Characterization of the Biosynthetic Pathway for Medicinal Monoterpenoid Indole Alkaloid

by

Jun Guo

BSc (Hons) Biology-Chemistry, University of New Brunswick, 2022

A Thesis Submitted in Partial Fulfillment of
the Requirements for the Degree of

Master of Science

in the Graduate Academic Unit of Chemistry

Supervisor: Yang Qu, Ph.D., Department of Chemistry

Examining Board: Gilles Villemure, Ph.D., Department of Chemistry, Chair
David Burns, Ph.D., Department of Chemistry
Bryan Crawford, Ph.D., Department of Biology

This thesis is accepted by the
Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

June 2023

ABSTRACT

Monoterpenoid indole alkaloids (MIA) are a complex and diverse class of alkaloids found in nature, boasting over 3000 reported structures. Many MIAs exhibit human medicinal properties, such as the anti-cancer drugs vinblastine and camptothecin. In this thesis, two enzymes involved in the biosynthesis of a plant derived MIA was elucidated and functionally characterized. To achieve this, a bioinformatics approach was used to shortlist candidate genes from the source plant, which were then cloned into a heterologous system to characterize the enzymes they encode. Experiment results show that one of the candidate genes codes for the required enzyme responsible for the biosynthesis of this MIA. The discovery from this work will allow assembly of the complete biosynthetic pathway in baker's yeast (Saccharomyces cerevisiae), facilitating *de novo* MIA synthesis.

ACKNOWLEDGEMENTS

I would like to express my most sincere gratitude to Professor Qu, whose guidance, support, and valuable insights have been instrumental in shaping my academic journey. Your commitment to excellence has inspired me to push my limits and achieve my goals. I cannot thank you enough for your unwavering encouragement and believe in me.

I would like to extend my gratitude to the Advisory Committee, Prof. Burns and Prof. Ignaszak, for all your helpful feedback throughout my research work. As well, I would like to thank current and alumni members of the Qu group, for the incredible amount of encouragement and support throughout my time here.

I would also like to thank the University of New Brunswick and the New Brunswick Innovation Foundation for providing me with an exceptional learning environment and financial support. The quality of education and resources offered here allowed me to develop both personally and professionally.

To my parents, I am grateful for your unconditioned love and encouragement throughout my academic journey. Despite being in different provinces/countries, you made every effort to ensure my wellbeing, and I will be forever grateful for that.

Last but not least, I would like to thank my partner Elisa for her love, support, and understanding during the ups and downs of my academic journey. Your constant motivation has been a source of strength, and I cannot thank you enough for always being there for me.

Once again, thank you all for being an integral part of my academic and personal growth.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

^{13}C-NMR = carbon-13 nuclear magnetic resonance

^{1}H-NMR = proton nuclear magnetic resonance

AAE = 17-*O*-acetylnorajmaline esterase

cDNA = complementary deoxyribonucleic acid

DHVR = dihydrovomilenine reductase

ESI = electrospray ionization

ESI-MS/MS = electrospray ionization mass spectrometry

GS = geissoschizine synthase

HMBC = heteronuclear multiple bond correlation

HSQC = heteronuclear single quantum coherence

LC-MS/MS = liquid chromatography with tandem mass spectrometry

MIA = monoterpenoid indole alkaloid

Rs (or *R. serpentina*) = *Rauwolfia serpentina*

NADPH = nicotinamide adenine dinucleotide phosphate

NNMT = norajmaline N-methyl transferase

NOSEY = Nuclear Overhauser Effect Spectroscopy

PCR = polymerase chain reaction

PNAE = polyneuridine aldehyde esterase

R_f = retention factor

RsCAD = *Rauwolfia serpentina* cinnamyl alcohol dehydrogenase

R_t = retention time

SAM = S-adenosyl methionine

SBE = sarpagan bridge enzyme

SD = chemical standard

SDS-PAGE = sodium dodecyl sulfate–polyacrylamide gel electrophoresis

SGD = strictosidine beta-glucosidase

STR = strictosidine synthase

TIC = total ion chromatogram

TLC = thin layer chromatography

TPM = transcript per million

UV = ultraviolet

VH = vinorine hydroxylase

VR = vomilenine reductase

VR2 = vomilenine reductase 2

VS = vinorine synthase

Chapter 1. Ajmaline and its Biosynthetic Pathway

1.1 Introduction to monoterpenoid indole alkaloids (MIA)

Humans have been using medicinal plants as natural remedies since ancient times. Perhaps the oldest written evidence for the use of medicinal plants is found in Nagpur, India, dating back 5000 years. Written on a clay slab, it comprised 12 recipes for drug preparation, some of which involved alkaloids such as poppy (Petrovska, 2012). As the human race evolved into modern times, advances in science allowed researchers to isolate the active compounds from medicinal plants and characterize them. Of the numerous classes of natural compounds reported, monoterpenoid indole alkaloids (MIA) are among the most complex and diverse class of alkaloids, with over 3000 structures reported thus far (Eng et al., 2022). MIAs are mostly found in the plant families Apocynaceae, Loganiaceae, and Rubiaceae, many of which have medically desirable properties. Some prominent MIAs used in modern medicine include quinine – an anti-malarial agent (O'Connor & Maresh, 2006), mitragynine – an analgesic agent acting on opioid receptors (Karunakaran et al., 2022), and vinblastine – an anti-cancer drug (Eng et al., 2022).

The structure of a typical MIA consists of two main components: the monoterpene moiety and the indole moiety. The monoterpene moiety includes a 9- or 10-carbon fragment derived from two isoprene units, while the indole moiety is an aromatic bicyclic structure comprised of a six-membered benzene ring joined with a five-membered pyrrole ring. Depending on how the monoterpene moiety is arranged, MIAs can be further categorized into several types: the *Corynanthe* type, the *Aspidosperma*

type, and the *Iboga* type (Dewick, 2002, p. 350), among many others. Figure 1 below shows the different arrangements of the monoterpene moiety in each class along with a representative MIA from each class. Ajmaline, which is the MIA of interest in this work, belongs to the Corynanthe type (Dewick, 2002, p.358). Like all other MIAs, the precursor secologanin (Fig. 2) provides the 10-carbon framework for the Corynanthe-type monoterpenoid moiety (highlighted in red; leftmost column). Through rearrangements of the Corynanthe skeleton, other MIA types, such as Iboga-type (center column, Fig. 1) and Aspidosperma-type (rightmost column, Fig. 1), will arise (Dewick, 2002, p. 351). The indole moieties (highlighted in blue, Fig. 1) are derived from tryptamine, which is from the indole amino acid tryptophan.

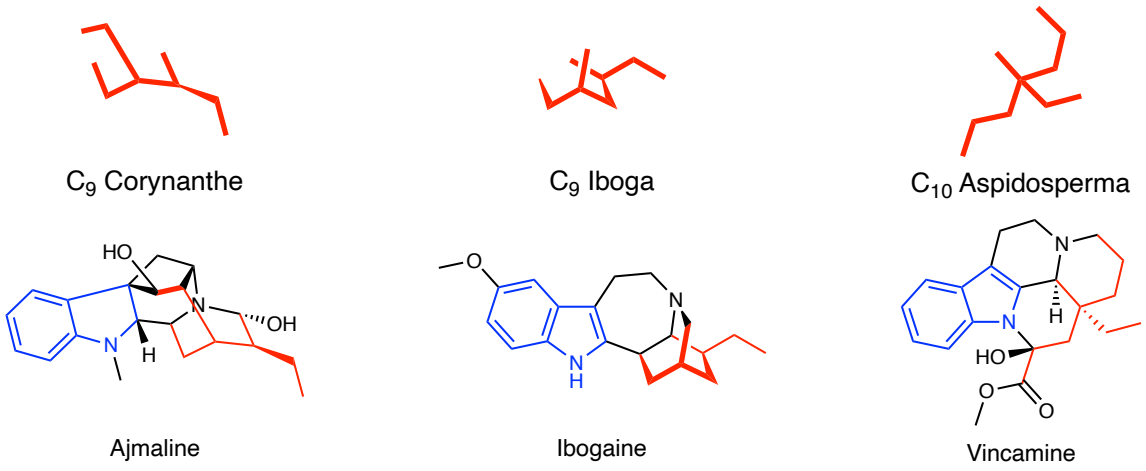


Figure 1. A diagram showcasing different types of monoterpenoid indole alkaloids based on the arrangement of the monoterpene moiety (top), along with a representative molecule from each class (bottom). The monoterpene moieties are highlighted in red, while the indole moieties are highlighted in blue. Adapted from Dewick, 2002, p. 358.

As briefly mentioned earlier, the general biosynthesis of monoterpene indole alkaloids starts from tryptamine (decarboxylated tryptophan) and secologanin (a

monoterpenoid derived from primary metabolites in the plastidial non-mevalonate pathway) (Dewick, 2002; O'Connor and Maresh, 2006). Figure 2 below shows the first committed step of MIA biosynthesis. This step is catalyzed by the enzyme strictosidine synthase (STR) condensing tryptamine and secologanin via a Pictet-Spengler type reaction to form strictosidine – an intermediate enroute to almost all MIAs (Ma et al., 2006). From here, strictosidine can undergo diverse modifications to form an incredibly wide range of compounds, thus making it hard to generalize the biosynthetic steps of all MIAs. The following chapters of this thesis will focus on the medicinal MIA ajmaline, its specific biosynthetic pathway (of what is reported to date), and how this work completes the decades-long research of ajmaline biosynthesis.

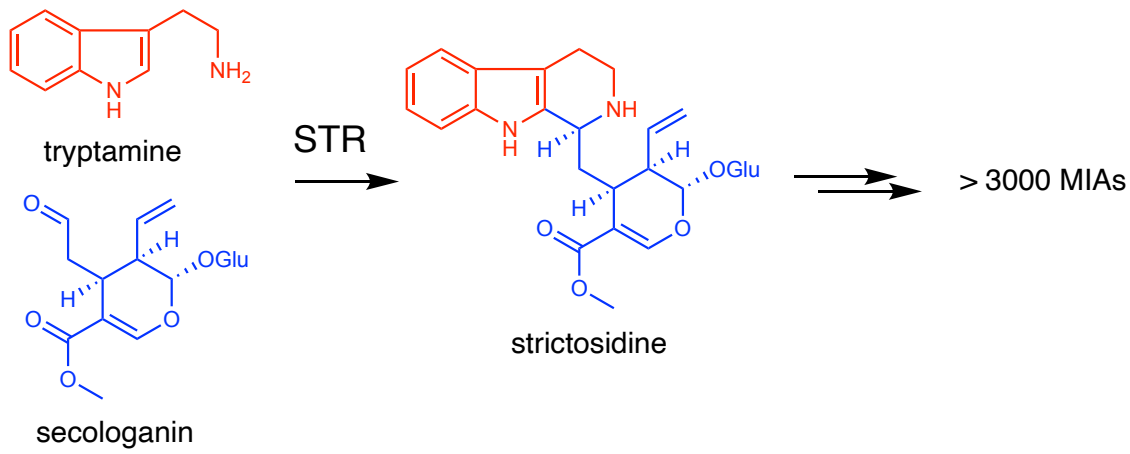


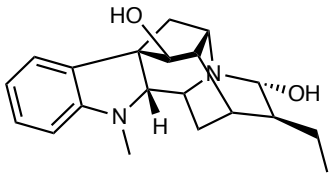
Figure 2. The structures of strictosidine and its precursors. Strictosidine is an intermediate enroute to almost all MIAs.

1.2 Ajmaline is a class Ia anti-arrhythmic MIA found in *Rauwolfia Serpentina*

The plant *Rauwolfia serpentina* (also known as Indian snakeroot) is a small shrub native to South and Southeastern Asia (Dewick, 2002). Historically, this plant has been used in Hindu culture as an Ayurvedic (ancient Indian medicine) under the names Sarpagandha, meaning snake repellent, and Chandrá, meaning moon and refers to the treatment of "moon's disease" or lunacy (Monachino, 1954). For this work, several *R. serpentina* plants were grown in greenhouse conditions locally in Canada; Fig. 3 shows one of the *R. serpentina* plants at the University of New Brunswick. As science advanced, many of the MIAs, which we now know are the main active ingredients responsible for the medicinal effects, have been extracted and characterized from *Rauwolfia serpentina*. Some notable indole alkaloids found in *R. serpentina* include reserpine that is used to treat high blood pressure (Monachino, 1954; Shamon & Perez, 2016), yohimbine that is used as a treatment for erectile disfunction (Ernst & Pittler, 1998; Tam et al., 2001), and ajmaline, which will be discussed in detail in the next paragraph.

Figure 3. One of the *Rauwolfia Serpentina* plants at the University of New Brunswick

Ajmaline in modern medicine is valued as a class Ia anti-arrhythmic drug and an important diagnostic agent for the rare genetic cardiovascular disease Brugada syndrome (Padrini et al., 1993; Rolf et al., 2003). It is available commercially under the trade names Gilurytmal, Ritmos, and Aritmina. Over the years, excellent work has demonstrated the total synthesis of ajmaline (Lewis, 2006; Edwankar et al., 2008; Chen et al., 2022). However, no commercially viable total chemical synthesis methods have been developed thus far. The difficulties in achieving ajmaline total synthesis can be explained by its structure. As seen from Figure 4, the structure of ajmaline involves an intricate bridge structure (known as a sarpagan bridge; Dang et al., 2018) and contains several stereocenters, thus making it difficult to synthesize this molecule efficiently and sustainably. Therefore, large-scale production of this medicinal MIA still relies on extraction from *R. serpentina* plant material. Though, because of the low abundancy of therapeutic alkaloids in *R. serpentina* (0.15-0.2% by plant weight; Dewick, 2002), paired with the low seed germination and survival rates (Goel et al., 2009), mass commercial cultivation of *R. serpentina* remains difficult and inefficient. Furthermore, *R. serpentina* is a designated endangered species in India owning to the over-exploitation of this medicinal plant (Goel et al., 2009). For all these reasons above, our group and other researchers are motivated to elucidate the biosynthetic steps of ajmaline in *R. serpentina*. Once the genes involved in the pathway are discovered, it is possible to clone them into a heterologous system such as bacteria (e.g., *Escherichia coli*) or Baker's yeast (*Saccharomyces cerevisiae*), thereby achieving heterologous production of ajmaline by fermentation. As of this work, all but two enzymes in the ajmaline biosynthesis pathway have been cloned and characterized by previous researchers.



Ajmaline

Figure 4. The structure of ajmaline

1.3 The 10-step biosynthesis of ajmaline: what is known and missing

Over the past four decades, most of the ajmaline biosynthetic pathway enzymes have been cloned and characterized, with only two reductases remaining to be elucidated. The ajmaline biosynthesis pathway enzymes and the reactions they catalyze are summarized in Figure 5 below. As mentioned in section 1.2, the first step of MIA biosynthesis begins with the enzyme strictosidine synthase (STR) condensing tryptamine and secologanin via a Pictet-Spengler type reaction to form strictosidine – an intermediate enroute to almost all MIAs (Ma et al., 2006). The enzyme strictosidine beta-glucosidase (SGD; Gerasimenko et al., 2002) then catalyzes the deglycosylation of strictosidine, generating several iminium aglycones in spontaneous equilibrium. These unstable aglycones are further reduced by a number of reductases to produce various stereoisomers (not shown in Fig. 5; O'Connor & Maresh, 2006; Szabó, 2008). One of the many reductases, geissoschizine synthase (GS; Qu et al., 2018), generates 19*E*-geissoschizine, which is further oxidatively cyclized to an unstable polyneuridine aldehyde by sarpagan bridge enzyme (SBE; Dang et al., 2018). From here, polyneuridine aldehyde esterase (PNAE; Dogru et al., 2000) hydrolyses the methoxyl group of polyneuridine aldehyde, forming 16-epi-vellosimine. Vinorine synthase (VS; Gerasimenko et al., 2004; Ma et al., 2005) then catalyzes the acetylation of this intermediate, leading to an intramolecular cyclization that forms vinorine. Next, a hydroxylation of vinorine by the enzyme vinorine hydroxylase (VH; Dang et al., 2017; Falkenhagen & Stöckigt, 1995) gives rise to vomilenine. Being the central intermediate in the ajmaline pathway, vomilenine is found in both the root and leaf tissues of *R. serpentina*, while ajmaline is only found in the root tissues (Chapter 3.1).

Continuing from vomilenine, in 2002 Gao and von Schumann used chromatographic methods to fraction crude *R. serpentina* proteins and discovered two putative reductases acting on vomilenine and its 1,2- reduced derivative. The first putative enzyme is named vomilenine reductase (VR), which carries out the reduction of vomilenine at its 1,2-imine double bond, forming 1,2(R)-dihydrovomilenine. Subsequent 19,20(S)- reduction of this 1,2- derivative by the second enzyme, dihydrovomilenine reductase (DHVR), yields 17-*O*-acetylnorajmaline. Because protein sequencing technology at the time was not as advanced, Gao (2002) and von Schumann (2002) were not able to elucidate the full amino acid sequence of VR and DHVR, which are needed to clone these two enzymes. After the two unidentified reductases converting vomilenine into 17-*O*-acetylnorajmaline, there are two additional modifications catalyzed by two already cloned and characterized enzymes. The penultimate step in this pathway is a deacetylation of 17-*O*-acetylnorajmaline by the enzyme 17-*O*-acetylnorajmaline esterase (AAE; Polz et al., 1987; Ruppert et al., 2005), forming norajmaline. Finally, the enzyme norajmaline *N*-methyltransferase (NNMT; Cázares-Flores et al., 2016) catalyzes an indole *N*-methylation of norajmaline to form the final product, ajmaline.

Although full amino acid sequencing for VR and DHVR was not possible, Gao (2002) and von Schumann (2002) used Edman degradation methods to reveal several partial peptide sequences believed to be part of VR and DHVR, allowing Geissler and co-workers in 2016 to search the now-available *R. serpentina* transcriptome and identify candidate reductases that contained these partial peptides. To their surprise, while Geissler's candidate enzymes did not perform the same functions as VR nor DHVR, one of their cinnamyl alcohol dehydrogenase (CAD)-like enzymes which they termed

vomilenine reductase 2 (VR2) instead reduced the 19,20- double bond of vomilenine. This suggests that there may be another enzyme that subsequently reduces the 1,2- double bond of 19,20-dihydrovomilenine into 17-*O*-acetylnorajmaline, signifying that vomilenine could go through an alternative reduction order (not shown in Fig. 5) rather than the strict 1,2- reduction followed by the 19,20- reduction (as suggested by Gao and von Schumann in 2002). Whichever way vomilenine may be reduced, the gene sequences encoding for the sequential reductases responsible for the conversion of vomilenine to 17-*O*-acetylnorajmaline remain undiscovered.

To summarize, in the classic ajmaline biosynthetic pathway, the universal MIA precursor strictosidine first undergoes several modifications by characterized enzymes to give rise to the intermediate vomilenine. From vomilenine, there are two unidentified reductases sequentially reducing the 1,2- and 19,20- double bond of vomilenine, though Geissler and coworkers (2016) suggest that vomilenine could go through an alternative reduction order. Via either or both reduction routes, the product 17-*O*-acetylnorajmaline is formed after the reductions, which then undergoes deacetylation and *N*-methylation catalyzed by two characterized enzymes to give rise to ajmaline. Of the many enzymes involved in the ajmaline pathway, the genes encoding for VR and DHVR remain undiscovered.

