21 h, and then diluted with ether and filtered through deactivated aluminum oxide. After evaporation under reduced pressure, the residue was purified by flash column chromatography (SiO$_2$ 6.2 x 2.5 cm, 7-mL fractions, hex/MTBE, 10:1). The pure fractions were combined and the solvent evaporated under reduced pressure to yield trichloroacetyl cycloadducts (107 mg, 0.26 mmol, 17%) as a pale yellow oil. A second compound was isolated and verified as the alternate (but undesired) cycloadduct (63 mg, 7%). Attempted rearrangement in acid (pTsOH in chloroform) did not transform this compound into the desired product with r$_f$ = 0.70 (hex/MTBE, 1:1). To a solution of this product (105 mg, 0.25 mmol) in anhydrous methanol (3 mL) under argon atmosphere was added DBU (100 mg) and the reaction was stirred for 20 min at room temperature. Most of the solvent was evaporated under reduced pressure, the yellow oil was dissolved in ether and filtered through a short silica gel column. The eluant was evaporated under reduced pressure and the residue contained analytically pure methyl ester **19** (81 mg, 0.25 mmol, 99%). *cis-19*: ^{1}H NMR (300 MHz,

185

CDCl$_3$): δ 7.42 (s, 1H), 5.07 (t, 1H, J = 2.7), 4.65 (dd, 1H, J = 5.6, 6.6), 3.65 (s, 3H),

3.68-3.40 (m, 4H), 3.52 (dd, 1H, J = 6.5, 9.1), 3.19 (dd, 1H, J = 6.4, 9.1), 2.66-2.57 (m,

1H), 2.11 (td, 1H, J = 2.3, 14.2), 2.02-1.70 (m, 4H), 1.19-1.10 (m, 6H), 0.84 (dd, 6H, J =

1.4, 6.7); ^{13}C NMR (125 MHz, CDCl$_3$): δ 167.99, 152.20, 110.60, 101.77, 98.81, 96.09,

61.93, 58.77, 51.28, 36.05, 29.11, 28.62, 24.29, 19.41, 15.47, 15.54. ***trans-19***: ^{1}H NMR

(300 MHz, CDCl$_3$): δ 7.42 (s, 1H), 4.94 (dd, 1H, J = 2.2, 9.0), 4.57 (t, 1H, J = 5.7), 3.65

(s, 3H), 3.68-3.40 (m, 5H), 3.25 (dd, 1H, J = 6.9, 9.3), 2.72 (dq, 1H, J = 3.5, 6.4), 2.02-

1.70 (m, 4H), 1.46 (ddd, 1H, J = 5.4, 9.7, 14.3), 1.19-1.10 (m, 6H), 0.87 (dd, 6H, J = 2.2,

6.7); ^{13}C NMR (125 MHz, CDCl$_3$): δ 167.84, 153.46, 109.78, 101.77, 98.81, 76.44,

62.33, 59.97, 51.28, 38.46, 31.40, 26.35, 19.35, 15.54, 15.47; ESI-MS(C$_{17}$H$_{30}$O$_6$$^+$) m/z

calculated: 353.1940 [M+Na]$^+$, found: 353.1957 [M+Na]$^+$.

Methyl 2-isobutoxy-4-(2-oxoethyl)-3,4-dihydro-2H-pyran-5-carboxylate **23**
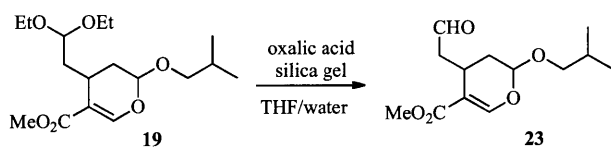
A slurry (1 mL water with 100 mg dissolved oxalic acid and 100 mg silica gel) was

added to a solution of acetal **19** (80 mg, 0.25 mmol) in THF (4 mL). The mixture was

stirred at 60 °C for 3 h. The reaction was quenched by addition of solid anhydrous

sodium bicarbonate (200 mg). After filtration, the solvent was evaporated under reduced

pressure. The residue was purified by flash column chromatography (SiO$_2$,

hex/MTBE/TEA, 5:1:0.01) to afford separated diastereomers of **23** (d.e. >95%): *cis*-

adduct (28 mg, 0.10 mmol, 54%); *trans*-adduct (20 mg, 0.078 mmol, 39 %). ***cis-23***: ^{1}H

186

NMR (500 MHz, CDCl$_3$): δ 9.75 (t, 1H, J = 1.3), 7.49 (d, 1H, J = 0.7), 5.11 (t, 1H, J = 2.5), 3.68 (s, 3H), 3.51 (dd, 1H, J = 6.7, 9.2), 3.21 (dd, 1H, J = 6.4, 9.2), 3.14-3.05 (m, 1H), 2.99 (ddd, 1H, J = 1.4, 9.6, 17.4), 2.80-2.70 (m, 1H), 2.03 (td, 1H, J = 2.1, 14.5), 1.93-1.84 (m, 1H), 1.80 (dd, 1H, J = 6.6, 13.4), 0.85 (d, 6H, J = 6.6); ^{13}C NMR (125 MHz, CDCl$_3$): δ 202.32, 167.65, 153.04, 108.84, 98.39, 76.26, 51.43, 48.00, 29.83, 28.58, 22.24, 19.46, 19.44.**trans-23**: ^{1}H NMR (500 MHz, CDCl$_3$): δ 9.74 (t, 1H, J = 1.8), 7.49 (d, 1H, J = 1.1), 4.95 (dd, 1H, J = 2.4 7.2), 3.67 (s, 3H), 3.61 (dd, 1H, J = 6.6, 9.2), 2.27 (dd, 1H, J = 6.8, 9.2), 3.24-3.18 (m, 1H), 2.88 (ddd, 1H, J = 1.5, 4.2, 17.0), 2.38 (ddd, 1H, J = 2.1, 9.0, 17.0), 1.96 (td, 1H, J = 6.4, 13.7), 1.84 (td, 1H, J = 6.7, 13.4), 1.76 (ddd, 1H, J = 2.4, 5.5, 13.9), 0.88 (dd, 6H, J = 3.4, 6.7); ^{13}C NMR (125 MHz, CDCl$_3$): δ 201.28, 167.51, 154.15, 108.29, 98.17, 76.32, 51.45, 48.65, 32.49, 28.62, 23.79, 19.42, 19.38; ESI-MS(C$_{13}$H$_{20}$O$_5$$^+$) m/z calculated: 279.1208 [M+Na]$^+$, found: 279.1209 [M+Na]$^+$.

Methyl 2-tert-butoxy-4-(2,2-diethoxyethyl)-3,4-dihydro-2H-pyran-5-carboxylate **20**
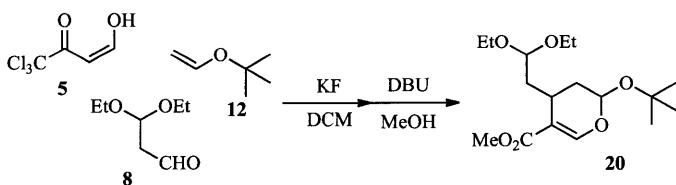
Potassium fluoride (0.17 g) was added to a solution of 3,3,-diethoxy-1-propanal **8** (0.44 g, 3.0 mmol, 1.5 eq) and trichloroacetyl **5** (0.46 g, 2.4 mmol, 1.2 eq) in anhydrous DCM (10 mL) under an argon atmosphere, accompanied by a color change from yellow to orange-red. *tert*-butyl vinyl ether **12** (0.20 g, 2.0 mmol, 1.0 eq) was added dropwise and the

reaction was allowed to stir at 22 °C. A precipitate formed initially, but dissolved after continued stirring. After a total of 21 h, the reaction was diluted with ether and filtered through water-deactivated aluminum oxide (eluted with ether). After evaporation of the solvents, the residue was partially purified by flash column chromatography (SiO$_2$, hex/MTBE/TEA, 8:1:0.01) to afford a mixture of cycloadducts as a yellow oil. The partially purified cycloadducts were then dissolved in anhydrous methanol (3 mL) and DBU (0.1 mL) was added. When TLC indicated a completed reaction (20 min), the solvent was evaporated under reduced pressure and the residue was purified by flash column chromatography (hex/MTBE/TEA, 5:1:0.01) to afford ester **20** as a clear oil (0.12 g, 0.28 mmol, 14% over 2 steps, 1.7:1.0 *cis* vs. *trans*). ***cis-20***: ^{1}H NMR (300 MHz, CDCl$_3$): δ 7.42 (s, 1H), 5.07 (t, 1H, J = 2.7), 4.65 (dd, 1H, J = 5.6, 6.6), 3.65 (s, 3H), 3.68-3.40 (m, 4H), 3.52 (dd, 1H, J = 6.5, 9.1), 3.19 (dd, 1H, J = 6.4, 9.1), 2.66-2.57 (m, 1H), 2.11 (td, 1H, J = 2.3, 14.2), 2.02-1.70 (m, 4H), 1.19-1.10 (m, 6H), 0.84 (dd, 6H, J = 1.4, 6.7); ^{13}C NMR (125 MHz, CDCl$_3$): δ 167.99, 152.20, 110.60, 101.77, 98.81, 96.09, 61.93, 58.77, 51.28, 36.05, 29.11, 28.62, 24.29, 19.41, 15.47, 15.54. ***trans-20***: ^{1}H NMR (300 MHz, CDCl$_3$): δ 7.42 (s, 1H), 4.94 (dd, 1H, J = 2.2, 9.0), 4.57 (t, 1H, J = 5.7), 3.65 (s, 3H), 3.68-3.40 (m, 5H), 3.25 (dd, 1H, J = 6.9, 9.3), 2.72 (dq, 1H, J = 3.5, 6.4), 2.02-1.70 (m, 4H), 1.46 (ddd, 1H, J = 5.4, 9.7, 14.3), 1.19-1.10 (m, 6H), 0.87 (dd, 6H, J = 2.2, 6.7); ^{13}C NMR (125 MHz, CDCl$_3$): δ 167.84, 153.46, 109.78, 101.77, 98.81, 76.44, 62.33, 59.97, 51.28, 38.46, 31.40, 26.35, 19.35, 15.54, 15.47. ESI-MS(C$_{17}$H$_{30}$O$_6{}^+$) *m/z* calculated: 353.1940 [M+Na]$^+$, found: 353.1957 [M+Na]$^+$.

Methyl 2-tert-butoxy-4-(2-oxoethyl)-3,4-dihydro-2H-pyran-5-carboxylate **24**
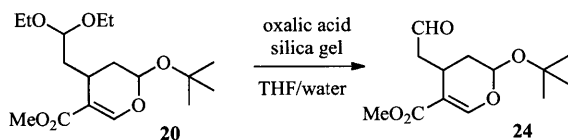
A slurry (1 mL water with 100 mg dissolved oxalic acid and 100 mg silica gel) was added to a solution of acetal **20** (0.12 g, 0.35 mmol) in THF (4 mL). The mixture was stirred at 60 °C for 90 min. The reaction was quenched by addition of 200 mg solid anhydrous sodium bicarbonate. After filtration, the solvent was evaporated under reduced pressure. The residue was purified by flash column chromatography (SiO_2, hex/MTBE/TEA, 5:1:0.01) to afford separated diastereomers of **24** (d.e. >95%): *cis*-adduct (39 mg, 0.15 mmol, 44%); *trans*-adduct (19 mg, 0.075 mmol, 22 %). **cis-24**: ^{1}H NMR (500 MHz, $CDCl_3$): δ 9.74 (t, 1H, J = 1.3), 7.49 (d, 1H, J = 0.6), 5.38 (t, 1H, J = 2.8), 3.67 (s, 3H), 3.09 (qd, 1H, J = 4.4, 9.0), 2.99 (ddd, 1H, J = 1.4, 9.5, 17.7), 2.77 (ddd, 1H, J = 1.3, 4.0, 17.7), 1.87 (dd, 2H, J = 3.0, 4.4), 1.20 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 202.6, 167.8, 153.6, 108.3, 92.9, 76.2, 51.4, 48.0, 31.4, 28.7, 22.8. **trans-24**: ^{1}H NMR (500 MHz, $CDCl_3$): δ 9.74 (dd, 1H, J = 1.8, 2.2), 7.49 (d, 1H, J = 1.0), 5.20 (dd, 1H, J = 2.5, 7.8), 3.67 (s, 3H), 3.24-3.18 (m, 1H), 2.86 (ddd, 1H, J = 1.5, 5.4, 16.7), 2.38 (ddd, 1H, J = 2.3, 8.9, 16.7), 1.89 (ddd, 1H, J = 5.9, 7.9, 13.8), 1.67 (ddd, 1H, J = 2.5, 3.7, 13.9), 1.24 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 201.4, 167.6, 154.7, 101.7, 92.6, 76.5, 51.4, 49.0, 33.7, 28.7, 24.3; ESI-MS($C_{13}H_{20}O_5$) *m/z* calculated: 279.1208 [M+Na]$^+$, found: 279.1200 [M+Na]$^+$.

Methyl-2-(cyclohexyloxy)-4-(2,2-diethoxyethyl)-3,4-dihydro-2H-pyran-5-carboxylate **21**
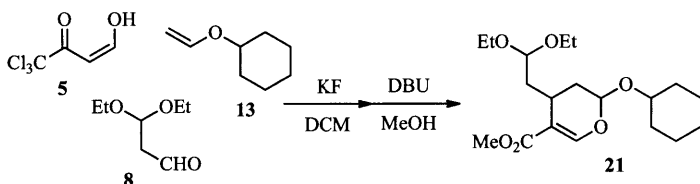
Potassium fluoride (0.20 g) was added to a solution of 3,3-diethoxy-1-propanal **8** (0.51 g, 3.5 mmol, 1.5 eq) and trichloroacetyl **5** (0.67 g, 3.5 mmol, 1.5 eq) in anhydrous DCM (10 mL) under an argon atmosphere, accompanied by a color change from yellow to orange-red. Cyclohexyl vinyl ether **13** (0.29 g, 2.3 mmol, 1.0 eq) was added dropwise and the reaction was allowed to stir at 22 °C for 21 h. The reaction was diluted with ether and filtered through water-deactivated aluminum oxide (eluted with ether). After evaporation of the solvents, the residue was partially purified by flash column chromatography (6.2 cm SiO$_2$, hex/MTBE, 10:1) to afford a crude mixture of cycloadducts as a yellow oil (2.1:1 *cis/trans* raio, r$_f$ = 0.75, hex/MTBE 1:1). The partially purified cycloadducts (0.38 g) were then dissolved in anhydrous methanol (3 mL) and DBU (0.1 mL) was added. When TLC indicated a completed reaction (10 min, r$_f$ = 0.67, hex/MTBE, 1:1), the solvent was evaporated under reduced pressure and the residue was purified by flash column chromatography (SiO$_2$, hex/MTBE,10:1) to afford ester **21** as a clear oil (0.19 g, 0.52 mmol, 15% over 2 steps). ***cis-21***: ^{1}H NMR (300 MHz, CDCl$_3$): δ 7.42 (d, 1H, J = 0.8), 5.32 (t, 1H, J = 3.1), 4.64 (dd, 1H, J = 4.7, 7.4), 3.65 (s, 3H), 3.67-3.57 (m, 2H), 3.50-3.41 (m, 2H), 2.03 (ddd, 1H, J = 4.1, 7.4, 13.8), 1.99-1.85 (m, 4H), 1.77 (ddd, 1H, J = 3.0, 6.6, 14.0), 1.20 (s, 9H), 1.19-1.13 (m, 6H); ^{13}C NMR (125 MHz, CDCl$_3$): δ 152.8, 110.2, 101.4, 93.3, 75.8, 61.8, 58.9, 51.2, 38.9, 36.1, 30.9, 28.7, 25.0, 15.5. ***trans-21***: ^{1}H

NMR (300 MHz, CDCl$_3$): δ 7.42-7.41 (m, 1H), 5.20 (dd, 1H, J = 2.4, 9.5), 4.55 (t, 1H, J = 5.5), 3.64 (s, 3H), 3.69-3.57 (m, 2H), 3.49-3.41 (m, 2H), 2.74-2.68 (m, 1H), 2.48 (ddd, 1H, J = 1.5, 5.6, 7.2), 1.98-1.85 (m, 2H), 1.49 (ddd, 1H, J = 5.7, 9.8, 14.2), 1.73-1.67 (m, 1H), 1.23 (s, 9H), 1.18-1.13 (m, 6H); ^{13}C NMR (125 MHz, CDCl$_3$): δ 154.0, 109.3, 102.2, 93.1, 76.3, 62.4, 60.4, 51.2, 39.8, 32.7, 28.7, 26.9, 15.5; ESI-MS(C$_{17}$H$_{31}$O$_6^+$) m/z calculated: 353.1935 [M+Na]$^+$, found: 353.1941 [M+Na]$^+$.

Methyl 2-(cyclohexyloxy)-4-(2-oxoethyl)-3,4-dihydro-2H-pyran-5-carboxylate **25**
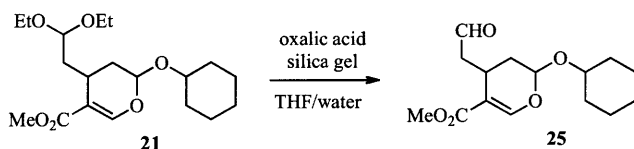
A slurry (1 mL water with 100 mg dissolved oxalic acid and 100 mg silica gel) was added to a solution of acetal **21** (0.18 g, 0.52 mmol) in THF (4 mL). The mixture was stirred at 60 °C for 90 min. The reaction was quenched by addition of 200 mg solid anhydrous sodium bicarbonate. After filtration, the solvent was evaporated under reduced pressure. The residue was purified by flash column chromatography (SiO$_2$, hex/MTBE/TEA, 5:1:0.01) to afford separated diastereomers of **25** (d.e. >95%). Major diastereomer, *cis*-adduct, 79 mg, 0.28 mmol, 54%. Minor diastereomer, *trans*-adduct, 25 mg, 0.09 mmol, 17%. **cis-25**: ^{1}H NMR (500 MHz, CDCl$_3$): δ 9.74 (t, 1H, J = 1.3), 7.49 (d, 1H, J = 0.6), 5.27 (t, 1H, J = 2.6), 3.67 (s, 3H), 3.65-3.59 (m, 1H), 3.10-3.05 (m, 1H), 3.02 (ddd, 1H, J = 1.3, 9.6, 17.5), 2.78-2.72 (m, 1H), 1.97 (td, 1H, J = 2.3, 14.4), 1.88 (ddd, 1H, J = 2.8, 6.5, 14.4), 1.82-1.72 (m, 2H), 1.68-1.58 (m, 2H), 1.48-1.41 (m, 1H),

191

1.40-1.13 (m, 5H); ^{13}C NMR (125 MHz, CDCl$_3$): δ 202.53, 167.73, 153.26, 108.70,

96.10, 76.55, 51.40, 48.02, 33.45, 31.59, 30.43, 25.72, 23.91, 23.74, 22.45; **trans-25**: ^{1}H

NMR (500 MHz, CDCl$_3$): δ 9.74 (t, 1H, J = 1.8), 7.48 (d, 1H, J = 1.1), 5.10 (dd, 1H, J =

2.4, 7.5), 3.67 (s, 3H), 3.68-3.62 (m, 1H), 3.21 (td, 1H, J = 5.1, 14.6), 2.86 (ddd, 1H, J =

1.5, 4.2, 16.9), 2.38 (ddd, 1H, J = 2.2, 8.9, 16.9), 1.93 (ddd, 1H, J = 6.1, 7.4, 13.7), 1.89-

1.82 (m, 2H), 1.74 (ddd, 1H, J = 2.4, 5.2, 13.9), 1.72-1.67 (m, 2H), 1.53-1.47 (m, 1H),

1.42-1.15 (m, 5H); ^{13}C NMR (125 MHz, CDCl$_3$): δ 201.36, 167.58, 154.37, 108.07,

96.15, 51.44, 48.80, 39.78, 33.61, 32.86, 32.03, 25.67, 24.26, 24.14, 24.00; ESI-

MS(C$_{15}$H$_{22}$O$_5^+$) *m/z* calculated: 283.1545 [M+H]$^+$, found: 283.1544 [M+H]$^+$.

*Methyl 4-((5,5-dimethyl-1,3-dioxan-2-yl)methyl)-2-(2-nitrobenzyloxy)-3,4-dihydro-2H-
pyran-5-carboxylate* **26**
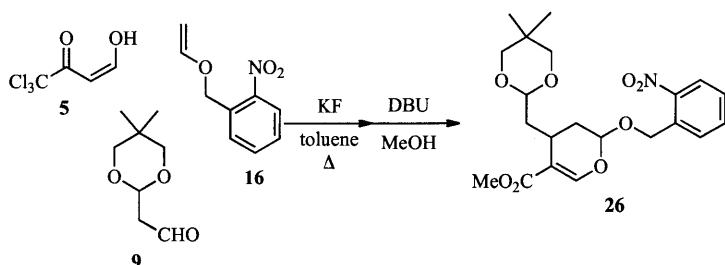
2-(5,5-dimethyl-1,3-dioxan-2-yl)ethanal **9** (0.68 g, 4.0 mmol, 2.0 eq) and 1,1,1-trichloro-

4-hydroxybut-3-en-2-one **5** (0.76 g, 4.0 mmol, 2.0 eq) were dissolved in toluene (8 mL)

and potassium fluoride was added (100 mg). A 2-mL solution of 2-nitrobenzyl vinyl ether

16 (0.36 g, 2.0 mmol, 1.0 eq) was added drop-wise to this solution and the mixture was

heated to 100 °C for 4 h, and then diluted with ether and filtered through water-

deactivated aluminum oxide. The solvents were evaporated under reduced pressure to

yield a yellow oil. Chloroform (20 mL) was added together with a few crystals of pTsOH. The flask was left standing 12 h and then filtered through SiO_2. After evaporation of the solvents, the residue was purified by flash column chromatography (16x 2.5 cm SiO_2, hex/MTBE, 25:1 to 15:1 to 9:1 to 4:1, 8-mL fractions) to give partially purified cycloadducts (fractions 7-33 were combined). The recovered product was then purified (DCM/hex, 2:1 to 8:3) to give a pale yellow foam of >90% pure cycloadducts (0.47 mg, 46%). ^{1}H NMR (600 MHz, CDCl$_3$): δ 8.08 (s, 1H, a), 8.08-8.01 (m, 1H, a,b), 7.78-7.71 (m, 1H, a,b), 7.67-7.54 (m, 1H, a,b), 7.48-7.36 (m, 1H, a,b), 5.41-5.37 (m, 1H, a,b), 5.36-5.00 (m, 2H, a,b), 4.57-4.46 (m, 1H, a,b), 3.63-3.25 (m, 4H, a,b), 3.15-2.89 (m, 1H, a,b), 2.42 (td, 0.55H, J = 1.7, 14.4, a), 2.33 (td, 0.45H, J = 2.5, 13.4, b), 2.14-1.60 (m, 3H, a,b), 1.15 (s, 1.6H, a), 1.12 (s, 1.4H, b), 0.68 (s, 1.6H, a), 0.67 (s, 1.4, b); ^{13}C NMR (125 MHz, CDCl$_3$): δ 167.82a, 167.64b, 152.87b, 151.72a, 134.43a, 134.25b, 134.02a, 133.91b, 128.86b, 128.77a, 128.38b, 128.22a, 124.94b, 124.89a, 110.95a, 110.33b, 102.04a, 101.51b, 98.55a, 93.39b, 68.35a, 67.94b, 56.52, 51.45a, 51.44b, 39.69b, 38.43a, 31.81b, 30.36a, 30.28b, 29.52a, 25.72b 23.75a, 23.31(2C), 22.10a, 22.08b, 25.72b, 23.75a, 23.31a, 22.10, 22.08; ESI-MS($C_{21}H_{24}NO_7Cl_3$) m/z calculated: 508.0691 [M+H]$^+$, found: 508.0675 [M+H]$^+$. The partially purified cycloadducts (81.9 mg, 0.16 mmol) were dissolved in anhydrous methanol (2 mL) and DBU was added (0.05 mL). The mixture was stirred for 10 minutes and then concentrated *in vacuo*. The residue was dissolved in hex/MTBE (1:1) and filtered through a short column of SiO_2 eluting with the same solvent mixture. Evaporation of the solvents *in vacuo* yielded methyl ester **26** as an inseparable mixture of diastereomers (66.3 mg, 0.16 mmol, 98%) in analytical purity.

The NMR signals corresponding to the major stereoisomer is indicated by 'a'; the minor stereoisomer is indicated by 'b'. ^{1}H NMR (400 MHz, CDCl$_3$): δ 8.13-8.09 (m, 1H, a,b), 7.82-7.76 (m, 1H, a,b), 7.69-7.62 (m, 1H, a,b), 7.50-7.43 (m, 2H), 5.36-5.01 (m, 3H, a,b), 4.65-4.52 (m, 1H, a,b), 3.73 (s, 1.7H, a), 3.72 (s, 1.3H, b), 3.66-3.56 (m, 2H, a,b), 3.51-3.38 (m, 2H), 2.98-2.92 (m, 0.4H, b), 2.88-2.83 (m, 0.6H, a), 2.40-2.34 (m, 0.6H, a), 2.27-2.21 (m, 0.4H, b), 2.20-1.80 (m, 2.6H, a,b), 1.66-1.59 (ddd, 0.4H, J = 4.7, 9.2, 14.0, b), 1.22 (s, 1.7H, a), 1.20 (s, 1.3H, b), 0.74 (s, 1.7H, a), 0.73 (s, 1.3H, b); ^{13}C NMR (75 MHz, CDCl$_3$): δ 167.81a, 167.63b, 152.86b, 151.71a, 134.42a, 134.24b, 134.01a, 133.90b, 128.84b, 128.76a, 128.37b, 128.21a, 124.93b, 124.87, 110.94a, 110.32b, 102.02a, 101.50b, 98.54a, 98.48b, 77.41a, 77.37b, 77.29(2C), 68.34a, 67.93b, 51.44a, 51.42b, 39.68b, 38.42a, 31.80b, 30.35a, 30.27b, 29.51a, 25.70b, 23.73a, 23.30(2C), 22.09, 22.07; ESI-MS(C$_{21}$H$_{27}$NO$_8$Na) m/z calculated: 444.1634 [M+Na]$^+$, found: 444.1624 [M+Na]$^+$.

Methyl 2-(2-nitrobenzyloxy)-4-(2-oxoethyl)-3,4-dihydro-2H-pyran-5-carboxylate **27**
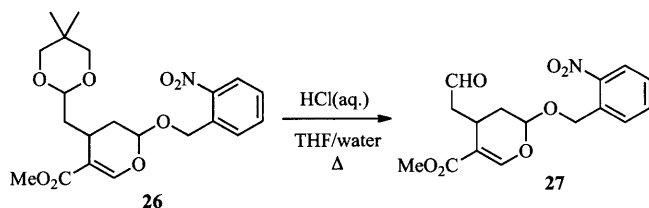
Acetal **26** (0.63 mmol) was added to 5 mL THF and 5 mL H$_2$O for a final concentration of 0.2 mM HCl and the mixture was stirred at 50 °C for 17 h. After that time, 5 drops of 10 M HCl were added, the temperature was increased to 65 °C, and the reaction was stirred for a total of 43 h. The solvents were evaporated *in vacuo* in a heated water bath

and the residue was dissolved in EtOAc. The EtOAc was washed with saturated aqueous solutions of sodium bicarbonate, water then brine; the aqueous layers were back extracted with EtOAc. The combined organic layers were evaporated and the residue was purified by flash column chromatography (SiO$_2$, 13x3 cm, 7-mL fractions, hex/MTBE, 2:1) to afford aldehyde **27** (29 mg, 14%) as an inseparable mixture of diastereomers. Starting material could also be recovered (110 mg, 53%). (major diastereomer, *cis*): ^{1}H NMR (500 MHz, CDCl$_3$): δ 9.78 (t, 1H, *J* = 1.4), 8.04 (dd, 1H, *J* = 1.1, 8.2), 7.74-7.59 (m, 2H), 7.50 (d, 1H, *J* = 0.9), 7.47-7.42 (m, 1H), 5.29 (t, 1H, *J* = 2.4), 5.15 (d, 1H, *J* = 14.9), 4.99 (d, 1H, *J* = 14.7), 3.69 (s, 3H), 3.19-3.12 (m, 1H), 2.91 (dd, 1H, *J* = 1.7, 9.6), 2.86-2.84 (m, 1H), 2.16-2.08 (m, 1H), 2.00 (dddd, 1H, *J* = 0.5, 3.1, 6.9, 14.6); ^{13}C NMR (125 MHz, CDCl$_3$): δ 200.07, 167.41, 152.66, 147.65, 134.01, 133.88, 128.90, 128.67, 125.05, 109.23, 98.18, 68.36, 51.55, 48.03, 29.82, 22.15; (minor diastereomer, *trans*): ^{1}H NMR (500 MHz, CDCl$_3$): δ 9.77 (t, 1H, *J* = 1.6), 8.07 (dd, 1H, *J* = 1.2, 8.2), 7.74-7.70 (m, 1H), 7.66-7.59 (m, 1H), 7.47-7.42 (m, 1H), 5.24 (d, 1H, *J* = 14.9), 5.18 (dd, 1H, *J* = 2.5, 6.8), 5.07 (d, 1H, *J* = 14.9), 3.69 (s, 3H), 3.30-3.23 (m, 1H), 2.98 (dd, 1H, *J* = 1.4, 4.1), 2.94 (dd, 1H, *J* = 1.6, 4.0), 2.82-2.80 (m, 1H), 2.15-2.08 (m, 1H), 1.85 (ddd, 1H, *J* = 2.5, 6.3, 13.9); ^{13}C NMR (125 MHz, CDCl^3): δ 200.96, 167.28, 153.59, 147.28, 133.95, 133.66, 128.85, 128.53, 125.00, 108.94, 97.84, 67.98, 51.52, 48.34, 32.29, 23.47; ESI-MS(C$_{16}$H$_{17}$NO$_7$) *m/z* calculated: 336.1078 [M+H]$^+$, found: 336.1078 [M+H]$^+$.