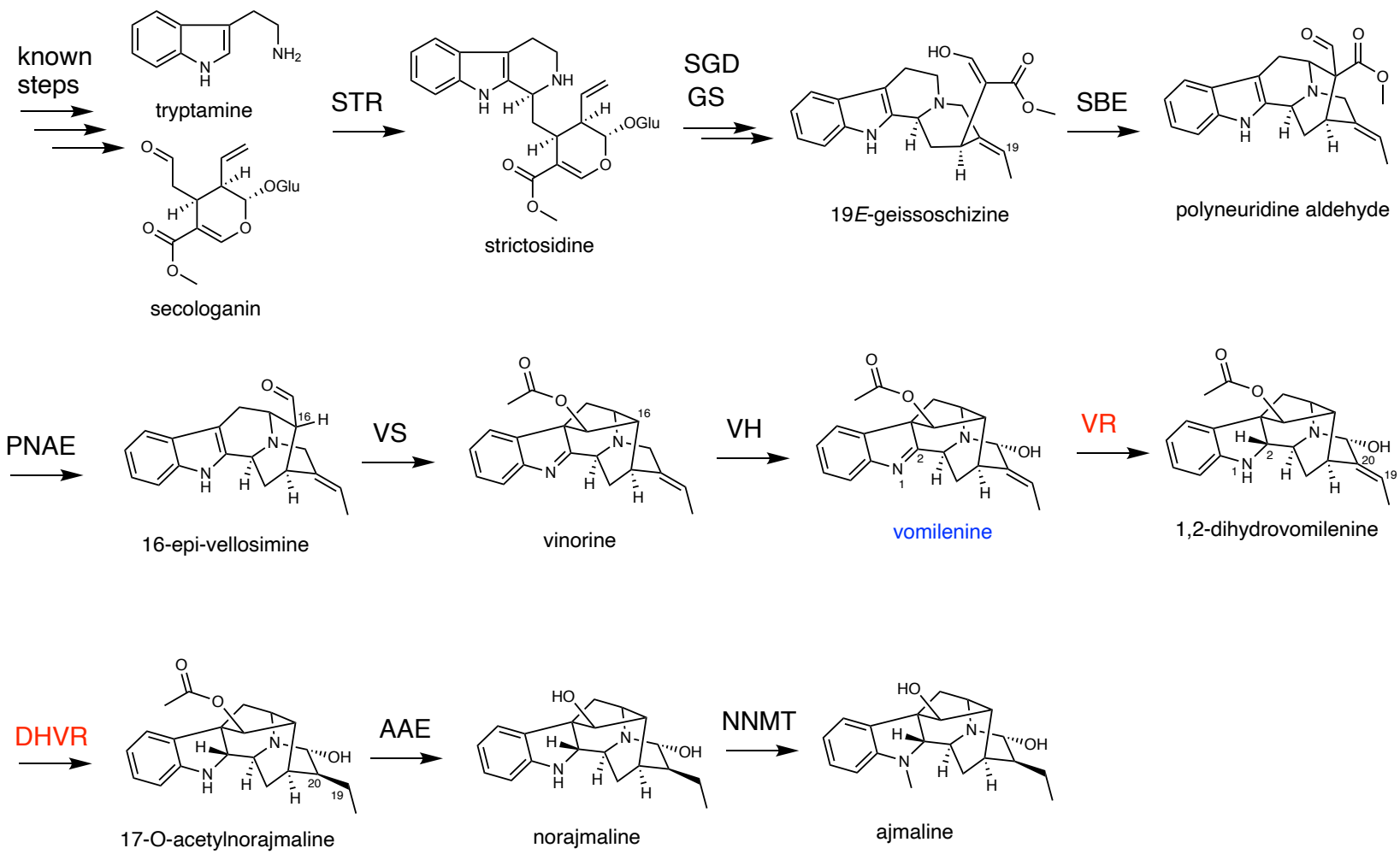


Figure 5. The biosynthetic pathway of ajmaline. Enzymes highlighted in red are unidentified (in terms of gene sequences). Vomilenine is highlighted in blue and is a central intermediate in this pathway.

1.4 Objective, experimental design, and rationale

The objective of this project is to identify and characterize the remaining reductases VR and DHVR from *R. serpentina*, which respectively and subsequently reduces vomilenine and its 1,2-dihydro- derivative into 17-*O*-acetylnorajmaline. A number of characterized reductases in MIA biosynthesis by our group and others belong to the NADPH dependent, cinnamyl alcohol dehydrogenase (CAD)-like enzymes, including the VR2 by Geissler and coworkers (2016). It is therefore hypothesized that the missing VR and DHVR are homologs of the known MIA CAD-like reductases. As such, interrogating the *R. serpentina* transcriptome for homologues of VR2 as well as other CAD-like reductases from different MIA pathways may generate a list of candidate genes for biochemical characterization.

To screen and characterize each reductase candidates, the enzymes will be cloned and expressed in *Escherichia coli (E. coli)*, which will then be assayed by feeding vomilenine substrate to the living cells. Using liquid chromatography tandem mass spectrometry (LC-MS/MS), VR enzyme activity would be evident from a change of vomilenine mass/charge ratio (*m/z*) = 351 to the reduced product dihydrovomilenine *m/z* = 353. The resulting dihydrovomilenine will be used as substrate in *E. coli* feeding experiments to screen for DHVR enzyme activity evident from a change of dihydrovomilenine mass/charge ratio (*m/z*) = 353 to the reduced product 17-*O*-acetylnorajmaline *m/z* = 355. The identity of the two reduced products will be verified by further including the already characterized AAE and NNMT for the production of ajmaline.

It is reasoned that coupled *in vitro* or *in vivo* biosynthetic pathway assembly will be the most effective method for verifying the intermediate (e.g., the *m/z* = 353 dihydrovomilenine) identity, since enzymatically produced intermediates will likely be of low quantity which does not allow formal NMR structural analysis, and none but ajmaline is commercially available as a standard. If the enzymatically produced product matches the characteristics of the commercial ajmaline standard, then this will undoubtedly confirm the stereochemistry for the dihydro-vomilenine product and the further reduced tetrahydro-vomilenine product.

In this thesis, both VR and DHVR (highlighted in red, Fig. 5) from *R. serpentina* have been identified and biochemically characterized to complete the ajmaline biosynthetic pathway. Enzyme assay results show that VR specifically reduces the 1,2-imine double bond of vomilenine and does not accept 19,20-dihydrovomilenine as a substrate. Furthermore, the previously reported VR2 is in fact DHVR. While it showed vomilenine 19,20- reductase activity, DHVR has 26-fold higher affinity for 1,2-dihydrovomilneine, indicating that vomilenine is not the preferred substrate for VR2/DHVR but rather 1,2-dihydrovomilenine. With the two reductases discovered, *in vivo* ajmaline biosynthesis was achieved in yeast by the enzymes VR, DHVR, AAE, and NNMT from vomilenine substrate.

Chapter 2. Materials and Methods

General considerations:

Commercial ajmaline standard was purchased from Toronto Research Chemicals (Toronto, ON, Canada). The norajmaline standard was a generous gift from Dr. Vincenzo De Luca at Brock University (Cázares-Flores et al., 2016). NMR spectra were recorded using an Agilent 400MHz spectrometer with $CDCl_3$ referenced at 7.26 ppm. LC-MS/MS measurements were made using an Agilent Ultivo Triple Quadrupole LC-MS equipped with an Avantor SuperC18 column (2.5 μm, 50x3 mm) and with the following solvent systems: solvent A, 29:71:2:398 (v/v) methanol:acetonitrile:1M ammonium acetate:water; solvent B, 130:320:0.25:49.7 (v/v) methanol:acetonitrile:1M ammonium acetate:water. The following linear gradient (8 min, 0.6 mL/min) were used: 0 min 80% A, 20% B; 0.5 min, 80% A, 20%B; 5.5 min 1% A, 99% B; 5.8 min 1% A, 99% B; 6.5 min 80% A, 20% B; 8 min 80% A, 20% B. The MS/MS was operated with gas temperature at 300 °C, gas flow of 10 L/min, capillary voltage 4 kV, fragmentor 135 V, collision energy 30 V with positive polarity. Qualitative Analysis 10.0 software by Agilent was used for all LC analyses.

Plant materials and RNA/cDNA synthesis

The plant *Rauwolfia serpentina* was grown in a greenhouse at 28 °C with 16/8 h photoperiod. Mature leaf (1 g) and root (1 g) tissues were collected for RNA extraction using standard TRIzol® RNA isolation reagent according to the manufacture's protocol (ThermoFisher Scientific). The resulting RNA was used to generate cDNA using the

LunaScript®RT SuperMix Kit according to the manufacture's protocol (New England Biolabs).

Vomilenine Purification from Plant Material

Mature leaves (600 g) from greenhouse grown *R. serpentina* were harvested for vomilenine extraction. The leaves were submerged in ethyl acetate for 30 mins and evaporated. The alkaloids were extracted first by 1 M HCl and ethyl acetate. The aqueous phase was basified with NaOH to pH >8, and subsequently extracted with ethyl acetate to afford total crude alkaloids. Total crude alkaloids were then separated by thin layer chromatography (SiliCycle; silica gel F254) employing a 9:1 ethyl acetate:methanol (v/v) mobile phase. Each distinct TLC bands were isolated from the plate and analyzed by LC-MS/MS. TLC harvested vomilenine was identified by LC-MS/MS, by comparing to enzymatically produced vomilenine, and by comparing its NMR spectra to literature reported vomilenine spectra (Dang et al., 2017; Ferreira Batista et al., 1996; Libot et al., 1980).

Cloning

RsCAD1-8, both truncated and full PhytoMetaSyn VR2, PNAE, VS, VH, AAE, and NNMT were amplified from *Rauwolfia serpentina* root and leaf combined cDNA using primers listed in Appendix II. Codon-optimized *Gelsemium serpemvirens* SBE was synthesized by Bio Basic Inc. (Markham, ON, Canada) and subcloned using primers listed in Appendix II. The following genes were cloned in pET30b+ vector in various restriction sites: RsCAD1, 2, 4-7, and both full/truncated VR2 were within BamHI/SalI

sites, RsCAD3 was within EcoRI/SalI sites, and RsCAD8 was within BamHI/XhoI sites. These vectors were then mobilized to *E. coli* BL21DE3 for expression. Because RsVR2 (full version) showed poor expression in BL21DE3, this vector was mobilized to BL21A1 instead. RsAAE was cloned in pESC-Ura yeast expression vector within EcoRI/NotI sites. RsCAD1-8, both full/truncated VR2, and other ajmaline pathway genes were also cloned into pESC yeast expression vectors of various selection markers (Leu, Ura, His, or Trp) at sites found in Appendix II. These vectors were mobilized to *S. cerevisiae* PEP4 for expression. The polymerases and restriction enzymes used for cloning were purchased from New England Biolabs, and the T4 DNA ligase was purchased from Promega Corporation; these products were used according to the manufacturer's protocol.

In vivo Biotransformation

E. coli strains BL21DE3 or BL21A1 containing various ajmaline pathway genes were inoculated in 1 ml LB media with appropriate antibiotics overnight at 37°C in a shaking incubator. The overnight cultures were used to inoculate 10 ml fresh LB media (1 in 100 dilution) with appropriate antibiotics, which were further grown at 37°C in a shaking incubator until OD600 reached 1.0. A final concentration of 0.1 mM IPTG (for BL21A1, also 0.1% (w/v) arabinose) was then added to the cultures and induced overnight in a shaking incubator at 15°C. The induced cells were collected and resuspended in 2 ml Tris HCl pH 7.5 supplemented with 10% (v/v) LB broth. Substrates including either 0.2 μg of vomilenine, 1,2-dihydrovomilenine, or 17-*O*-acetylnorajmaline was added to the biotransformation mixture, which was incubated in a shaking incubator

at 15°C for 24 hr. The reactions were terminated with equal volumes of methanol and used for LC-MS/MS analyses.

Yeast strain PEP4 containing various biosynthetic genes were inoculated in 1 ml drop-out (Leu, His, Ura, and/or Trp) SC media with 2% glucose (w/v) overnight at 30°C in a shaking incubator. The cells were collected by centrifugation, washed twice with water, and resuspended in 1 ml drop-out SC media with 2% galactose (w/v) for 24 hr at 30°C in a shaking incubator. The induced cells were collected and resuspended in 1 ml Tris-HCl pH 7.5 with 0.2 ug substrate vomilenine, 1,2-dihydrovomilenine, or 19*E*-geissoschizine. The biotransformation mixtures were incubated in a shaking incubator at 30°C overnight, then equal volume of methanol was added for LC-MS/MS analyses.

1,2-dihydrovomilenine, 19,20-dihydrovomilenine, and 17-*O*-acetylnorajmaline enzymatic production

E. coli strains BL21DE3 containing CAD2 or VR2 were inoculated in 2 ml LB media with appropriate antibiotics overnight at 37°C in a shaking incubator. The overnight cultures were used to inoculate 200 ml fresh LB media (1 in 100 dilution) with appropriate antibiotics, which were further grown at 37°C in a shaking incubator until OD600 reached 0.6. A final concentration of 0.1mM IPTG was then added to the cultures and induced overnight in a shaking incubator at 15°C. The induced cells were collected and resuspended in 20 ml Tris HCl pH 7.5 supplemented with 10% (v/v) LB. For 1,2-dihydrovomilenine, 50 μg vomilenine was added to cells expressing CAD2. For 19,20-dihydrovomilenine, 50 μg vomilenine was added to cells expressing tVR2. For

17-*O*-acetylnorajmaline, 50 µg vomilenine was added to an equal mixture of cells expressing CAD2 and VR2. The reactions were then incubated in a shaking incubator at 15°C for 24 hr and monitored for product formation by LC-MS/MS. After sufficient product has been formed, the alkaloid products were extracted by EtOAc and dried under vacuum. The resulting alkaloid solids were then resuspended in MeOH and purified by TLC (see vomilenine purification). A pre-determined vomilenine standard curve was used to quantify 1,2-dihydrovomilenine, 19,20-dihydrovomilneine, and 17-*O*-acetylnorajmlaine.

Recombinant Protein Expression and Purification

An overnight culture (2 ml) of *E. coli* BL21DE3 or BL21A1 strains containing ajmaline pathway genes in pET30b+ vector was used to inoculate 200 mL lysogeny broth (LB) media, which were cultured at 200 rpm at 37°C until OD600 reached 0.6-0.7. The cultures were induced with 0.1 mM IPTG at 15°C, 200 rpm overnight. For full VR2, an overnight culture (2 ml) of *E. coli* BL21A1 was used instead to inoculate 200 ml LB media, which were cultured at 200 rpm and 37°C until OD600 reached 0.6-0.7. The cultures were induced with 0.1 mM IPTG and 0.1% (w/v) arabinose at 15°C, 200 rpm overnight. The induced cells were pelleted, resuspended in 10 mL sample buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 10% (v/v) glycerol) with 20 mM imidazole, sonicated (10 s x 4 rounds), centrifuged (10,000 g, 10 mins), and the resulting supernatant were subjected to Ni-NTA affinity chromatography protein purification (HisPur™, Thermo Scientific) according to the manufacturer's protocol. After eluting with 250 mM imidazole in sample buffer, the purified recombinant proteins were desalted using a

PD-10 desalting column (GE Health Sciences) according to manufacturer's protocol into sample buffer and stored at -80°C.

Because RsAAE is a glycoprotein and requires proper folding through eukaryotic protein glycosylation pathway, the *Saccharomyces cerevisiae* PEP4-strain was chosen to functionally express and purify this protein. An overnight culture (5 mL) of *S. cerevisiae* PEP4 strain carrying pESC-Ura-RsAAE were used to inoculate 500 mL of synthetic complete media without uracil (SC-Ura) with 2% glucose (w/v) overnight at 30°C in a shaking incubator. The cells were collected by centrifugation, washed twice with water, and resuspended in 500 ml SC-Ura media with 2% galactose (w/v) for 24 hr at 30°C in a shaking incubator. The induced cells were collected, lysed in ice-cold sample buffer using a microtube homogenizer and glass beads (180 s x 3 rounds; 5 mins pause in between rounds; Bead Bug), and subsequently purified via the Ni-NTA methods above.

Obtaining *R. serpentina* total root proteins

Greenhouse grown *R. serpentina* mature roots (50 g) were harvested and frozen in liquid nitrogen. After the root tissues were frozen solid, a mortar and pestle was used to crush the roots into a fine powder. The resulting powder was resuspended in 40 mL sample buffer along with 1 mM dithiothreitol (DTT) and 5% (w/v) polyvinylpyrrolidone (PVP), mixed well, and filtered through a cheese cloth. The filtrate was then centrifuged (15,000 rpm, 25 mins, 4°C), and the resulting supernatant was twice desalted via the same protocol as above to obtain *R. serpentina* total root proteins.

In vitro Assays and Kinetics

A standard *in vitro* reaction (50 µl) included 20 mM Tris-HCl pH 7.5 and some or all these components: 100 µM SAM, 1 mM NADPH, 1 µg each of various ajmaline pathway enzymes, or 20 ug *R. serpentina* total root enzymes. After feeding 0.2 µg of the substrate (vomilenine or other intermediates in the ajmaline pathway), the reaction was incubated at 30°C for 1 h and terminated by adding 150 µL methanol. The kinetics assays (50 µl) were run in triplicates and included 20 mM Tris-HCl pH7.5, 1 mM NADPH, 1 µg CAD2 or tVR2, and substrate vomilenine at 4, 6.6, 10, 20, 30, 40, and 80 µM for RsCAD2 and 40, 66, 100, 200, 300, 400, and 800 µM for tVR2. As well, we measured enzyme kinetics of tVR2 with substrate 1,2-dihydrovomilenine at 4, 6.6, 10, 20, 30, 40, and 80 µM. The kinetics assays were performed at 30°C for 2 min before they were terminated by adding 150 µL methanol to the reactions. The products were quantified using a standard curve to generate the enzyme velocity. The kinetics parameters and saturation curves were approximated using the software Prism 9.5.0 (GraphPad Software, LLC.).

Chapter 3. Experimental Results and Discussion

3.1 Purification and identification of vomilenine in *R. serpentina* plant material

Vomilenine is the substrate needed to test our candidate enzymes for VR and DHVR activity. Since it is not available commercially, LC-MS/MS was used to examine the total alkaloids from *R. serpentina* grown in the greenhouse for the existence of this MIA. A major MIA exists in both root and leaf tissues, which matches the basic characteristics of vomilenine by *m/z* 351 (Fig. 6) and by UV maximum absorption at 260 nm indicating an indolenine chromophore (Fig. 7). As well, both ajmaline and

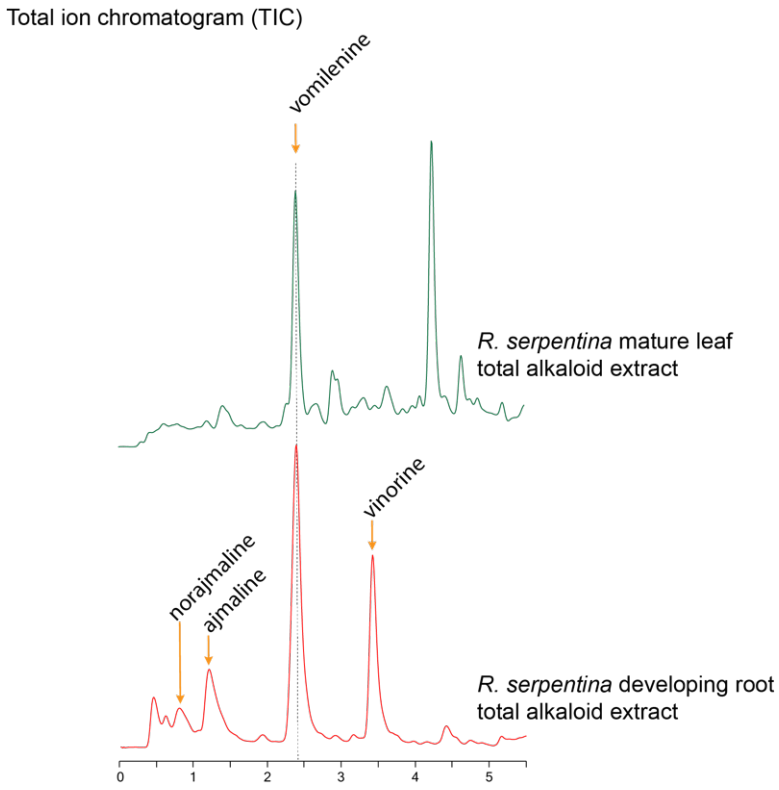


Figure 6. LC-MS/MS total ion chromatogram (TIC) showing vomilenine accumulation in both *R. serpentina* root and leaf tissues. Norajmaline, ajmaline, and vinorine were also presented in root tissue.

norajmaline were identified in *R. serpentina* root tissues by comparing them to the authentic standards (Fig. 6). With the experiments described below (Fig. 14), a major peak in root representing vinorine was also identified (Fig. 6).

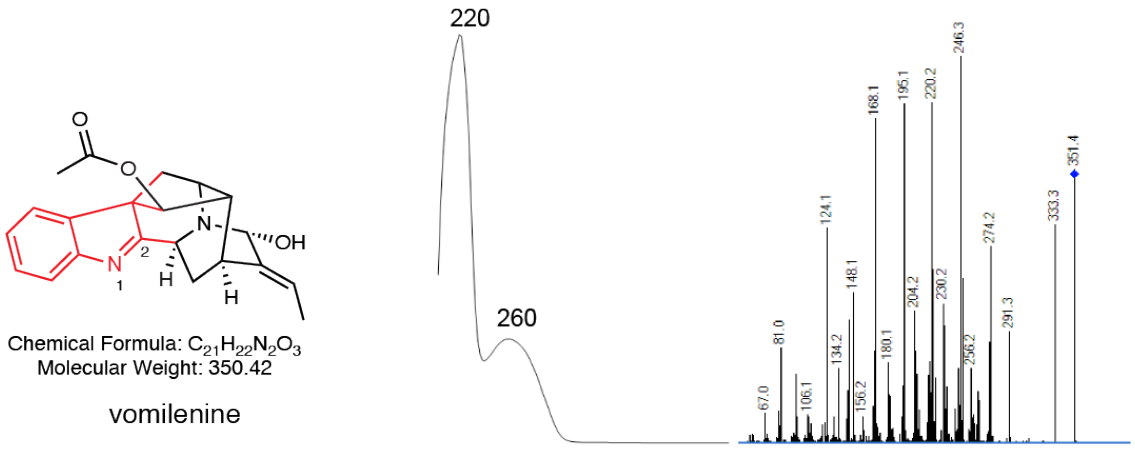


Figure 7. The UV absorption profile (center) and electrospray ionization mass spectrometry (ESI-MS/MS) ion fragmentation patterns (right) of vomilenine. The indolenine chromophore of vomilenine is highlighted in red (left) and shows a characteristic UV absorption at 260nm (von Schumann, 2002), consistent with the observation.

With the initial identification of vomilenine, this MIA was purified from total alkaloids by thin layer chromatography (TLC) (Fig. 8). The band representing vomilenine (labeled in Fig. 8) showed a retention factor of 0.27 under the 9:1 EtOAc:MeOH mobile phase. From 600g of *R. serpentina* leaves, 2 mg of the suspected vomilenine (*m/z* 351) was extracted and purified by TLC for nuclear magnetic resonance (NMR) analysis. Table 1 lists the chemical shifts for both proton (^{1}H) and carbon-13 (^{13}C) NMR of the *m/z*

351 MIA (Fig. 9-10), which matched well with literature reported vomilenine values (Table 1; Dang et al., 2017; Ferreira Batista et al., 1996; Libot et al., 1980). To further confirm the structure, Heteronuclear Multiple Bond Correlation (HMBC), Heteronuclear Single Quantum Coherence (HSQC), and Nuclear Overhauser Effect Spectroscopy (NOSEY) spectra for the *m/z* 351 MIA (Fig. 11-13) were recorded and analyzed, which were consistent with 1) literature reported spectra and 2) what is expected for the vomilenine structure. Therefore, the extracted *m/z* 351 MIA was identified as vomilenine.

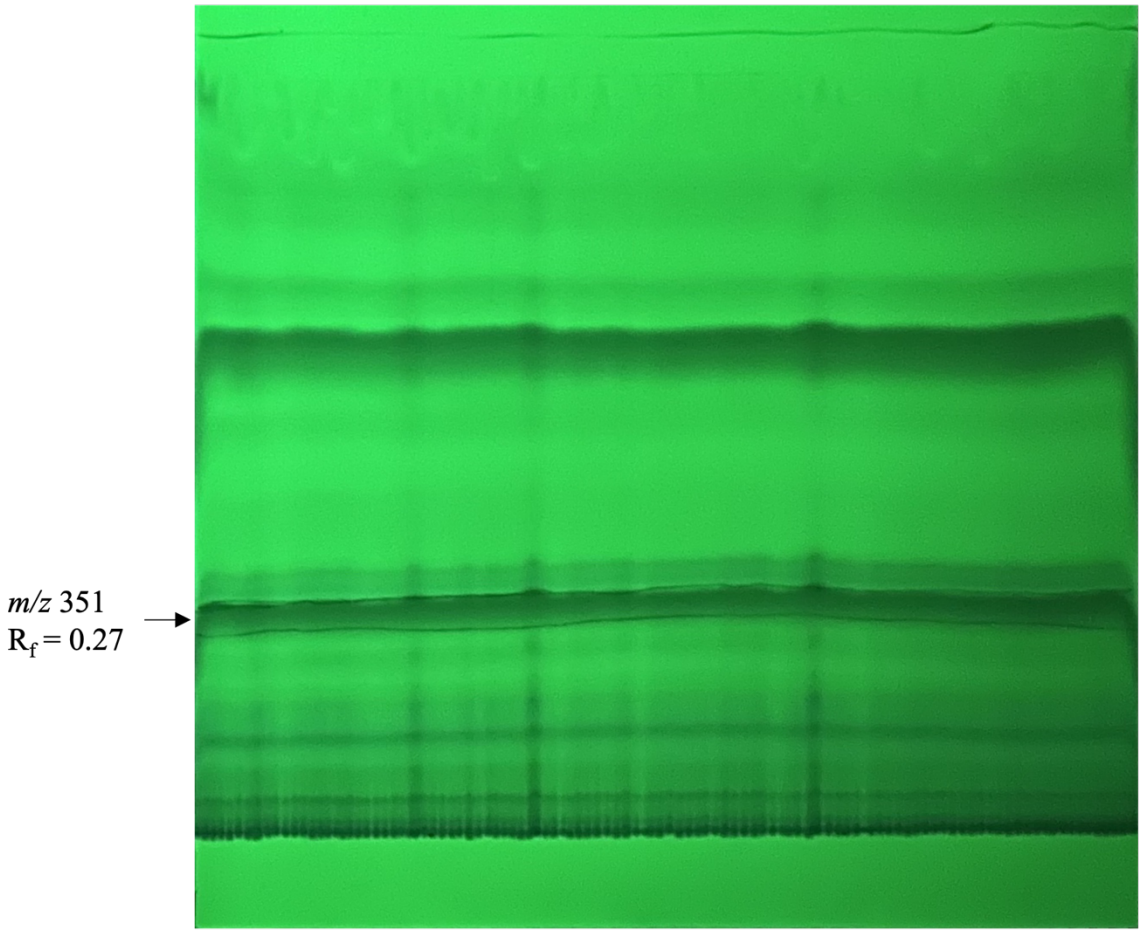


Figure 8. *R. serpentina* total alkaloids separated by TLC and visualized under UV light. The m/z 351 band was harvested and subjected to NMR analysis.