Methyl-4-(2,2-diethoxyethyl)-2-(4,5-dimethoxy-2-nitrophenoxy)-3,4-dihydro-2H-pyran-5-carboxylate **33**
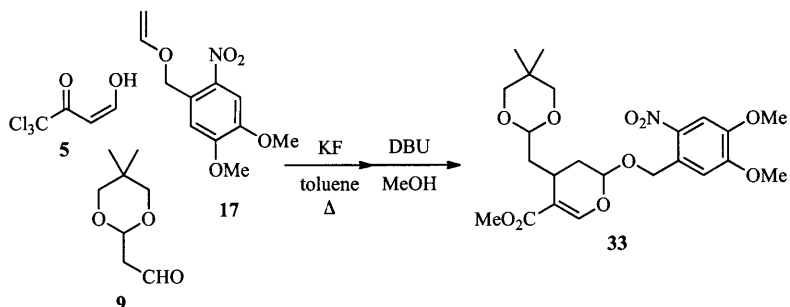
2-(5,5-dimethyl-1,3-dioxan-2-yl)ethanal **9** (0.15 g, 1.0 mmol, 2.0 eq) and 1,1,1-trichloro-4-hydroxybut-3-en-2-one **5** (0.25 g, 1.3 mmol, 2.6 eq) were dissolved in DCM (5 mL) and potassium fluoride was added (50 mg). To this mixture 4,5-dimethoxy-2-nitrobenzyl vinyl ether **17** (0.12 g, 0.51 mmol, 1.0 eq) was added drop-wise. The mixture was stirred at room temperature for 20 h, then diluted with ether and filtered through water-deactivated aluminum oxide. The solvents were evaporated under reduced pressure and the residue was purified by flash column chromatography (SiO$_2$, hex/MTBE/TEA, 5:1:0.01 to 100:1 MTBE/TEA) to yield partially purified cycloadducts. The combined fractions were concentrated to give a residue (26 mg), which was dissolved in anhydrous MeOH (2 mL). DBU was added (0.10 mL) to the solution, the mixture was stirred at room temperature for 10 min, and the MeOH was then evaporated *in vacuo*. The residue was purified by flash column chromatography (SiO$_2$, hex/MTBE/TEA, 2:1:0.01) to give methyl ester **33** (25 mg, 0.055 mmol, 11% over 2 steps) as an inseparable mixture of diastereomers. The NMR signals corresponding to the major stereoisomer is indicated by 'a'; the minor stereoisomer is indicated by 'b'. ^{1}H NMR (400 MHz, CDCl$_3$): δ 7.72-7.66

196

(m, 1H, a,b), 7.46-7.42 (m, 1H, a,b), 7.23-7.13 (m, 1H, a,b), 5.43-4.97 (m, 3H, a,b), 4.75-4.51 (m, 1H, a,b), 3.96-3.89 (m, 6H, OMe-a,b), 3.69 (s, 1.8H, a), 3.68 (s, 1.2H, a), 3.68-3.61 (m, 2H, a,b), 3.55-3.39 (m, 2H, a,b), 2.86-2.66 (m, 1H), 2.37-2.27 (m, 1H), 2.20-1.87 (m, 3H), 1.25-1.09 (6H); ESI-MS($C_{23}H_{32}NO_{10}{}^+$) *m/z* calculated: 482.2021 [M+H]$^+$, found: 482.2014 [M+H]$^+$.

Methyl-2-(4,5-dimethoxy-2-nitrophenoxy)-4-(2-oxyethyl)-3,4-dihydro-2H-pyran-5-carboxylate **34**
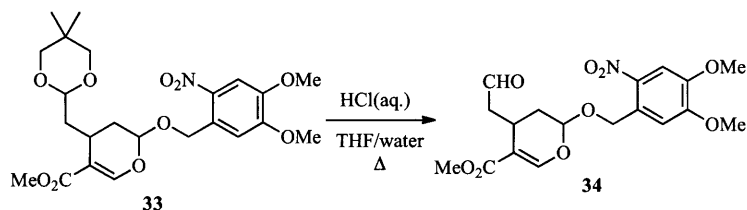
A slurry of silica gel (100 mg) and oxalic acid (100 mg) in water (1 mL) was prepared and added (0.5 mL) to a solution of acetal **33** (25 mg, 0.055 mmol) in tetrahydrofuran (THF, 2 mL). The resulting mixture was stirred at 65-70 °C for 3 h and then quenched by addition of solid sodium bicarbonate (100 mg). After filtration, the solvent was evaporated under reduced pressure and the residue was purified by flash column chromatography (SiO$_2$, hex/MTBE/TEA: 2:1:0.01) to yield aldehyde **34** as unseparable diastereomers (3.4 mg, 0.0086 mmol, 25%). The NMR signals corresponding to the major stereoisomer is indicated by 'a'; the minor stereoisomer is indicated by 'b'. ^{1}H NMR (400 MHz, CDCl$_3$): δ 9.80-9.75 (m, 1H, a,b), 7.70 (s, 0.2H, b), 7.70 (s, 0.8H, a), 7.54-7.51 (m, 0.7H, a), 7.49 (d, 0.2H, *J* = 1.2), 5.40-4.99 (m, 3H, a,b), 3.38-3.13 (m, 1H), 3.97 (s, 2.5H, b), 3.94 (s, 3.5H, a), 3.71 (s, 2.1H, a), 3.69 (s, 0.9H, b), 3.38-3.13 (m, 1H, a,b),

3.07-2.84 (m, 1.7H, a,b), 2.42 (ddd, 0.3H, J = 1.7, 9.2, 17.2, b), 2.20-2.11 (m, 1H, a,b), 2.04 (ddd, 0.8H, J = 3.2, 6.9, 14.7, a), 1.87 (ddd, 0.3H, J = 2.4, 6.2, 13.9, b); ESI-MS($C_{18}H_{22}NO_9^+$) *m/z* calculated: 396.1289 [M+H]$^+$, found: 396.1292 [M+H]$^+$.

5.5 References

1. Tietze, L.-F. "Secologanin, a biogenetic key compound - synthesis and biogenesis of the iridoid and secoiridoid glycosides", *Angew. Chem. Int. Ed. Engl.* **1983**, *22*, 828-841.

2. O'Connor, S.E.; Maresh, J.J. "Chemistry and biology of monoterpene indole alkaloid biosynthesis", *Nat. Prod. Rep.* **2006**, *23*, 532-547.

3. Galan, M.C.; O'Connor, S.E. "Semisynthesis of secologanin analogues", *Tetrahedron Lett.* **2006**, *47*, 1563-1565.

4. Galan, M.C.; McCoy, E.; O'Connor, S.E. "Chemoselective derivatization of alkaloids in periwinkle", *Chem. Comm.* **2008**, *18*, 3095-3098.

5. Tietze, L.F.; Meier, H.; Nutt, H. "The tandem Knoevenagel hetero-Diels-Alder reaction with a formylacetic acid equivalent. Synthesis of dihydropyrancarboxylates", *Chem. Ber.* **1989**, *122*, 643-650.

6. Tietze, L.F.; Meier, H.; Nutt, H. "Synthesis of (±)-secologanin aglucone *O*-ethyl ether and derivatives by tandem Knoevenagel heter Diels-Alder reaction", *Lieb. Ann. Chem.* **1990**, 253-260.

7. Bernhardt, P.; O'Connor, S.E. "Synthesis and biochemical evaluation of des-vinyl secologanin *O*-analogs with alternate stereochemistry", *Tetrahedron Lett.* **2009**, *50*, 7118-7120.

8. Ma, X,; Panjikar, S.; Koepke, J.; Loris, E.; Stöckigt, J. "The structure of *Rauvolfia serpentina* strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins", *Plant Cell* **2006**, *23*, 532-547.

9. Tietze, L.F.; Meier, H.; Voss, E. "Highly efficient syntheses of alkyl 3,3-dialkoxypropanoates, alkyl 4-ethoxy-2-oxo-3-butenoates, and monoprotected malonaldehydes", *Synthesis* **1988**, 274-277.

10. Botteghi, C.; Soccolini, F. "Malonaldehyde, succinaldehyde, and glutaraldehyde monoacetals: synthesis and applications", *Synthesis* **1985**, 592-604.

11. Frigeria, M.; Santagostino, M.; Sputore, S. "A user-friendly entry to 2-iodoxybenzoic acid (IBX)", *J. Org. Chem.* **1999**, *64*, 4537-4538.

12. Dess, B.D.; Martin, J.C. "Readily accessible 12-I-5 oxidant for the conversion of primary and secondary alcohols to aldehydes and ketones", *J. Org Chem.* **1983**, *48*, 4155-4156.

13. Dess, B.D.; Martin, J.C. "A useful 12-I-5 triacetoxyperiodinane (the Dess-Martin periodinane) for the selective oxidation of primary or secondary alcohols and a variety of related 12-I-5 species", *J. Am. Chem. Soc.* **1991**, *113*, 7277-7287.

14. Ireland, R.E.; Liu, L.J. "An improved procedure for the preparation of the Dess-Martin periodinane", *J. Org. Chem.* **1993**, *58*, 2899.

15. Meyer, S.D.; Schreiber, S.L. "Acceleration of the Dess-Martin oxidation by water", *J. Org. Chem.* **1994**, *59*, 7549-7552.

16. Okimoto, Y.; Sakaguchi, S.; Ishii, Y. "Development of a highly efficient catalytic method for synthesis of vinyl ethers", *J. Am. Chem. Soc.* **2002**, *124*, 1590-1591.

17. Gopalaiah, K. "Oxalic acid: a very useful Brønsted acid in organic synthesis", *Synlett* **2004**, 2838-2839.

18. Bochet, C.G. "Photolabile protecting groups and linkers", *J. Chem. Soc., Perkin Trans. 1* **2002**, 125-142.

19. Bernhardt, P.; Usera, A.R.; O'Connor, S.E. "Strictosidine synthase from *Ophiorrhiza pumila* is a stereoselective and promiscuous biocatalyst for the Pictet-Spengler reaction", *submitted* **2009**.

20. Bernhardt, P.; Yerkes, N.; O'Connor, S.E. "Bypassing stereoselectivity in the early steps of alkaloid biosynthesis", *Org. Biomol. Chem.* **2009**, *7*, 4166-4168.

21. Chen, S.; Galan, M.C.; Coltharp, C.; and O'Connor, S.E. "Redesign of a central enzyme in alkaloid biosynthesis", *Chem. Biol.* **2006**, *13*, 1137-1141.

22. Bernhardt, P.; McCoy, E.A.; O'Connor, S.E. "Rapid identification of enzyme variants for reengineered alkaloid biosynthesis in periwinkle", *Chem. Biol.* **2007**, *14*, 888-897.

CHAPTER 6

IDENTIFYING AND BYPASSING STEREOSELECTIVITY IN

THE EARLY STEPS OF ALKALOID BIOSYNTHESIS

6.1 Introduction

Stereo configuration often determines how a natural product interacts with biological systems. For example, different stereosiomers can have different pharmacological profiles.[1,2] Biosynthetic enzymes, however, typically allow the production of only one natural product stereoisomer.[3] Fermentation of natural products with alternate stereochemistry must therefore bypass the stereochemical restrictions of biosynthetic pathway enzymes.

Several monoterpene indole alkaloids (MIAs) produced by *Catharanthus roseus* have important biological activities (Chapter 1).[4,5] Some MIAs are single isomers; others are stereoisomer mixtures. The heteroyohimbine alkaloids, a subclass of MIAs, exist in the plant cells as different stereoisomers where each stereoisomer has a different bioactivity. Ajmalicine (raubasine) **5a** acts as a smooth muscle relaxant and as an α1 anti-adrenergic, while tetrahydroalstonine **5b** acts as an α2 anti-adrenergic (Figure 6.1).[6-8] Access to alternative stereoisomers of **5a** and **5b**, and alternative stereoisomers of other MIAs, would provide diversified compound libraries for bioactivity screening.

MIA biosynthesis starts with a strictosidine synthase (STS)-catalyzed reaction between tryptamine **1** and secologanin **2** to yield strictosidine **3a** (Figure 6.1).[9-11] The stereogenic centers of strictosidine **3a**, the key intermediate for all MIAs,[9] derive from the densely functionalized secologanin dihydropyran (C15, C20, and C21) or the prochiral aldehyde carbon, which becomes the C3 stereogenic center of strictosidine (Figure 6.1).
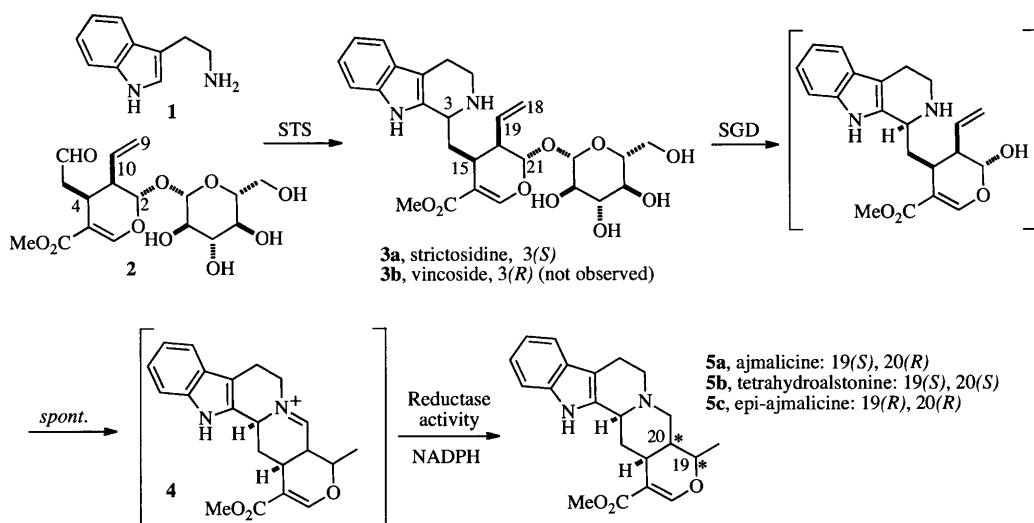
Figure 6.1 Early steps of monoterpene indole alkaloid biosynthesis in *Catharanthus roseus*.

The second step of the MIA pathway is the strictosidine-β-D-glucosidase (SGD) catalyzed deglucosylation of **3a** to form hemiacetal **4**.[12,13] Compound **4** is then channeled into one or several MIA pathway branches, though the rearrangements and enzyme-catalyzed steps of these branching reactions are poorly understood (Chapter 1).[14] Entry into the heteroyohimbine class of alkaloids (**5a-c**) is likely controlled by one or several NADPH-dependent reductases (Figure 6.1).[15-17]

This chapter that describes the total synthesis of *glucosylated* des-vinyl strictosidine stereoisomers (**12a** and **12**, Figure 6.2) and the identification of stereochemical restrictions in the early steps of MIA biosynthesis,[18] that if bypassed, may allow the generation of libraries of heteroyohimbine stereoisomers from heterologously expressed protein.[19]

6.2 Results and discussion

6.2.1 SGD-catalyzed deglucosylation of C3 stereoisomers

SGD promotes not only the deglucosylation of its natural substrate, 3*(R)*-strictosidine **3a**, but also the epimer, 3*(S)*-vincoside **3b**.[20] SGD has an 80-fold higher catalytic efficiency (V_{max}/K_M) with **3a** compared to **3b** (Table 6.1).[18] The change in V_{max}/K_M is due to a change in the maximal rate (V_{max}); the Michaelis constant (K_M) is the same for the two epimers (Table 6.1).[18]

205

Table 6.1 Kinetic constants for deglucosylation and reduction of strictosidine **3a** and vincoside **3b**.

SGD[a]	V_{max} [U mg^{-1}][b]	K_M [mM][b]	V_{max}/K_M [U mg^{-1} mM^{-1}][b]
3a	14.0 ± 0.7	0.15 ± 0.03	93 ± 30
3b	0.17 ± 0.02	0.15 ± 0.04	1.1 ± 0.5
Reductase[b]	V_{max} [U mg^{-1}][b]	K_M [mM]	V_{max}/K_M [U mg^{-1} mM^{-1}][b]
Deglucosylated **3a**	0.027 ± 0.001	0.10 ± 0.02	0.26 ± 0.04
Deglucosylated **3b**	0.012 ± 0.0003	0.66 ± 0.04	0.019 ± 0.001

[a] Activity quantified by monitoring the disappearance of starting material at pH 6.0, 30 °C. [b] 1 U = 1 μmol product formed per minute at pH 7.0 and 30 °C.

The x-ray crystal structure of SGD from the closely related medicinal plant *Rauvolfia serpentina* (PDB: 2FJ6) shows that the C3 carbon atom of strictosidine **3a** is near the surface of the protein while the glucose moiety is in the interior of the binding site.[13] I hypothesize that the glucose moiety of **3a** and **3b** binds in an identical fashion in the active site of SGD, while the binding pocket for the more distal regions of the substrate, including the C3 carbon, is sufficiently flexible to accommodate **3b**.[21-23] An x-ray crystal structure of SGD in complex with vincoside **3b** may clarify the molecular basis of the relaxed stereoselectivity.

6.2.2 Synthesis of des-vinyl strictosidine

Chapter 5 described the synthesis and biochemical evaluation of des-vinyl secologanin analogs without glucosylation, and therefore, it was only possible to study the stereoselectivity of STS. To examine how different stereoisomers of des-vinyl analogs function in a biosynthetic context, I developed a method to synthesize *glucosylated* des-vinyl secologanin.

I accessed alternative configurations at C2 and C4 by applying the previously reported synthesis of acetal-protected des-vinyl secologanin aglycones.[24-26] The key step is, as for the previously described *O*-analogs (Chapter 5), Tietze's tandem Knoevenagel-hetero-Diels-Alder (KHDA) reaction, which assembles a 2,4-disubstituted dihydropyran from 7, 8, and an electron rich dienophile (*e.g.* vinyl ether). I adopted this strategy to obtain, for the first time, *O*-glucosylated des-vinyl strictosidine 12a and its stereoisomers 12 (Figure 6.2). Ir-catalyzed vinyl transfer from vinyl acetate to the corresponding alcohol formed the KHDA substrate, 2,3,4,6-tetrabenzyl vinyl-glucose 6.[27] The KHDA reaction of 6, 7, and 8, in the presence of potassium fluoride, afforded cycloadducts, and after partial purification of the desired regioisomers, methanolysis generated methyl ester 9 (section 6.4.2 and Figure 6.2, steps a-b). Pd-catalyzed debenzylation of 9 led to acetal-protected secologanin 10 as a mixture of inseparable stereoisomers (Figure 6.2, step c), and after a challenging acetal hydrolysis step, des-vinyl secologanin 11 was carried partially purified into either an enzymatic or a non-enzymatic Pictet-Spengler reaction to generate tetrahydro-β-carboline products 12a or 12 (Figure 6.2 and section 6.4.3).

Enzymatic synthesis by STS using 1 and 11 as substrates produced a single isomer 12a (Figure 6.2, step e) which was purified by preparative HPLC and characterized by ¹H NMR (section 6.4.3 and Appendix D). The doublet of a doublet at 5.8 ppm suggested a dihydropyran 2,4-*trans* configuration,[24,25] and the doublet at 4.8 ppm with a *J*-coupling constant of 7.8 Hz indicated a β-anomeric configuration for the glycoside linkage;[28] the
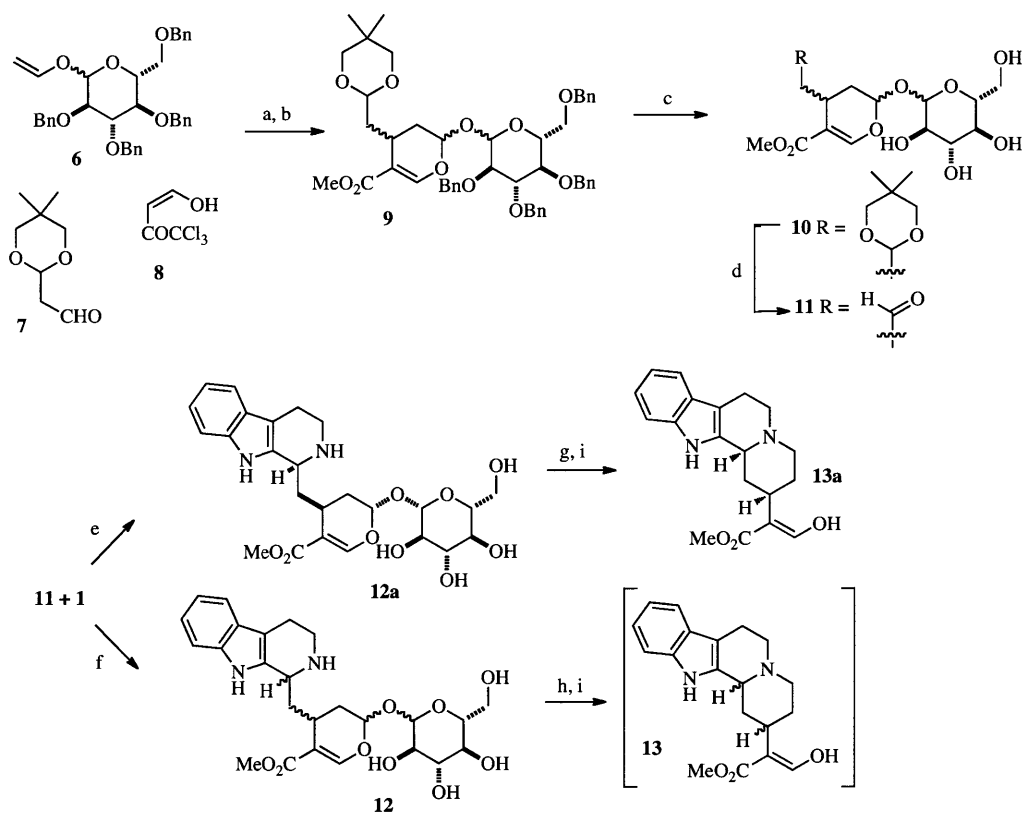
Figure 6.2 Synthesis of **12**, **12a**, **13**, and **13a**. (a) KF, toluene; (b) DBU, MeOH, 38% over 2 steps; (c) Pd/C, MeOH, 77%; (d) PPTS (2 eq, 0.2 M), acetone/H₂O (2:1), ~10%; (e) tryptamine **1**, *C. roseus* STS, NaP$_i$ (0.05 M, pH 7.0), 89%; (f) tryptamine **1**, maleic acid buffer (0.01 M, pH 2.0); (g) *C. roseus* SGD, citrate-phosphate buffer (0.15 M, pH 6.0); h) *B. stearothermophilus* α-glucosidase, almond β-glucosidase, SGD, citrate phosphate buffer (0.15 M, pH 6.0); (i) reductase activity from *C. roseus* cell suspension culture, NADPH, sodium phosphate buffer (0.05 M, pH 7.0).

208

same relative configuration as strictosidine **3a** (Figure 6.1). Therefore, STS stringently prefers the 2,4-*trans*-β-glucoside of **11** when challenged with an array of stereoisomers.

Chemical synthesis by acid catalysis (maleic acid buffer, pH 2.0) afforded isomers **12**, from **1** and **11** (Figure 6.2, step f), which were isolated by preparative HPLC and characterized by ^{1}H NMR (section 6.4.3 and Appendix D). UPLC-MS resolved six peaks, each with a mass consistent with des-vinyl strictosidine **12** (*m/z* 505) (Figure 6.3A (i): pk 1-6). Therefore, chemical synthesis of **12** bypasses the stringent stereospecificity of STS and provides stereoisomers for assessing stereochemical restrictions of subsequent biosynthetic steps.

6.2.3 Deglucosylation of des-vinyl strictosidine

The second enzyme in the pathway, SGD, turns over both **3a** and **3b** (section 6.2.1), and as expected, SGD efficiently catalyzed the deglucosylation of the single tetrahydro-β-carboline isomer **12a**, suggesting that the vinyl group is not essential for deglucosylation activity.

When SGD was incubated with all isomers of **12**, only two of the six separable peaks decreased in area (Figure 6.3A (ii): pk 1 and pk 3), suggesting that SGD, in addition to STS, imposes stereochemical restrictions in the MIA biosynthetic pathway. Pk 3 co-elutes with **12a** (not shown) and the minor decrease in peak area suggests turnover of only the H15/H21-*trans* configuration of **3**. Pk 1, however, contains two stereoisomers
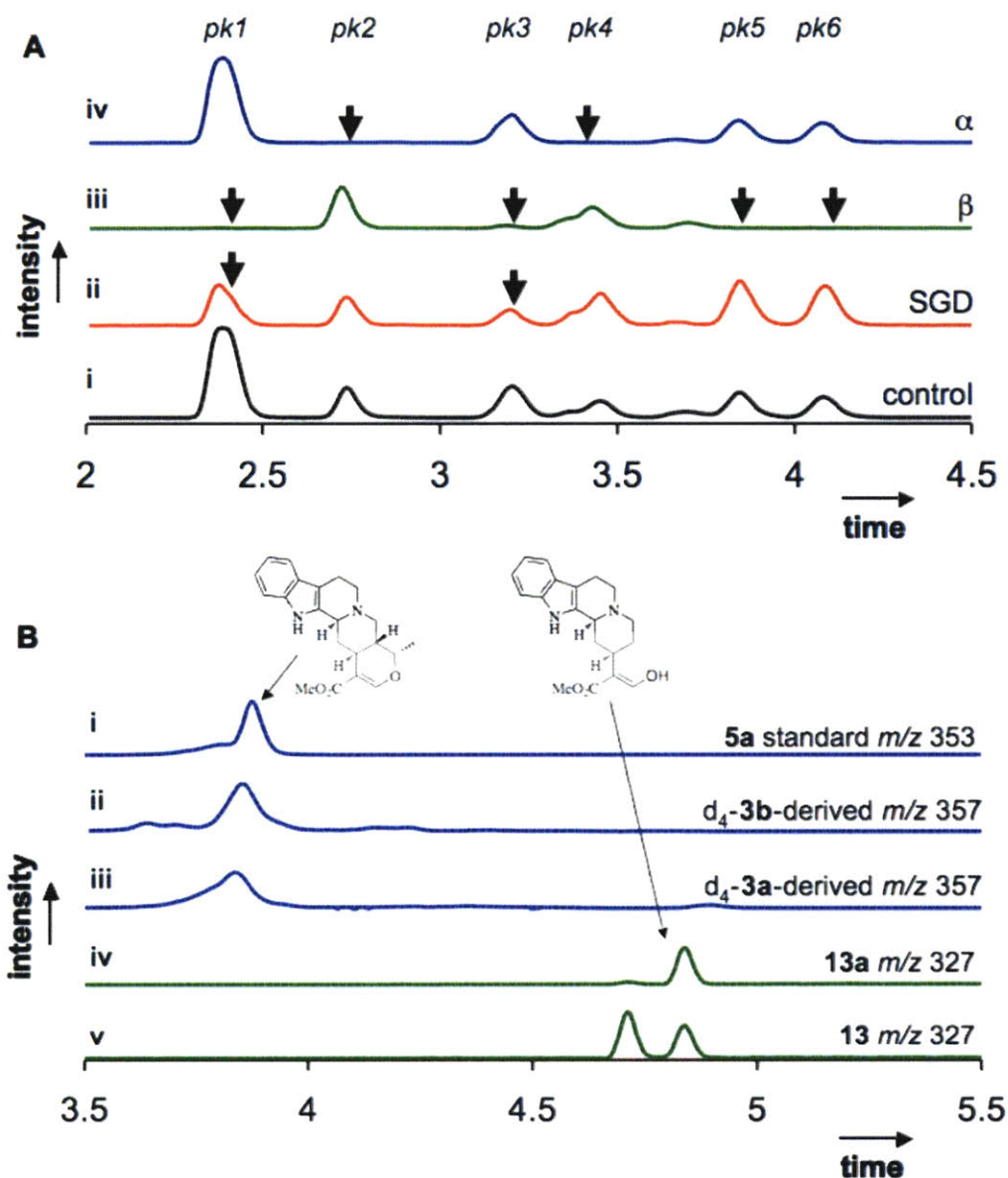
Figure 6.3 LC-MS traces that show: (A) Remaining starting material after deglucosylation of **12** (m/z 505.1 ± 0.5) by *C. roseus* SGD (**ii**), almond β-glucosidase (β, **iii**), and *B. stearothermophilus* α-glucosidase (α, **iv**). (**i**) no-enzyme control. NMR spectra of isolated peaks are found in Appendix D. (B) Products of the reductase assay. (**i**) ajmalicine **5a** standard (m/z 353); (**ii**) reduction of deglucosylated d_4-vincoside **3b** (m/z 357); (**iii**) reduction of deglucosylated d_4-strictosidine **3a** (m/z 357); (**iv**) reduction of deglucosylated **12a** yields a single product, **13a**; (**v**) reduction of **12** yields two separatable stereoisomers **13**.

with H15/H21-*cis* configuration, as evidenced by the triplet signals at 5.9 and 5.7 ppm assigned to H21 (section 6.4.3 and Appendix D).[24,25] The two isomers in pk 3 have β-anomeric configuration since the *J*-coupling constant between H1' and H2' is 8.0 Hz.[28] Therefore, although SGD accommodates alternatives stereoisomers, the anomeric configuration must be β for turnover to occur.