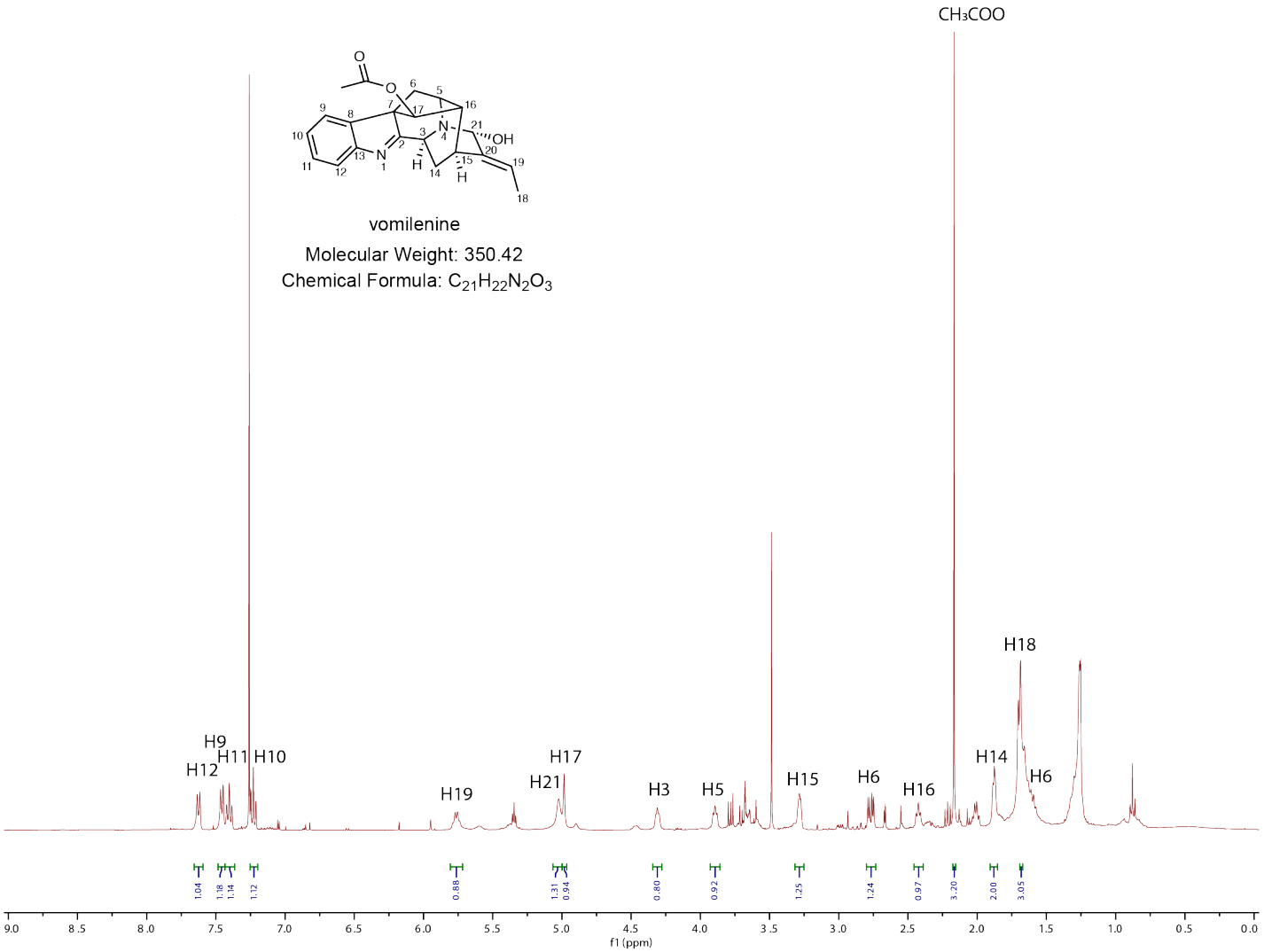


Figure 9. ^{1}H-NMR spectrum of vomilenine in $CDCl_3$

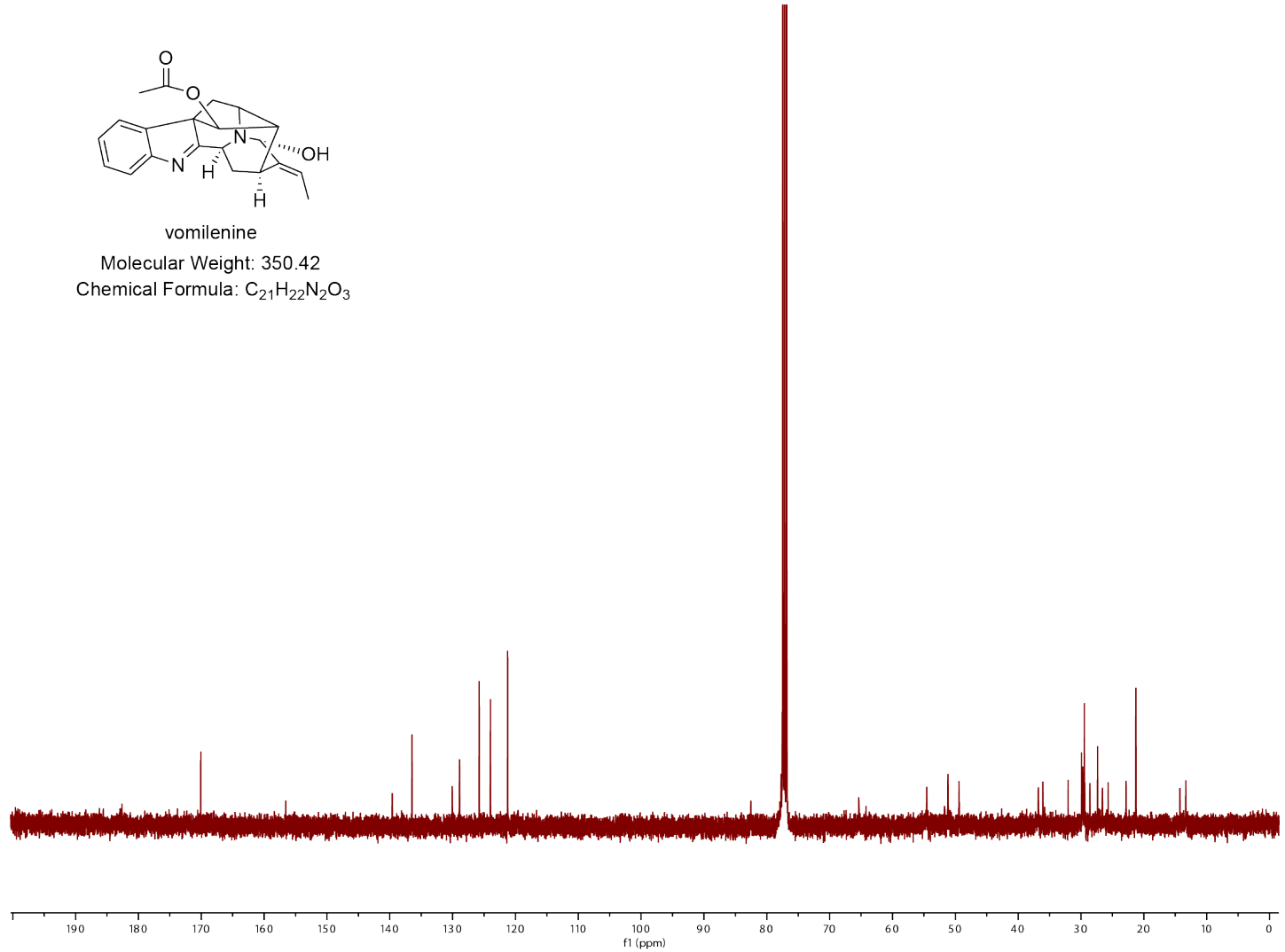


Figure 10. 13C-NMR spectrum of vomilenine in $CDCl_3$

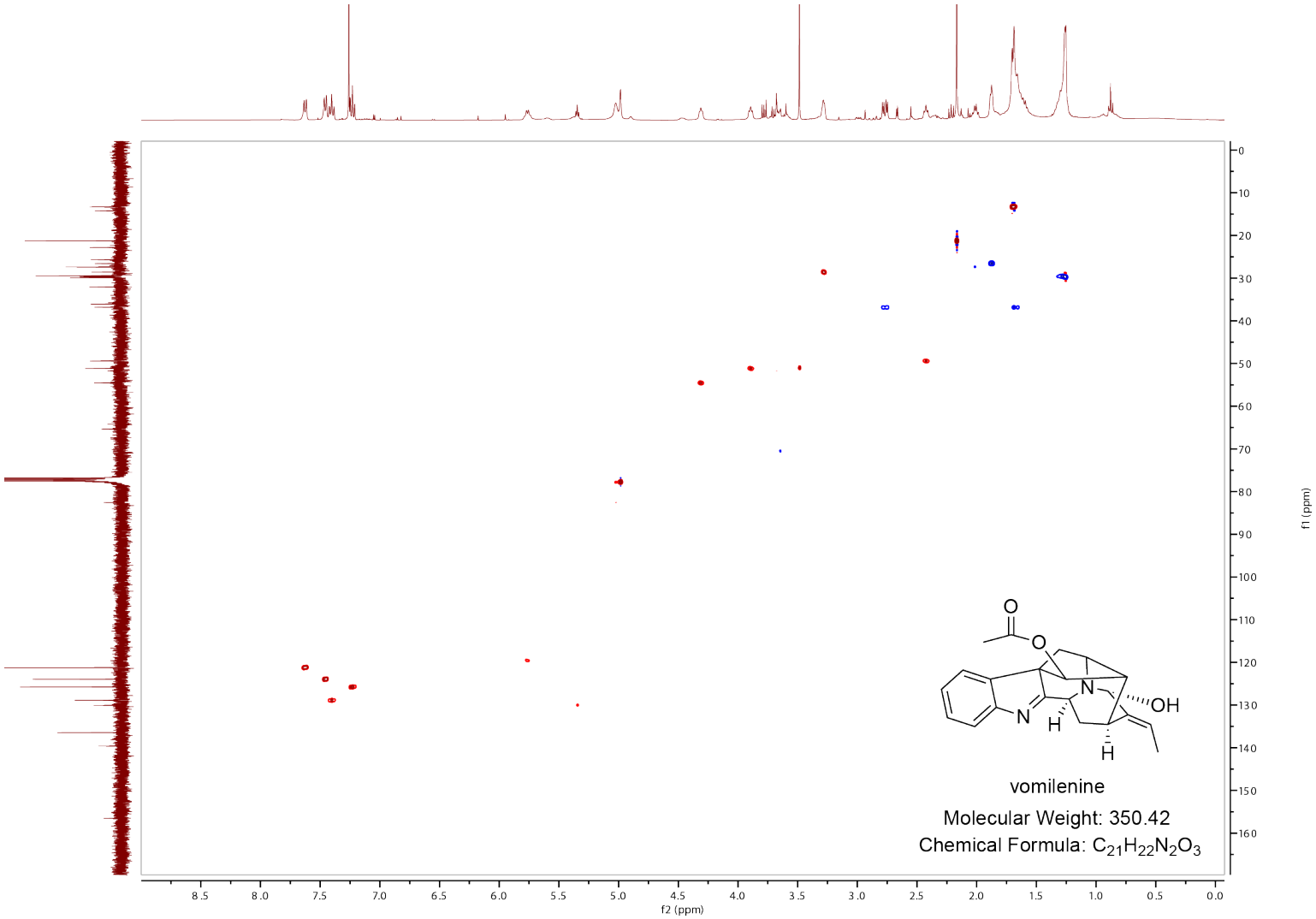


Figure 11. HSQC spectrum of vomilenine in $CDCl_3$

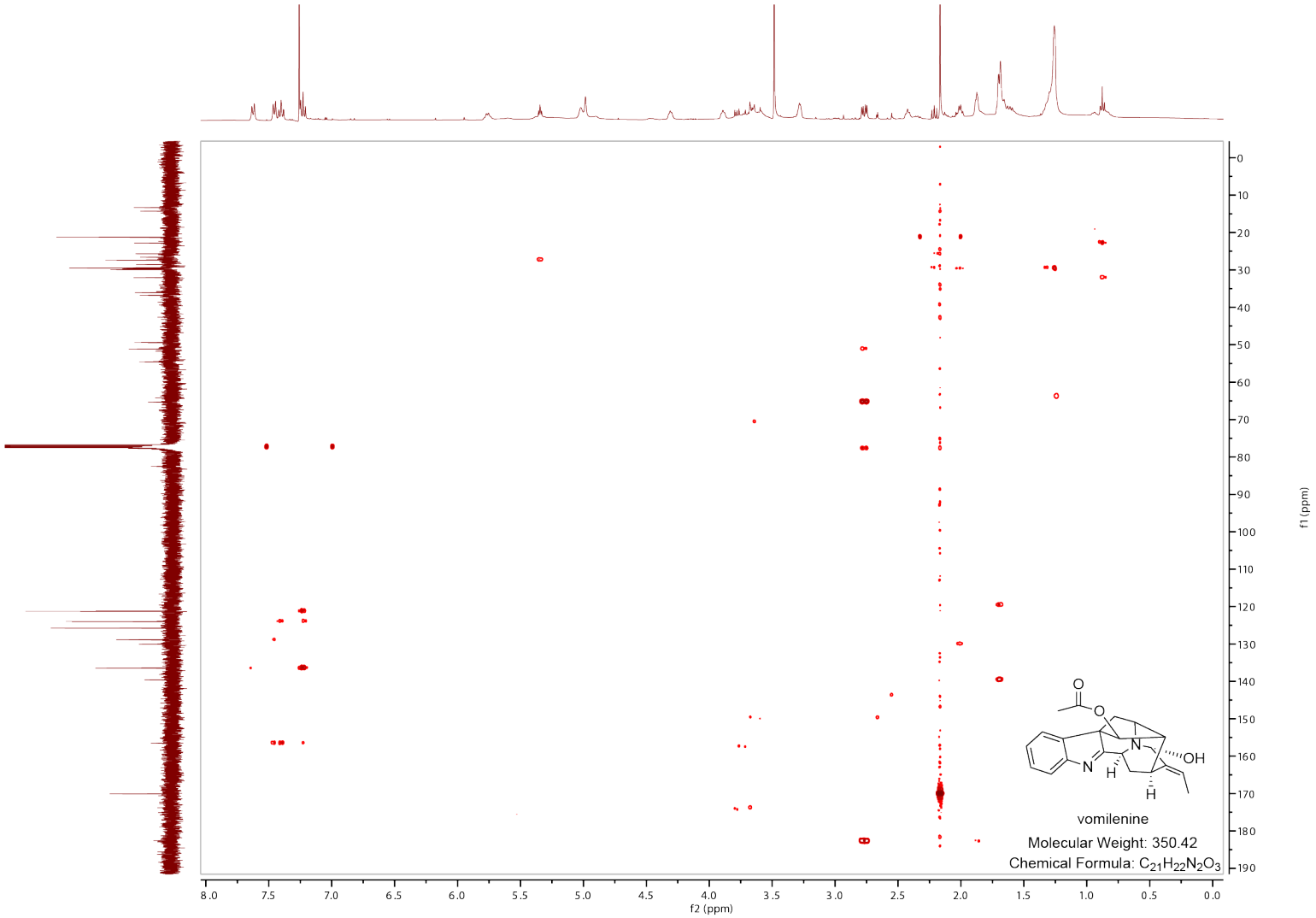


Figure 12. HMBC spectrum of vomilenine in $CDCl_3$

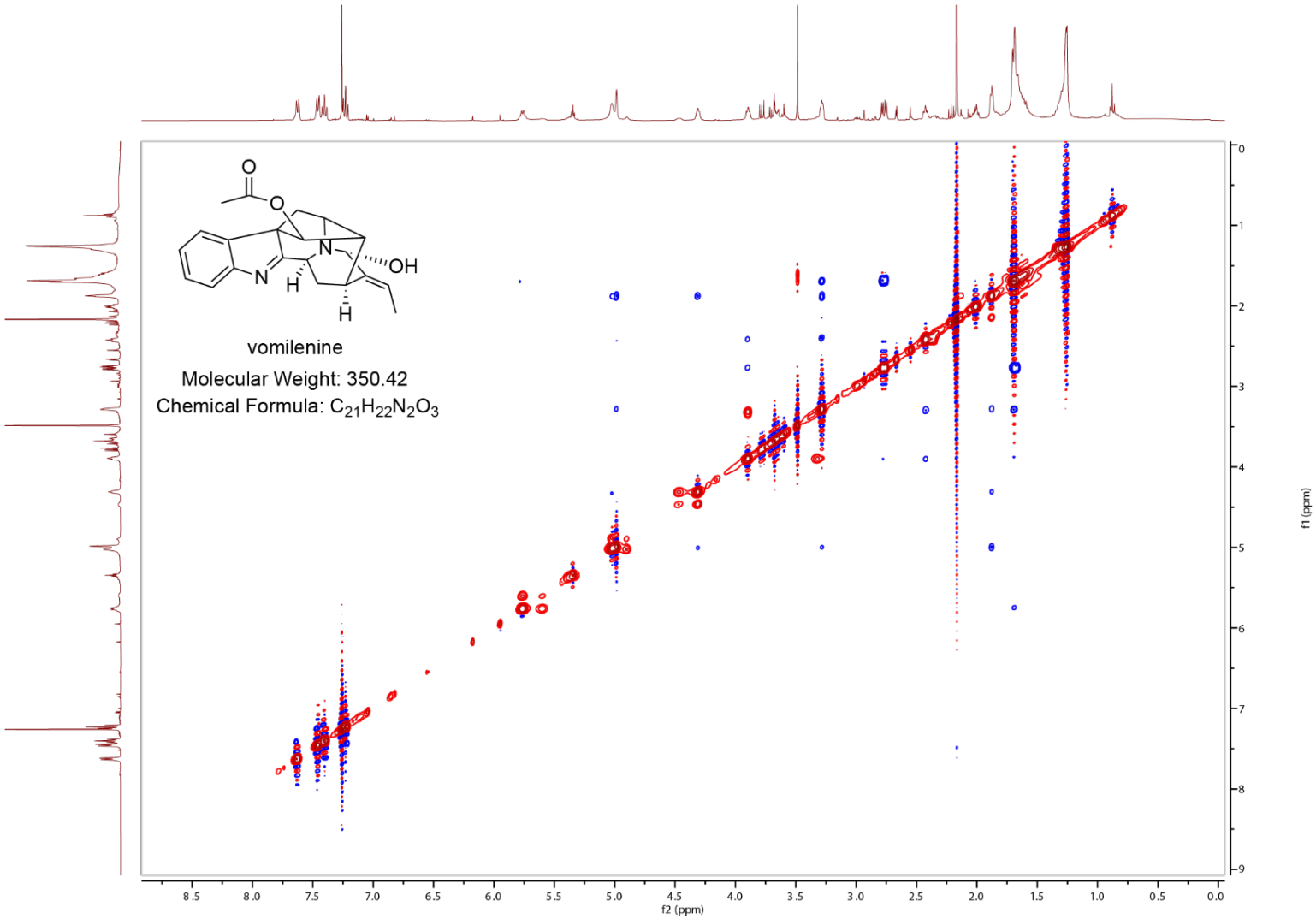
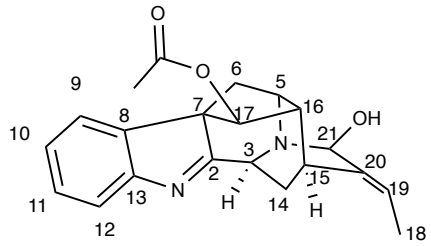


Figure 13. NOSEY spectrum of vomilenine in $CDCl_3$

Table 1. ^{1}H- and ^{13}C- NMR chemical shifts of vomilenine from this study and from literature. Both proton and carbon-13 NMR chemical shifts matched well with values reported by Dang *et al.,* 2017.



Chemical Formula: $C_{21}H_{22}N_2O_3$
Molecular Weight: 350.42

Carbon/ Proton #	^{1}H, This Study ($CDCl_3$)	^{1}H, Dang *et al.*, 2017. ($CDCl_3$)	^{13}C, This Study ($CDCl_3$)	^{13}C, Dang *et al.* 2017. ($CDCl_3$)
2	-	-	182.5	181.7
3	4.32 (dd)	4.39 (br s)	54.5	n.d.
5	3.90 (dd)	3.93 (br s)	51.2	n.d.
6	2.78 (dd), 1.67 (m)	2.79 (dd), 1.71 (m)	36.8	36.2
7	-	-	65.3	65.0
8	-	-	136.2	136.1
9	7.46 (d)	7.47 (d)	123.9	123.8
10	7.23 (t)	7.24 (ddd)	128.9	125.7
11	7.41 (t)	7.41 (ddd)	125.8	128.9
12	7.62 (d)	7.64 (d)	121.2	121.2
13	-	-	156.3	156.5
14	1.88 (m)	1.90 (m)	26.6	26.2
15	3.28 (m)	3.33 (t)	28.6	28.0
16	2.43 (dd)	2.48 (t)	49.4	48.8
17	4.98 (s)	4.99 (s)	77.6	77.4
18	1.69 (s)	1.73 (dd)	13.3	13.2
19	5.76 (q)	5.8 (br q)	119.6	120.2
20	-	-	139.4	138.6
21	5.02 (br s)	5.12 (s)	82.6	83.1
CO	-	-	169.6	169.9
CH_3CO	2.17 (s)	2.18 (s)	21.2	21.1

Vomilenine was enzymatically produced to further confirm the identity of the plant-extracted vomilenine. To do this, VH (Dang et al., 2017) was cloned and co-expressed in yeast with previously characterized SBE from *Gelsemium serpemvirens* (Dang et al., 2018), PNAE (Dogru et al., 2000), and VS (Ma et al., 2005) from *R. serpentina*. The enzymes SBE, PNAE, and VS were previously cloned by other members in the Qu group. After feeding the substrate 19*E*-geissoschizine (Qu et al., 2018) to the yeast co-expressing SBE, PNAE, and VS, a MIA product with *m/z* 335 for vinorine (Fig. 14) was produced, as expected from previous research. When VH was added to the co-expression, vomilenine was produced at the expense of vinorine (Fig. 14). The LC-MS/MS retention time and ESI-MS/MS fragmentation pattern (Fig. 7) for the enzymatically produced vomilenine is identical with the plant extracted vomilenine, which confirmed the identification of this MIA. Furthermore, the ESI-MS/MS fragmentation pattern (Appendix I) and retention time were used to identify the peak of vinorine in *R. serpentina* root (Fig. 6).

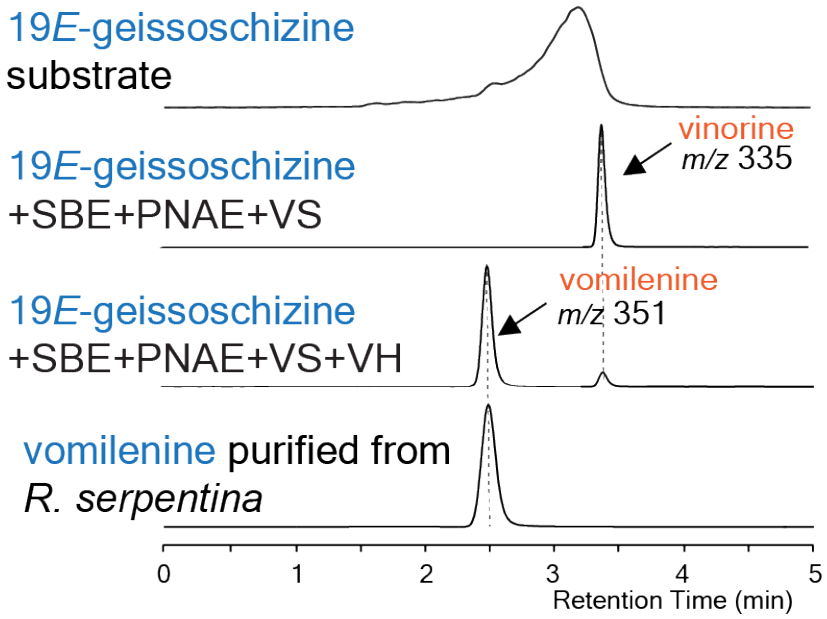


Figure 14. **LC-MS/MS chromatograms comparing enzymatically produced vomilenine (top 3) to plant extracted vomilenine (bottom). In both cases they exhibit identical LC retention times (2.5 min) and fragmentation patterns (see Figure 7).**

3.2 A NADPH-dependent enzyme reduces vomilenine into a dihydro derivative

Previously Gao (2002) and von Schumann (2002) partially purified VR and DHVR from *R. serpentina* root by chromatography. Edman degradation revealed several peptide fragments believed to be VR and DHVR. More than a decade later, Geissler *et al.* (2016) used these peptide fragments to search the now-available *R. serpentina* transcriptomes and identified several reductase candidates. While the candidate enzymes did not show VR and DHVR activity, one of their enzymes (RsRR4) showed 19,20-reduction with vomilenine substrate. The RsRR4 enzyme was thus re-named to vomilenine reductase 2 (VR2).

It is interesting to note that VR2 falls under the CAD enzyme superfamily. Our group and other researchers have previously characterized a number of CAD-like reductases involved in MIA biosynthesis such as GS (Qu et al., 2018) and heteroyohimbine synthase (HYS; Stavrinides et al., 2016), among others. It is therefore reasoned that VR and DHVR are likely members of the CAD-like reductases. With this lead information, the public *R. serpentina* leaf/root transcriptomes (the PhytoMetaSyn project, https://bioinformatics.tugraz.at/phytometasyn/) was searched for homology with GS, HYS, and other CAD-like MIA reductases. The top 8 highly expressed CAD-like reductases in *R. serpentina* root were then chosen as the candidate genes, which were labeled RsCAD1-8 (Table 2). A list of their gene sequences can be found in Appendix II along with the PCR primer sequences used to amplify them from *R. serpentina* cDNA. After cloning and expressing RsCAD1-8 in *E. coli*, vomilenine was fed to the cells and monitored for VR activity using LC-MS/MS (Fig. 15).

Table 2. Top 11 highly expressed CAD-like reductases in *R. serpentina root*. Rows highlighted in red are the candidate enzymes for this study. TPM = transcript per million.

Query	Gene length	Name	Lowest E-value	Accession (E-value)	Greatest identity %	Accession (identity %)	Root TPM	Leaf TPM
TRINITY_DN464_c0_g1_i1	1702	RsGS	0	CrGS	90.4	CrGS	2600.9	483.3
TRINITY_DN9553_c0_g1_i28	**1627**	**RsCAD1**	**3.7E-145**	**Cr10HGO**	**67.3**	**Cr10HGO**	**1471.2**	**224.1**
TRINITY_DN204_c0_g1_i53	**1721**	**RsCAD2**	**5.3E-150**	**Cr10HGO**	**58.8**	**Cr10HGO**	**730.7**	**0**
TRINITY_DN9370_c1_g1_i1	**1692**	**RsCAD3**	**1.8E-95**	**Cr10HGO**	**52.0**	**Cr10HGO**	**594.3**	**84.3**
TRINITY_DN4810_c0_g1_i8	2218	RsVR2	0	RsVR2	92.8	RsVR2	335.3	3.7
TRINITY_DN9063_c0_g1_i3	**1358**	**RsCAD4**	**0**	**CrTHAS**	**81.5**	**CrTHAS**	**470.6**	**42.0**
TRINITY_DN352_c0_g1_i1	1785	RsRR6-2	0	Cr10HGO	90.3	Cr10HGO	199.9	1.3
TRINITY_DN352_c0_g1_i9	**1868**	**RsCAD5**	**0**	**Cr10HGO**	**74.8**	**Cr10HGO**	**109.1**	**11.4**
TRINITY_DN6753_c0_g1_i4	**1667**	**RsCAD6**	**1.2E-146**	**Cr10HGO**	**71.5**	**Cr10HGO**	**116.5**	**75.6**
TRINITY_DN204_c0_g1_i56	**1637**	**RsCAD7**	**1.8E-140**	**Cr10HGO**	**59.7**	**Cr10HGO**	**66.1**	**43.1**
TRINITY_DN2461_c0_g1_i8	**1495**	**RsCAD8**	**8.4E-103**	**CrRedox1**	**53.9**	**Cr10HGO**	**57.2**	**0.6**

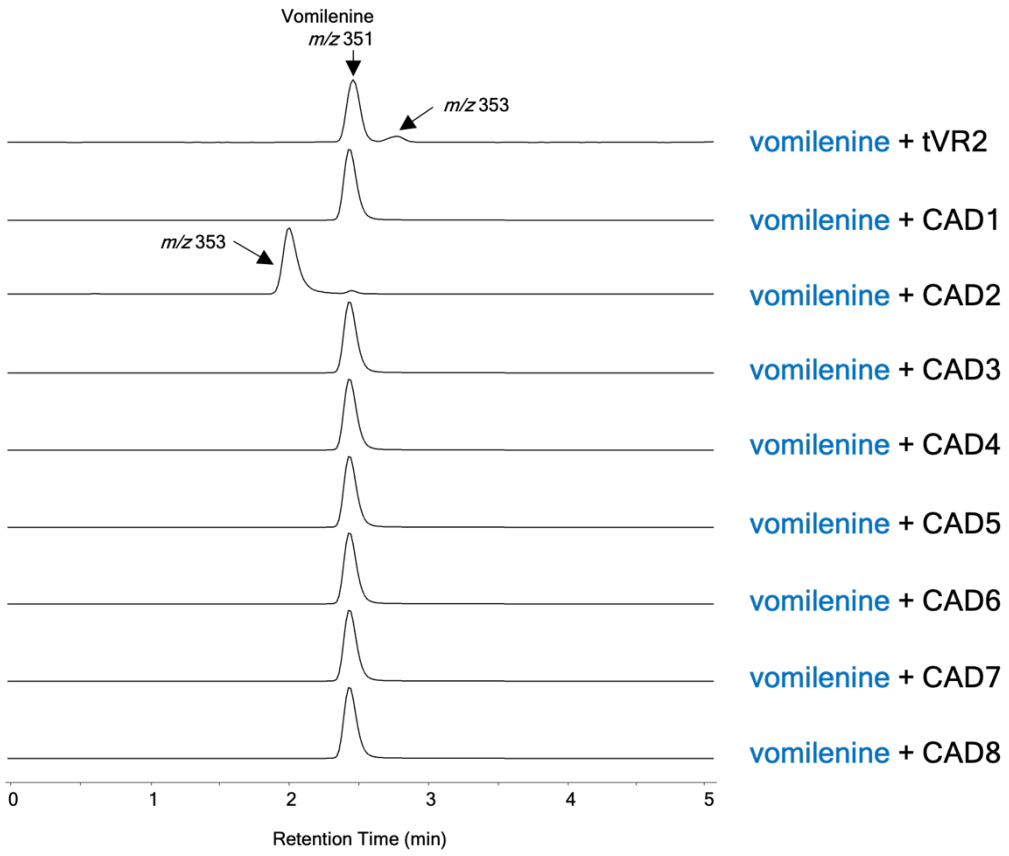


Figure 15. LC-MS/MS chromatograms showing enzyme activity with vomilenine. Truncated VR2 (tVR2; more on this in Chapter 3.3) reduced vomilenine into its 19,20-dihydro derivative as expected from previous research. Among the candidate enzymes, CAD2 clearly reduced vomilenine into an unknown dihydro product with a different retention time (R_t) than the VR2 product.

In these *E. coli* feeding experiments, VR2 reduced small amounts of vomilenine into 19,20-dihydrovomilenine (*m/z* 353) as expected from previous research (Geissler et al., 2016). Amongst the candidate enzymes, RsCAD2 clearly showed vomilenine reduction into an unknown dihydro product with a different R_t than the VR2 product. The UV absorption profiles for vomilenine and the unknown CAD2 dihydrovomilenine product (Fig. 16) were compared, which showed a shift in maximum absorption from 260 nm (vomilenine) to 238/287 nm (dihydrovomilenine), indicating that the vomilenine

indolenine ring was reduced into an indoline-product (von Schumann, 2002). Considering vomilenine only has two double bonds that can be reduced (either 1,2-indolenine or 19,20-), it is not ambitious to assume that the CAD2 product is 1,2-dihydrovomilenine, though the stereochemistry still needed to be confirmed.

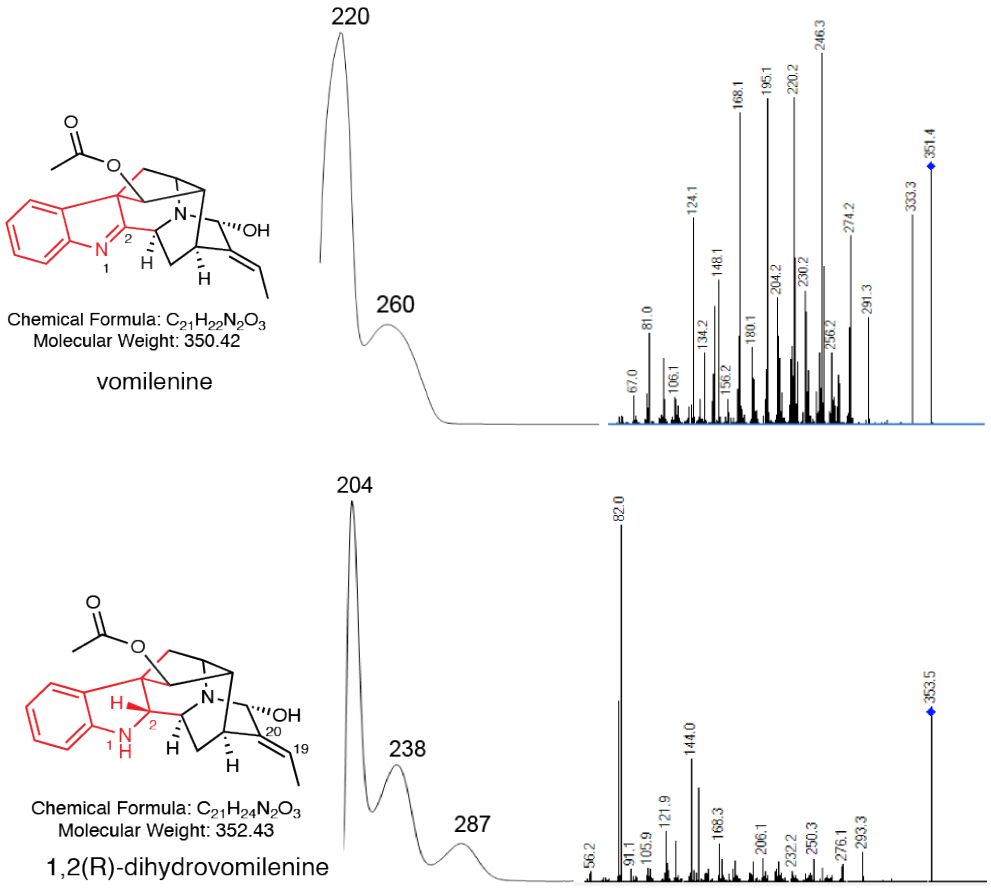


Figure 16. UV absorption profiles and ESI-MS/MS ion fragmentation patterns for vomilenine (top) and the unknown dihydrovomilenine (bottom). A shift in UV absorption from 260 nm to 238/287 nm indicated an indolenine -> indoline conversion. The stereochemistry identification for the dihydrovomilenine product will be explained in Chapter 3.4.

With this finding, RsCAD2 was purified from *E. coli* using standard His-tag affinity chromatography and the purification was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE; Fig 17). As seen from the figure, RsCAD2 (left, Fig. 17) was purified to relative purity compared to the induced total proteins. Furthermore, the RsCAD2 band is not seen in the uninduced lanes, meaning this band is indeed RsCAD2 and not a native *E. coli* protein. Comparing to the protein size ladder, RsCAD2 is approximately 50 kDa in molecular mass. The previously reported VR2 was also purified to relative purity (right, Fig. 17).