The complete deglucosylation of **12** is necessary for a comprehensive assessment of the stereoselectivity of the NADPH-dependent reductase – the subsequent enzyme in the heteroyohimbine alkaloid biosynthetic pathway. SGD only deglucosylates a few stereoisomers, but I hypothesized that bacterial glucosidases may have broader substrate scopes,[29] which would enable the deglucosylation of all stereoisomers. Therefore, I assayed **12** with the commercially available enzymes: *Bacillus stearothermophilus* α-glucosidase and almond β-glucosidase. Almond β-glucosidase is considerably more permissive than SGD, catalyzing the deglucosylation of isomers of **12** in four out of the six peaks that could be resolved by LC (Figure 6.3A (iii): pk 1, 3, 5, and 6; and Appendix D); *B. stearothermophilus* α-glucosidase facilitated the removal of the other two peaks (Figure 6.3A (iv): pk 2 and 4; and Appendix D). Hence, by using glucosidases from different, unrelated, metabolic pathways, it is possible to bypass the native biosynthetic pathway to fully deglucosylate **12**.

6.2.4 NADPH-dependent reduction of C3 stereoisomers

In the heteroyohimbine biosynthetic pathway, one or several reductases catalyze the NADPH-dependent reduction of deglycosylated strictosidine **3a** to form MIAs such as **5a-c** (Figure 6.1). A cell-free extract from *C. roseus* cell suspension cultures, obtained by N. Yerkes in the O'Connor lab, reconstituted the reductase activity. Controls showed that **3a**, SGD, NADPH, and the reductase activity are each a necessary component for the formation of a reduced product; the product elutes near an authentic standard of ajmalicine **5a** (*m/z* 353, Figure 6.3B). N. Yerkes determined the steady-state kinetics for reduction of the natural substrate, deglucosylated strictosidine **3a**, and the unnatural substrate, deglucosylated vincoside **3b** (Table 6.1). The reductase shows a 15-fold catalytic preference (V_{max}/K_M) for the 3*(S)* configuration of strictosidine **3a** (0.26 U M^{-1} mg^{-1}) over the 3*(R)* stereochemistry of vincoside **3b** (0.019 U M^{-1} mg^{-1}). This is mainly due to a 7-fold difference in K_M (0.1 mM compared to 0.7 mM with **3a** and **3b**, respectively). The catalytic differentiation between **3a** and **3b** suggests that the reductase activity derives from a MIA pathway. However, since the enzyme is a cleared cell-lysate we cannot rigorously exclude the involvement of additional reductases in the turnover of deglucosylated **3a** and **3b**.

6.2.5 NADPH-dependent reduction of deglucosylated des-vinyl strictosidine

The reductase activity, provided by N. Yerkes, converts deglucosylated des-vinyl strictosidine **12a** into **13a** (Figure 6.2, section 6.4.5, and Appendix D), and **12** into two separatable reduced products **13**. One separatable peak of **13** co-elutes with **13a** (Figure

6.3B). Since the reduced product contains two stereogenic centers and only two diastereomers separate under the chromatographic conditions, we concluded that the second peak contains the second diastereomer. Since **12** is completely deglucosylated by the mixture of glucosidases, and because all deglucosylated isomers **12** are completely consumed upon addition of the reductase activity (Figure 6.4), the reductase turns over all secologanin dihydropyran configurations. Therefore, heteroyohimbine alkaloid biosynthesis appears to contain reductases that are capable of reducing a variety of substrates; this promiscuity may prove useful in the chemoenzymatic synthesis of compound libraries.

6.3 Conclusions

The first step of the heteroyohimbine alkaloid biosynthetic pathway, catalyzed by STS, has a stringent substrate specificity, including stereoselectivity, to ensure the integrity of the first committed intermediate in MIA biosynthesis, strictosidine **3a**. SGD shows strict stereocontrol, but accepts at least one unnatural dihydropyran stereoisomer with H15/H21-*cis* relative stereo configuration. Chemical synthesis and the recruitment of glucosidases from unrelated metabolic pathways highlighted the potential to bypass stereochemical restrictions in MIA biosynthesis and to diversify alkaloid production. *C. roseus* cell culture extracts provided a reductase fraction, which harbors at least one enzyme that converts unnatural stereoisomers to reduced alkaloids *in vitro*. Once cloned, this enzyme can be useful in heterologous expression systems to yield novel heteroyohimbine alkaloid derivatives.
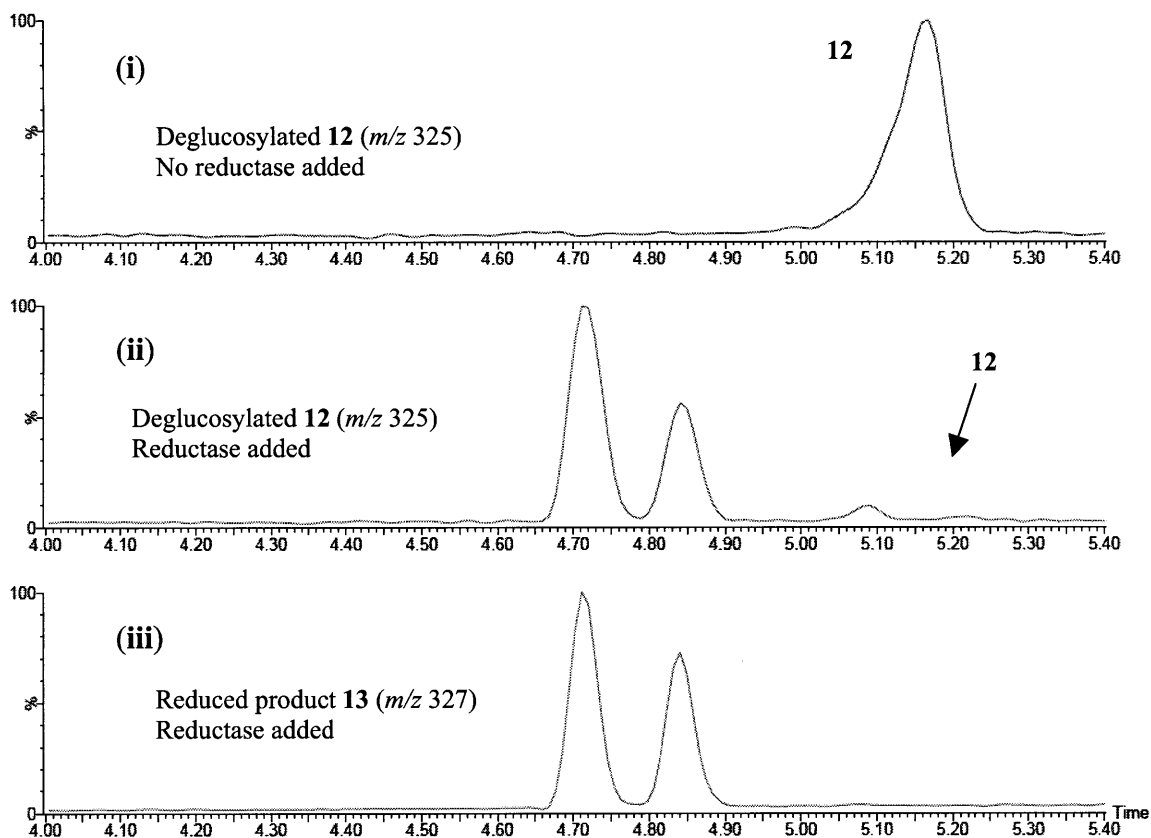
Figure 6.4 LC-MS traces show that deglucosylated des-vinyl strictosidine **12** (**i**, m/z = 325) is completely converted by the reductase activity to form **13** (**iii**, m/z = 327), as evidenced by the absence of m/z 325 in the reductase-supplemented reaction in LC-MS analysis (**ii**, m/z 325). Trace (**i**) shows the product of incubation of **12** with *B. stearothermophilus* α-glucosidase, almond β-glucosidase, and *C. roseus* strictosidine-β-D-glucosidase, without reductase. Traces (**ii**) and (**iii**) are extracted masses from the same reaction, where reductase has been added. The two peaks with m/z 325 observed in trace (**ii**) correspond to the isotope peaks of m/z 327.

214

6.4 Experimental methods

6.4.1 General methods and analytical techniques

All chemicals were used as received from the supplier without further purification. All glassware were oven dried, evacuated, and filled with argon. Secologanin was purified from *Lonicera tatarica* as previously described.[30] Strictosidine **3a** and strictosidine analogs were either prepared enzymatically using CrSTS or using chemical synthesis in aqueous maleic acid buffer (0.1 M, pH 2.0) and isolated by preparative HPLC. STS and SGD were expressed in *E. coli* BL21(DE3) and purified by affinity chromatography using a hexa-His tag as previously described.[31] α-Glucosidase from *B. stearo-thermophilus* (104 U mg^{-1}, where 1 U = 1 μmol D-glucose from *p*-nitrophenyl α-D-glucoside per minute at pH 6.8 at 37 °C) and β-glucosidase from almonds (26 U mg^{-1}, where 1 U = 1 μmol of salicin per minute at pH 5.0 at 37 °C) were purchased from Sigma-Aldrich (St. Louis, MO).

UPLC analysis was performed using an Acquity Ultra Performance BEH C18 column with a 1.7 μm particle size, 2.1 x 100 mm dimension, and a flow rate of 0.6 mL min^{-1}. The column elution was coupled to MS analysis carried out using a Micromass LCT Premier TOF Mass Spectrometer with an ESI source. Both modules are from Waters Corporation (Milford, MA). The capillary and sample cone voltages were 3000 V and 30 V, respectively. The desolvation and source temperatures were 300 and 100 °C, respectively. The cone and desolvation gas flow rates were 60 and 800 L h^{-1}. Analysis was performed with MassLynx 4.1 and integrations were carried out using QuantLynx.

Proton nuclear magnetic resonance (^{1}H NMR and ^{13}C NMR) spectra were recorded on Varian 300 MHz, Varian 500 MHz, or Brucker 400 MHz spectrometers.

6.4.2 Synthesis of acetal-protected des-vinyl secologanin 10

2-(5,5-dimethyl-1,3-dioxan-2-yl)ethanal 7
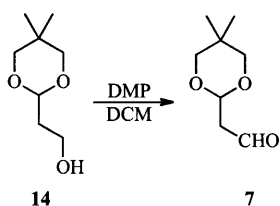
Dess-Martin periodinane (DMP, 5.1 g, 12 mmol, 1.2 eq) was added to a solution of 2-(5,5-dimethyl-1,3-dioxan-2-yl)ethanol **14** (1.6 g, 10 mmol, 1.0 eq) in DCM (100 mL) at 0 °C. The reaction was stirred at 0 °C for 5 min, warmed to room temperature and then stirred for an additional 45 min. The reaction was diluted with DCM (100 mL) and quenched by the successive addition of a saturated aqueous solution of sodium bicarbonate (50 mL) and a concentrated aqueous solution of sodium thiosulfate (50 mL). The layers were separated and the aqueous layer was extracted with DCM (3 x 20 mL). The combined organic layers were washed with a (1:1) mixture of sodium bicarbonate/sodium thiosulfate and dried over anhydrous sodium sulfate. After evaporation of the solvent in vacuo, the residue was purified by flash column chromtography (SiO$_2$, hex/ether 3:1) to afford 2-(5,5-dimethyl-1,3-dioxan-2-yl)ethanal **7** (1.5 g, 9.7 mmol, 97%) as a clear oil.[32] ^{1}H NMR (CDCl$_3$): δ 9.81 (t, 1H, J = 2.4), 4.86 (t, 1H, J = 4.5), 3.61 (d, 2H, J = 10.0), 3.45 (d, 2H, J = 10.5), 2.68 (dd, 1H, J = 2.4, 4.6), 1.17 (s, 3H), 0.72 (s, 3H); ^{13}C NMR (CDCl$_3$): δ 199.57, 97.81, 77.17, 48.48, 29.99,

22.84, 21.75; ESI-MS($C_8H_{15}O_3^+$): *m/z* calculated: 159.1016 [M+H]$^+$, found: 159.1014 [M+H]$^+$.

α/β-vinyl-2,3,4,6-tetra-O-benzyl-glucopyranose 6
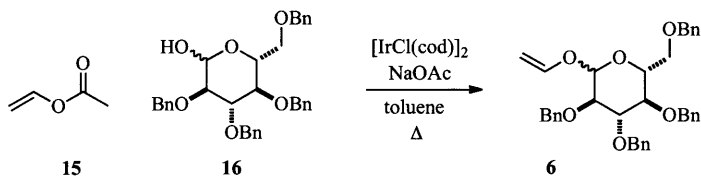
In a Schlenk flask under a stream of argon, [IrCl(cod)]$_2$ (0.050 g, 0.074 mmol, 0.02 eq) was added to a toluene solution (7.5 mL) of vinyl acetate **15** (0.64 g, 7.4 mmol, 2.0 eq), 2,3,4,6-tetra-*O*-benzylglucopyranose **16** (2.0 g, 3.7 mmol, 1.0 eq), and sodium acetate (0.36 g, 4.4 mmol, 1.2 eq). The mixture was heated to 95 °C and stirred at this temperature for 90 min. A second portion of vinyl acetate **15** was then added (0.32 g, 3.7 mmol, 1.0 eq) and the reaction was allowed to continue for a total of 2 h. The reaction was diluted with wet ether (40 mL) and filtered through SiO$_2$ eluting with ether. After evaporation of the solvent, the residue was purified by flash column chromatography (SiO$_2$, hex/EtOAc, 4:1) to afford vinyl-2,3,4,6-tetra-*O*-benzyl-glucopyranose **6** (1.7 g, 3.0 mmol, 80%) as an oil (0.62:0.38 mixture of α- and β-anomers), which solidified on standing. ^{1}H NMR (CDCl$_3$): δ 7.44-7.05 (m, 20H), 6.48 (dd, 0.62H, J = 6.4, 13.9), 6.36 (dd, 0.38H, J = 6.5, 14.1), 5.10-4.39 (m, 10H), 4.24 (ddd, 1H, J = 1.7, 6.5, 16.4), 3.87-3.41 (m, 6H); ESI-MS($C_{36}H_{39}O_6^+$): *m/z* calculated: 567.2741 [M+H]$^+$, observed: 567.2752 [M+H]$^+$.

217

Methyl 4-((5,5-dimethyl-1,3-dioxan-2-yl)methyl)-2-(α/β-glucopyranose)-3,4-dihydro-2H-pyran-5-carboxylate **9**
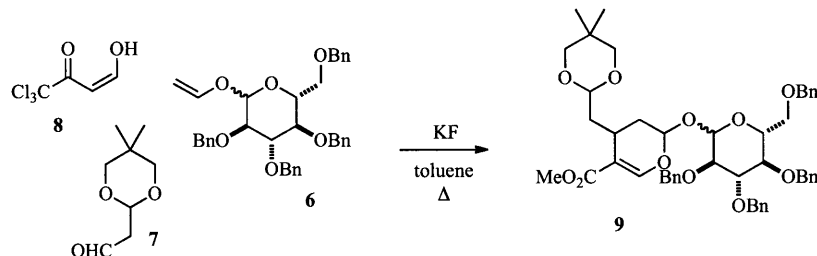
1,1,1-trichloro-4-hydroxybut-3-en-2-one **8** (1.4 g, 7.4 mmol, 2.0 eq), synthesized as described previously[33] and 2-(5,5-dimethyl-1,3-dioxan-2-yl)ethanal **7** (1.3 g, 8.2 mmol, 2.2 eq) were dissolved in toluene (7.5 mL) at 0 °C. Potassium fluoride was added (0.19 g) and the resulting mixture quickly changed color to orange. 1-Vinyl-2,3,4,6-tetra-*O*-benzyl-glucopyranose **6** (2.1 g, 3.7 mmol, 1.0 eq) was then added drop-wise and the resulting reaction was heated to 100 °C. After a total reaction time of 8 h, the reaction was diluted with ether and filtered through water-deactivated aluminum oxide. Evaporation of the solvent and purification of the residue by flash column chromatography (SiO$_2$, hex/MTBE, 8:1 to 5:1) afforded partially purified cycloadduct (six stereoisomers by NMR, four major stereoisomers). The cycloadducts were dissolved in anhydrous methanol (10 mL) and DBU was added (0.1 mL). After 20 min stirring at room temperature, the solvent was evaporated and the residue was purified by flash column chromatography (SiO$_2$, hex/MTBE, 5:1 to 3:1) to afford analytically pure methyl 4-((5,5-dimethyl-1,3-dioxan-2-yl)methyl)-2-(α/β-glucopyranose)-3,4-dihydro-2H-pyran-5-carboxylate **9** (1.2 g, 1.4 mmol, 38%) as a mixture of diastereomers. ^{1}H NMR (CDCl$_3$), overlapping signals: δ 7.48-7.06 (m, 1H), 7.37-7.06 (m, 20H), 5.58 (t, 0.20H, J

= 2.9), 5.51-5.45 (m, 0.31H), 5.40 (t, 0.17H, J = 2.9), 5.34-5.27 (m, 0.42H), 5.02-4.64 (m, 4H), 4.64-4.39 (m, 5H), 3.80-3.22 (m, 8H), 2.93-2.75 (m, 1H), 2.37-1.75 (m, 4H), 1.66-1.50 (m, 1H), 1.25-1.03 (m, 3H), 0.74-0.52 (m, 3H); ESI-MS($C_{48}H_{57}O_{11}^{+}$): *m/z* calculated: 809.3895 [M+H]+, found: 809.3880 [M+H]+.

Methyl 4-((5,5-dimethyl-1,3-dioxan-2-yl)methyl)-2-(α/β-glucopyranose)-3,4-dihydro-2H-pyran-5-carboxylate 10

To a solution of the tetrabenzylated starting material **9** (1.1 g, 1.4 mmol) in MeOH (50 mL) was added water (1.5 mL) and then Pd/C (cat.). The flask was evacuated and filled with an H_2-atmosphere. The resulting heterogenous mixture was stirred for 2 h at room temperature or until TLC (DCM/MeOH, 9:1) indicated complete conversion. Filtration through celite and evaporation of the solvents afforded a yellow oil that was subjected to flash column chromatography (SiO$_2$, DCM/MeOH, 9:1) to afford the debenzylated product **10** as a white foam (0.45 g, 1.0 mmol, 74%); ^{1}H NMR (MeOD): δ 7.51-7.45 (m, 1H), 5.67 (t, 0.26H, J = 2.5), 5.59-5.54 (m, 0.38H), 5.50 (t, 0.22H, J = 2.5), 5.47 (dd, 0.11H, J = 2.3, 9.2), 5.26 (d, 0.12H, J = 3.7), 5.16 (d, 0.21H, J = 3.7), 4.72 (d, 0.17H, J = 7.9), 4.67-4.60 (m, 1H), 4.57 (t, 0.38H, J = 4.3), 4.52 (d, 0.24H, J = 7.8), 3.92-3.77 (m, 1H), 3.72-3.68 (m, 3H), 3.69-3.40 (m, 6H), 3.26-3.09 (m, 1H), 2.94-2.73 (m, 1H), 2.37-

1.79 (m, 4H), 1.62 (ddd, 0.3H, J = 4.8, 9.7, 14.2), 1.17 (s, 3H), 0.73 (s, 3H); ESI-MS($C_{20}H_{33}O_{11}^+$): *m/z* calculated: 449.2019 [M+H]+, found: 449.2025 [M+H]+.

6.4.3 Synthesis of des-vinyl strictosidine **12a** and **12** and deglucosylation of **12**

Deprotection of acetal-protected des-vinyl secologanin 10
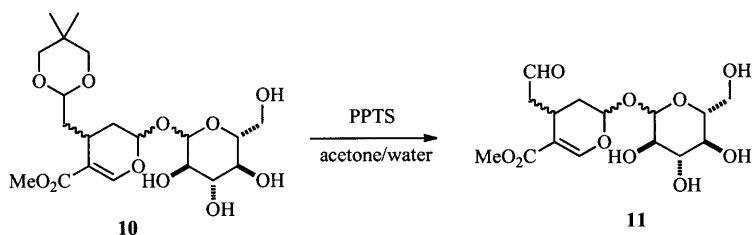
PPTS (1.56 mmol, 2 eq) was added to a solution of methyl 4-((5,5-dimethyl-1,3-dioxan-2-yl)methyl)-2-(glucopyranose)-3,4-dihydro-2H-pyran-5-carboxylate **10** (0.35 g, 0.78 mmol) to make a 0.2 M solution of PPTS in acetone/water (2:1). The mixture was heated to 65 °C for 35 h, after which it was quenched by addition of solid sodium carbonate (0.5 g). After filtration, the solvents were evaporated in the presence of toluene. Purification by normal phase flash column chromatography (SiO₂, DCM/MeOH, 85:15) followed by reverse-phase flash column chromatography (C_{18}-SiO₂, packed in MeOH, equilibrated to H₂O, run with H₂O/MeOH, 99:1 to 8:2) afforded partially purified aldehyde **11** (68 mg, est. 20% yield). NMR signals at δ 9.80-9.60 (m) are indicative of the aldehyde functional group; signals associated with the dioxane protective group are absent; and, the aldehyde exists at least partially in its hydrated form due to the low intensity of the aldehyde proton signal (NMR spectrum in Appendix D). ESI-MS($C_{15}H_{23}O_{10}^+$) *m/z* calculated: 363.1286

[M+H]$^+$, found: 363.1287 [M+H]$^+$. Partially purified **11** was used in the Pictet-Spengler reactions with tryptamine **1**.

*Enzymatic synthesis of des-vinyl strictosidine **12a***
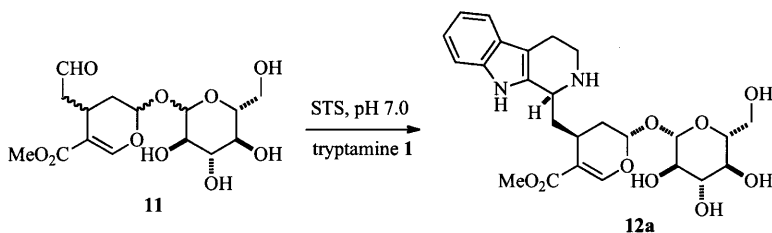
To a solution of the partially purified des-vinyl secologanin **11** (68 mg, 0.17 mmol, 8.5 mM, 1.0 eq) in sodium phosphate buffer (20 mL, 50 mM, pH 7.0) was added tryptamine **1** hydrochloride (44 mg, 0.22 mmol, 11 mM, 1.3 eq) and *C. roseus* STS (0.25 mL, 1 mol%). The reaction was followed by HPLC until no further product was formed. MeOH was added (200 mL) and the resulting precipitate was removed by filtration. After evaporation of most of the solvent the enzymatic product was purified by preparative HPLC (detection at 239 nm; solvent system: 25-50% MeCN in water with 0.1% TFA over 20 min; $r_{t,SM}$ = 4.5 min, $r_{t,product}$ = 8.5 min). The mobile phase was evaporated in vacuo and dried to yield an amorphous solid (2 mg, 0.004 mmol), which consisted of a single diastereomer **12**. ^{1}H NMR (MeOD): δ 7.76 (s, 1H, H-17), 7.45 (d, 1H, J = 7.8, H-12), 7.30 (d, 1H, J = 8.1, H-9), 7.13 (dt, 1H, J = 1.0, 7.1, H-10), 7.02 (dt, 1H, J = 0.9, 8.0, H-11), 5.75 (dd, 1H, J = 2.9, 9.2, H-2), 4.77 (d, 1H, J = 7.8, H-1'), 4.65 (d, 1H, J = 8.1, H-3), 3.93 (dd, 1H, J = 1.9, 11.7), 3.75 (s, 3H, OMe), 3.59 (dd, 1H, J = 7.0 11.8), 3.52-2.95 (m, 9H), 2.48 (ddd, 1H, J = 3.7, 10.0, 13.9), 2.06-1.92 (m, 3H, H-14); ESI-MS($C_{25}H_{33}N_2O_9^+$) *m/z* calculated: *m/z* 505.2181 [M+H]+, found: 505.2180 [M+H]+.

221

Chemical synthesis and deglucosylation of des-vinyl strictosidine isomers 12
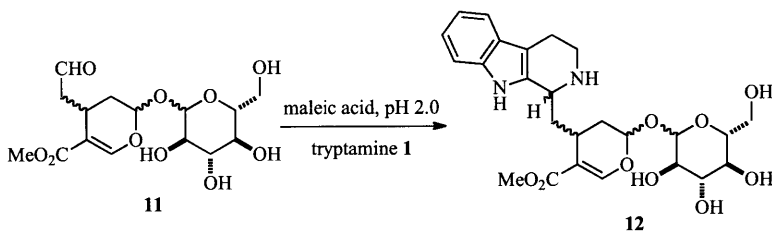
Des-vinyl secologanin **11** was reacted with tryptamine **1** hydrochloride in maleic acid buffer (10 mM, pH 2.0) for 48 h at 37 °C. The reaction products were collectively purified by preparative HPLC as described for the enzymatic product to afford an amorphous solid. Since the reaction products were complex mixtures of multiple diastereomers, complete structural characterization could not be performed. However, several representative signals in each ^{1}H NMR spectrum could be used to assign certain structural elements. (For NMR and LC-traces, see Appendix D.) To monitor the deglucosylation reaction of the different des-vinyl secologanin stereoisomers, either almond β-glucosidase (5 U) or *B. stearothermophilus* α-glucosidase (15 U) was added to **12** (0.24 mM final concentration) in citrate-phosphate buffer (0.15 M, pH 6.0) and aliquots (5 μL) were quenched at appropriate time points, analyzed by LC-MS, and integrated and normalized to an internal standard (yohimbine). The LC-MS chromatograms are shown in Appendix D and summarized as follows.

Pk 1: Two stereoisomers: 2,4-*cis* (t, δ 5 ppm), β-anomers (d, J = 8.0 Hz).[28] Pk 1 is turned over by SGD and almond β-glucosidase. Pk 2: One stereoisomer: 2,4-*cis* (t, δ 5 ppm), and α-anomer (d, J = 3.8 Hz). Pk 2 is turned over by the α-glucosidase. Pk 3 and 4:

Several stereoisomers, 2,4-*trans* and 2,4-*cis* configurations, as evidenced by the br-s and dd in the 5-6 ppm region, and a mixture of alpha and beta anomers. Pk 3 is turned over by SGD and almond-β-glucosidase; Pk 4 is turned over by α-glucosidase. Pk 5 and 6: Several stereoisomers, 2,4-*trans* and 2,4-*cis* configurations, as evidenced by the br-s and dd in the 5-6 ppm region. 2,4-*cis* appears to be dominating in these two peaks. Both α and β anomers are present. Pk 5 is converted by both α- and β-glucosidases; Pk 6 likely contains mainly β-anomers, since it is converted mainly by almond β-glucosidase.