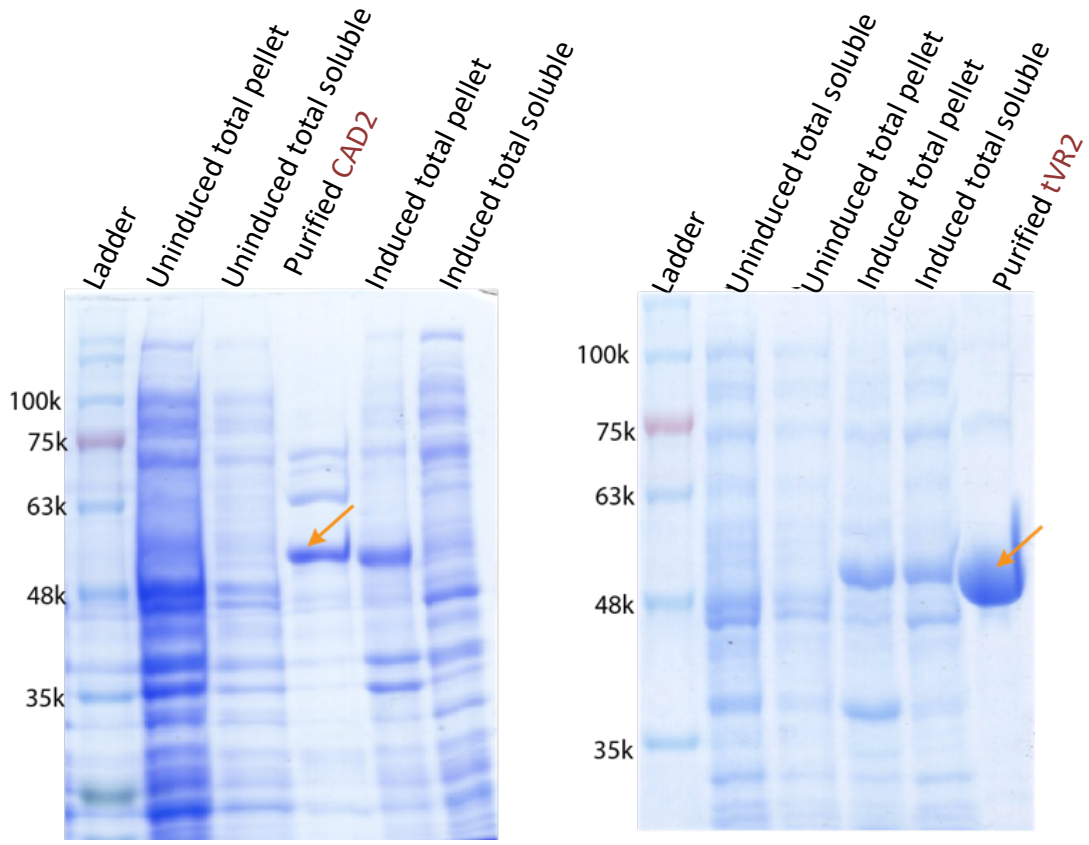


Figure 17. SDS-PAGE of purified, His-tagged recombinant enzymes for RsCAD2 and RsVR2 (tVR2 = truncated VR2; more on this in 3.3). Arrows indicate the purified enzymes. Lanes 2-3 and 5-6 of each gel show either induced or uninduced total soluble or pelleted *E. coli* proteins, as labeled.

With the purified RsCAD2, *in vitro* enzyme assays were set up with the co-factor NADPH. Figure 18 shows RsCAD2 activity on vomilenine with and without NADPH. Clearly, in the absence of NADPH, no vomilenine reduction activity was observed; adding NADPH to the RsCAD2 reaction mixture led to the production of 1,2-dihydrovomilenine. The *in vitro* study was consistent with the *in vivo* feeding experiment and confirmed that RsCAD2 is a NADPH-dependent reductase as originally reported. In addition, RsCAD2 did not accept the VR2 product 19,20-dihydrovomilenine (data not shown), indicating that RsCAD2 specifically reduces vomilenine, as originally reported (von Schumann, 2002).

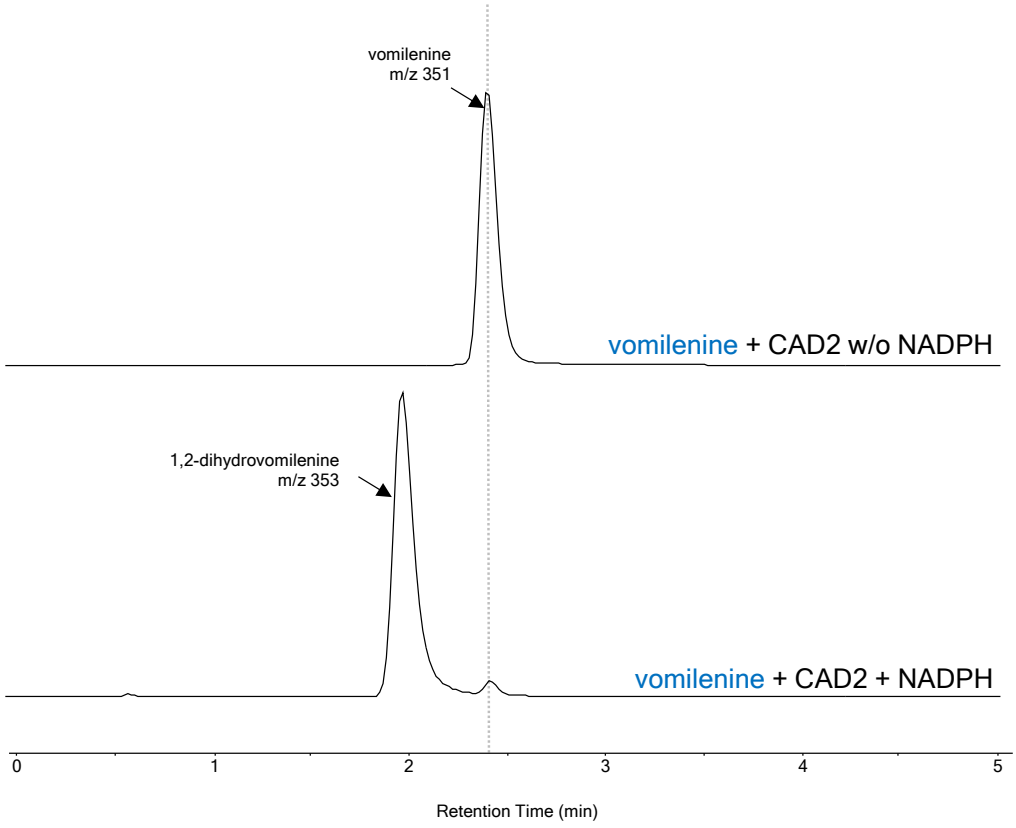


Figure 18. LC-MS/MS chromatograms showing RsCAD2 activity against vomilenine substrate in the presence (bottom) and absence (top) of NADPH.

3.3 A previously reported enzyme further reduces 1,2-dihydrovomilenine

Reduction of vomilenine 1,2-double bond can theoretically result in either 2R or 2S stereochemistry, and ajmaline and intermediates leading to ajmaline all adopt the 2R stereochemistry (Fig. 5). While it may be possible to enzymatically produce larger amounts of RsCAD2 reduction product to conduct formal NMR structure analysis and identify the C_2 stereochemistry, an alternative method was proposed. That is, by conducting coupled *E. coli* feeding experiments, it is possible to screen for the second reductase, DHVR. If DHVR were to be found, it is possible to convert vomilenine to the final product ajmaline using the characterized AAE and NNMT. If ajmaline could be produced by these enzymes, the stereochemistry would be evident.

Specifically, vomilenine was fed to *E. coli* culture mixtures of RsCAD2 and another RsCAD candidate, and the products were examined by LC-MS/MS. Remarkably, the combination of RsCAD2 and RsVR2 (Geissler et al., 2016) led to the formation of an unknown *m/z* 355 tetrahydrovomilenine with the expected mass for 17-*O*-acetylnorajmaline (Fig. 19). It is likely that a sequential 1,2- and 19,20-reduction occurred to account for the increased masses, and RsCAD2 and RsVR2 are likely the expected RsVR and RsDHVR.

It is important to note that there are gene sequence discrepancies between our identified VR2 sequence and the previously reported sequence (Geissler et al., 2016). While the exact reported transcript of VR2 could not be found in PhytoMetaSyn transcriptome, a VR2 homolog containing extra 50 amino acids in front of the N-terminus of reported VR2 was identified as one the top expressing CAD enzymes (Table 2). The remaining aligned 363 amino acids were 95% identical to VR2 (Fig. 21). Additionally,

the exact reported VR2 sequence were not found in another public *R. serpentina* transcriptome (Medicinal Plant Genomics Resources, MPGR, http://mpgr.uga.edu); only the PhytoMetaSyn version of VR2 could be found in MPGR. While the reasons for the discrepancies remain to be identified, both the full 413 amino acid PhytoMetaSyn/MPGR VR2 and the truncated 363 amino acid PhytoMetaSyn/MPGR VR2 (tVR2) were cloned for this study. There were no differences in enzyme function between the PhytoMetaSyn/MPGR VR2 and tVR2 (Fig. 22), therefore tVR2 (95% identical to reported VR2 at amino acid level) was used for the experiments reported above and in future sections.

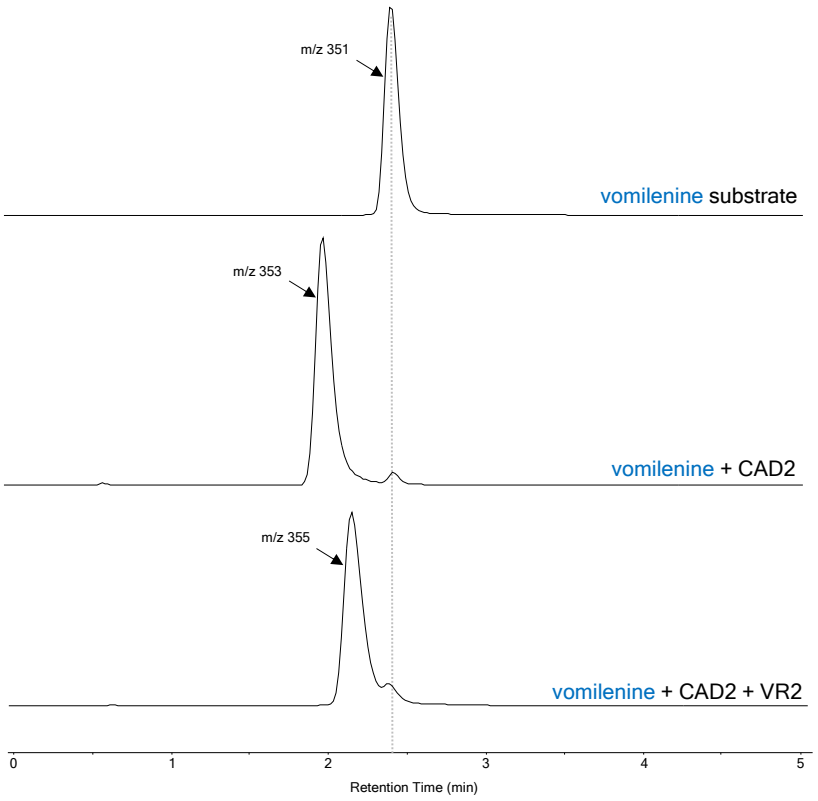


Figure 19. LC-MS/MS Chromatograms showing CAD2 + VR2 twice reducing vomilenine (top) into first 1,2-dihydrovomilenine (middle), then an unknown *m/z* 355 tetrahydrovomilenine (bottom).

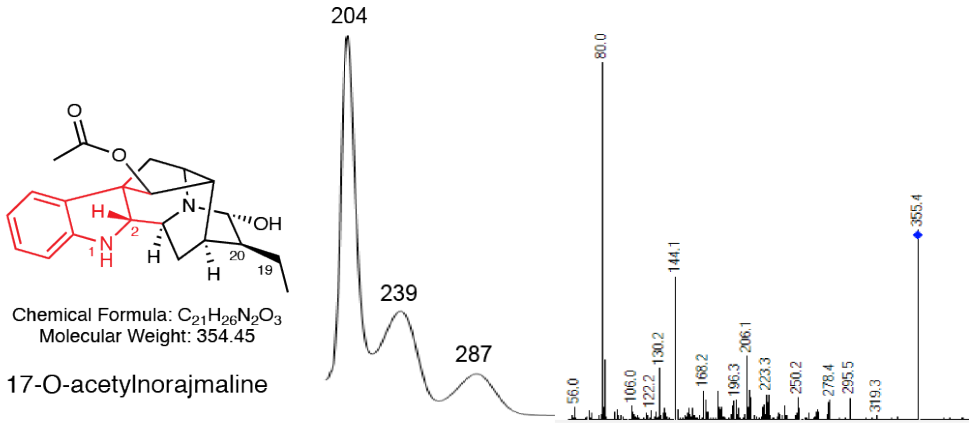


Figure 20. The UV absorption profile and ESI-MS/MS ion fragmentation pattern of the tetrahydrovomilenine produced by CAD2 + VR2 with vomilenine as substrate.

Figure 21. Protein alignments showing the amino acid differences between VR2 from this study and reported VR2. Note the extra 50 amino acids in our version of VR2. Both the full 413 amino acids PhytoMetaSyn/MPGR VR2 and the truncated 363 amino acids PhytoMetaSyn/MPGR VR2 were cloned for this study.

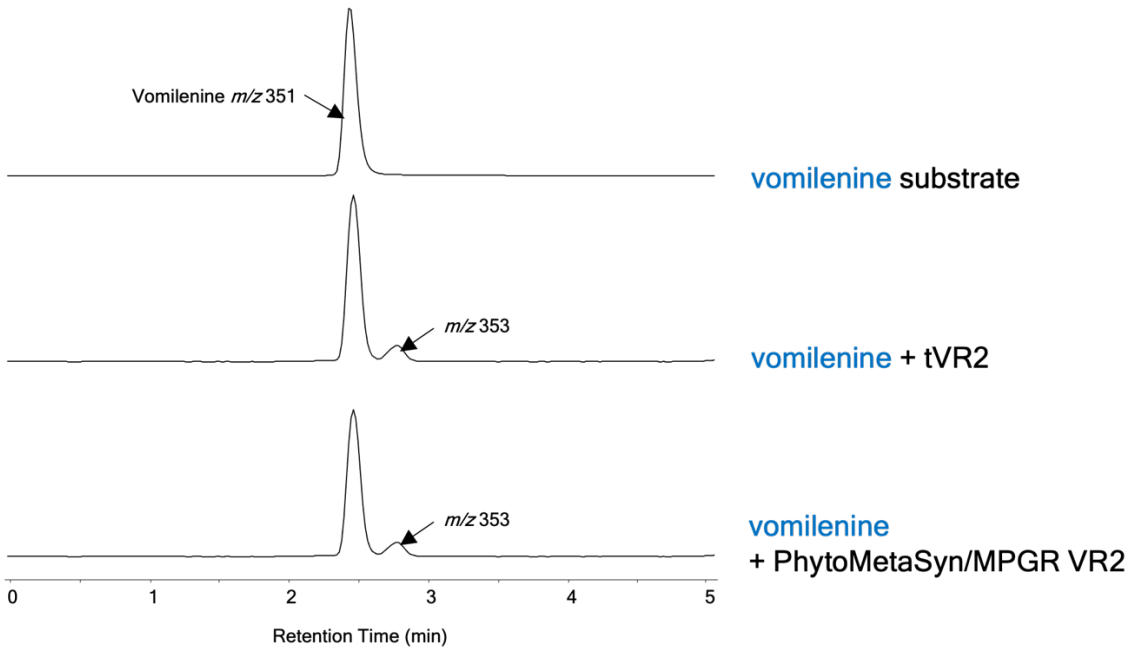


Figure 22. LC-MS/MS chromatograms showing truncated VR2 (tVR2) and full PhytoMetaSyn/MPGR VR2 serving identical enzyme functions. In both enzymes, vomilenine is reduced into the reported *m/z* 353 19,20-dihydrovomilenine product.

3.4 *In vivo* and in *vitro* production of ajmaline

After 17-*O*-acetylnorajmaline is formed, there are two more modifications by AAE (Polz et al., 1987; Ruppert et al., 2005) and NNMT (Cázares-Flores et al., 2016) to form ajmaline. If the unknown tetrahydrovomilenine product produced by CAD2 and VR2 were 17-*O*-acetylnorajmaline, AAE and NNMT should convert it to ajmaline. To test this, the unknown tetrahydrovomilenine was fed to 1) *R. serpentina* root enzymes and 2) yeasts co-expressing AAE and NNMT. Additionally, vomilenine was fed to 3) yeasts co-expressing CAD2, VR2, AAE, and NNMT. The products produced from these experiments was then compared to commercial ajmaline standards by LC retention time, UV spectrometry, and fragmentation patterns. However, it is important to note that if the product is an enantiomer of ajmaline, the LC retention time would appear identical for the enantiomer pairs. Though, it is unlikely that the product formed is the enantiomer of ajmaline, considering VR2 specifically reduces vomilenine into 19,20(S)- configuration, and 1,2-reduction shall not theoretically produce a pair of enantiomers.

3.4.1 Using *R. serpentina* root enzymes

In theory, if the CAD2 and VR2 pair reduced vomilenine into 17-*O*-acetylnorajmaline, then *R. serpentina* total root proteins should convert the tetrahydrovomilenine product into first norajmaline (catalyzed by AAE), followed by ajmaline (catalyzed by NNMT). After enzymatically producing more of the unknown tetrahydrovomilenine, this product was purified with TLC and incubated with *R. serpentina* total root proteins, at first without the methyl-supplying cofactor S-adenosyl methionine (SAM) for the production of norajmaline.

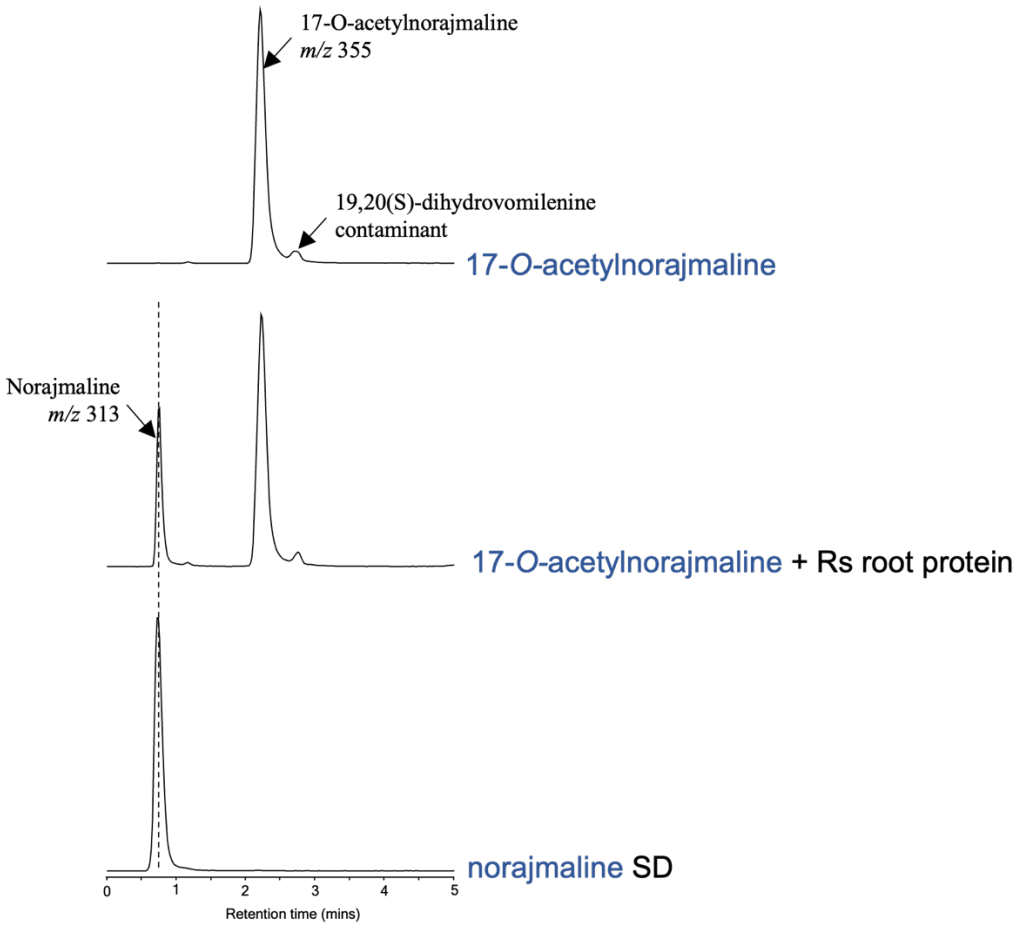


Figure 23. LC-MS/MS chromatograms showing the unknown tetrahydrovomilenine (labeled as 17-*O*-acetylnorajmaline) was converted by *R. serpentina* total root proteins into a *m/z* 313 product having the same LC retention time as the norajmaline standard (SD; generously provided by Prof. De Luca).