6.4.4 Isolation of reductase activity from cell suspension cultures

Cells were recovered from two-week-old *C. roseus* PC510 cell suspension cultures (DSMZ, Germany) by filtration through cheesecloth. The cells were suspended in borate buffer (0.1 M, pH 7.0) containing polyvinylpyrrolidone (PVP, 0.3% w/v) and β-mercaptoethanol (20 mM) and ground in a blender (4 °C, 10 min). The cell extract was centrifuged (10min, 6,000g, 4 °C) and subjected to ammonium sulfate precipitation near 0 °C. The 40-70% fraction was recovered and desalted by passing the solution through a Sephadex G25 column and eluting with sodium phosphate buffer (30 mM, pH 7.0), supplemented with sodium chloride (50 mM) and glycerol (10% v/v). The desalted eluent was concentrated using a centrifugal filter device (AmiconUltra 10-kDa cut-off, Millipore Corp., Billerica, MA), aliquoted for single-use fractions, and stored at -20 °C.

6.4.5 Deglucosylation and reduction of 12 and 12a to form 13 and 13a, respectively

12 or 12a   **13 or 13a**
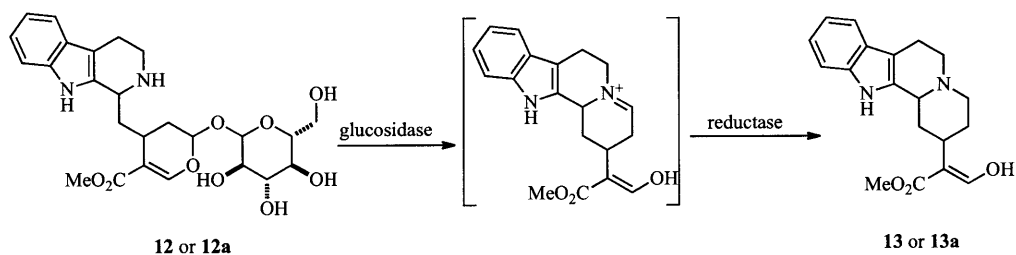
Des-vinyl secologanin **12a**, a single isomer, was incubated with *C. roseus* SGD in citrate-phosphate buffer (0.15 M, pH 6.0) for 30 min at 30 °C and then with the reductase fraction (section 6.4.4). The resulting product **13a** was purified by semi-preparative HPLC and analyzed by NMR. ^{1}H NMR (MeOD): δ 7.59 (s, 1H), 7.45 (d, J = 7.9, 1H), 7.34 (td, J = 0.8, 8.1, 1H), 7.11 (ddd, J = 1.2, 7.1, 8.2, 1H), 7.04 (ddd, J = 1.0, 7.1, 8.0, 1H), 4.62-4.55 (m, 1H), 3.77-3.72 (m, 1H), 3.62 (s, 3H), 3.61-3.53 (m, 3H), 2.89-2.73 (m, 3H), 2.39 (ddd, J = 1.9, 3.7, 13.3, 1H), 1.69-1.64 (m, 2H), 1.57 (ddd, J = 5.0, 12.1, 13.2, 1H); ESI-MS($C_{19}H_{23}N_2O_3^+$) *m/z* calculated: 327.1703 [M+H]$^+$, found: 327.1701 [M+H]$^+$. Des-vinyl secologanin **12** (0.24 mM final concentration), the mixture of all isomers, was incubated with both almond β-glucosidase (5 U) and *B. stearothermophilus* α-glucosidase (15 U) in citrate-phosphate buffer (0.15 M, pH 6.0) for 30 min at 30 °C, which resulted in the complete deglucosylation of **12**. After complete deglucosylation the reductase fraction (section 6.4.4) was added; after overnight incubation, all deglucosylated products had disappeared (Figure 6.3) and the mixture of reduced products **13** was separated by LC-MS (Figure 6.3)

6.4.6 Kinetics

Steady-state kinetic constants for strictosidine **3a** (15-1000 μM) or vincoside **3b** (30 to 1100 μM) using SGD (73 μg) was determined in citrate-phosphate buffer (0.15 M, pH 6.0) at 30 °C (0.200 mL reaction volume). Aliquots (10 μL) were quenched in 1 mL methanol, containing yohimbine (0.5 μM) as an internal standard, at appropriate time points. The samples were centrifuged (1 min, 13,000g, 4 °C) to remove particulates and then analyzed by LC-MS. The starting material and products were separated using a gradient of 10-40% MeCN in water with 0.1% formic acid over 5 minutes. The disappearance of strictosidine **3a** or vincoside **3b** was monitored by peak integration and normalized to the area of the internal standard. Seven concentrations were analyzed for each substrate, and the amount of **3a** or **3b** remaining after reaction at each time point was correlated to peak area by comparing to a standard curve. Each concentration was assayed at least three times and the averaged value is reported with standard deviations. The data were fitted using non-linear regression to the Michaelis-Menten or Northrop equations using OriginPro 7 (OriginLab, Northhampton, MA).

Steady-state kinetic constants using the reductase activity were obtained by performing a coupled assay with isotopically labeled strictosidine d_4-**3a** or vincoside d_4-**3b** in the presence of SGD. The isotope-labeled versions of the substrates were used to ensure that no trace amounts of alkaloid present in the cell free extract would interfere with kinetic analyses. Briefly, d_4-**3a** or d_4-**3b** (5 to 3000 μM) were diluted in citrate-phosphate buffer (0.15 M, pH 6.0, 0.05 mL final volume) and SGD (2.5 μg) was added. Under these

conditions, the reactions with the highest concentration of deuterated SGD substrate is completely deglucosylated after 30 min at 30 °C. After 30 min, deglucosylated d_4-**3a** or d_4-**3b** (50 μL volume) was added to a solution containing NADPH (2.4 mM final concentration) and cell free extract (prepared as described above, 2.4 mg total protein mL^{-1}) to a total volume of 200 μL. The reaction mixture was incubated at 30 °C and aliquots (15 μL) were quenched at appropriate time points and treated as described for the SGD kinetic analysis. The increase in the area under the product peak was used to obtain the initial rates of the reduction reaction. Each concentration was assayed at least three times and the averaged value is reported with standard deviations. The data were fitted by the same method as described for SGD kinetics.

6.5 References

1. De Camp, W.H. "The FDA perspective on the development of stereoisomers", *Chirality* **1989**, *1*, 2-6.

2. Crossley, R., Ed. "Chirality and biological activity of drugs", CRC Press **1995**.

3. Fergus, S.; Bender, A.; Spring, D.R. "Assessment of structural diversity in combinatorial synthesis", *Curr. Opin. Chem. Biol.* **2005**, *9*, 304-309.

4. van der Heijden, R.; Jacobs, D.I.; Snoeijer, W.; Didier, H.; and Verpoorte, R. "The *Catharanthus* alkaloids: pharmacognosy and biotechnology", *Curr. Med. Chem.* **2004**, *11*, 607-628.

5. Ghisalberti, E.L.; Pennacchio, M.; Alexander, E. "Survey of secondary plant metabolites with cardiovascular activity", *Pharmaceutical Biology* **1998**, *36*, 237-279.

6. Brevetti, G.; Chiariello, M.; Verrienti, S.; Spena, M.; Desiderati, M.; Condorelli, M. "Beneficial effect of papaverine plus raubasine in peripheral arterial insufficiency", *Angiology* **1983**, *34*, 517-526.

7. Li, S.; Long, J.; Ma, Z.; Xu, Z.; Li, J.; Zhang, Z. "Assessment of the therapeutic activity of a combination of almitrine and raubasine on functional rehabilitation following ischaemic stroke", *Curr. Med. Res. Opin.* **2004**, *20*, 409-415.

8. Roquebert, J.; Demichel, P. "Inhibition of the alpha1- and alpha2-adrenoceptor-mediated pressor response in pithed rats by raubasine, tetrahydroalstonine and akuammigine", *Eur. J. Pharmacology* **1984**, *106*, 203-205.

9. Stöckigt, J.; Zenk, M.H. "Strictosidine (isovincoside): the key intermediate in the biosynthesis of monoterpene indole alkaloids", *J. Chem. Soc., Chem. Commun.* **1977**, 646-648.

10. Treimer, J.F.; Zenk, M.H. "Purification and properties of strictosidine synthase, the key enzyme in indole alkaloid formation", *Eur. J. Biochem.* **1979**, *101*, 225-233.

11. Ma, X,; Panjikar, S.; Koepke, J.; Loris, E.; Stöckigt, J. "The structure of *Rauvolfia serpentina* strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins", *Plant Cell* **2006**, *23*, 532-547.

12. de Waal, A.; Meijer, A.H.; Verpoorte, R. "Strictosidine synthase from *Catharanthus roseus*: purification and characterization of multiple forms", *Biochem. J.* **1995**, *306*, 571-580.

13. Barleben, L.; Panjikar, S.; Ruppert, M.; Koepke, J.; Stöckigs, J. "Molecular architecture of strictosidine glucosidase: the gateway to the biosynthesis of the monoterpene indole alkaloid family", *Plant Cell* **2007**, *19*, 2886-2897.

14. Heinstein, P,; Höfle, G.; Stöckigt, J. "Involvement of cathenamine in the formation of N-analogues of indole akaloids", *Planta Med.* **1979**, *37*, 349–357.

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

16. Stöckigt, J.; Hemscheidt, T.; Höfle, G.; Heinstein, P.; Formacek, V. "Steric course of hydrogen transfer during enzymatic formation of 3.alpha.-heteroyohimbine alkaloids", *Biochemistry* **1983**, *22*, 3448-3452.

17. O'Connor, S.E.; Maresh, J.J. "Chemistry and biology of monoterpene indole alkaloid biosynthesis", *Nat. Prod. Rep.* **2006**, *23*, 532-547.

18. Bernhardt, P.; Yerkes, N.; O'Connor, S.E. "Bypassing stereoselectivity in the early steps of alkaloid biosynthesis", *Org. Biomol. Chem.* **2009**, *7*, 4166-4168.

19. Ulaczyk-Lesanko, A.; Hall, D.G. "Wanted: new multicomponent reactions for generating libraries of polycyclic natural products", *Curr. Opin. Chem. Biol.* **2005**, *9*, 266-276.

20. Yerkes, Y.; Wu, J.X.; McCoy, E.; Galan, M.C.; Chen, S.; O'Connor, S.E. "Substrate specificity and diastereoselectivity of strictosidine glucosidase", *Bioorg. Med. Chem. Lett.* **2008**, *15*, 3095-3098.

21. Mezzetti, A.; Schrag, J.; Cheong, C.; Kazlauskas, R.J. "Mirror-image packing in enantiomer discrimination: molecular basis of the enantioselectivity of *B. cepacia* lipase toward 2-methyl-3-phenyl-1-propanol", *Chem. Biol.* **2005**, *12*, 427-437.

22. Klebe, G. "Differences in binding stereoisomers to protein active site in supramolecular structure and function", Springer USA **2004**, pp. 31-52.

23. Mugford, P.F.; Wagner, U.G.; Jiang, Y.; Faber, K.; Kazlauskas, R.J. "Enantiocomplementary enzymes: classification, molecular basis for their enantiopreference, and prospects for mirror-image biotranformations", *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 8782-8793.

24. Tietze, L.F.; Meier, H.; Nutt, H. "Inter- and intramolecular hetero-Diels-Alder reactions, XXV. The tandem Knoevenagel hetero-Diels-Alder reaction with a

formylacetic acid equivalent. Synthesis of dihydropyrancarboxylates", *Chem. Ber.* **1989**, *122*, 643-650.

25. Tietze, L.F.; Meier, H.; Nutt, H. "Synthesis of (±)-secologanin aglucone *O*-ethyl ether and derivatives by tandem Knoevenagel hetero Diels-Alder reaction", *Lieb. Ann. Chem.* **1990**, 253-260.

26 Tietze, L.F. "Stereoselective synthesis of iridoid glycosides", *Angew. Chem. Int. Ed. Engl.* **1983**, *22*, 888.

27. Okimoto, Y.; Sakaguchi, S.; Ishii, Y. "Development of a highly efficient catalytic method for synthesis of vinyl ethers", *J. Am. Chem. Soc.* **2002**, *124*, 1590-1591.

28. Sinnott, M.L. *Carbohydrate chemistry and biochemistry: structure and mechanism*, Royal Society of Chemistry **2007**.

29. Zhengwu, S.; Eisenreich W.; Kutchan, T.M. "Bacterial biotransformation of 3α*(S)*-strictosidine to the monoterpenoid indole alkaloid vallesiachotamine", *Phytochemistry* **1998**, *48*, 293-296.

30. Galan, M.C.; O'Connor, S.E. "Semi-synthesis of secologanin analogues", *Tetrahedron Lett.* **2006**, *47*, 1563-1565.

31. Bernhardt, P.; McCoy, E.A.; O'Connor, S.E. "Rapid identification of enzyme variants for reengineered alkaloid biosynthesis", *Chem. Biol.* **2007**, *14*, 888-897.

32. Shi-Qi, P.; Winterfeldt, E. "Synthesis of malonaldehyde monoacetals", *Liebigs Ann. Chem.* **1989**, 1045-1047.

33. Tietze, L.F.; Meier, H.; Voss, E. "Highly efficient syntheses of alkyl 3,3-dialkoxypropanoates, alkyl 4-ethoxy-2-oxo-3-butenoates, and monoprotected malonaldehydes", *Synthesis* **1988**, 274-277.

CHAPTER 7

CONCLUSIONS AND FUTURE DIRECTIONS

7.1 Conclusions

Chapter 2 describes the development of two new screening methods to visually detect STS activity.[1,2] In the first method, a yellow colored pigment forms between a deglucosylated strictosidine intermediate and tryptamine.[1] The second method also depends on the deglucosylation of strictosidine, but addition of the dye BTB results in a blue to yellow color change.[2] The color change may result from formation of an ion pair between the positively charged deglucosylated intermediate and the negatively charged form of BTB.[2] The adaptation of the BTB-containing method for detecting STS activity on agar plates could facilitate screening of thousands of STS producing *S. cerevisiae* colonies.

Chapter 3 describes the use of the pigment-based assay to screen saturation mutagenesis libraries of STS in multi-well plates.[1] Screening of four libraries with mutations in the tryptamine binding site identified two enzyme variants, Val214Met and Phe232Leu STS, which catalyze the formation of new strictosidine analogs with substitution on the tetrahydro-β-carboline moiety, including chlorinated and brominated derivatives.[1] The Val to Met mutation allows substitutions at the 5-position of tryptamine; the Phe to Leu mutation increases the volume of the tryptamine binding pocket to tolerate 2'-substitutions. *C. roseus* hairy-root cultures derived from seedlings transfected with the Val214Met mutant of STS are able to form chlorinated MIA analogs not formed in the wild-type plant.[3] Therefore, Val214Met STS functions in the context of a metabolic pathway, despite its low activity with unnatural substrates. The agar plate adapted

screening method might be useful for the screening of multi-site saturation mutagenesis libraries for expanded substrate scopes or increased catalytic activity with substrate analogs.

Chapter 4 describes the characterization of OpSTS, which has broader amine and aldehyde substrate scopes compared to its closest homologs, CrSTS and RsSTS.[4] For example, OpSTS catalyzes the asymmetric biocatalytic formation of a variety of tetrahydro-β-carbolines; the first reported biocatalytic route to these molecules. The high enantiomeric excess (>98%) of the product that forms in the reaction between tryptamine and achiral aldehydes indicates that the seco-iridoid portion of the aldehyde, while important for an optimal catalytic rate, is not essential for maintaining the stereoselectivity of the Pictet-Spengler reaction. OpSTS also catalyzes the Pictet-Spengler reaction between secologanin and a variety of tryptamine analogs. For example, 2'*(R)*-tryptophanol, a substrate for Phe232Leu STS, and 5-methoxytryptamine that is not a substrate for CrSTS or any mutant of CrSTS, are substrates for OpSTS. As I mentioned in section 4.2.5, efforts to interconvert the substrate specificities of CrSTS and OpSTS failed, but future experiments may provide a molecular basis for the broader substrate scopes (section 7.2). It is possible that the electronics at the entry to the active site determines which substrates are accepted, by guiding the appropriate substrates into the binding pocket. Therefore, mutagenesis to the surface residues may change the substrate scope. Chapter 4 also describes a model for STS stereoselectivity.[4] In this model, a Glu residue acts as a general base to deprotonate one out of the four possible intermediates –

the isomer that results in the 3*(S)* stereogenic center of strictosidine. The model predicts that the stereoselectivity can be altered by inverting the active site, thereby relocating the catalytic base to the opposite face of the indole ring.

Section 7.2 describes potential use of the BTB-containing agar plate screening method to increase the activity of Val214Met STS, to investigate the molecular basis of the broad substrate specificity of OpSTS, to create a vincoside synthase, and to evolve a more thermostable OpSTS.

Chapters 5 and 6 describe the total synthesis of a series of des-vinyl secologanin *O*-analogs using a tandem Knoevenagel hetero-Diels-Alder reaction as the key step.[5,6] When secologanin is replaced by smaller acetal groups, CrSTS accepts not only the 2,4-*trans* dihydropyran configuration of secologanin, but also the unnatural 2,4-*cis* configuration.[5] The activity is dramatically lower when the glucose is replaced with a smaller group, likely due to the loss of productive non-covalent interactions between the glycoside and the enzyme; mutagenesis of the secologanin binding site may increase the activity and specificity with the unnatural aldehydes.

Total synthesis also affords a stereoisomer mixture of (glucosylated) des-vinyl secologanin.[6] Incubation of this stereoisomer mixture with a reconstituted heteroyohimbine alkaloid pathway that can generate ajmalicine, or an isomer of ajmalicine, identified stereochemical restrictions in the early steps of alkaloid
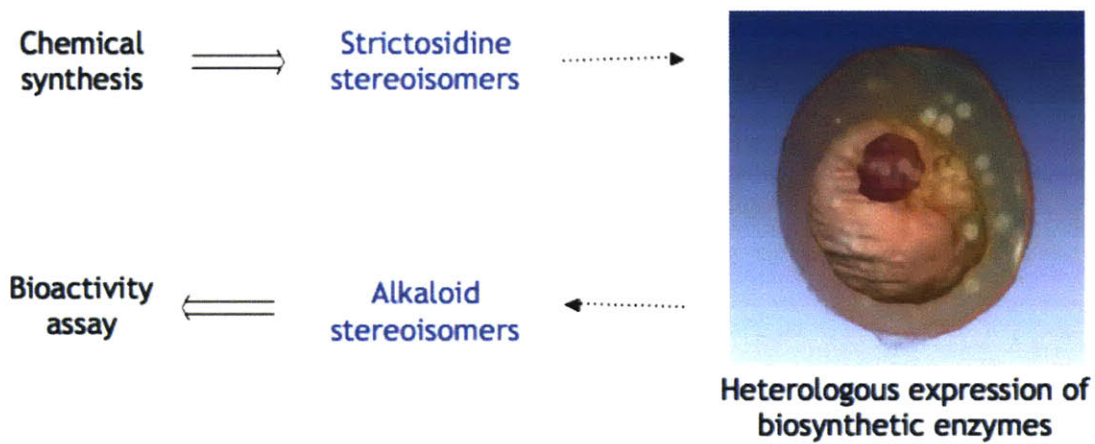
Figure 7.1 Chemo-enzymatic synthesis of monoterpene indole alkaloid stereoisomers libraries for biological evaluation.

biosynthesis (Chapter 6).[6] For example, when the glucose moiety is present, CrSTS catalyzes the formation of only one product when presented with an aldehyde substrate that is a mixture of multiple stereoisomers. Furthermore, SGD requires a β-glycoside linkage but catalyzes the deglucosylation of at least one unnatural (2,4-*cis*) stereoisomer. Chemical synthesis and the recruitment of enzymes from unrelated biosynthetic pathways bypasses the stereochemical restrictions and allowed an experiment that shows that the NADPH-dependent reductase converts all tested stereoisomers. By using a combination of glucosidases and a synthetic mixture of strictosidine stereoisomers, it may be possible to generate novel alkaloid libraries for bioactivity evaluation (Figure 7.1). Cloning of the plant NADPH-reductase and expression in a heterologous system will likely increase chemo-enzymatic product titers.

7.2 Future directions

The pH-indicator assay system adapted to agar plates, described in section 2.2.5, should ions allow screening of larger enzyme libraries (1,000-10,000 variants), which enables the pursuit of an answer to the following questions: What is the molecular basis of the broader substrate specificity of OpSTS? How can we improve the activity of STS with tryptamine analogs? How can we obtain a vincoside-forming enzyme? Is it possible to improve the activity of OpSTS of CrSTS with tryptamine analogs?

7.2.1 Molecular basis of the broader substrate scope with OpSTS

Enzyme specificity is challenging to manipulate because it often depends on the small contribution of many, including distant, side-chain and main-chain interactions. The intriguingly broad amine and aldehyde scopes of OpSTS are not easily explained by comparing amino acid residues in the substrate binding pockets of OpSTS and RsSTS (Chapter 4). However, a promising strategy may be to gradually modify the sequence of CrSTS to resemble that of OpSTS until the desired switch in specificity occurs. Since some part of the DNA sequence of *opsts* confers the broader substrate specificity, different stretches of DNA in the *opsts* and *crsts* genes could be swapped. For example, DNA shuffling of the *C. roseus* and *O. pumila* strictosidine synthase genes by treating with DNaseI, followed by self-reassembly and splicing-overlap extension PCRs leads to a random shuffled library (Figure 7.2A).[7] Alternatively, defined overlapping gene fragments could be obtained by PCR and then reassembled to give defined chimerae (Figure 7.2A). The resulting random and defined chimerae may then be evaluated for activity with natural substrates using the BTB-containing agar plate screen, to eliminate inactive library members. The active chimerae can then be individually expressed and assayed in 96-well plates similar to methods described in Chapter 3 (Figure 7.2B).

7.2.2 Optimizing the activity of STS toward substrate analogs

Optimizing the activity of OpSTS or CrSTS toward substrate analogs may result in a useful biocatalyst for synthetic transformations and an improved production titer of MIA alkaloids in transgenic plant cultures. Amino acid substitutions can lead to the

A

opsts ....   DNA shuffling
                                                      ──────────────→
crsts ──────────────────   (random)

opsts/crsts library

│ PCR amplification
↓

opsts ....   Chimerae construction
                                                      ──────────────→   defined chimerae
crsts ─── ─── ─── ───   (semi-rational)

overlapping gene fragments
of known identity

B

hits
──────→

Library   1° screen   2° screen
            (agar plate)   (substrate analogs)
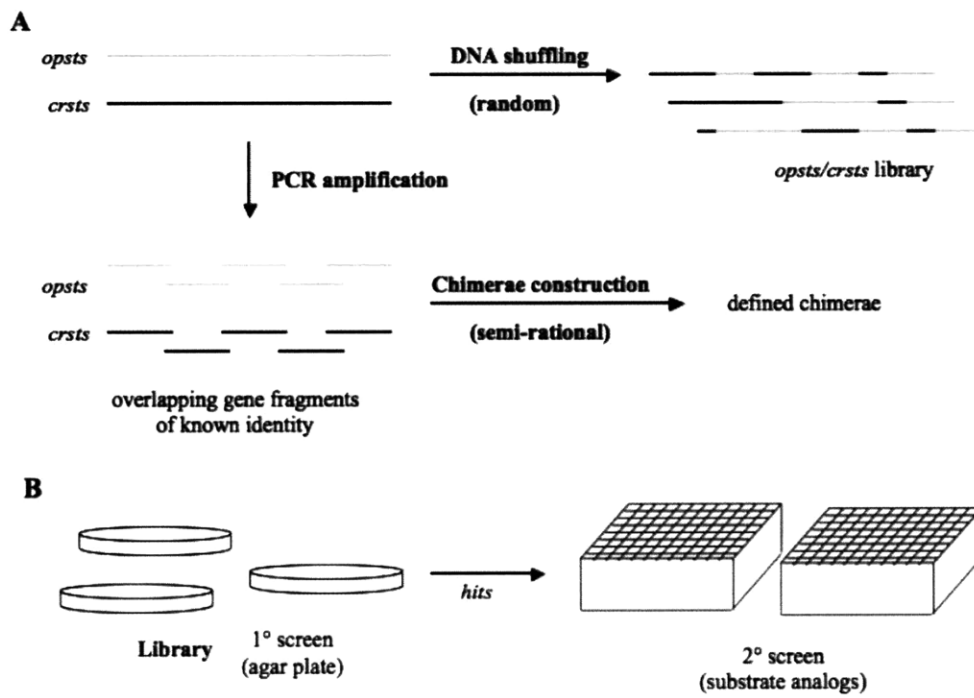
Figure 7.2 (A) Library construction based on DNA shuffling that results in a random or defined library of chimerae. (B) Gene libraries introduced into *S. cerevisiae* can be screened for activity toward the natural substrates to eliminate inactive members. The active chimerae can then be carried into a secondary screen using substrate analogs.

introduction of repulsive interactions with adjacent residues or substrates. This may be counteracted by mutagenesis of the residues adjacent to the substitution. For the Val214Met STS mutant, the two adjacent residues (213 and 215) could be subjected to saturation mutagenesis using a single complementary primer pair. The resulting library, transformed into *S. cerevisiae* could then be screened using the BTB-containing agar plate method for activity toward 5-methyl-, 5-chloro-, or 5-bromotryptamine. Remote substitutions could also have unexpected effects on activity with substrate analogs and error-prone PCR may be a potentially useful complement to more focused strategies.

7.2.3 Reengineering the stereoselectivity of STS

The stereochemical model of STS presented in Chapter 4 predicts that relocation of the catalytic Glu residue may allow the formation of the 3*(R)* stereogenic center found in vincoside. A successful redesign would be an compelling example of rational design and would validate this stereoselectivity hypothesis. The ability to form the alternative stereoisomer also has a practical application, as it would result in an asymmetric biocatalyst complementary to the 3*(S)*-forming OpSTS. Site-directed mutants of STS can be designed based on the model described in Chapter 4, and will likely include Phe232 and Glu315. One strategy is to switch the locations of the residues that interact with the top and the bottom faces of the indole ring, using a combination of targeted and random mutagenesis. Although computer modeling may serve as a guide, this project will likely present considerable challenges.

7.2.4 Evolving a thermostable STS

Since the asymmetric synthesis catalyzed by STS is effectively irreversible, elevated temperatures should not affect the enantiomeric excess of the tetrahydro-β-carboline product. The biocatalytic application of OpSTS could therefore benefit from a higher tolerance to increased reaction temperatures. The directed evolution of thermostability in OpSTS is currently in progress. Briefly, error-prone PCR libraries will be transformed into *S. cerevisiae*. As the colonies form on indicator plates, individual OpSTS variants are exported to the media around each colony. After incubation at an elevated temperatures the plates will be screened for activity as described in Chapter 4. Any hits will be used as a template for second round of error-prone PCR. The evolved thermostable mutants will be characterized using circular dichroism and inactivation kinetics and compared to the wild-type enzyme. Finally the thermostable mutants will be tested for biocatalytic applications.

7.2.5 Converting STS into paraoxonase

Sequence alignments show that there are hundreds, if not thousands, of STS-like proteins. However, only a few of these proteins catalyze the Pictet-Spengler reaction. It is likely that the sequence similarity merely identifies a superfamily with a common protein structure class – the six-bladed β-propeller fold. For example, several β-propellers with significant strictosidine synthase sequence identity (~30%) have phosphatase activity (*e.g.* human paraoxonase, PON). Other superfamily members are without enzyme activity

241

(*e.g.* hemomucin). The sequence requirements of STS activity are unknown, but it is interesting to observe that the same structural scaffold can support different activities.

The use of protein structures as scaffolds to support different natural *or* unnatural catalytic activities offers tremendous potential for the development of new biocatalysts.[8,9] Understanding how to alter protein function without disrupting protein structure is key for the successful creation of engineered enzymes. One ongoing project is to introduce PON activity into STS. This would require the introduction of Ca-binding site. The strategy is based on assembly of genes that gradually modify the gene encoding RsSTS into the gene encoding PON1. Assembly of genes by a multi-primer self-assembly PCR reaction followed by cloning of the resulting artificial genes into an expression vector may be followed by expression and characterization of the chimeric proteins. Hydrolase activity toward aryl esters and lactones, which are substrates for PON1 but not STS, and the dependence on calcium for activity, will be signs for the successful conversion of activity. In the case that designed proteins are not active, the artificial genes may be recombined using DNA shuffling in the presence of the gene encoding PON1. A successful outcome would cast light on the minimum sequence requirements for Ca-binding and hydrolase activity in STS. X-ray structure determination of any active variants may be of considerable interest for the understanding of how structure evolves as sequences change.