Figure 23 clearly showed that the unknown tetrahydrovomilenine was converted by *R. serpentina* total root enzymes into norajmaline, which was evident for with the same *m/z* = 313 and retention time (0.7 mins) as norajmaline standard. Furthermore, this *m/z* = 313 product exhibited the same UV maximum absorption and ion fragmentation patterns as the norajmaline standard (Fig. 27). These findings provided initial evidence that the unknown tetrahydrovomilenine is 2β-(R)-17-*O*-acetylnorajmaline, since

norajmaline is of the 2β-(R) configuration, and that the *m/z* 313 product produced from the unknown tetrahydrovomilenine exhibited identical LC retention time as norajmaline. Furthermore, VR2 (Geissler et al., 2016) specifically reduces vomilenine into the 19,20(S)- configuration. This suggests that the unknown tetrahydrovomilenine must also exhibit the 20α-(S) configuration, ruling out the possibility that the enzymatically produced *m/z* 313 MIA is an enantiomer [2β-(S)-20α-(R)] of norajmaline. This finding was further confirmed by pure recombinant AAE (Polz et al., 1987; Ruppert et al., 2005) in the next section.

The next step is to add the methyl-supplying cofactor SAM and produce ajmaline. While *R. serpentina* root proteins supplemented with SAM alone only produced norajmaline (data not shown), adding the previously characterized NNMT (Cázares-Flores et al., 2016) to the reaction produced a *m/z* 327 MIA with identical retention time (Fig. 24), UV maximum absorption and fragmentation patterns (Fig. 27), as ajmaline. This finding suggested that the unknown tetrahydrovomilenine was indeed 17-*O*-acetylnorajmaline.

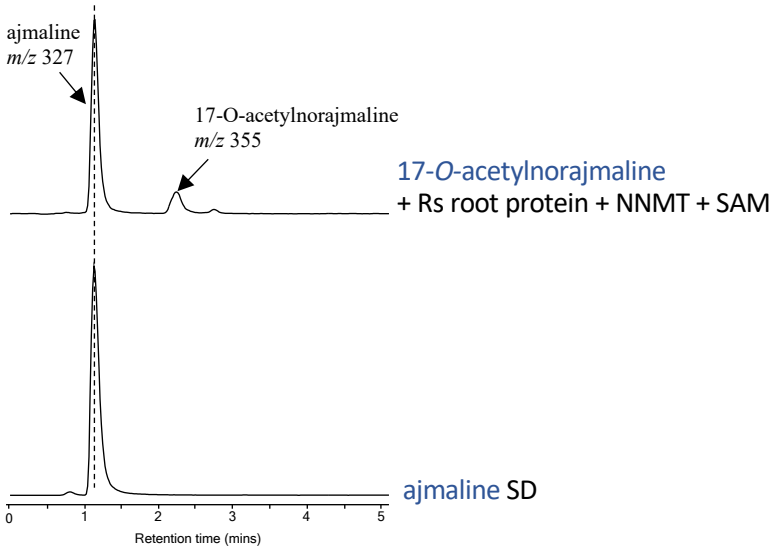


Figure 24. LC-MS/MS chromatograms showing that the unknown tetrahydrovomilenine is converted by *R. serpentina* total root proteins and purified NNMT into a m/z 327 product having the same LC retention time (1.2 mins) as the commercial ajmaline standard.

3.4.2 Using recombinant proteins

To further confirm the findings above, the experiments were repeated replacing *R. serpentina* root proteins with yeast expressing either one or a combination of CAD2, VR2, AAE (Polz et al., 1987; Ruppert et al., 2005), and NNMT (Cázares-Flores et al., 2016). The reason yeast was chosen is because AAE is a glycoprotein and requires proper folding through the eukaryotic protein glycosylation pathway, which the prokaryote *E. coli* lacks. Previously, AAE was expressed and characterized in the tobacco plant (*Nicotiana benthamiana*) via *Agrobacterium* gene delivery (Ruppert et al., 2005).

The yeast expressing AAE was fed with 17-*O*-acetylnorajmaline produced by CAD2 and VR2. As expected, 17-*O*-acetylnorajmaline was consumed by the yeasts,

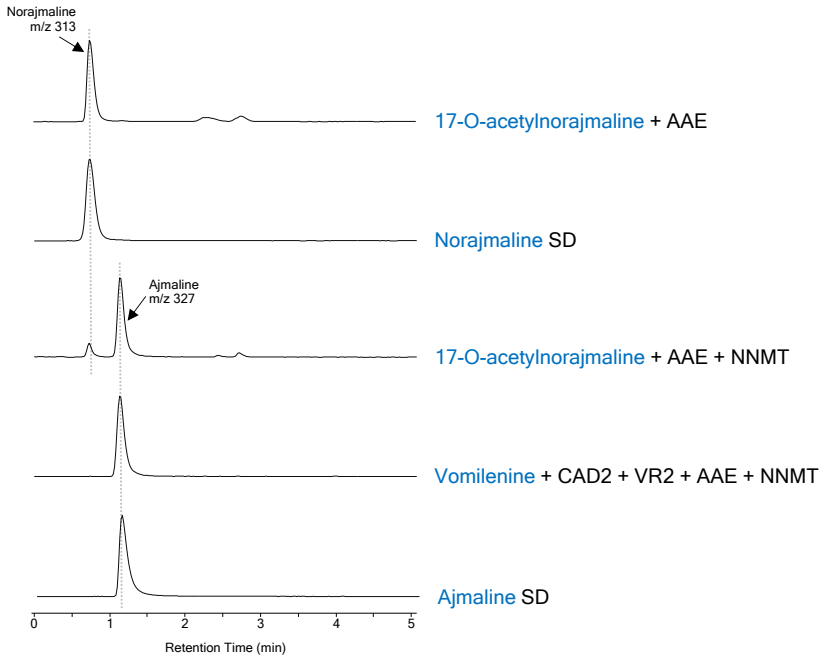


Figure 25. LC-MS/MS chromatograms showing yeasts co-expressing AAE and NNMT converting the enzymatically produced 17-*O*-acetylnorajmaline into a *m/z* 327 product having the same LC retention time as commercial ajmaline standard (SD). When vomilenine is used as a substrate and reacted with CAD2 + VR2 + AAE + NNMT, the same product results.

and a new product with *m/z* = 313 was formed. This product appears to be norajmaline based on the identical retention time (Fig. 25), UV maximum absorption, and ion fragmentation patterns (Fig. 27) as authentic norajmaline. When 17-*O*-acetylnorajmaline was fed to yeasts co-expressing AAE and NNMT, ajmaline was produced at the expense of norajmaline (Fig. 25), which was evident from the new product matching the LC retention time (Fig. 25), UV maximum absorption, and ion fragmentation patterns (Fig. 27) of commercial ajmaline standard. As a further confirmation, vomilenine was fed to yeasts co-expressing CAD2, VR2, AAE, and NNMT. The same ajmaline production was observed (Fig. 25).

Furthermore, the above experiments were repeated with purified, recombinant AAE from yeast and NNMT from *E. coli*. AAE purification from yeast (Fig. 26) was made possible because a C-terminal His-tag was incorporated into the AAE sequence (see Appendix II, AAE PCR primer sequence) during cloning. With the *in vitro* enzyme assays, the same ajmaline production was observed when enzymatically produced 17-*O*-acetylnorajmaline was fed to pure recombinant AAE and NNMT along with the co-factor SAM. The same was seen when vomilenine was fed into pure recombinant CAD2, VR2, AAE, and NNMT along with the co-factors NADPH and SAM (data not shown as it is identical with Fig. 25). Same as the products produced by *R. serpentina* root proteins, the *m/z* 313 MIA and the *m/z* 327 MIA produced by both the *in vivo* yeast feeding experiments and *in vitro* recombinant enzyme assays exhibit the same UV absorption profiles and ESI-MS/MS ion fragmentation patterns as authentic norajmaline and ajmaline standards, respectively (Fig. 27).

AAE exhibits exceptionally high substrate specificity in that it exclusively deacetylates 17-*O*-acetylnorajmaline of the 2β-(R)- configuration and not the 2α-(S)-configuration (Polz et al., 1987). This enzyme also does not deacetylate the closely related vomilenine. These observations undoubtedly confirm the stereochemistry of the tetrahydrovomilenine produced by CAD2 and VR2 to be those of 2β-(R)-17-*O*-acetylnorajmaline, since AAE would not be active if the tetrahydrovomilenine product was of another stereochemistry. Paired with the previous observations from LC retention time, UV maximum absorption, and ion fragmentation, the *m/z* 313 product is concluded to be norajmaline, and the *m/z* 327 product is concluded to be ajmaline. These findings confirm the hypothesis that CAD2 is VR catalyzing the 1,2(R)- reduction of vomilenine,

and the previously reported VR2 is in fact DHVR catalyzing the subsequent 19,20(S)-reduction of 1,2(R)-dihydrovomilenine, forming the tetrahydro product 17-*O*-acetylnorajmaline. Therefore, CAD2 was re-named as VR, and VR2 was renamed as DHVR.

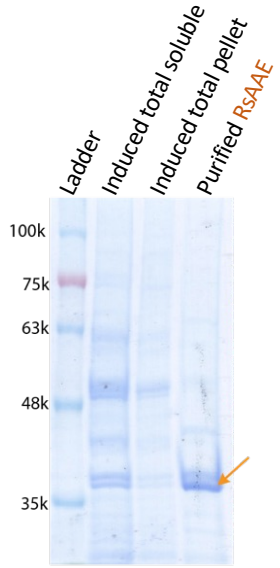


Figure 26. SDS-PAGE of purified, His-tagged recombinant enzyme for AAE. Arrow indicates the purified enzyme. Lanes 1-2 show induced total soluble or pelleted *S. cerevisiae* proteins.

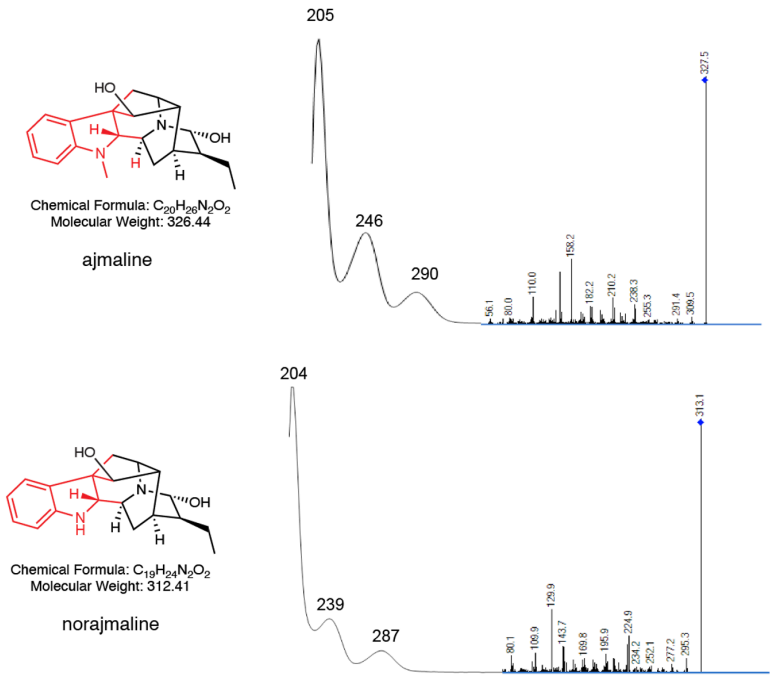


Figure 27. The UV absorption profiles and ESI-MS/MS ion fragmentation patterns of the enzymatically produced ajmaline and norajmaline. They are identical with commercial ajmaline and norajmaline standards.

It is important to note that similar to the VR2 gene sequence discrepancy, the originally reported AAE is not found in the public PhytoMetaSyn nor MPGR *R. serpentina* transcriptome. For the above experiments, the PhytoMetaSyn version of AAE was used, which is 86% identical at the amino acid level to the reported AAE (Figure 28; Ruppert et al., 2005). While there are no confirmed explanations for the discrepancies, it is important to point out that as Geissler and coworkers (2016) mentioned, *Rauwolfia* is a genus of 60-70 species, which poses a challenge in distinguishing these morphologically similar plants. It is therefore possible that the gene discrepancies stem from the fact that different *Rauwolfia* species were misidentified as *R. serpentina*, either in this study or in prior studies. Regardless of the reason for the discrepancy, the results show that the

PhytoMetaSyn/MPGR version of the ajmaline pathway enzymes remain functional despite their sequence differences.

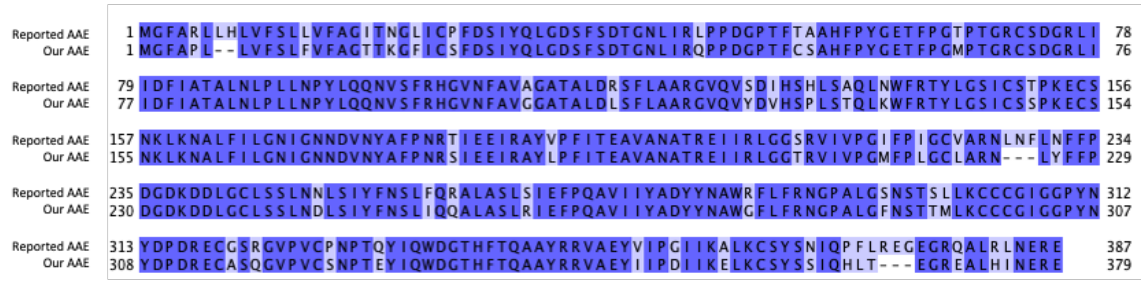


Figure 28. Protein alignments showing the amino acid differences between AAE from this study and reported AAE.

3.5 Enzyme kinetic studies reveal vomilenine reduction order in ajmaline pathway

Having completed the ajmaline pathway, there is one more question left to solve. That is, Chapter 1.3 discussed about the alternative reduction order in which the two double bonds of vomilenine may be reduced. It is therefore necessary to find out whether the 1,2- double bond of vomilenine is reduced first, or that the 19,20- double bond is reduced first. From *in vitro* enzyme assays it is evident that CAD2 (referred to as VR from now on) would not accept the VR2 (referred to as DHVR from now on) 19,20-dihydrovomilenine product. This provides initial evidence that the VR-mediated vomilenine 1,2- double bond reduction proceeds first and is followed by DHVR-mediated reduction as originally proposed (Gao et al., 2002; von Schumann, 2002). However, for an additional verification, enzyme kinetic studies were conducted to test this hypothesis (Table 3; Figure 29). Since VR and DHVR both accept vomilenine as a substrate, their relative affinities for this substrate would reveal the reduction order; if VR has higher affinity for vomilenine relative to DHVR, then this will provide evidence that the VR-mediated vomilenine 1,2- reduction happens first, since VR has higher affinity for vomilenine and would reduce this substrate before DHVR does.

Table 3. Michaelis-Menton enzyme kinetics for RsVH and RsDHVR.

	RsVH (RsCAD2)			RsDHVR (RsVR2, N-terminal truncated)		
	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)
vomilenine	41.6	1.70	4.09×10^4	1089	0.20	183
1,2-dihydrovomilenine	-	-	-	32.0	10.4	3.26×10^5
19,20-dihydrovomilenninne	-	-	-	-	-	-

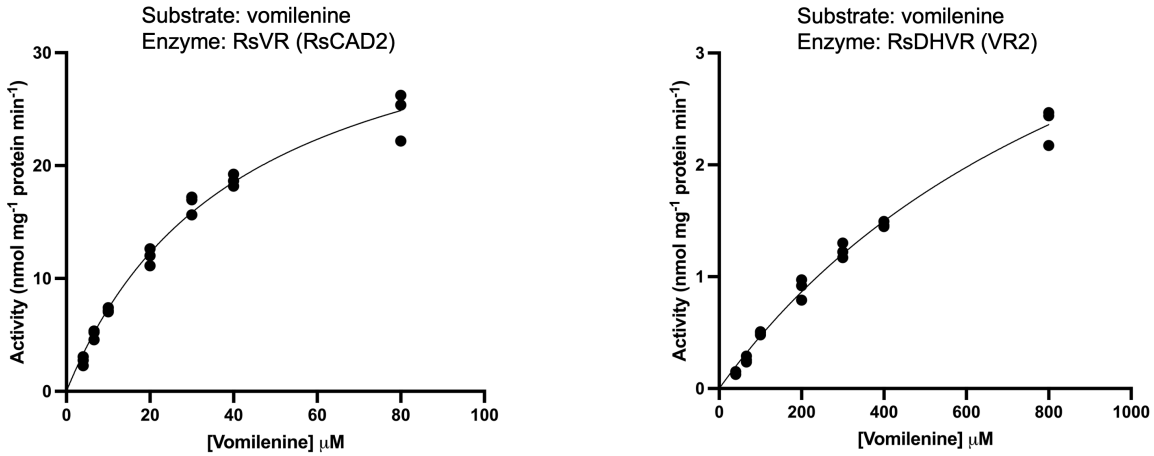


Figure 29. Michaelis-Menton kinetics of RsVR (left) and RsDHVR (right) with vomilenine. The data was generated from triplicates (•) at each substrate concentrations. Note that the x-axis scale (vomilenine concentration) for DHVR is 10x larger.

The enzyme kinetics data show that VR has a 26-fold higher affinity (K_M = 41.6 μM) for vomilenine than DHVR (K_M = 1089 μM). A difference of this magnitude provides unequivocal evidence that VR, rather than DHVR, is the correct enzyme to accept vomilenine. DHVR having a K_M in the mM range suggests that its vomilenine 19,20- reduction activity could be "accidental" (non-substrate-specific activity). In other words, DHVR is not adapted to accept vomilenine as substrate. This is more evident when examining the substrate affinity of DHVR for 1,2-dihydrovomilenine (Fig. 30), which showed a more moderate K_M value at 32 μM (compared to vomilenine K_M = 1089 μM), indicating that DHVR prefers 1,2-dihydrovomilenine than vomilenine as a substrate. This suggests that VR, rather than DHVR, reduces vomilenine at the 1,2-double bond to form 1,2-dihydrovomilenine, followed by DHVR subsequently reducing the 19,20- double bond of 1,2-dihydrovomilenine. The above findings agree with the original proposal by Gao (2002) and von Schumann (2002) that in ajmaline biosynthesis, VR reduction of vomilenine proceeds first before DHVR.

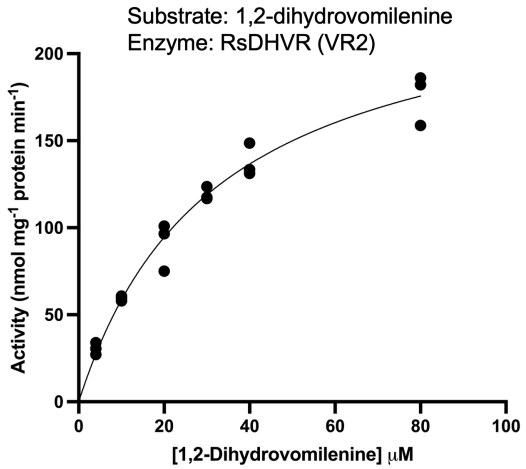


Figure 30. Michaelis-Menton kinetics of RsDHVR with 1,2-dihydrovomilenine substrate. Compared to the K_M for vomilenine (Figure 28, right), this shows that 1,2-dihydrovomilenine, rather than vomilenine, is the rightful substrate for DHVR.

Chapter 4. Conclusion

Ajmaline is a MIA-type class Ia anti-arrhythmic agent naturally found in the plant *Rauwolfia serpentina*. The commercial production of ajmaline relies on cultivating the plant *R. serpentina*. However, large-scale ajmaline production is greatly hindered by low seed germination and survival rates (Goel et al., 2009). This therefore motivates researchers, such as our group, to elucidate the ajmaline biosynthesis pathway and potentially achieve ajmaline production by bacteria/yeast fermentation. Over the past four-decades, all but two key enzymes have been cloned and characterized. In 2002, Gao and von Schumann fractioned *R. serpentina* proteins by chromatography and discovered two putative enzymes, VR and DHVR, which respectively and sequentially reduces the 1,2- double bond of vomilenine and the 19,20- double bond of 1,2-dihydrovomilenine. While a full amino acid sequencing was not possible, Gao and von Schumann used Edman degradation methods to reveal several partial peptide sequences believed to be part of VR and DHVR. More than a decade later, Geissler and coworkers (2016) used these partial peptide sequences to search the now-available *R. serpentina* transcriptome and identified several candidate enzymes. Although their candidate enzymes did not turn out to be VR and DHVR, one enzyme (VR2) surprisingly reduced the 19,20- double bond of vomilenine. Geissler *et al.* (2016) therefore suggested that vomilenine could go through an alternative reduction order through which the 19,20- double bond is reduced first followed by the 1,2- double bond reduction. Whichever way vomilenine may be reduced, the gene sequences for VR and DHVR remain the only two missing to complete the ajmaline biosynthesis pathway.