7.3 References

1. Bernhardt, P.; McCoy, E.; O'Connor, S.E. "Rapid identification of enzyme variants for reengineered alkaloid biosynthesis in periwinkle", *Chem. Biol.* **2007**, *14*, 888-897.

2. Bernhardt, P.; Giddings, L.-A.; Loh, K.; O'Connor, S.E. "pH-indicator based assay for strictosidine synthase activity and its application in enzyme engineering", *manuscript in preparation.*

3. Runguphan, W.; O'Connor, S.E. "Metabolic reprogramming of periwinkle plant cell culture", *Nat. Chem. Biol.* **2009**, *5*, 151-153.

4. Bernhardt, P.; Usera, A.R.; O'Connor, S.E. "Strictosidine synthase from *Ophiorrhiza pumila* is a stereoselective and promiscuous biocatalyst for the Pictet-spengler reaction", *submitted* **2009**.

5. Bernhardt, P.; O'Connor, S.E. "Synthesis and biochemical evaluation of des-vinyl secologanin *O*-analogs with alternate stereochemistry", *Tetrahedron Lett.* **2009**, *50*, 7118-7120.

6. Bernhardt, P.; Yerkes, N.; O'Connor, S.E. "Bypassing stereoselectivity in the early steps of alkaloid biosynthesis", *Org. Biomol. Chem.* **2009**, *7*, 4166-4168.

7. Stemmer, W.P.C "Rapid evolution of a protein *in* vitro by DNA shuffling", *Nature* **1994**, *370*, 389-391.

8. Jiang, L.; Althoff, E.A.; Clemente, F.R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J.L.; Betker, J.L.; Tanaka, F.; Barbas, C.F.; Hilvert, D.; Houk, K.N.;

Stoddard, B.L., Baker D. "De novo computational design of retro-aldol enzymes", *Science* **2008**, *319*, 1387-1391.

9. Röthlisberger, D.; Khersonsky, O.; Wollacott, A.M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J.L.; Althoff, E.A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K.N.; Tawfik, D.S.; Baker, D. "Kemp elimination catalysts by computational enzyme design", *Nature* **2008**, *453*, 190-195.

CHAPTER 3 - APPENDIX A

LC-MS TRACES: FEEDING OF STRICTOSIDINE ANALOGS **8m** AND **8o**

TO *C. ROSEUS* HAIRY-ROOT CULTURES
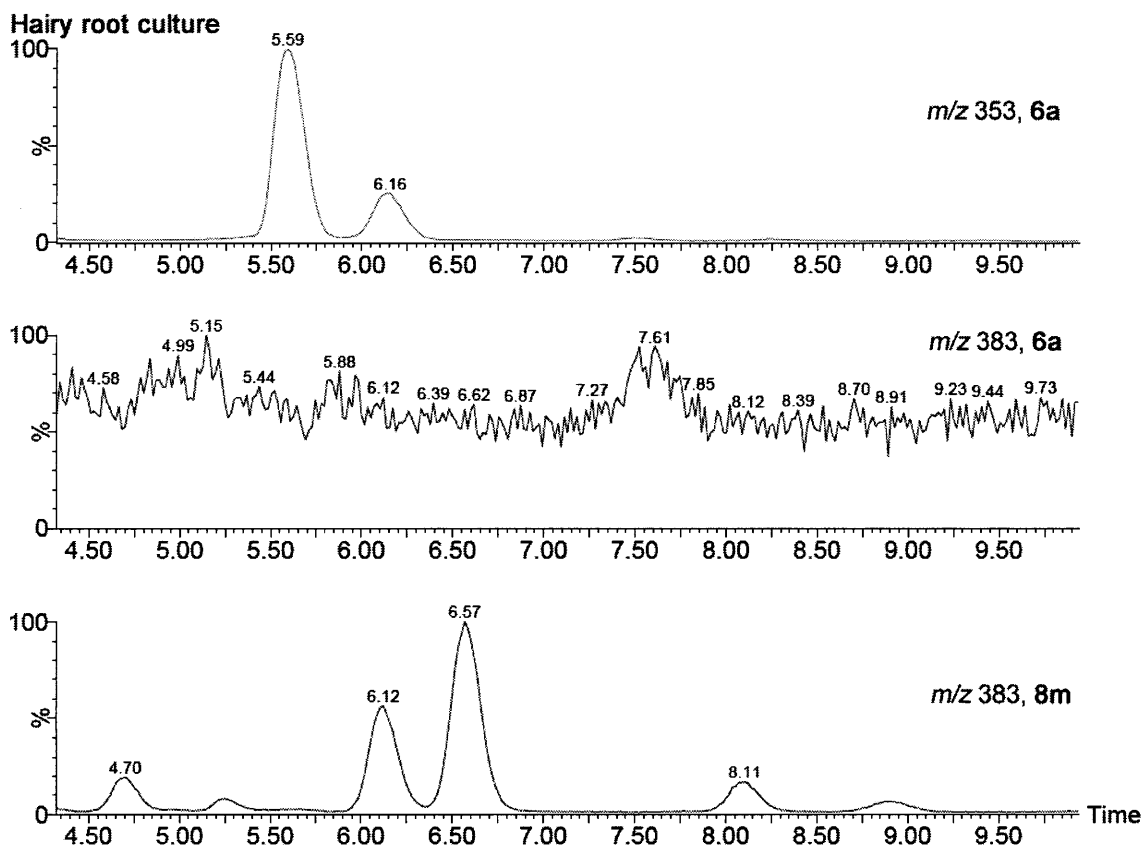
Figure 3A.1 Chapter 3. LC-MS trace of putative ajmalicine **1** (or isomer of **1**) analog. Peaks for *m/z* 383 ($C_{22}H_{27}N_2O_4$ calc *m/z* 383.1971 $[M+H]^+$, observed *m/z* 383.1955 $[M+H]^+$) exist in the **8m**-supplemented culture that are not present in the control (supplemented with **6a**). The corresponding natural MIA, ajmalicine **1** (*m/z* 353), is present in the control culture.
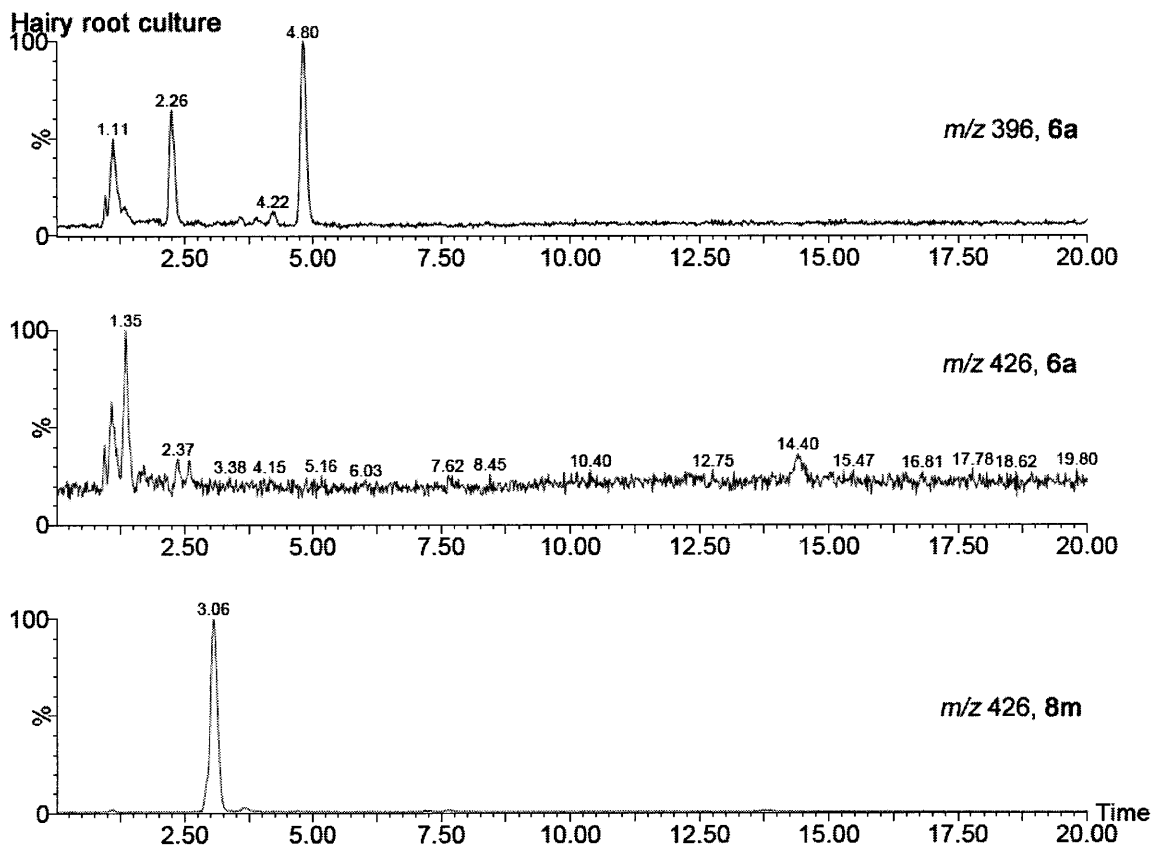
Figure 3A.2 Chapter 3. LC-MS trace of unknown compound **14** (retention time, 3.1 min), which is formed when hairy roots are supplmented with **8m**. NMR spectroscopy shows that the compound is an indole alkaloid. The alkaloid with *m/z* 426 does not exist in the control culture (supplemented with **6a**). The peak at 4.8 min in the control culture (top trace) is the isotope signal of the unidentified compound, *m/z* 395, which is produced in hairy root culture.
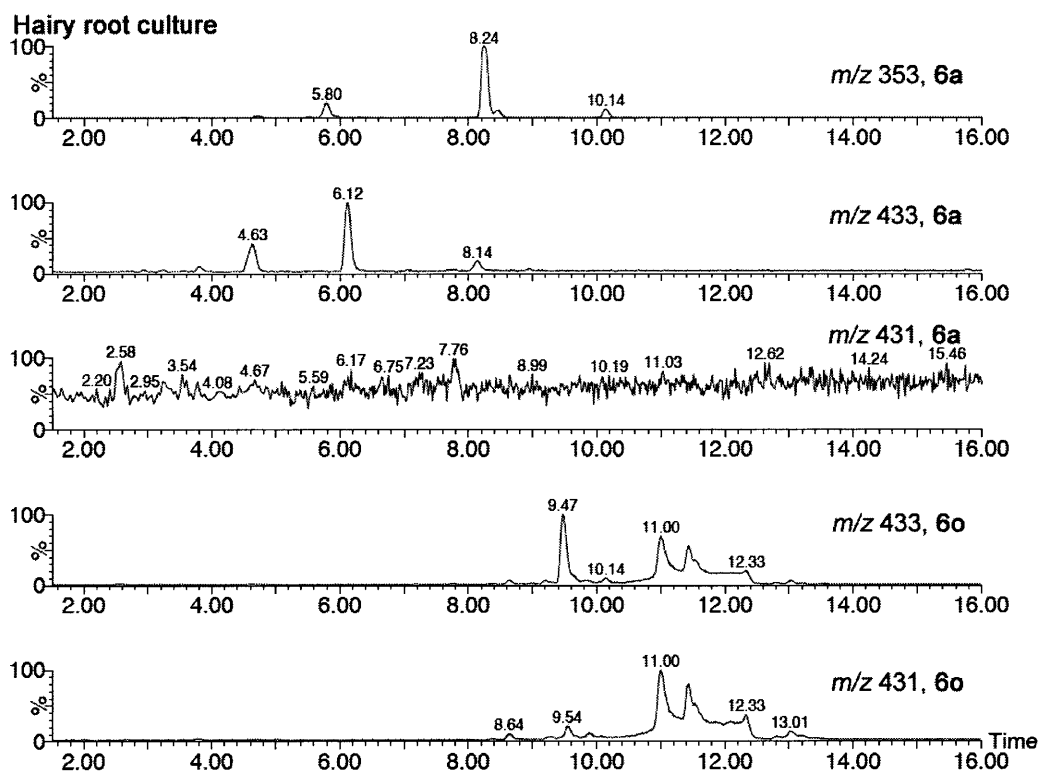
Figure 3A.3 Chapter 3. LC-MS trace of a brominated ajmalicine analog with stereoisomers. Both isotope peaks are shown (*m/z* 431 and 433) and bromine isotopes contain 1:1 mixture of M and M+2. The culture supplemented with **8o** contain two peaks (retention times of 11.0 min and 11.4 min) that are not present in the control culture (derived from **6a**). The peak at 8.2 min in the top trace co-elutes with an authentic standard of ajmalicine (not shown).
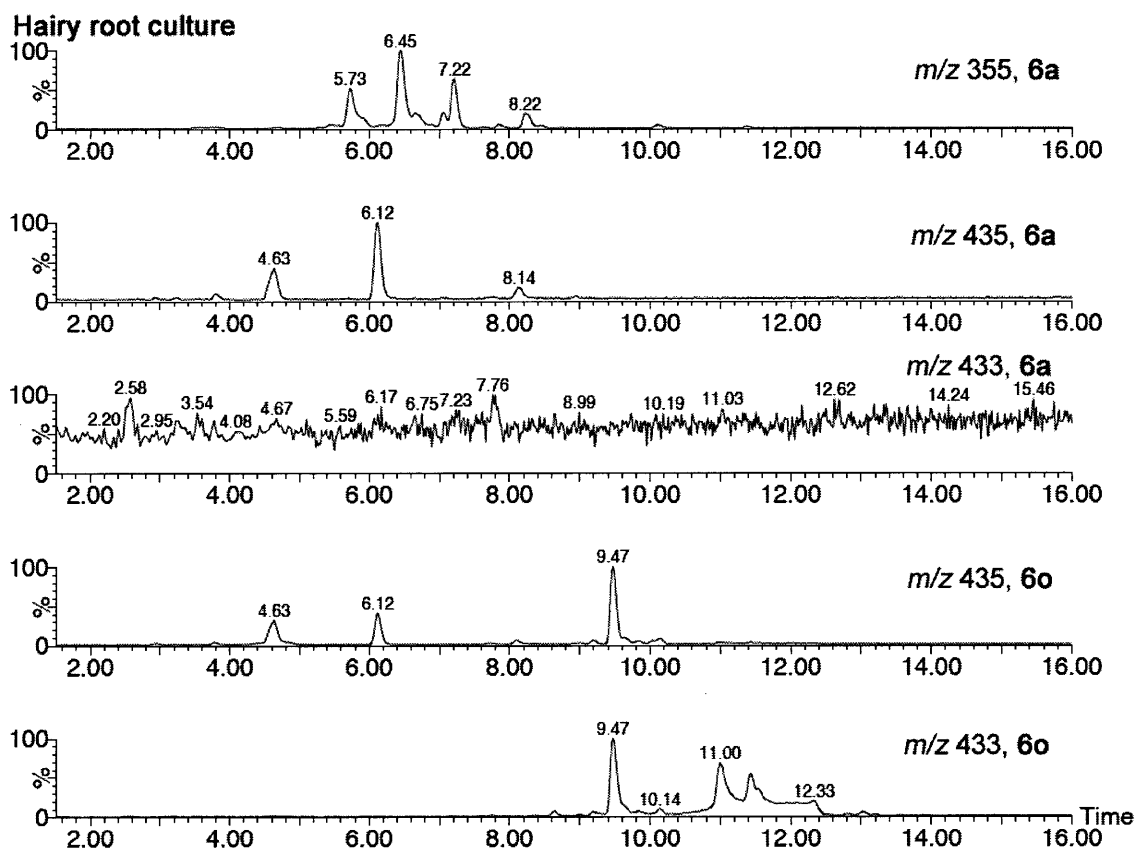
Figure 3A.4 Chapter 3. LC-MS trace of putative brominated yohimbine-type alkaloid (retention time, 9.5 min). Based on the presence of the two characteristic isotope signals expected from a brominated molecule. The corresponding signals are not observed in the control culture (derived from **6a**). The natural MIA yohimbine elutes at 6.5 min using this LC-method (top trace). The isotope peaks of the putative ajmalicine analogs (*m/z* 433, compare to Figure A3) are seen in the bottom trace.
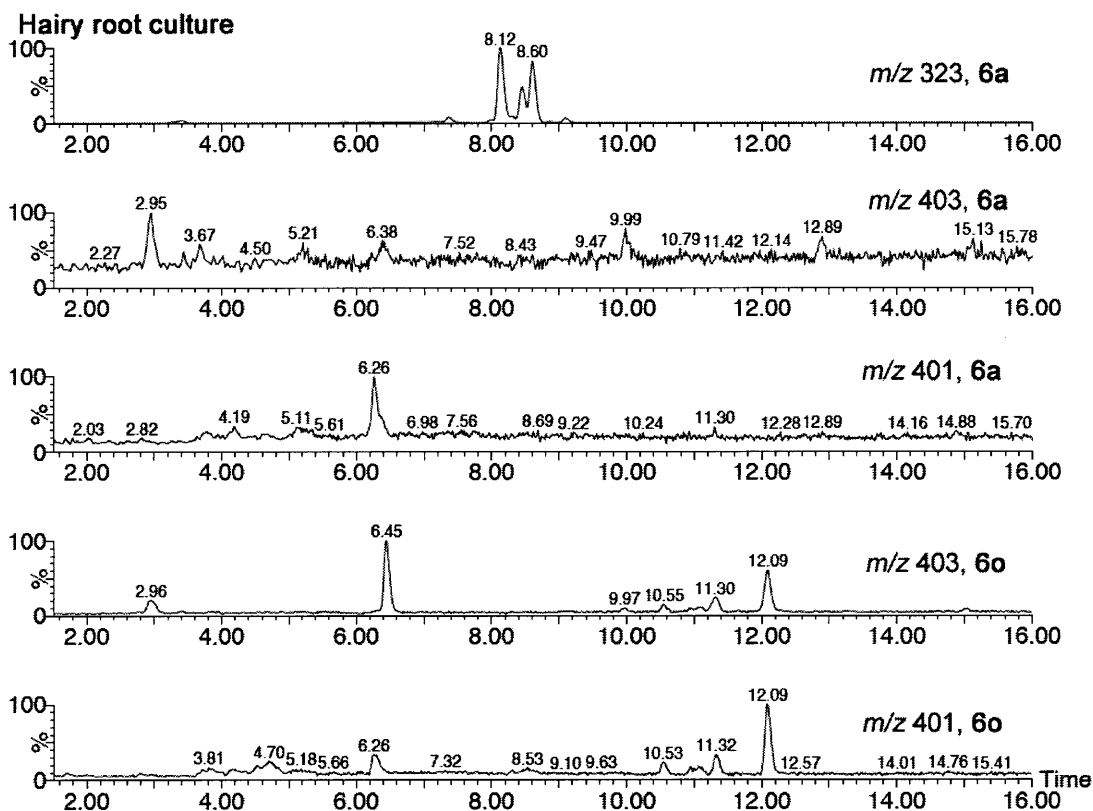
Figure 3A.5 Chapter 3. LC-MS trace of a putative brominated alkaloid analog of akuammacine 11 (m/z 323,r_t = 8.1 min, NMR evidence in McCoy, E.; O'Connor, S.E. "Directed biosynthesis of alkaloid analogs in the medicinal plant periwinkle", *J. Am. Chem. Soc.* **2006**, *128*, 14276-14277). Based on the presence of the two characteristic isotope signals expected from a brominated molecule, we assigned m/z 401 and m/z 403 at 12.1 min as any structural isomer of akuammicine (**11o**); the corresponding signals are not observed in the control culture.

CHAPTER 4 - APPENDIX B

CHIRAL HPLC TRACES AND NMR SPECTRA:

O. PUMILA STRICTOSIDINE SYNTHASE

i. Product of **1** and **10**

racemic standard

OpSTS

ii. Product of **1** and **11**

racemic standard

OpSTS

iii. Product of **1** and **13**

racemic standard

OpSTS
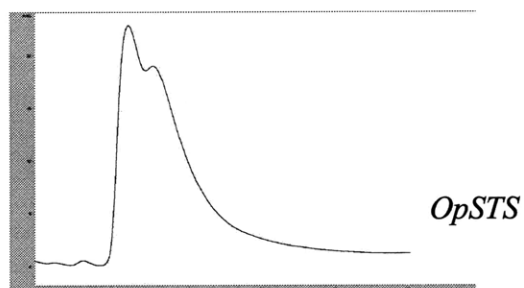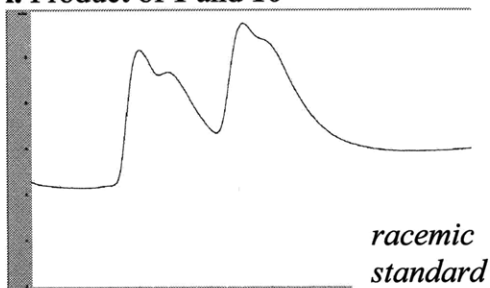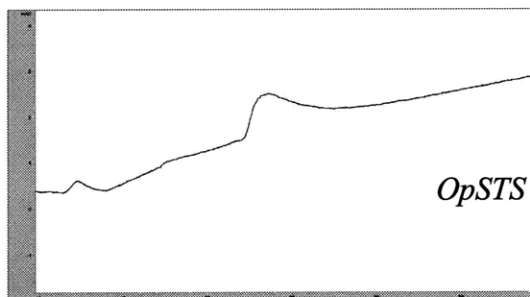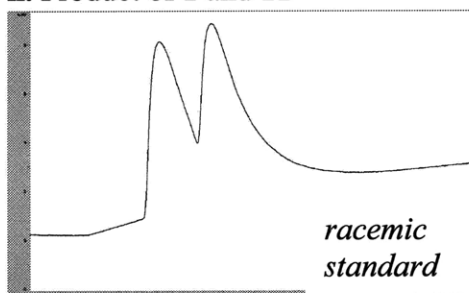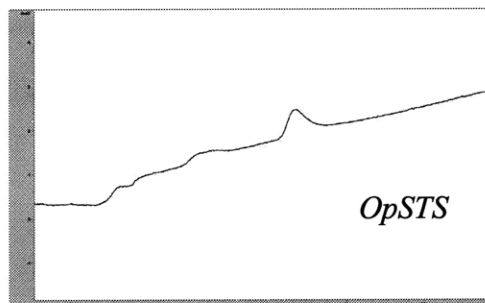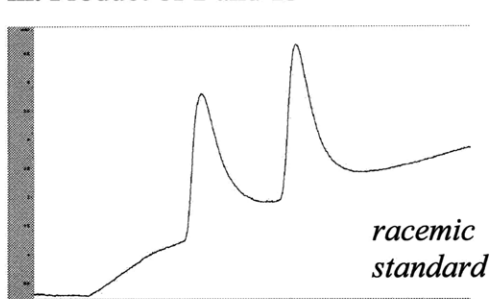
iv. Product of **1** and **14**

racemic standard

OpSTS
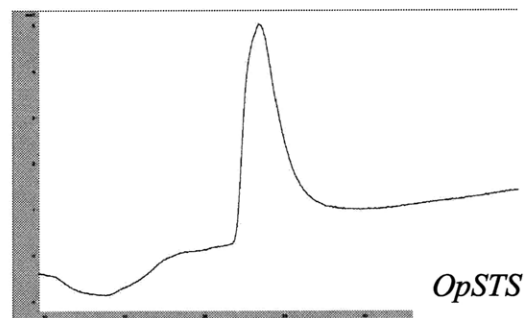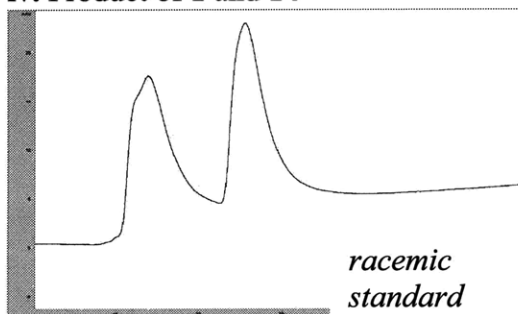
Figure 4B.1 Chapter 4. Chiral HPLC traces of tetrahydro-β-carbolines. The scale on each of the traces is 10 min to 40 min.

ppm (f1)

137.612
129.583
126.732
122.798
119.912
118.344
111.635
106.545

68.032
67.879

51.081

42.391
39.943
33.937
32.400
31.093

18.810

200   150   100   50   0

ppm (f1)

ppm (f1)

137.589
129.951
126.752
122.690
119.841
118.309
111.642
106.394

51.573
42.455
40.532
34.565
33.589
32.414
26.737
26.536
26.155
18.810

ppm (t1)   150   100   50   0

ppm (f1)

137.219
131.636
126.975
122.744
120.074
119.092
112.460
106.814
51.507
41.944
24.671
24.532
22.511
19.159

150   100   50   0
ppm (t1)

135.675
135.819
130.101
129.873
129.206
127.708
126.260
122.361
119.543
118.593
111.898
106.716

54.201

41.900
37.757

18.445

150   100   50   0

9.0   8.0   7.0   6.0   5.0   4.0   3.0   2.0   1.0   0.0
ppm (t1)

1.88 1.88   1.02   1.01 1.01 2.03   1.01 1.01 4.04 3.03
                                                                   2.04

138.355
130.317
127.462
123.476
120.594
119.100
112.392
107.180

55.170

43.194

33.382
32.786
25.919
23.481
19.544
14.369

150   100   50   0

ppm (f1)

138.297
130.139
127.417
123.460
120.568
119.130
112.382
107.246
56.186
43.154
26.331
19.522
9.927

150 100 50 0

9.0   8.0   7.0   6.0   5.0   4.0   3.0   2.0   1.0   0.0
ppm (f1)

137.644
128.268
126.714
122.756
119.836
118.312
111.654
107.405

59.191

43.146
40.320

30.123
27.070
26.822
26.431
26.407
18.698

150   100   50   0

ppm (t1)

CHAPTER 5 - APPENDIX C

LC-MS TRACES:
C. ROSEUS STRICTOSIDINE SYNTHASE-CATALYZED REACTIONS BETWEEN
TRYPTAMINE AND DES-VINYL SECOLOGANIN AGLYCONE *O*-ANALOGS

C. roseus STS catalyzed reaction between **1** and **22** (*O*-Et)

4 h time-point

C. roseus STS catalyzed reaction between **1** and **25** (*O*-Cy)
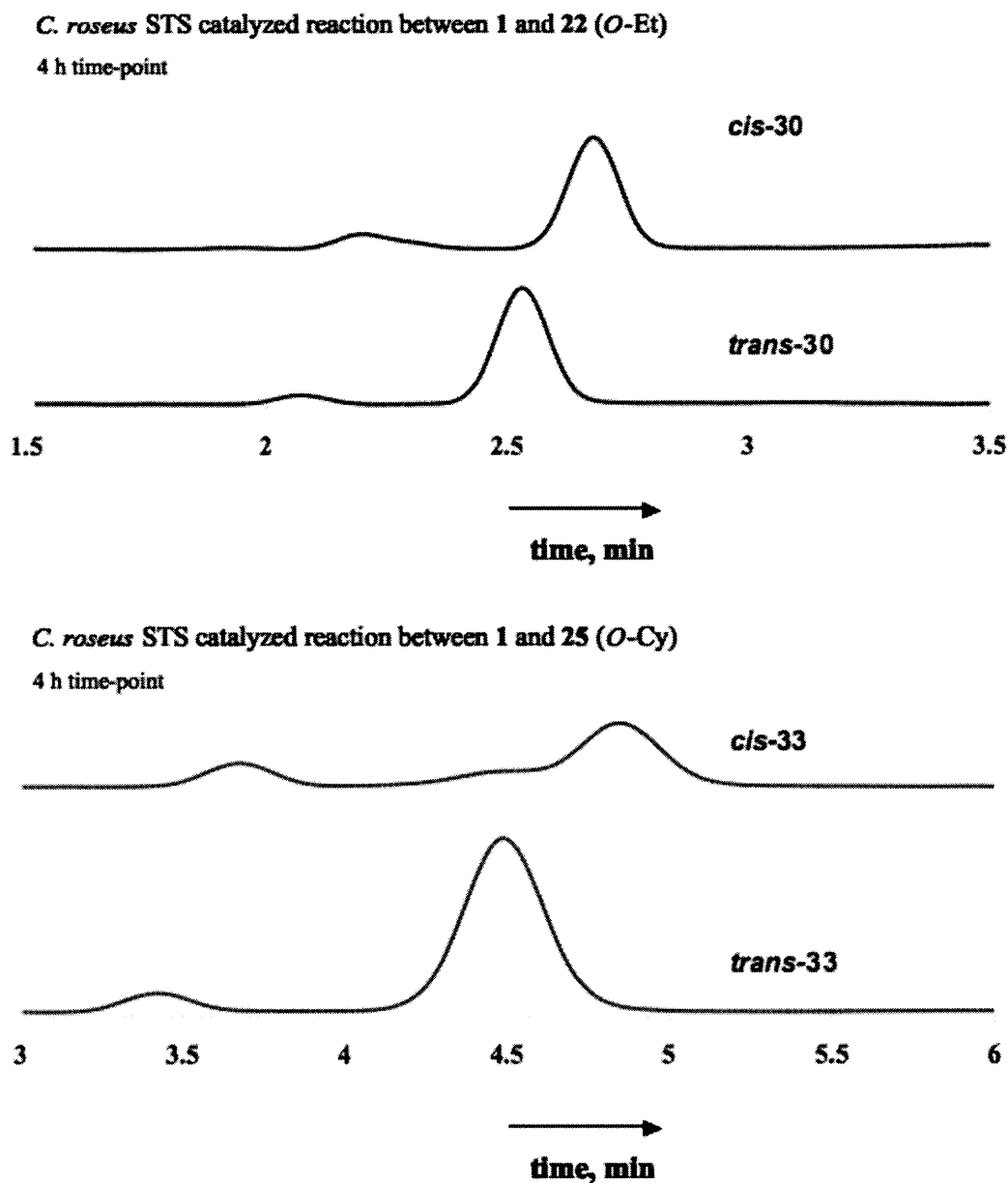
4 h time-point

Figure 5C.1 Chapter 5. LC-MS traces of the 4 h time-point in kinetic experiments that were used to determine the diastereoselectivity of CrSTS. CrSTS has a higher *trans*/*cis*-diastereoselectivity with compound **33** (*O*-Cy) than compound **30** (*O*-Et).
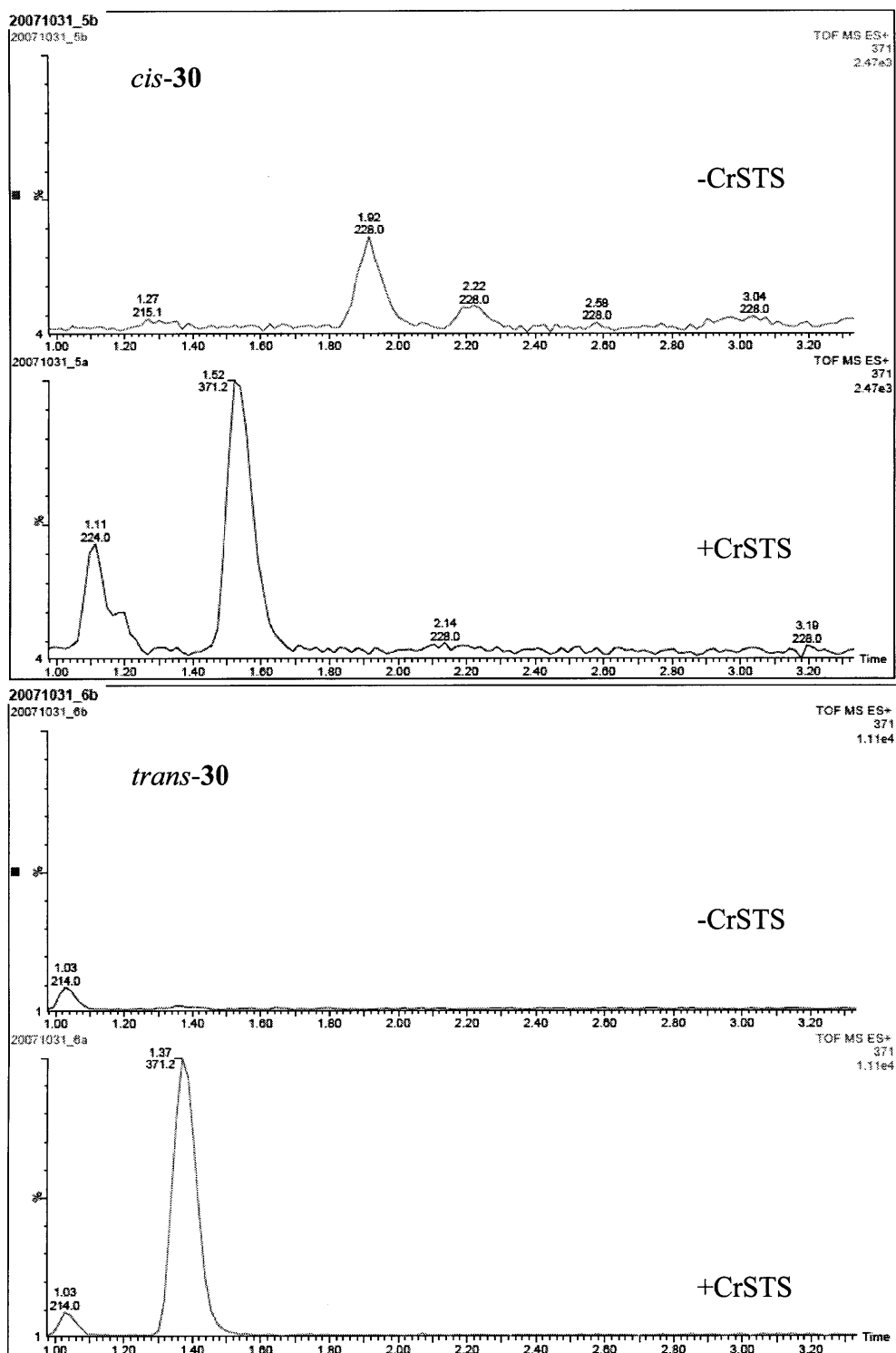
Figure 5C.2 Chapter 5. LC-MS traces for the CrSTS-catalyzed formation of the *O*-Et analog of des-vinyl strictosidine **30**. For each diastereomer, the no-enzyme control (-CrSTS) is shown in the top trace.
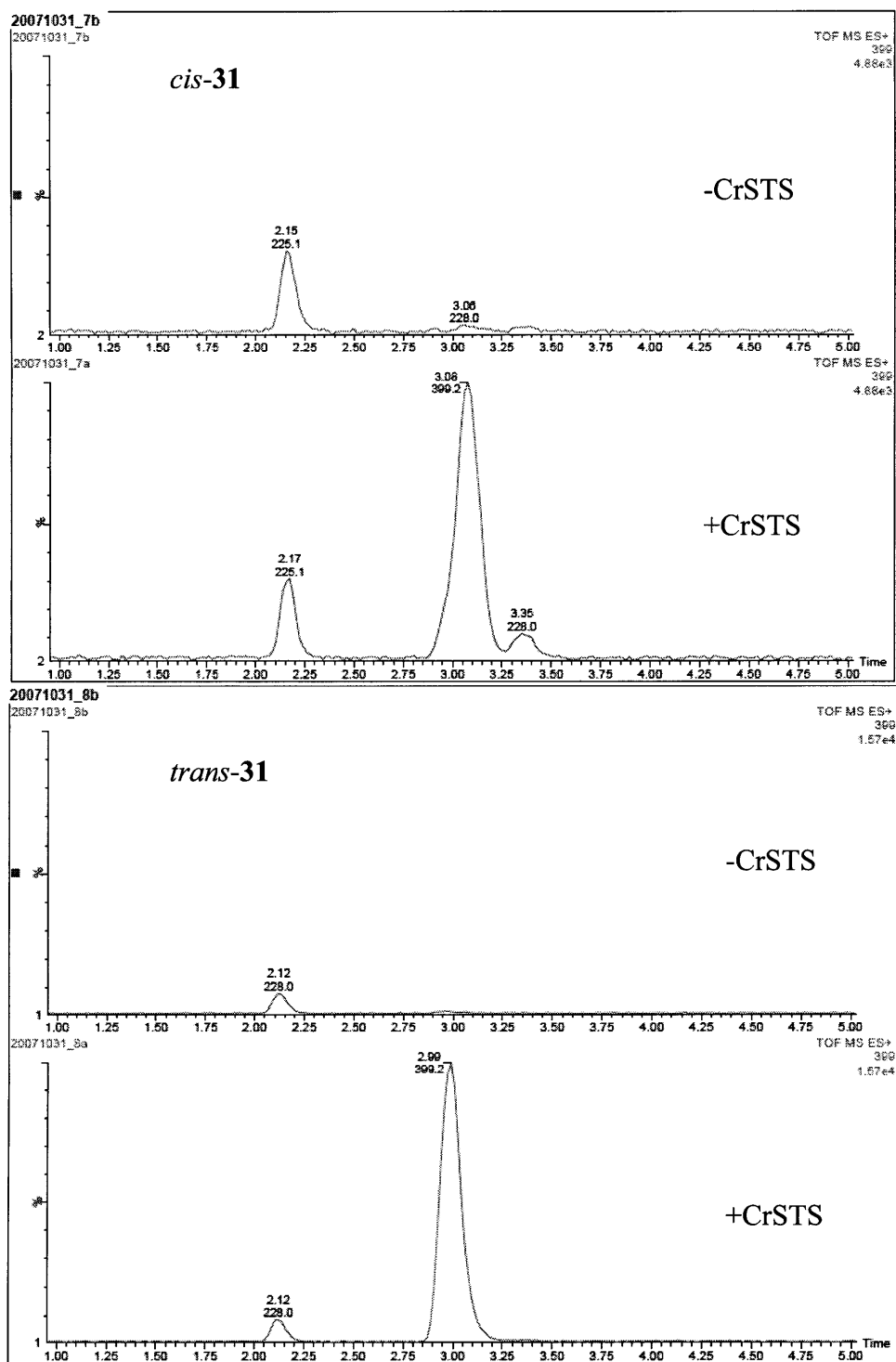
Figure 5C.3 Chapter 5. LC-MS traces for the CrSTS-catalyzed formation of the *O-i*-Bu analog of des-vinyl strictosidine **31**. For each diastereomer, the no-enzyme control (-CrSTS) is shown in the top trace.
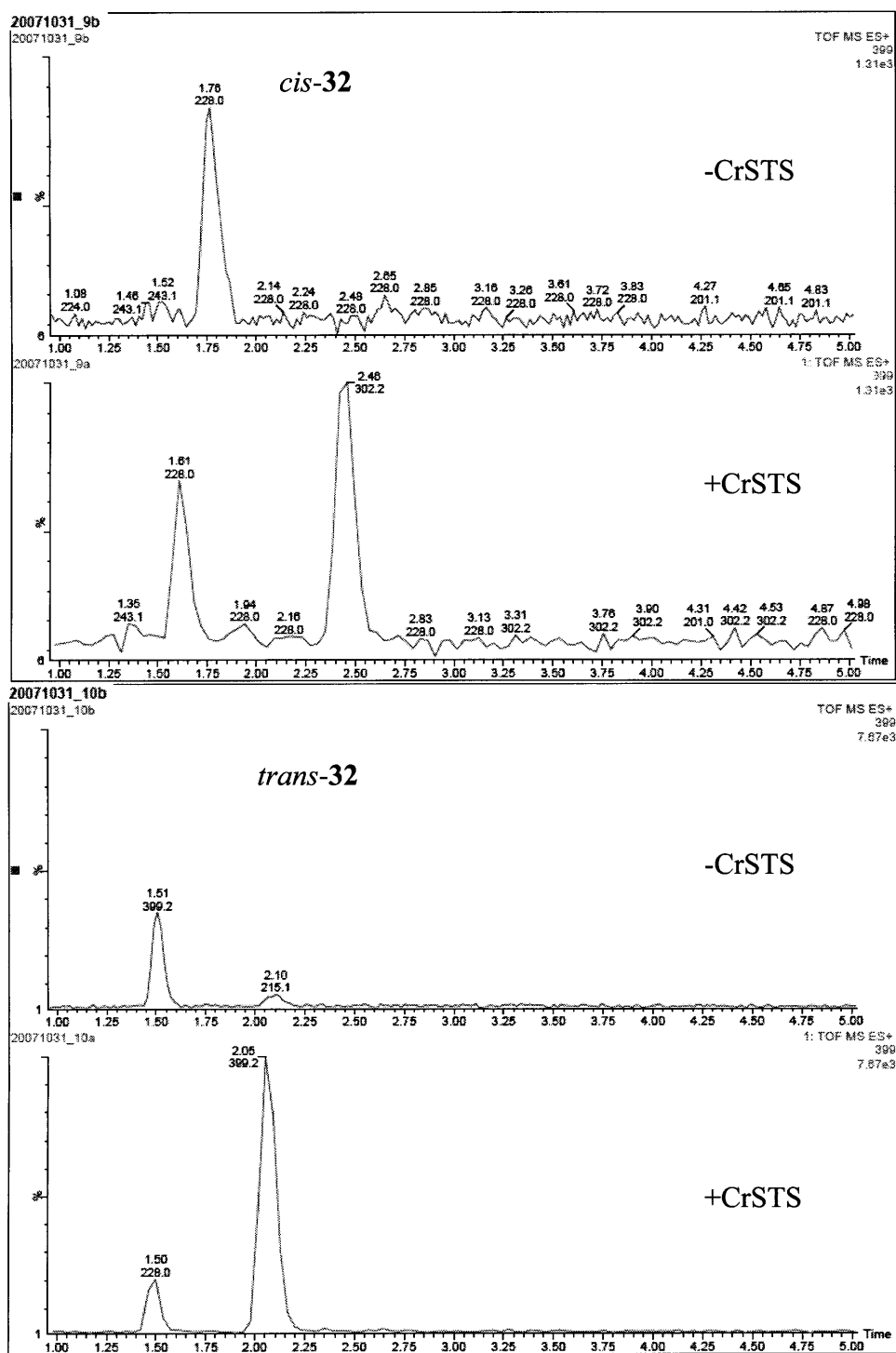
Figure 5C.4 Chapter 5. LC-MS traces for the CrSTS-catalyzed formation of the *O-t*-Bu analog of des-vinyl strictosidine **32**. For each diastereomer, the no-enzyme control (-CrSTS) is shown in the top trace.
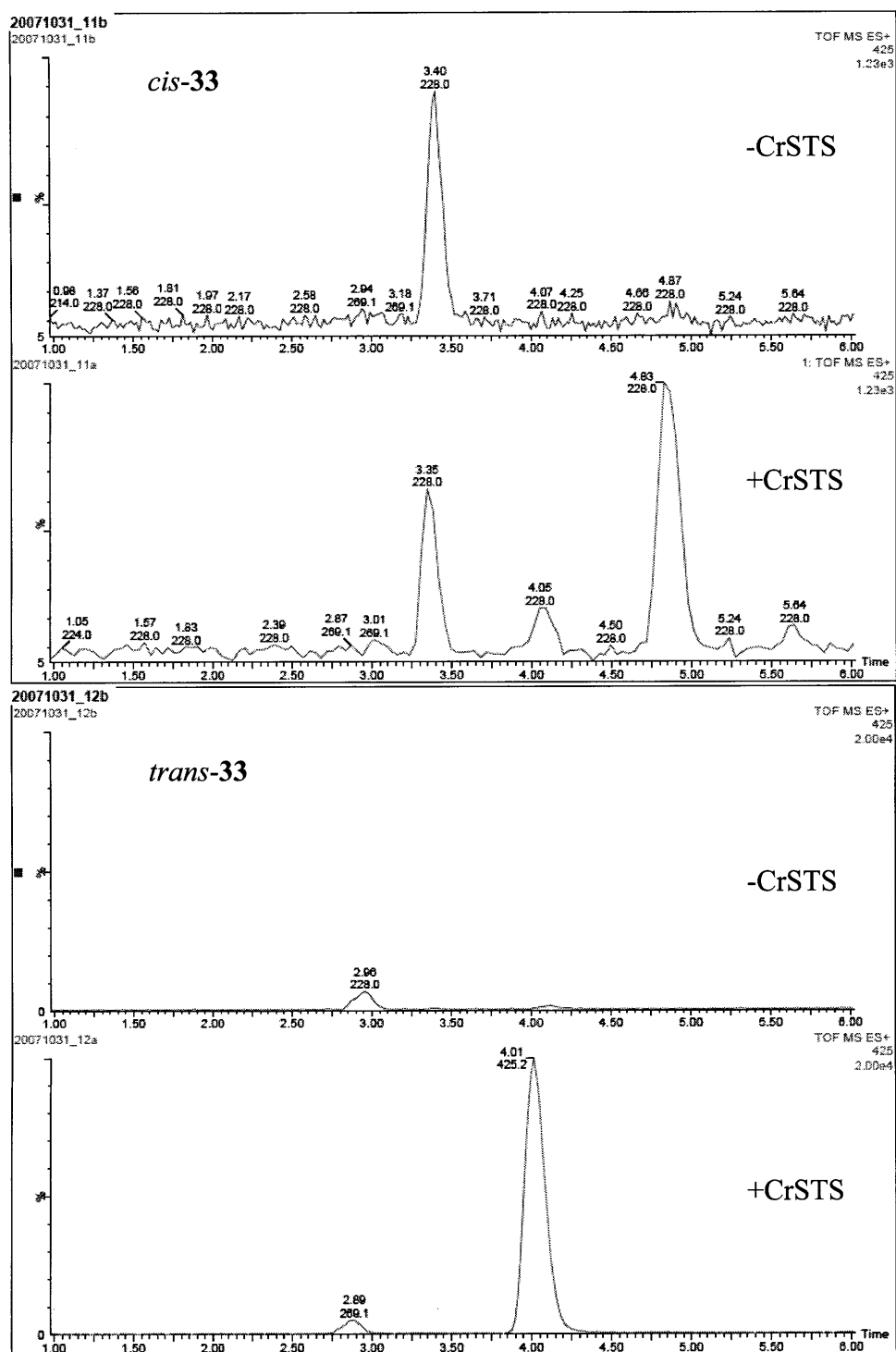
Figure 5C.5 Chapter 5. LC-MS traces for the CrSTS-catalyzed formation of the *O*-Cy analog of des-vinyl strictosidine **33**. For each diastereomer, the no-enzyme control (-CrSTS) is shown in the top trace.
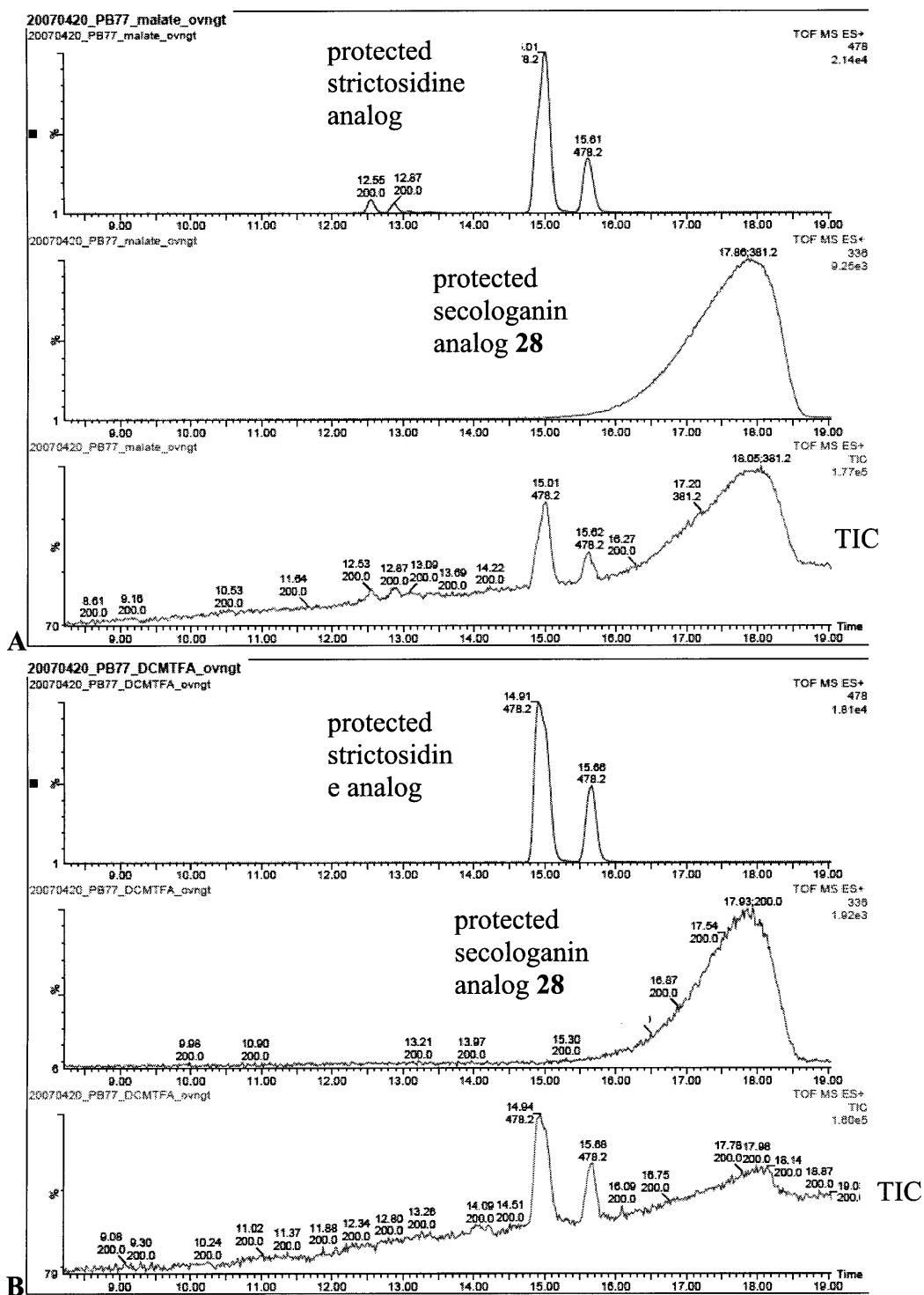
Figure 5C.6 Synthesis of *O*-nitrobenzyl strictosidine analog from **28** and **1** in the presence of either (A) maleic acid buffer (10 mM, pH 2.0) or (B) trifluoroacetic acid (TFA, pH 1) in DCM. The TFA-catalyzed reaction is higher-yielding.

273
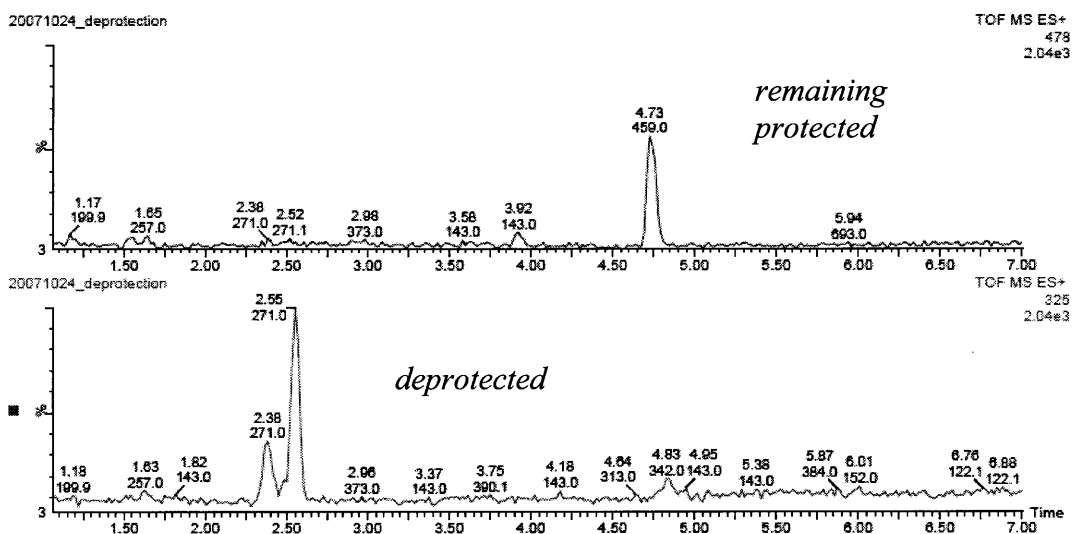
Figure 5C.7 Irradiation of *o*-nitrobenzyl des-vinyl strictosidine using UV-light (320 nm). After 10 minutes of irradiation some starting material remains (*m/z* 478, 4.7 min, above), while the expected rearranged deprotected product is formed (*m/z* 325, 2.4 and 2.6 min, below).

CHAPTER 6 - APPENDIX D

LC-MS TRACES AND NMR SPECTRA:

SYNTHESIS AND CHARACTERIZATION OF DES-VINYL SECOLOGANIN
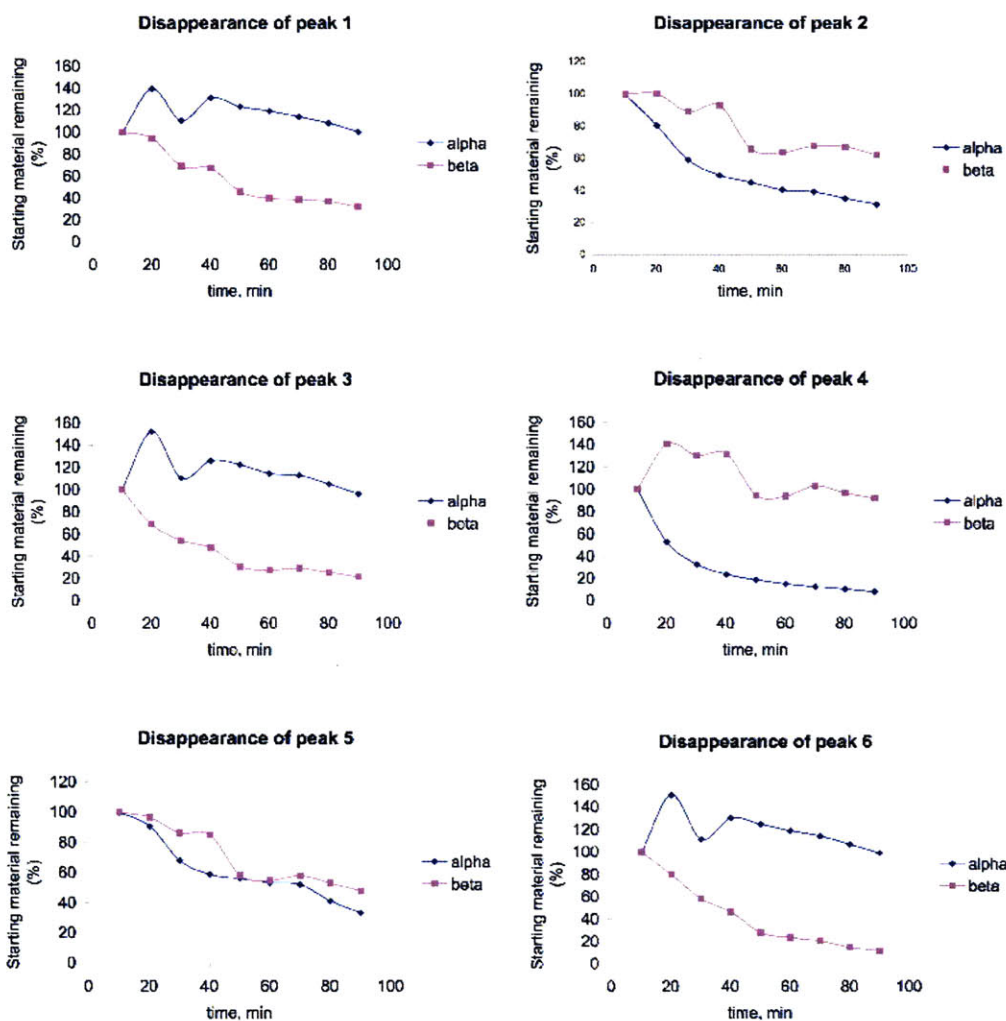
Figure 6D.1 Chapter 6. Deglucosylation of des-vinyl strictosidine isomers **12**. Graphs showing the relative rate of disappearance of UPLC peaks 1-6 using *B. stearothermophilus* α-glucosidase (alpha, blue points) or almond β-glucosidase (beta, pink points). Based on these data, pk 1, pk 3, and pk 6 contain mainly β-anomers, while pk 2 and pk 4 contain mainly α-anomers. Pk 5 appears to contain both α- and β-anomers. Each data point is the average value of three experiments. The error of each point is <20%.