In this project, the gene sequences for both VR and DHVR was elucidated. To achieve this, the public PhytoMetaSyn/MPGR *R. serpentina* transcriptome database was searched for homologues of CAD-like MIA reductases including VR2 (Geissler et al., 2016), GS (Qu et al., 2018), HYS (Stavrinides et al., 2016), among others, which generated a list of candidate genes. Upon cloning and expressing the candidate genes in *E. coli*, VR and DHVR activity was detected by feeding the cells with vomilenine purified from *R. serpentina* plant material. LC-MS/MS and UV spectrometry results show that one of the candidate enzymes, RsCAD2, reduces the 1,2-imine bond of vomilenine in the presence of NADPH, forming 1,2-dihydrovomilenine. It is also observed that CAD2 does not accept 19,20-vomilenine as substrate. Additionally, co-expression of CAD2 and VR2 converts vomilenine into 17-*O*-acetylnorajmaline product, which, upon further reaction with AAE (Polz et al., 1987; Ruppert et al., 2005) and NNMT (Cázares-Flores et al., 2016), forms norajmaline and ajmaline. CAD2 was thus re-named to VR, and VR2 was re-named to DHVR. Enzyme kinetic studies reveal that VR has a 26-fold higher affinity for vomilenine than DHVR, which suggests that the VR-mediated vomilenine 1,2- reduction proceeds first, followed by the DHVR-mediated 19,20- reduction of 1,2-dihydrovomilenine, as originally proposed by Gao (2002) and von Schumann (2002). The findings suggest that the VR and DHVR discovered in this study complete the classic ajmaline biosynthesis pathway.

Chapter 5. Future Work

With the discovery of the two remaining reductases, it now becomes possible to reconstitute the complete ajmaline biosynthesis pathway in yeast and achieve *de novo* synthesis. This will allow the biosynthesis of ajmaline in yeast from glucose, without any dependency on the plant *Rauwolfia serpentina*. To achieve this, the *de novo* synthesis will involve the cloning of all ajmaline biosynthetic pathway enzymes into yeast. However, it is challenging to accomplish this with the cloning methods described in this study, because the yeast expression vectors used limit the maximum number of genes that can be cloned into yeast (usually 2 genes per vector, 4 vectors per yeast cell-line), and it is possible that transformed yeast cells will lose the vectors over a few generations. Therefore, a more robust and permanent metabolic engineering method such as CRISPR/Cas9 should be exploited. Once we achieve *de novo* synthesis, we can focus on optimizing ajmaline yield, time-efficiency, scaling up, *et cetera*. Additionally, another direction we can take is to discover novel enzymes that can modify ajmaline and increase its bioavailability. For example, prajmaline is a semi-synthetic anti-arrhythmic derivative of ajmaline with a higher bioavailability, because of the additional propyl group at the Nβ location. Prajmaline having a higher bioavailability would allow this drug to be administered orally, rather than solely intravenously as in the case of ajmaline (Rolf et al., 2003; Sowton et al., 1984). Furthermore, ajmaline Nβ-methyltransferase (ANMT) installs a methyl group at the Nβ location of ajmaline (Cázares-Flores et al., 2016). It is therefore not unreasonable to propose that we can discover new enzymes or modify the existing ANMT to install a Nβ-propyl group for ajmaline, thereby achieving biosynthesis of prajmaline. Overall, there are multiple directions we can take after the completion of

the ajmaline biosynthetic pathway, but the *de novo* synthesis of ajmaline remains a priority target.

Bibliography

Cázares-Flores, P., Levac, D., & De Luca, V. (2016). Rauvolfia serpentina N-methyltransferases involved in ajmaline and Nβ -methylajmaline biosynthesis belong to a gene family derived from γ-tocopherol C-methyltransferase. The Plant Journal: For Cell and Molecular Biology, 87(4), 335–342. https://doi.org/10.1111/tpj.13186

Chen, W., Ma, Y., He, W., Wu, Y., Huang, Y., Zhang, Y., Tian, H., Wei, K., Yang, X., & Zhang, H. (2022). Structure units oriented approach towards collective synthesis of sarpagine-ajmaline-koumine type alkaloids. Nature Communications, 13(1), Article 1. https://doi.org/10.1038/s41467-022-28535-x

Dang, T.-T. T., Franke, J., Carqueijeiro, I. S. T., Langley, C., Courdavault, V., & O'Connor, S. E. (2018). Sarpagan bridge enzyme has substrate-controlled cyclization and aromatization modes. Nature Chemical Biology, 14(8), Article 8. https://doi.org/10.1038/s41589-018-0078-4

Dang, T.-T. T., Franke, J., Tatsis, E., & O'Connor, S. E. (2017). Dual Catalytic Activity of a Cytochrome P450 Controls Bifurcation at a Metabolic Branch Point of Alkaloid Biosynthesis in Rauwolfia serpentina. Angewandte Chemie, 129(32), 9568–9572. https://doi.org/10.1002/ange.201705010

Dewick, P. M. (2002). Medicinal natural products: A biosynthetic approach (2nd ed). Wiley.

Dogru, E., Warzecha, H., Seibel, F., Haebel, S., Lottspeich, F., & Stöckigt, J. (2000). The gene encoding polyneuridine aldehyde esterase of monoterpenoid indole alkaloid biosynthesis in plants is an ortholog of theα/β hydrolase super family. European

Journal of Biochemistry, 267(5), 1397–1406. https://doi.org/10.1046/j.1432-1327.2000.01136.x

Edwankar, C. R., Edwankar, R. V., Rallapalli, S., & Cook, J. M. (2008). General Approach to the Total Synthesis of Macroline-Related Sarpagine and Ajmaline Alkaloids1. Natural Product Communications, 3(11), 1934578X0800301114. https://doi.org/10.1177/1934578X0800301114

Eng, J. G. M., Shahsavarani, M., Smith, D. P., Hájíček, J., De Luca, V., & Qu, Y. (2022). A Catharanthus roseus Fe(II)/α-ketoglutarate-dependent dioxygenase catalyzes a redox-neutral reaction responsible for vindolinine biosynthesis. Nature Communications, 13(1), 3335. https://doi.org/10.1038/s41467-022-31100-1

Ernst, E., & Pittler, M. H. (1998). Yomhimbine for Erectile Dysfunction: A Systematic Review and Meta-Analysis of Randomized Clinical Trials. The Journal of Urology, 159(2), 433–436. https://doi.org/10.1016/S0022-5347(01)63942-9

Falkenhagen, H., & Stöckigt, J. (1995). Enzymatic Biosynthesis of Vomilenine, a Key Intermediate of the Ajmaline Pathway, Catalyzed by a Novel Cytochrome P 450-Dependent Enzyme from Plant Cell Cultures of Rauwolfia serpentina. Zeitschrift Für Naturforschung C, 50(1–2), 45–53. https://doi.org/10.1515/znc-1995-1-208

Ferreira Batista, C. V., Schripsema, J., Verpoorte, R., Beatriz Rech, S., & Henriques, A. T. (1996). Indole alkaloids from Rauwolfia sellowii. Phytochemistry, 41(3), 969–973. https://doi.org/10.1016/0031-9422(95)00666-4

Gao, S., Schumann, G., & Stöckigt, J. (2002). A Newly-Detected Reductase from Rauvolfia Closes a Gap in the Biosynthesis of the Antiarrhythmic Alkaloid Ajmaline. Planta Medica, 68, 906–911. https://doi.org/10.1055/s-2002-34935

Geissler, M., Burghard, M., Volk, J., Staniek, A., & Warzecha, H. (2016). A novel cinnamyl alcohol dehydrogenase (CAD)-like reductase contributes to the structural diversity of monoterpenoid indole alkaloids in Rauvolfia. Planta, 243(3), 813–824. https://doi.org/10.1007/s00425-015-2446-6

Gerasimenko, I., Ma, X., Sheludko, Y., Mentele, R., Lottspeich, F., & Stöckigt, J. (2004). Purification and partial amino acid sequences of the enzyme vinorine synthase involved in a crucial step of ajmaline biosynthesis. Bioorganic & Medicinal Chemistry, 12(10), 2781–2786. https://doi.org/10.1016/j.bmc.2004.02.028

Gerasimenko, I., Sheludko, Y., Ma, X., & Stöckigt, J. (2002). Heterologous expression of a Rauvolfia cDNA encoding strictosidine glucosidase, a biosynthetic key to over 2000 monoterpenoid indole alkaloids. European Journal of Biochemistry, 269(8), 2204–2213. https://doi.org/10.1046/j.1432-1033.2002.02878.x

Goel, M. K., Mehrotra, S., Kukreja, A. K., Shanker, K., & Khanuja, S. P. S. (2009). In Vitro Propagation of Rauwolfia serpentina Using Liquid Medium, Assessment of Genetic Fidelity of Micropropagated Plants, and Simultaneous Quantitation of Reserpine, Ajmaline, and Ajmalicine. In S. M. Jain & P. K. Saxena (Eds.), Protocols for In Vitro Cultures and Secondary Metabolite Analysis of Aromatic and Medicinal Plants (Vol. 547, pp. 17–33). Humana Press. https://doi.org/10.1007/978-1-60327-287-2_2

Karunakaran, T., Ngew, K. Z., Zailan, A. A. D., Mian Jong, V. Y., & Abu Bakar, M. H. (2022). The Chemical and Pharmacological Properties of Mitragynine and Its Diastereomers: An Insight Review. Frontiers in Pharmacology, 13. https://www.frontiersin.org/articles/10.3389/fphar.2022.805986

Lewis, S. E. (2006). Recent advances in the chemistry of macroline, sarpagine and ajmaline-related indole alkaloids. Tetrahedron, 62(37), 8655–8681. https://doi.org/10.1016/j.tet.2006.06.017

Libot, F., Kunesch, N., & Poisson, J. (1980). Structure complete de la raucaffrinoline et filiation avec la vomilenine. Phytochemistry, 19(5), 989–991. https://doi.org/10.1016/0031-9422(80)85163-6

Ma, X., Koepke, J., Panjikar, S., Fritzsch, G., & Stöckigt, J. (2005). Crystal structure of vinorine synthase, the first representative of the BAHD superfamily. The Journal of Biological Chemistry, 280(14), 13576–13583. https://doi.org/10.1074/jbc.M414508200

Ma, X., Panjikar, S., Koepke, J., Loris, E., & Stöckigt, J. (2006). The Structure of Rauvolfia serpentina Strictosidine Synthase Is a Novel Six-Bladed β-Propeller Fold in Plant Proteins. The Plant Cell, 18(4), 907–920. https://doi.org/10.1105/tpc.105.038018

Monachino, J. (1954). Rauvolfia serpentina: Its History, Botany and Medical Use. Economic Botany, 8(4), 349–365.

O'Connor, S. E., & Maresh, J. J. (2006). Chemistry and biology of monoterpene indole alkaloid biosynthesis. Natural Product Reports, 23(4), 532. https://doi.org/10.1039/b512615k

Padrini, R., Piovan, D., Javarnaro, A., Cucchini, F., & Ferrari, M. (1993). Pharmacokinetics and Electrophysiological Effects of Intravenous Ajmaline: Clinical Pharmacokinetics, 25(5), 408–414. https://doi.org/10.2165/00003088-199325050-00006

Petrovska, B. B. (2012). Historical review of medicinal plants' usage. Pharmacognosy Reviews, 6(11), 1–5. https://doi.org/10.4103/0973-7847.95849

Polz, L., Schübel, H., & Stoekigt, J. (1987). Characterization of 2β(R)-17-O-AcetylajmaIan: Acetylesterase — a Specific Enzyme Involved in the Biosynthesis of the Rauwolfia Alkaloid Ajmaline. Zeitschrift Für Naturforschung C, 42(4), 333–342. https://doi.org/10.1515/znc-1987-0403

Qu, Y., Thamm, A. M. K., Czerwinski, M., Masada, S., Kim, K. H., Jones, G., Liang, P., & De Luca, V. (2018). Geissoschizine synthase controls flux in the formation of monoterpenoid indole alkaloids in a Catharanthus roseus mutant. Planta, 247(3), 625–634. https://doi.org/10.1007/s00425-017-2812-7

Rolf, S., Bruns, H.-J., Wichter, T., Kirchhof, P., Ribbing, M., Wasmer, K., Paul, M., Breithardt, G., Haverkamp, W., & Eckardt, L. (2003). The ajmaline challenge in Brugada syndrome: Diagnostic impact, safety, and recommended protocol. European Heart Journal, 24(12), 1104–1112. https://doi.org/10.1016/S0195-668X(03)00195-7

Ruppert, M., Woll, J., Giritch, A., Genady, E., Ma, X., & Stöckigt, J. (2005). Functional expression of an ajmaline pathway-specific esterase from Rauvolfia in a novel plant-virus expression system. Planta, 222(5), 888–898. https://doi.org/10.1007/s00425-005-0031-0

Shamon, S. D., & Perez, M. I. (2016). Blood pressure-lowering efficacy of reserpine for primary hypertension. Cochrane Database of Systematic Reviews, 2016(12). https://doi.org/10.1002/14651858.CD007655.pub3

Sowton, E., Sullivan, I. D., & Crick, J. C. P. (1984). Acute haemodynamic effects of ajmaline and prajmaline in patients with coronary heart disease. European Journal of Clinical Pharmacology, 26(2), 147–150. https://doi.org/10.1007/BF00630278

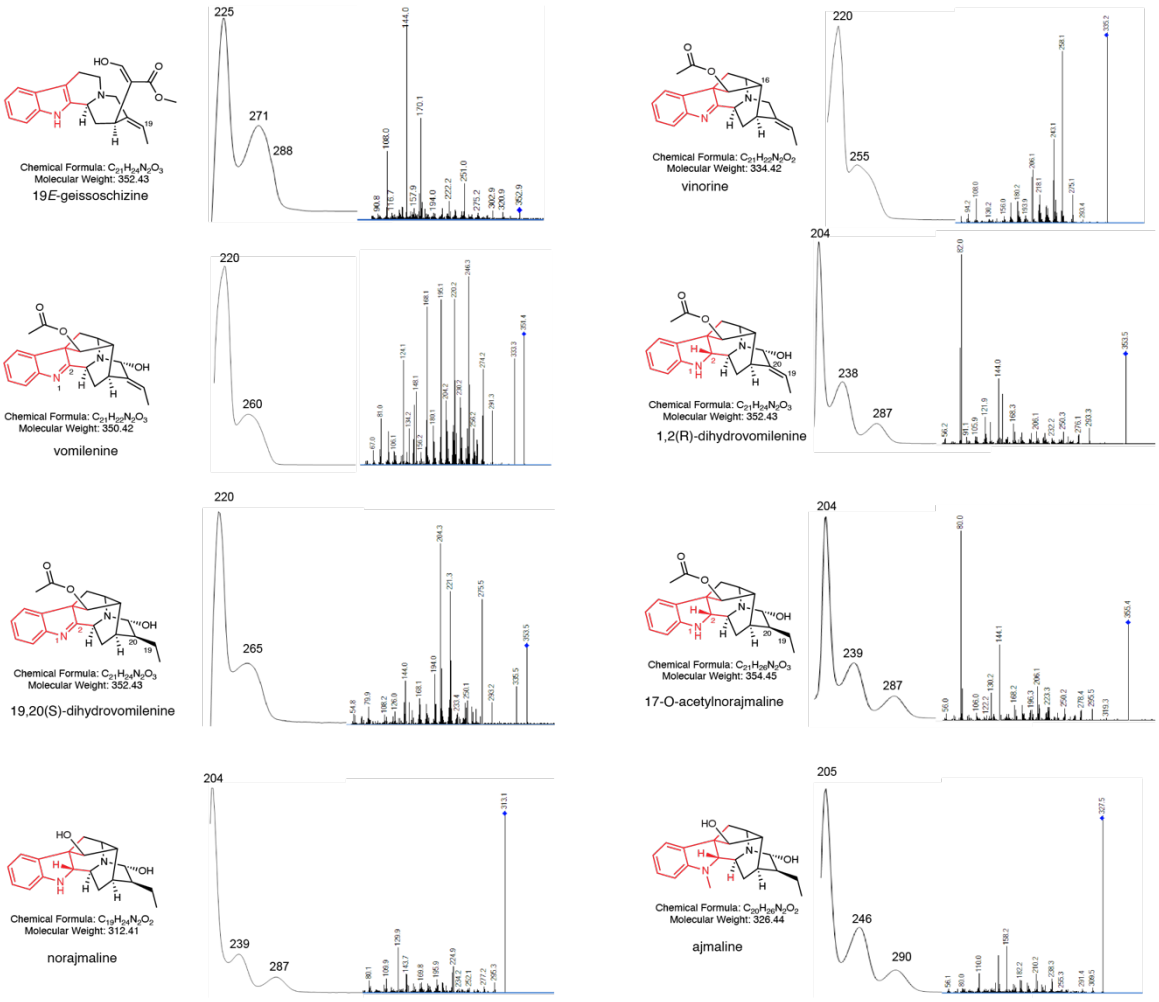
Stavrinides, A., Tatsis, E. C., Caputi, L., Foureau, E., Stevenson, C. E. M., Lawson, D. M., Courdavault, V., & O'Connor, S. E. (2016). Structural investigation of heteroyohimbine alkaloid synthesis reveals active site elements that control stereoselectivity. Nature Communications, 7(1), 12116. https://doi.org/10.1038/ncomms12116

Szabó, L. F. (2008). Rigorous biogenetic network for a group of indole alkaloids derived from strictosidine. Molecules (Basel, Switzerland), 13(8), 1875–1896. https://doi.org/10.3390/molecules13081875

Tam, S. W., Worcel, M., & Wyllie, M. (2001). Yohimbine: A clinical review. Pharmacology & Therapeutics, 91(3), 215–243. https://doi.org/10.1016/S0163-7258(01)00156-5

von Schumann, G. (2002). Vomilenine Reductase—A novel Enzyme catalyzing a crucial Step in the Biosynthesis of the Therapeutically applied Antiarrhythmic Alkaloid Ajmaline. Bioorganic & Medicinal Chemistry, 10(6), 1913–1918. https://doi.org/10.1016/S0968-0896(01)00435-7

Appendix I UV absorption profiles and ESI-MS/MS ion fragmentation patterns of the alkaloids in this study



Appendix Figure 1. The UV absorption profiles and ESI-MS/MS ion fragmentation patterns of the alkaloids in this study. The structures in red are the chromophores responsible for their unique UV absorption profiles, which are distinct for indole (geissoschizine), dihydroindole (1,2-dihydrovomilenine, 17-*O*-acetylnorajmaline, norajmaline and ajmaline), and indolenine (vinorine, vomilenine, 19,20-dihydrovomilenine) chromophores. Norajmaline N-methylation results in absorption shift to longer wavelengths.

Appendix II Gene and PCR primer sequences

Below is a list of all the ajmaline pathway enzyme gene sequences and PCR primers used in this study. F = forward primer; R = reverse primer. Red texts indicate start/stop codons; blue texts indicate gene sequence.