Compound 7

0.90

0.97

2.23
2.18

2.00

3.23

3.10

10.0

5.0

0.0

199.573

97.813

77.171

48.278

29.986
22.839
21.746

CHO

0.62:0.38
β/α-anomers

Compound **8**

250   200   150   100   50

Compound 6

21.95   1.00   10.24   1.91   6.16

8.0   7.0   6.0   5.0   4.0   3.0

Compound **10**

dioxane -CH₃

acetal region

partially purified mixture
of stereoisomers

CH₃ missing
-> deprotected

aldehyde -CHO

10.0   5.0   0.0
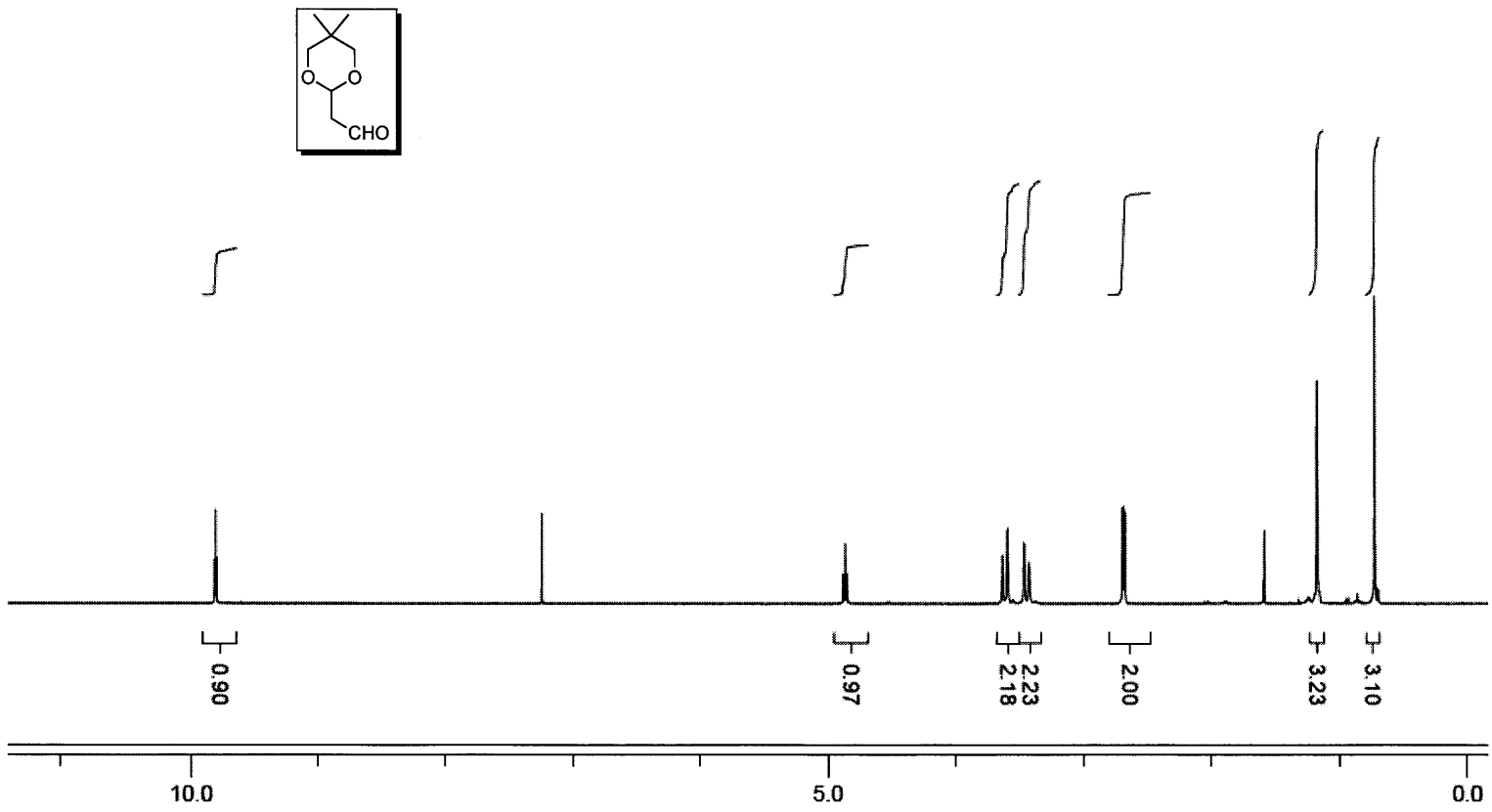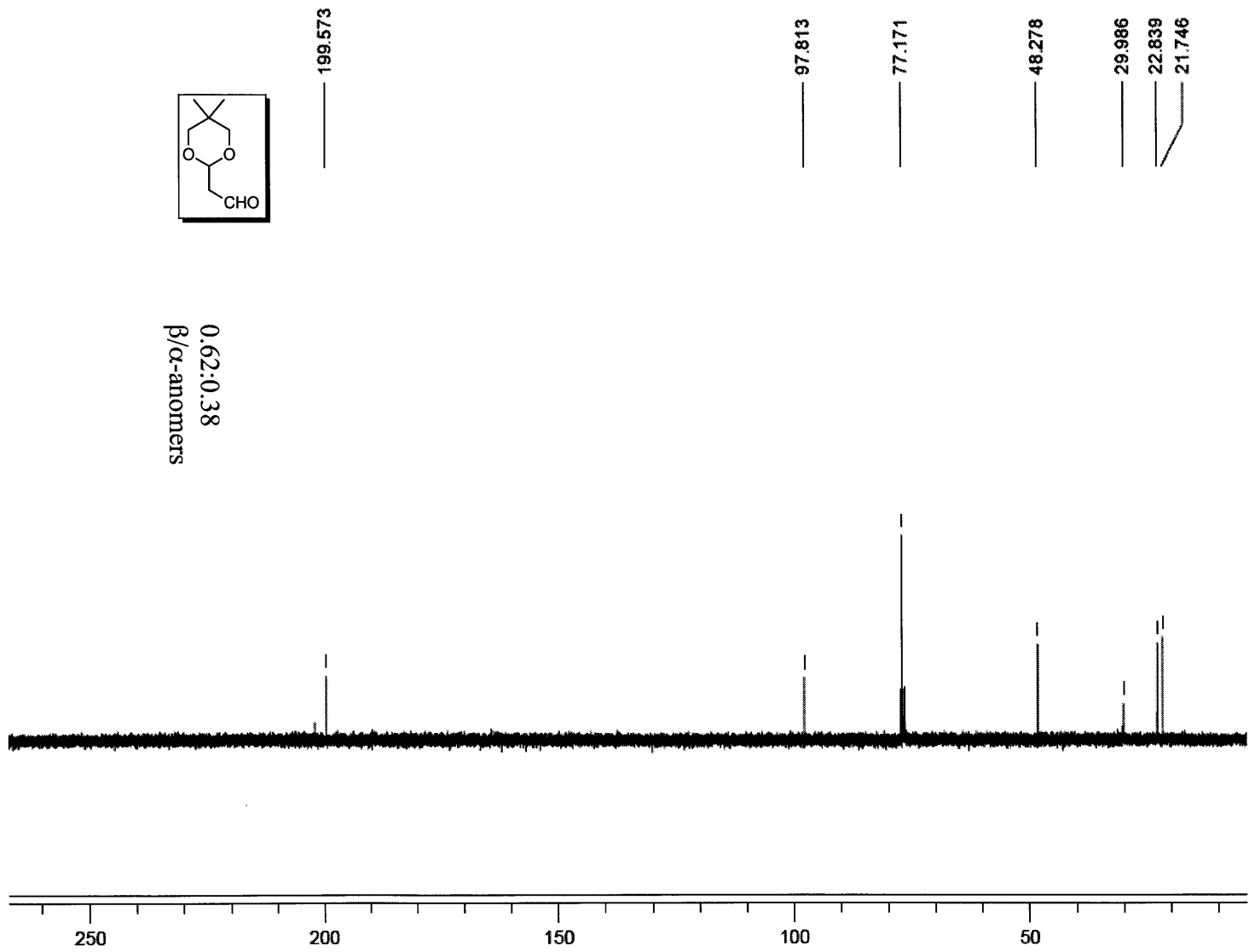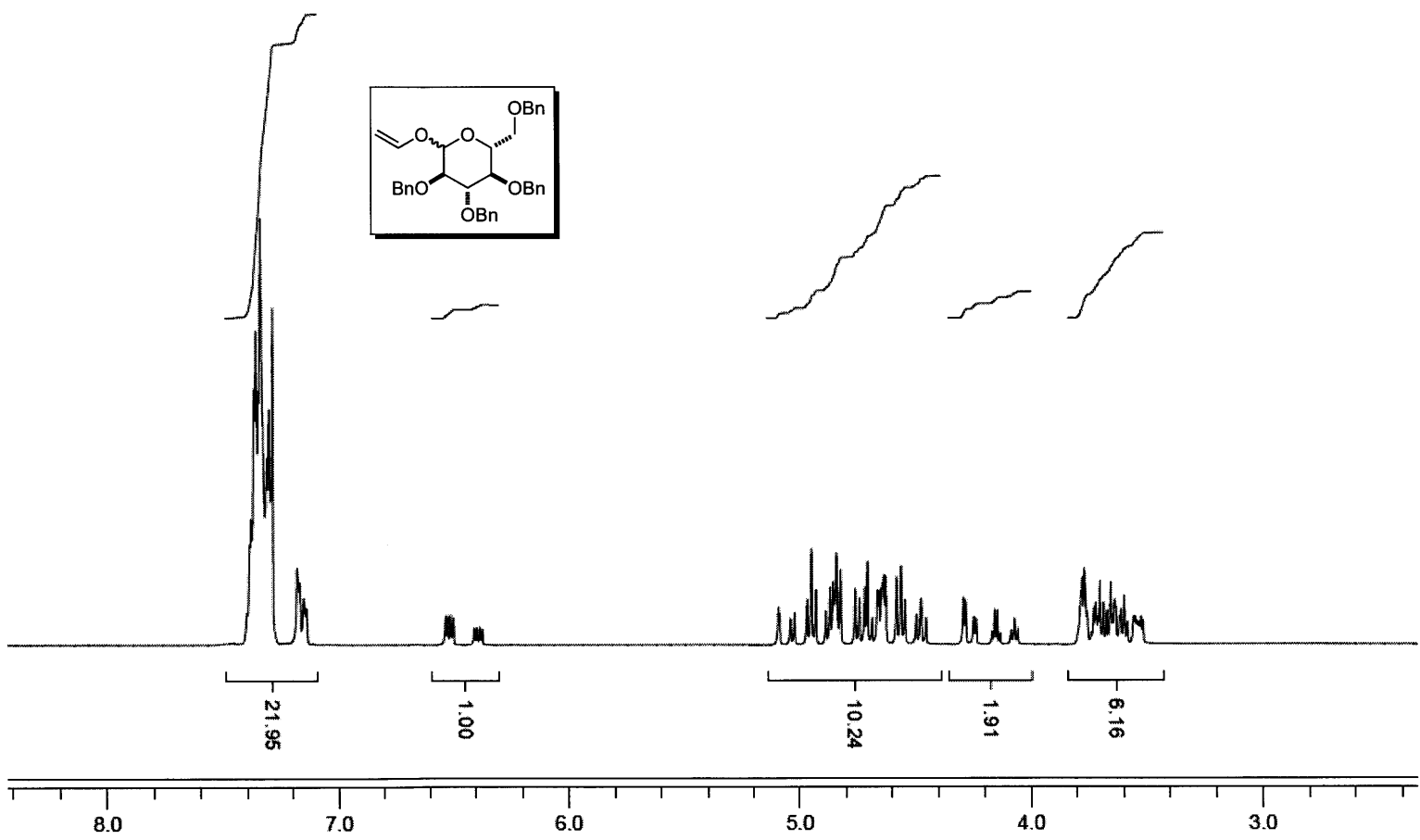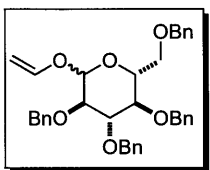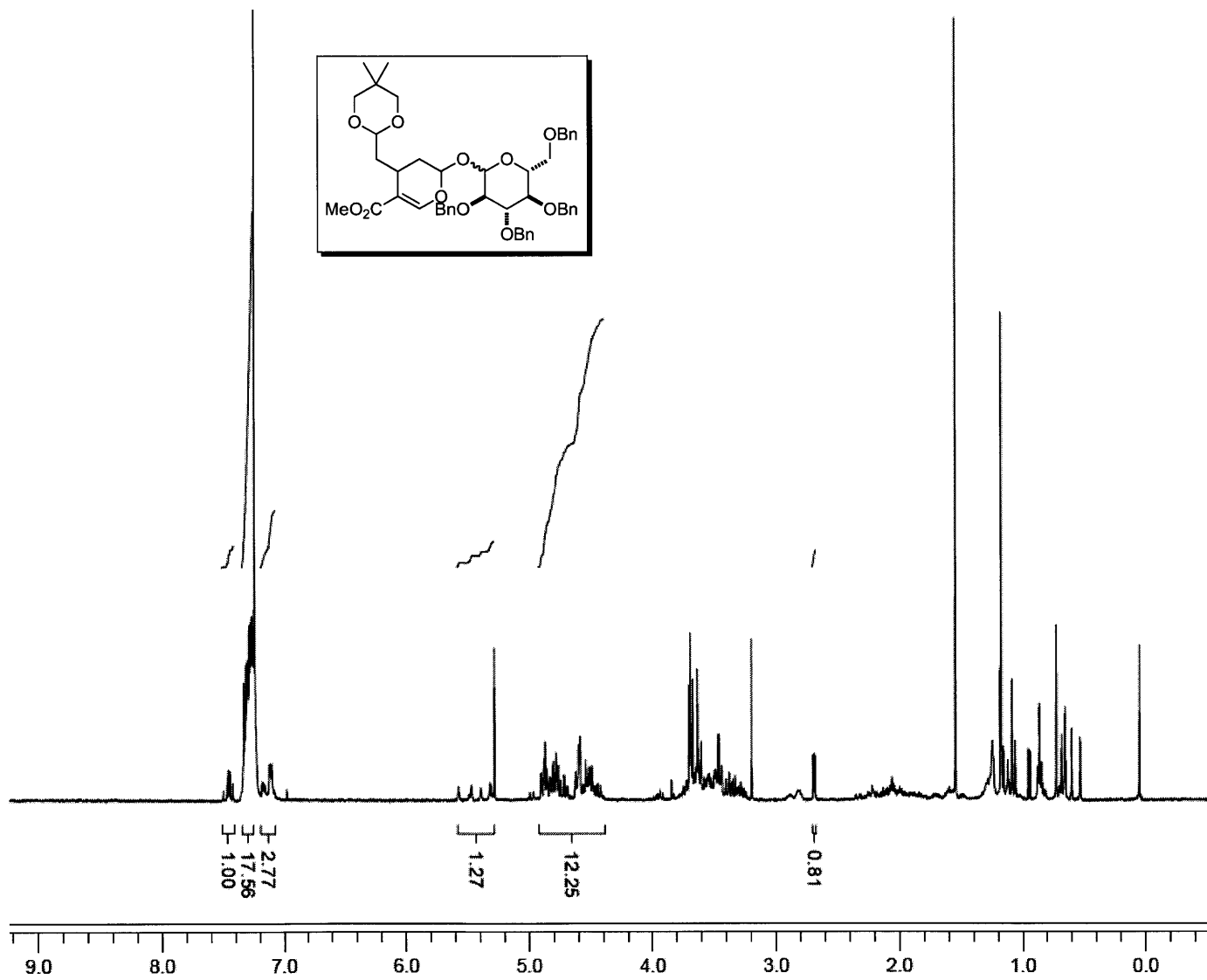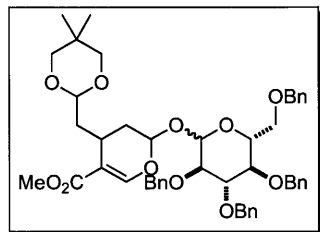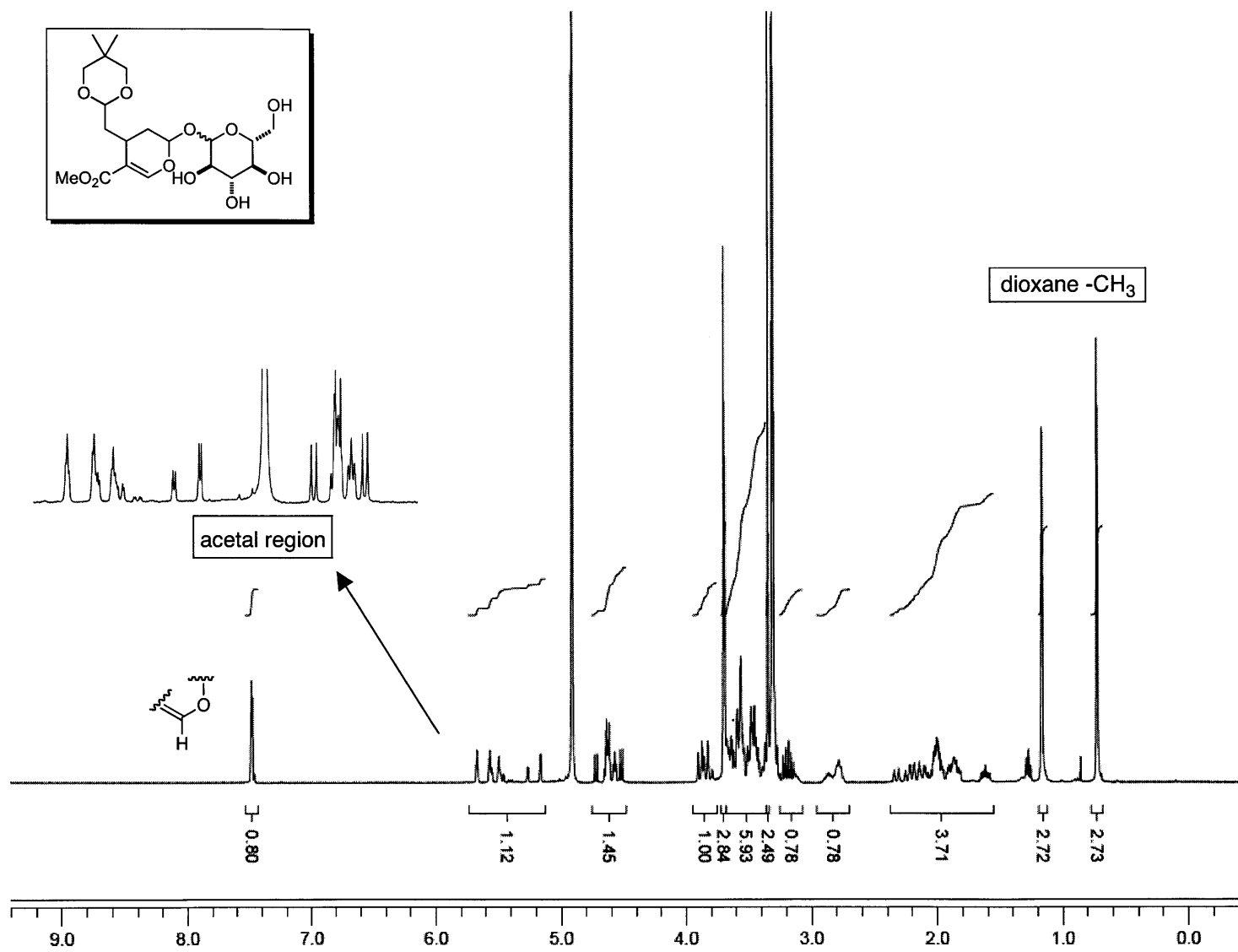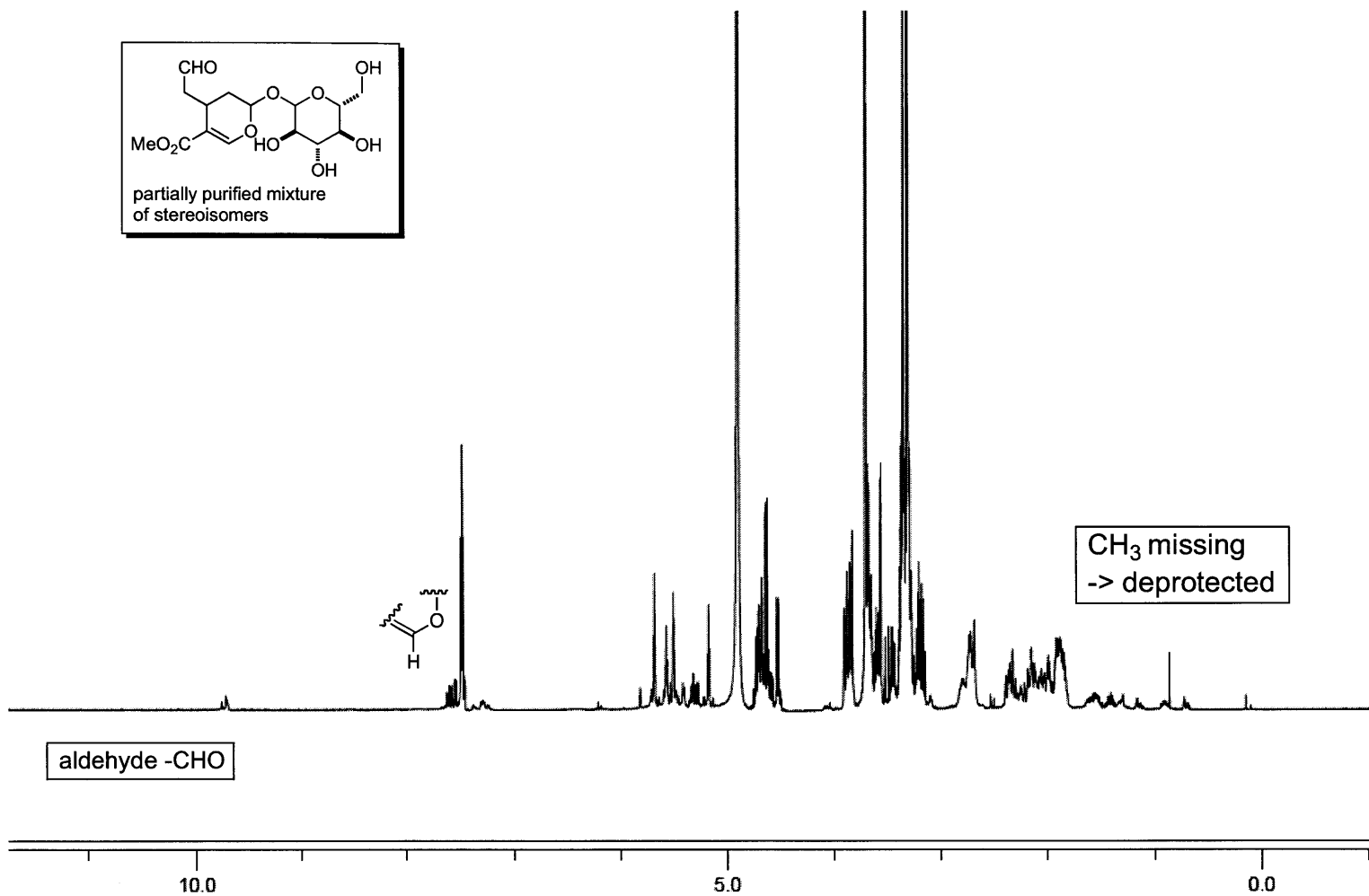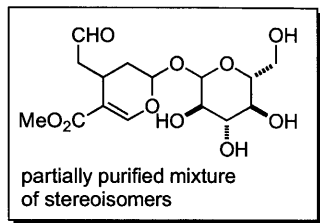
Compound **12a** (generated enzymatically from CrSTS, **1**, and **11**)

5.8 ppm,
dd (2,4-*trans*)

4.8 ppm,
d (*J* = 8 Hz)
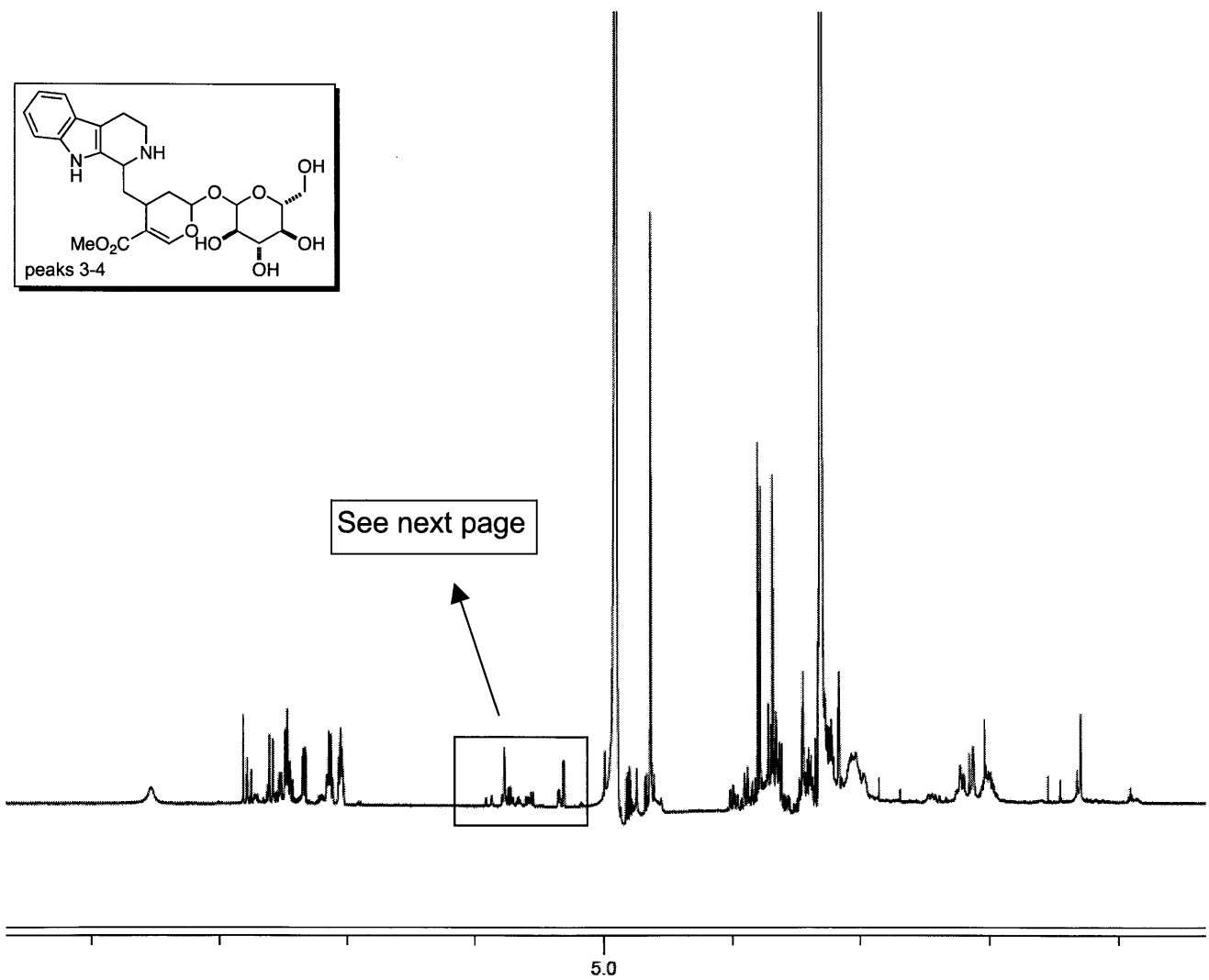(β-anomer)

peak 1

br-s
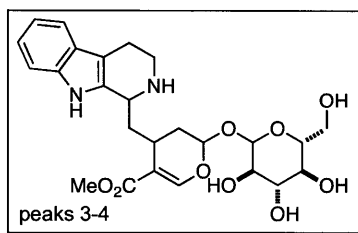(2,4-*cis*)

d 4.6-4.9 ppm
β-anomers
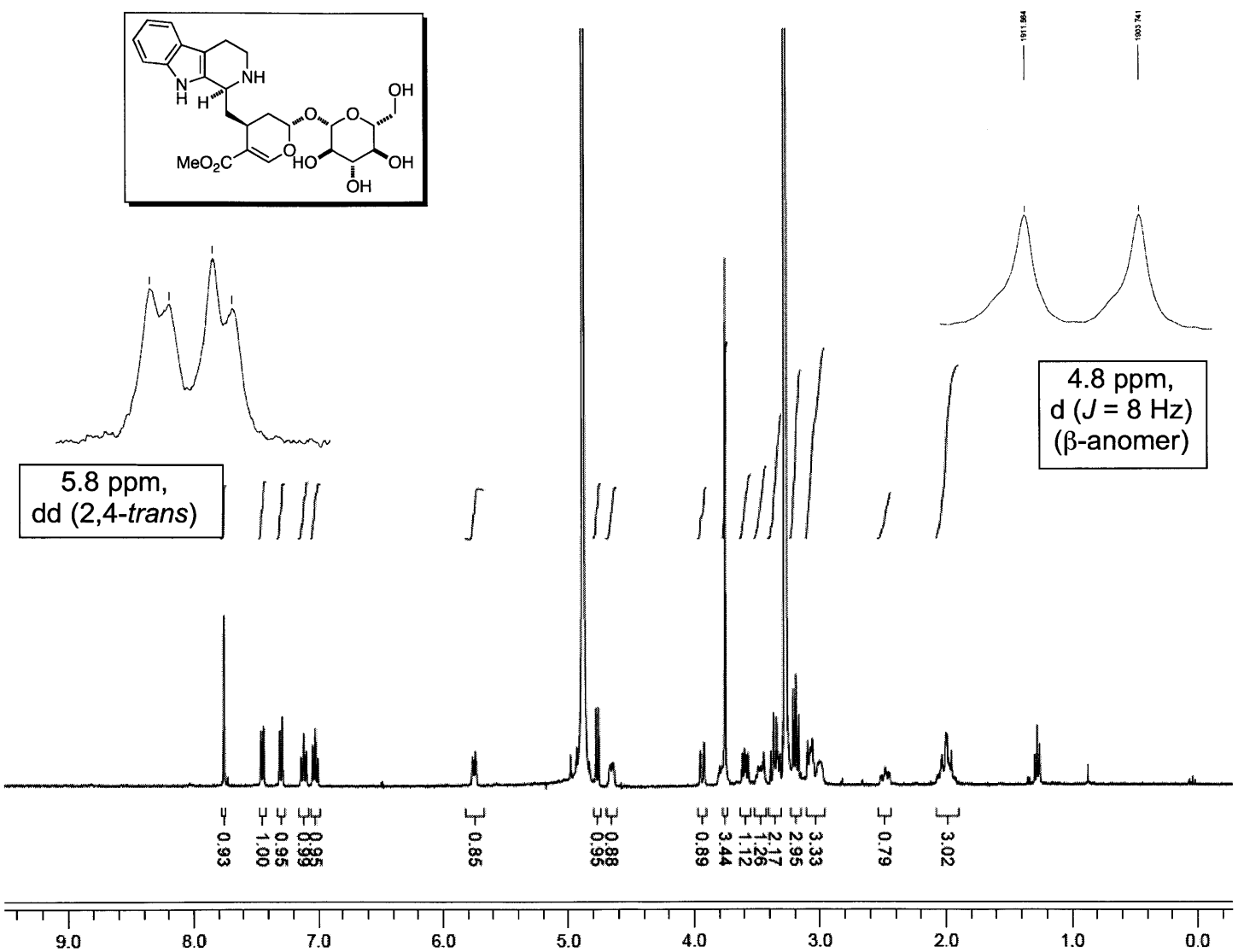
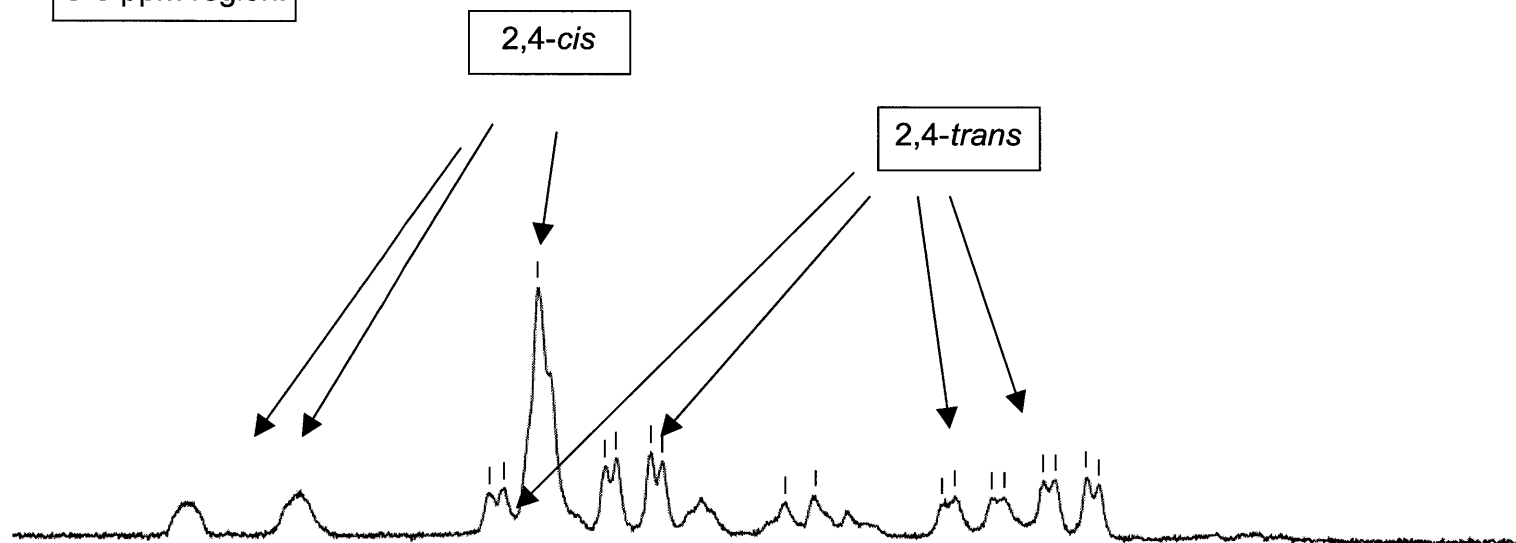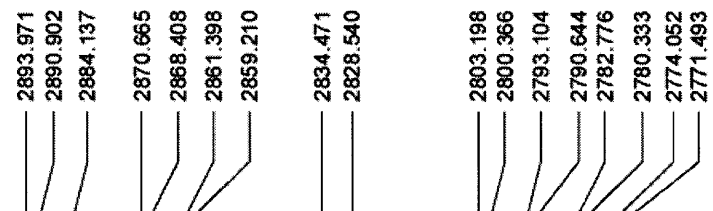Compound **12a** pk1 (generated from chemical reaction of **1** and **11**)
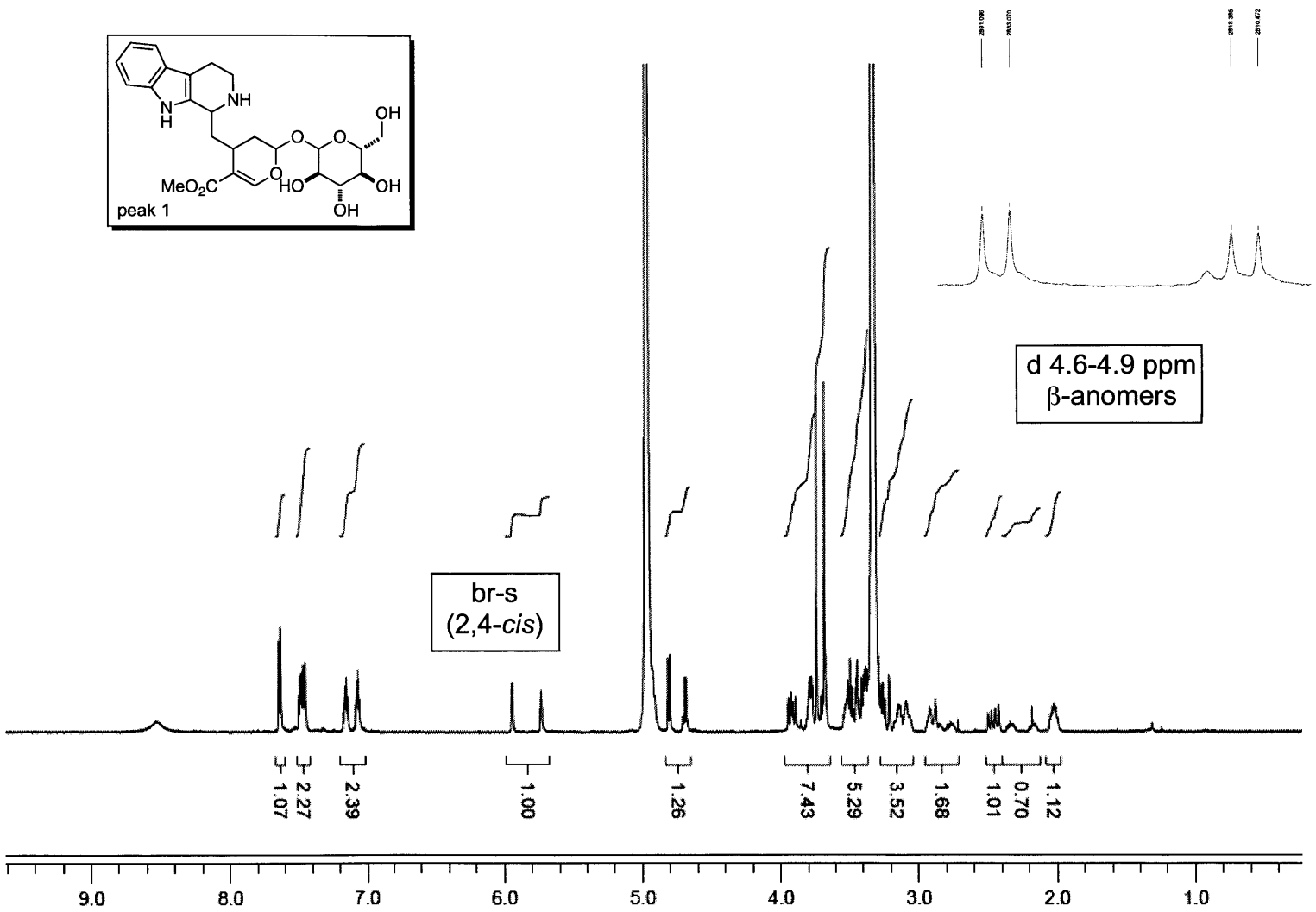
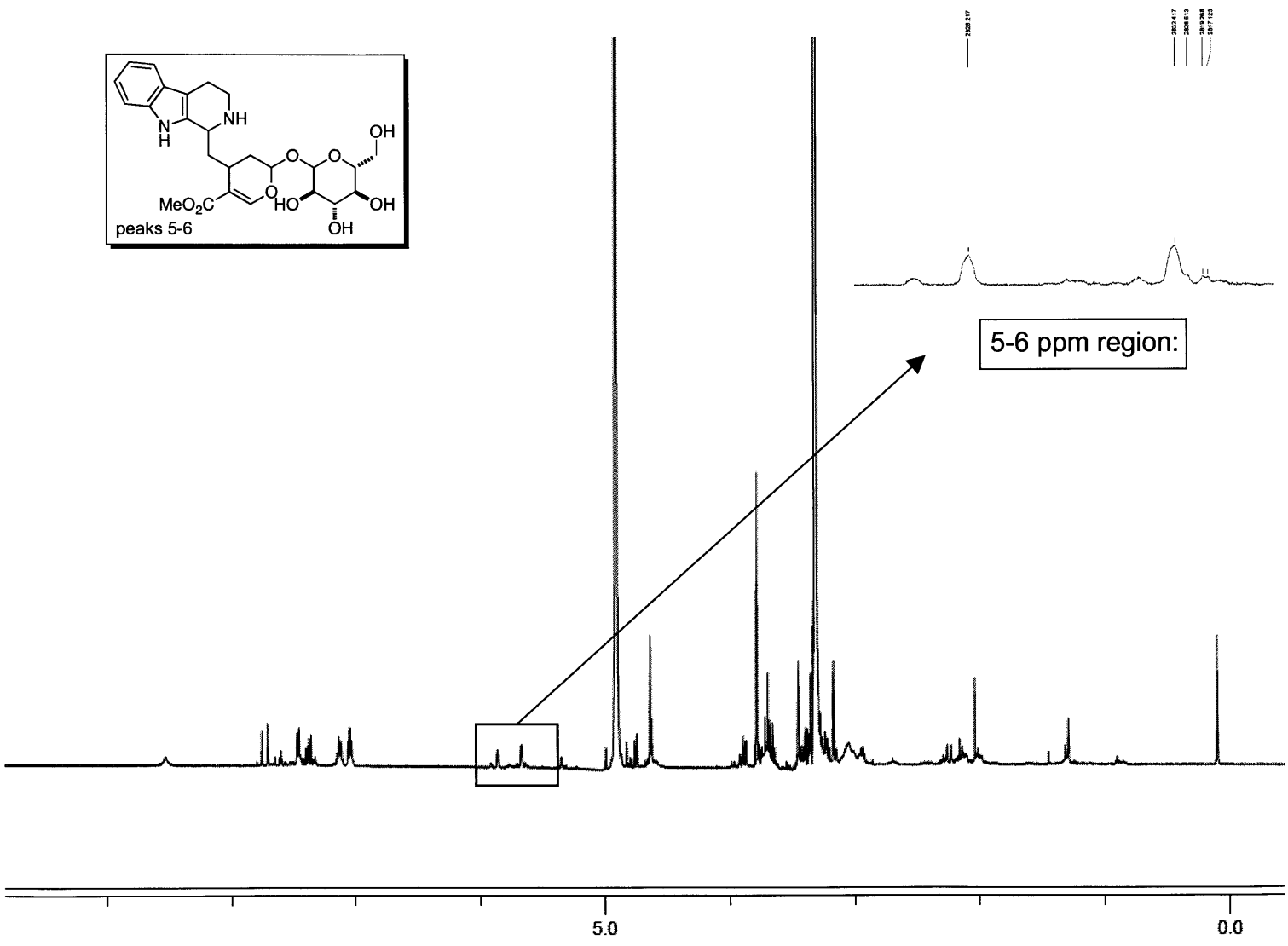Compound **12a** pk2 (generated from chemical reaction of **1** and **11**)

peak 2

br-s
(2,4-*cis*)

d (3.8 Hz)
(α-anomer)

2668.079
2664.245

1.12
1.00
1.39
1.32
0.97
0.95
1.86
0.99
9.74
9.45

9.0 8.0 7.0 6.0 5.0 4.0 3.0 2.0 1.0

peaks 3-4

See next page

5.0

peaks 3-4

5-6 ppm region:

2,4-*cis*

2,4-*trans*

2893.971
2890.902
2884.137
2870.665
2868.408
2861.398
2859.210
2834.471
2828.540
2803.198
2800.366
2793.104
2790.644
2782.776
2780.333
2774.052
2771.493

peaks 5-6
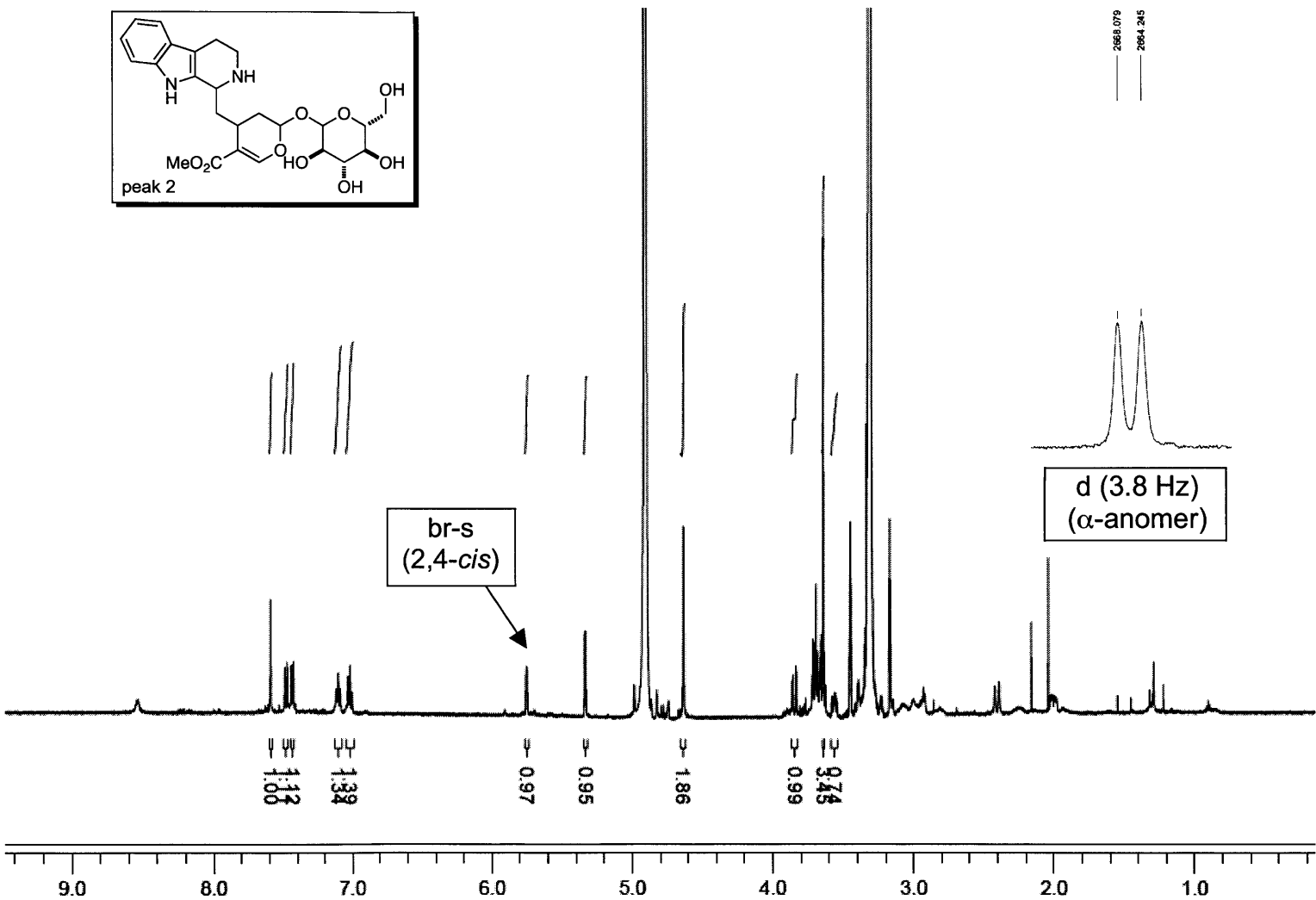
5-6 ppm region:

Compound **12a** pk5-6 (generated from chemical reaction of **1** and **11**)
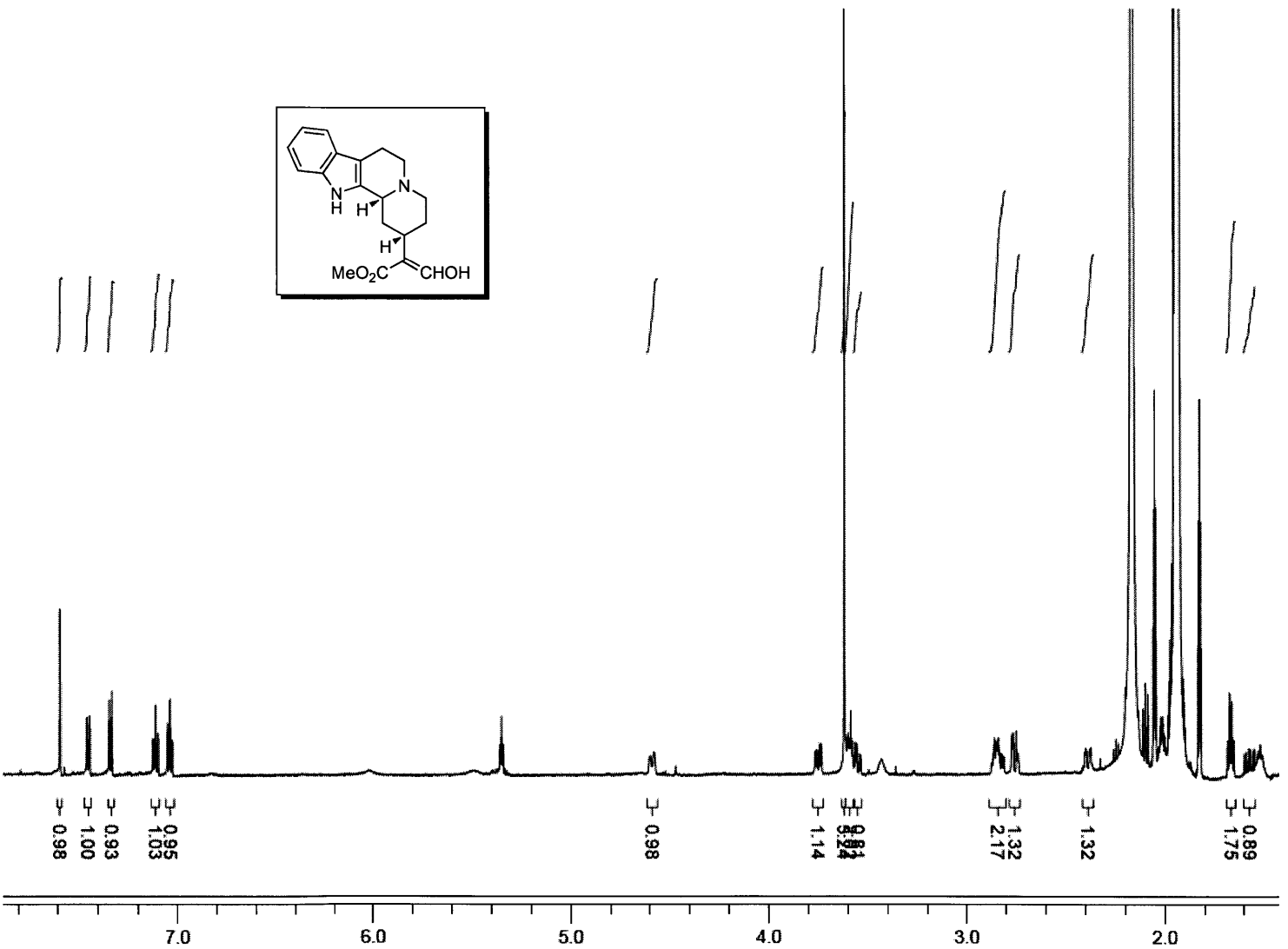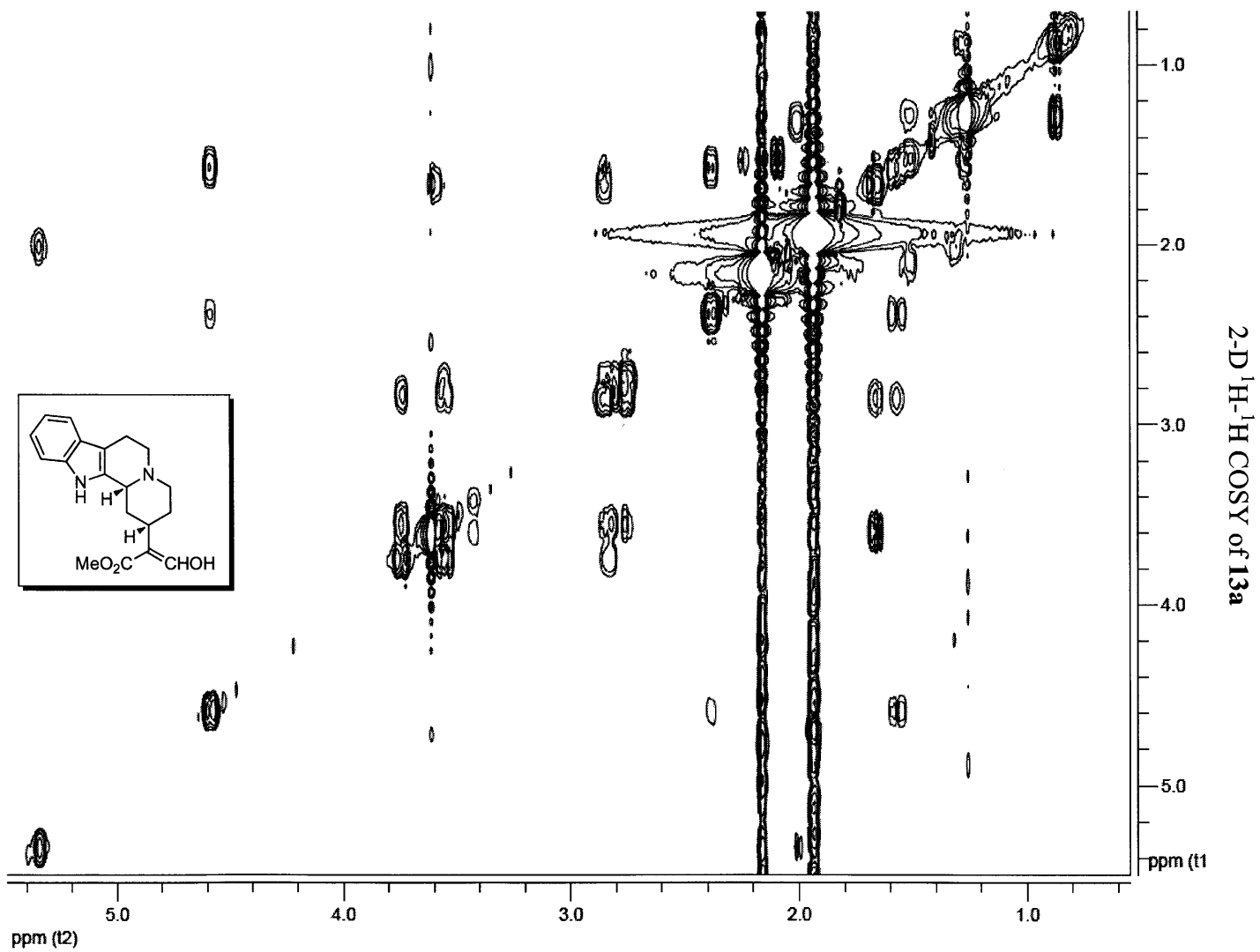
5.0

0.0

Compound **13a** (deglucosylated and reduced product of **12a**)

2-D ^{1}H-^{1}H COSY of **13a**

CV - Peter Bernhardt

Education

2005-2009   PhD Chemistry, MIT

2000-2004   MSc Biotechnology, Royal Institute of Technology (KTH)

Research Experience

2010-   Postdoctoral Scholar, Scripps Institution of Oceanography

2005-2009   Research Assistant, MIT

2004-2005   Research Specialist, University of Minnesota

2003-2004   Research Technician, Royal Institute of Technology (KTH)

Publications and Patents

1. P. Bernhardt, K. Hult, R.J. Kazlauskas, "Molecular basis of perhydrolase activity in serine hydrolases", *Angew. Chem., Int. Ed. Engl.* 2005, *44*, 2742-2746.

2. P. Bernhardt, E.A. McCoy, S.E. O'Connor, "Rapid identification of enzyme variants for reengineered alkaloid biosynthesis in periwinkle", *Chem. Biol.* 2007, *14*, 888-897.

3. R.J. Kazlauskas, P. Bernhardt, and T. Yin, "Increasing perhydrolase activity in esterases and/or lipases", 2008, US patent: 7,384,787.

4. P. Bernhardt, S.E. O'Connor, "Opportunities for enzyme engineering in natural product biosynthesis", *Curr. Opin. Chem. Biol.* 2009, *13*, 35-42.

5. P. Bernhardt, N. Yerkes, S.E. O'Connor, "Bypassing stereoselectivity in the early steps of alkaloid biosynthesis", *Org. Biomol. Chem.* 2009, *7*, 4166-4168.

6. P. Bernhardt, S.E. O'Connor, "Synthesis and evaluation of des-vinyl secologanin aglycones with alternate stereochemistry", *Tetrahedron Lett.* 2009, *50*, 7118-7120.

7. P. Bernhardt, A.R. Usera, S.E. O'Connor, "Strictosidine synthase from *Ophiorrhiza pumila* is a stereoselective and promiscuous biocatalyst for the Pictet-Spengler reaction", *submitted* 2009.

8. T. Yin, P. Bernhardt, K.L. Morley, Y. Jiang, J.D. Cheeseman, J.D. Schrag, R.J. Kazlauskas, "Switching catalysis from hydrolysis to perhydrolysis in *P. fluorescens* esterase.", *submitted* 2009.

Honors and Awards

2009   Daniel S. Kemp Graduate Fellowship

2009   Morse Travel Grant

2008   Amgen Graduate Fellowship

2005-2009   MIT Society of Presidential Fellows

2005-2006   Robert T. Haslam Presidential Graduate Fellowship

2004   Travel stipend for diploma thesis work, Royal Institute of Technology (KTH)

Selected Presentations

2009   Poster, GRC: Enzymes, Co-enzymes, and Metabolic Pathways, Waterville, NH

2008   Research talk, The Broad Institute of Harvard and MIT, Cambridge, MA

2007   Poster, Enzyme Engineering XIX Conference, Vancouver, BC

2007   Research talk, MIT Department of Biology, Cambridge, MA

2005   Research talk, Royal Institute of Technology, Stockholm, Sweden