GsSBE (this study)

ATGGAAGTTATGCAATTGTCTTTCTCCTACCCAGCCTTGTTTTTATTTGTTTTCTTCTTGTTCAT
GTTGGTTAAGCAATTACGTAGACCAAAAAATCTGCCACCAGGTCCAAACAAATTGCCAATTAT
TGGTAACTTGCATCAATTAGCCACTGAATTGCCACATCATACTTTGAAGCAATTGGCTGATAA
GTATGGTCCAATTATGCACTTACAATTCGGTGAAGTTAGTGCTATTATTGTTTCTTCTGCTAAG
TTAGCCAAAGTTTTTCTAGGTAATCACGGTTTGGCTGTTGCTGACAGACCAAAAACAATGGTC
GCTACTATTATGTTATATAATTCTTCCGGTGTCACTTTCGCTCCATACGGTGACTACTGGAAGC
ACTTGAGACAAGTCTACGCTGTTGAATTGTTGTCACCTAAATCTGTTAGATCTTTCTCAATGAT
TATGGATGAAGAAATTAGTTTGATGCTTAAGAGAATACAATCTAACGCTGCCGGTCAACCTTT
GAAGGTTCATGATGAAATGATGACTTACTTGTTCGCTACTTTGTGTAGAACTTCAATAGGTTCT
GTTTGTAAAGGTAGAGACTTATTGATTGATACTGCTAAGGATATATCTGCTATTTCCGCTGCAA
TTAGAATTGAAGAATTGTTCCCTTCTTTGAAAATTTTACCATACATCACTGGTTTACATAGACA
ATTGGGCAAGTTGAGTAAACGTTTGGATGGTATTTTAGAGGACATTATTGCCCAAAGAGAAAA
GATGCAAGAATCATCTACCGGTGACAACGATGAAAGAGATATCTTGGGTGTCTTGTTAAAGTT
GAAGAGATCTAATTCCAATGATACAAAGGTCAGAATTAGAAATGATGATATCAAAGCTATTG
TTTTTGAATTGATCTTAGCCGGTACCTTGTCTACCGCCGCTACCGTTGAATGGTGTTTATCTGA
ATTGAAAAAGAACCCAGGTGCTATGAAGAAAGCTCAAGATGAAGTTAGACAAGTTATGAAAG
GTGAAACAATTTGTACTAACGATGTCCAAAAGTTAGAATATATTAGAATGGTTATTAAGGAAA
CCTTCAGAATGCACCCACCTGCCCCATTGTTGTTTCCAAGAGAATGTCGTGAACCAATTCAAG
TCGAAGGTTACACCATTCCTGAAAAGTCTTGGTTGATTGTTAATTATTGGGCTGTTGGTAGAG
ATCCAGAATTGTGGAATGATCCAGAAAAGTTTGAACCAGAAAGATTTCGTAACTCTCCAGTTG
ATATGTCTGGTAACCATTACGAATTAATTCCATTCGGTGCTGGTAGAAGAATTTGTCCAGGTA
TTTCTTTCGCTGCTACTAACGCTGAATTATTGTTAGCTTCTTTAATCTACCACTTCGATTGGAAG
TTACCAGCTGGTGTTAAAGAATTAGATATGGATGAATTGTTTGGTGCTGGTTGTGTTAGAAAG
AACCCATTACACTTGATTCCTAAGACTGTTGTTCCATGTCAAGAT

Gene-Site-Direction: primer sequence
GeSBE-ApaI-F: AGGGCCCATGGAAGTTATGCAATTGTCTT
GsSBE-SalI-R: AGAAGTCGACGTCCTGGCATGGAACAACAGTC

RsPNAE

ATGCATTCTGCTGCAAACGCCAAGCAACAAAAGCATTTTGTTCTGGTACACGGCGGATGTCTC
GGAGCTTGGATCTGGTACAAGCTCAAGCCGCTGCTCGAGTCAGCCGGACATAAGGTCACCGC
CGTTGACCTGTCGGCCGCCGGCATCAACCCAAGAAGGCTCGATGAGATTCACACATTTCGGGA
CTACTCGGAGCCCTTGATGGAAGTCATGGCTAGTATTCCTCCTGATGAGAAGGTTGTTCTTCTT
GGCCATAGCTTTGGTGGCATGAGTTTGGGTCTTGCCATGGAAACCTACCCAGAGAAGATATCA
GTTGCTGTTTTTATGTCTGCAATGATGCCTGATCCTAACCACTCACTAACTTATCCGTTTGAGA
AGTACAATGAGAAGTGTCCGGCAGATATGATGTTGGACTCACAGTTTTCAACCTACGGAAACC
CAGAGAACCCAGGAATGTCAATGATTCTTGGACCTCAGTTTATGGCCCTCAAAATGTTCCAGA
ATTGCTCAGTCGAGGACCTTGAATTAGCCAAAATGTTGACTCGACCAGGTTCGTTATTTTTCCA
AGATTTGGCCAAGGCCAAAAAGTTCTCAACCGAGAGGTACGGTTCGGTGAAGCGAGCTTATA

TCTTTTGCAATGAAGATAAATCATTTCCAGTTGAGTTTCAGAAATGGTTTGTTGAAAGTGTTGG
AGCTGATAAAGTAAAAGAAATCAAAGAAGCAGATCATATGGGAATGCTTTCGCAGCCAAGGG
AAGTTTGCAAGTGCCTGCTTGATATATCAGATTCATAA

Gene-Site-Direction: primer sequence
RsPNAE-BamHI-F: ATAGGATCCCATGCATTCTGCTGCAAACGC
RsPNAE-SalI-R: GAAGTCGACTGAATCTGATATATCAAGCAGGCA

RsVS

ATGGCACCCCAGATGGAGAAAGTATCGGAGGAGCTGATTCTACCATCATCTCCAACACCCCA
AAGCTTGAAATGCTATAAAATTTCCCACCTAGATCAACTGTTATTAACGTGTCACATCCCTTTT
ATTCTCTTCTATCCAAATCCGTTAGACTCAAACCTCGATCCTGCCCAGACATCTCAGCACCTGA
AACAATCTTTGTCCAAAGTGTTAACTCACTTTTACCCTCTAGCTGGAAGGATCAACGTAAATTC
TTCCGTAGACTGTAATGATTCTGGAGTTCCTTTTGTCGAAGCTCGGGTTCAAGCTCAACTCTCA
CAGGCAATTCAGAACGTCGTCGAGTTAGAAAAACTCGATCAATACCTTCCGTCCGCAGCTTAT
CCCGGCGGGAAAATTGAGGTGAACGAGGATGTTCCCCTGGCTGTCAAAATCAGTTTCTTTGAG
TGTGGAGGCACGGCCATTGGTGTCAACTTATCGCATAAGATAGCTGATGTATTGTCCCTGGCC
ACCTTCCTCAACGCATGGACTGCCACATGCCGTGGGGAAACGGAGATTGTGCTACCTAATTTT
GACTTGGCAGCACGTCATTTTCCGCCCGTGGACAACACCCCGTCTCCTGAATTGGTACCGGAT
GAAAACGTTGTGATGAAAAGATTCGTATTTGATAAAGAAAAAATAGGAGCCCTCAGAGCACA
AGCTTCCTCGGCCTCAGAGGAGAAGAATTTCAGTCGGGTACAGCTTGTTGTTGCTTATATATG
GAAGCACGTCATTGACGTGACCCGGGCAAAATATGGTGCTAAAAACAAGTTTGTGGTAGTTC
AAGCAGTGAACCTGAGGTCAAGAATGAATCCGCCCCTTCCTCACTATGCTATGGGGAACATCG
CCACACTATTATTCGCGGCTGTAGATGCAGAGTGGGACAAAGATTTTCCGGATCTCATCGGTC
CGTTGAGAACCAGCCTAGAAAAAACTGAGGACGACCATAACCACGAATTACTAAAGGGAATG
ACTTGTTTGTATGAACTGGAACCTCAAGAACTTTTGTCTTTCACCAGTTGGTGTAGGCTTGGCT
TTTATGACTTGGATTTCGGCTGGGGGAAGCCTCTTTCAGCGTGCACAACAACTTTTCCCAAGA
GGAACGCGGCGCTTTTGATGGATACAAGATCCGGAGATGGAGTGGAAGCATGGCTCCCAATG
GCAGAAGATGAAATGGCGATGCTTCCTGTTGAATTGCTGTCACTTGTAGACAGCGATTTTAGC
AAGTGA

Gene-Site-Direction: primer sequence
RsVS-NotI-F: AGCGGCCGCATGGCACCCCAGATGGAGAAA
RsVS-SpeI-tga-R: GGACTAGTCACTTGCTAAAATCGCTGTCT

RsVH

ATGGATCTCCTCCAGATTCTGCTGGCTATTGCTGGGCTGCTGGCCATCCTCTTGCTTCAAAAGC
AATGGAGAACAAAGACATCTCCTGGAGCAAAAGCTGGGCGCAAGTTACCACCAGAACCAGCA
GGTGCCTGGCCTGTAATAGGCCACCTTCACAAACTCGGCGGCCCTAACCCCATATACCGGAAT
CTGGCGGAGTGGTCTGATAAATATGGTCCGGTCATGACACTCAAGCTGGGAATGCAAAATGC
AGTGGTGGTGAGCGACCGCGAAGCAATTAAAGAATGCTTCACCACCAACGATAAGGCCCTCG
CCGACCGCCCACCTTCCAGTATCGGCTTACACCTTGGCTTCAACTATGCGGCTATTGGTGCTGC
ACCTTATGGTCCGTACTGGCGTGATATGCGGAAGTTGGTTTTGTTAGAAGTCCTTTCGAGTCGG
AGGCTCGAGATGCTCAGGAACGTCAGGATCTCCGAGATAGGAACCAGCATCAAAGAATTATA
CTCGAATATTATCAGGAGCAGTGGGGGTTCAGGTCCAGCAAAAGTGGTGATCAGCCACTGGA
TCGAGCAATTGACCTTGAATTACATTTTGAGGACAATTGCTGGGAGGAGATTCAGCGATGACT
CAAGCAAGGATGCACAGTATGTCAAAGGAGTAATCAACGACTTCATGTATTTTGCAGGACAA
TTTGTGGTATCGGATGTGATTCCAATTCCATTGTTGAGATGGCTCGATCCCCAGGGACATCTTA
AAGGGATGAAGCGCGTAGCTAAAGAGGTTGATACTATGTGTGAAGCTTGGATCCAAGAACAC
GTGCAACGAAGGATGAGGGAAAAGCCGGGACCTGGGCAGGAGCAAGACTTCATAGATGTGCT
GTTGAATAATAGGGATGTTATGAGGAAAGCTCAAGAGGAAATAGATAACCATGTTGGTAAAG

AAAGATGGGTAGACGAAACTGATCTCAAACACCTGGTGTACCTCCAGGCAATAGTGAAAGAA
GGGTTGCGATTGTATCCTCCAGGCCCTCTCGGCGCCCCTCACCGAGCCATAGAAGACTGCCAG
GTAGGTGGTTACTTCATTCCGAAAGGCACCCAATTGTTGGTAAATGTTTGGAAGCTGCACCGA
GACCCGAGAGTTTGGTCGGAACCAGAAAAATTCATGCCTGAAAGATTTTTAACAAGGCAAGC
AGAGGTGGATGTGTTCGGCCATCATTTTGAACTGCTCCCTTTTGGCTCTGGACGACGAGCGTG
TCCAGGAATAACATTTGCAGTACAAGTAATGCACTTGACGGTTGCTCGACTGCTTCAAGGTTT
CGACATGACAACGCCATCAAACTTGCCAGTGGACATGACGGAAGGCCCAGGCGTCACCATGC
CTAAGGCCCATCCAGTGGAGGTATTGATGATGCCGCGACTTCCTTCAGCACTTTATGAACCAT
AG

Gene-Site-Direction: primer sequence
RsVH-ApaI-F: AGGGCCCATGGATCTCCTCCAGATTCTGC
RsVH SalI-R: ATAGTCGACTGGTTCATAAAGTGCTGAAGGA

RsVR (RsCAD2)

ATGGGTGCAGCATGTGCAGAAACAGCAAAGCCAATTGAGGCCTACGGATGGGCAGCCAGAGA
CGCATCTGGAGTTCTCTCTCCATTCAAGTTCCAAAGAAGGGCTACGGGAAAGCACGATGTGCA
GCTCAAAGTGTTGTACTGTGGGATGTGCGATTGGGATCTACTTGTAGTCAAGAATTTGCTTGG
CACTACTAAATATCCCATTGTACCTGGGCATGAGGTGGTGGGAGTGGTGACTGAGATCGGTAG
CAAGGTGCAAAAGTTCAAGGTTGGGGACAGGGTAGGTGTTGGTCACTACGTTCAAACATGTC
GTAAATGTGAGAGATGTGAAGAAGGTCTTGACAGTTATTGTCCAAACTTGGTAATAGCGGATG
GAACTTCTTTTAGTGATGGAAAGGACCTATATTTCTACGATCCAAATGACACAGAGAGCAAGA
TGTACGGTGCCTATTCCAACATCACGGTTGTCGATGAGTACTACGTAATCCGTTGGCCGGAAA
ACTTTCCTTTGGCTGCCGGCGTACCTCTTTTATGTGCTGGTGTAGTTCCCTACAGCCCCATGAG
ATACTACGGCTTTGATAAACCCGGAATTCATATTGGCGTCGTTGGACTTGGTGGGATGGGCAG
ATTAACCGTGAAATTTGCTAAGGCTTTTGGAGCAAAAATTACAGTAATCAGTACATCCATTGA
CAAGAAGCAAGAAGCTATTGAGAAATATGGTGCAGATAAATTTTTACTCAGCAAAAAACCTG
AGCAGCTGCAGGCCGCGATTGGTACGCTGGATGGCATCATTGACACAGTCCCCAGAGTTCACC
CCCTTCGTGAATTGATCAAATTGTTGAAATTCGACGGCACTCTTGTTTTGCTAGGAGCACCCCC
GGAGCCATATGAGTTGCCGGCCTCTCCACTGCTTGTAGGGAGGAAGAAGGTGGTTGGAAGTG
GTGGTGCGAGTATCAAAGAAACACAAGAGATGATGGATTTTGCAGCGAAGCACAATATAGTC
GCAGATACAGAGATCATTCCAATGGGTTATGCAAACACTGCAATCGAGCGGATAGAAAAGGG
TGATTTCAGAAAACGATTCGTGATTGATATAGAGAATACATTGAAATATTCTGCTTAG

Gene-Site-Direction: primer sequence
RsCAD2-BamHI-F: ATAGGATCCGGAAATGGGTGCAGCATGTG
RsCAD2-SalI-tga-R: GAAGTCGACCCCGGCCTATTCTAAGCAGA

RsDHVR (full PhytoMetaSym VR2)

ATGCGTTTTGCTAACTTTGAGGGCTCCAATTATTATTGTATTGACACCACTCCTTATTCTTCTTC
TGCTCTGTGTATGTGTATGATTATCTCATTCTCAAATCACTACTTTGTGATCAGTTTTCCCTCAT
TCTCTCGTTTTGTATTCGAAATGGCTGGAAAATCTCCAGAAGAGCAACACCCAGTTAAGGCAT
ATGGATGGGCAGCAAGAGACTCATCTGGGATTCTTTCTCCCTTCAAGTTCTCCAGAAGGGCAA
CAGGTGATCATGATGTCAGAGTAAAGATTCTCTACGCTGGTGTTTGTCATTCTGACCTTCAATC
TGCCAGGAATGACATGGGCTGCTTTACATATCCTCTTGTGCCCGGGTTCGAGACGGTAGGCAT
AGCGACTGAAGTAGGAAGCAAGGTCACAAAAGCGAGAGTCGGCGATAAAGTTGCAGTGGGA
ATCATGGTGGGATCATGCGGCAAATGCCACGAGTGCGTCAATGACCATGAATGTTACTGCCCA
GAGGTGATCACATCTTATGGTCGAATGTACCATGATGGAACTCCCACTTACGGAGGTTTCTCC
AATGAGACAGTAGTGAGTGAGAAATTCGTTTTTCGTTTTCCTGAAAAACTTCCAATGGCTGCT
GGTGCTCCACTGCTCAGTGCTGGAGTCTCTGTGTACAGTGCAATGAGATTTTATGGCCTGGAT
AAGCCAGGGATGCACCTGGGAGTTGTAGGGCTTGGTGGACTTGGTCATTTAGCGGTCAAGTTC

GCCAAGGCTTTTGGGGTCAAAGTCACTGTGATTAGTACCTCTACAAGCAAGAAGGATGAAGCT
ATCAATGATCTTGGCGCTGATGCATTCTTGGTCAGTACTGATGATGAACAAATGCAGGCTGGC
TCTGGAACCTTGGATGGAATTCTTGACACCGTACCTGTTGTCCATCCTATTGGGGCCTTGCTAG
GTCTACTGAAGAATCACACGAAGCTTGTATTGGTGGGAGCTACAATGGGCTCATTTGAGTTGC
CAATTCTTCCTTTAGGAGTGGGCAGGAAAAGTGTGGTTTCAACTATTGGAGGAAGTACGAAGG
AGACTCAAGAGATGCTCGATTTTGCAGCAGAACACGATATCACCGCCAATGTTGAGATTATTC
CGATGGACTATATAAATACAGCAATGGAACGCATTGAGAAGCGCGATGTTCGATATCGATTTG
TGATTGACATCGGCAACACCTTAACTCCACCTGAGTCCTAA

Gene-Site-Direction: primer sequence
RsDHVR(full)-BamHI-F: ATAGGATCCAATGCGTTTTGCTAACTTTGAGG
RsDHVR(full)-SalI-tga-R: ATAGTCGACTTAGGACTCAGGTGGAGTTAAGG

RsDHVR (truncated PhytoMetaSyn VR2)

ATGGCTGGAAAGTCTCCAGAAGAGCAACACCCAGTTAAGGCTTATGGATGGGCGGCTACAGA
CTCATCTGGGATTCTTTCACCCTTCAAGTTTTCCAGAAGGGCAACAGGGGATCATGATGTCAG
AGTAAAGATTCTCTATGCTGGTGTTTGTCATTCTGACCTTCAATCCGCCAGGAATGACATGGG
CTGCTTTACATATCCTCTTGTGCCAGGGTTCGAGACAGTAGGCACAGCAACTGAAGTCGGAAG
CAAGGTCACAAAAGTGAAAGTCGGCGATAAAGTTGCAGTGGGAATCATGGTGGGATCATGCG
GCAAATGCGATGAGTGTGTCAATGATCGTGAATGTTACTGCCCAGAGGTGATCACATCTTATG
GTCGCATAGACCATGACGGAACTCCCACATATGGAGGCTTCTCCAGTGAGACTGTAGCAAATG
AGAAATTCGTTTTTTGTTTTCCTGAAAAACTTCCAATGGCTGCTGGTGCTCCACTGCTCAATGC
TGGAGTCTCCGTGTACAGTGCAATGAGATTTTATGGCCTGGATAAGCCAGGGATGCACTTGGG
AGTTGTAGGGCTTGGTGGACTTGGTCATTTAGCTGTCAAGTTCGCCAAGGCTTTTGGGGTCAA
AGTCACTGTGATTAGCACCTCTACAAGCAAGAAGGGTGAAGCTATCAATGATCTTGGTGCTGA
TGCATTCTTGGTTAGCACTGATGCTGAACAAATGCAGGCTGGCTCTGGAACCTTGGATGGGAT
TCTTGATACCGTGCCTGTTGTTCATCCTATTGAGGCCCTGCTAGGTCTACTGAAGAATCACACG
AAGCTTGTATTGGTGGGAGCTACGATGGGCTCATTTGAGTTGCCCATTCTTCCTTTGGGAGTGG
GCAGGAAAAGTGTGGTTTCAACCATTGGAGGAAGTACGAAGGAGACTCAAGAGATGCTTGAT
TTTGCAGCAGAACACGATATAACCGCGAGTGTTGAGATTATTCCGATGGACTATGTAAATACA
GCAATGGAACGCATTGAGAAGGGCGATGTTCGATATCGATTTGTGATTGACATCGGCAACACC
TTAACTCCACCTGAGTCCTAA

Gene-Site-Direction: primer sequence
RsDHVR(truncated)-BamHI-F: ATAGGATCCTATGGCTGGAAAGTCTCCAGAAG
RsDHVR(truncated)-SalI-tga-R: ATTGTCGACTTAGGACTCAGGTGGAGTTAAGG

RsAAE (this study; PhytoMetaSyn version)

ATGGGTTTTGCTCCGCTTTTAGTTTTTTCTCTCTTTGTTTTTGCAGGGACAACCAAAGGGTTCAT
ATGCTCCTTCGATTCAATATATCAGTTAGGTGATTCATTTTCGGATACGGGCAATCTTATCCGT
CAACCACCTGACGGTCCGACGTTTTGCTCTGCACATTTTCCTTATGGAGAAACTTTTCCAGGAA
TGCCTACAGGTCGTTGCTCAGATGGTCGTTTGATAATAGATTTTATTGCTACGGCTCTCAATCT
ACCGTTGCTTAATCCTTATCTACAACAGAATGTTTCCTTCCGACACGGTGTCAACTTTGCTGTT
GGCGGGGCTACAGCGCTAGATCTTTCTTTCTTAGCAGCAAGAGGTGTTCAAGTCTACGATGTC
CATTCTCCCCTAAGTACTCAGTTGAAGTGGTTTCGAACATATCTTGGTTCCATCTGTTCTTCAC
CAAAAGAATGTTCAAACAAGCTCAAGAACGCTCTTTTCATCCTTGGTAATATTGGAAACAATG
ATGTTAATTATGCATTCCCGAACAGAAGTATTGAAGAGATCCGAGCTTATTTACCATTCATAA
CTGAGGCCGTTGCCAATGCAACAAGAGAAATTATCCGTCTGGGTGGAACTCGAGTAATCGTTC
CTGGAATGTTTCCCCTTGGCTGCTTGGCCAGGAACTTGTATTTCTTCCCGGATGGTGATAAGGA
TGATCTGGGCTGCTTGAGCAGTTTGAATGATCTTTCAATATACTTCAACAGTCTCATCCAACAA
GCTTTGGCTTCTCTCAGGATCGAATTTCCTCAGGCAGTCATAATTTATGCTGACTATTACAATG

CTTGGGGATTCCTTTTTCGGAATGGACCTGCTCTTGGTTTTAACTCAACAACCATGCTAAAATG
TTGCTGTGGGATTGGAGGGCCTTATAACTATGATCCAGACCGAGAATGTGCATCTCAAGGAGT
GCCTGTTTGCTCTAATCCAACAGAATATATTCAATGGGACGGTACTCATTTTACACAAGCTGCT
TACCGTCGCGTCGCAGAATATATTATTCCTGACATTATCAAAGAACTCAAATGCTCTTATAGC
AGTATTCAGCATCTGACAGAGGGAAGAGAAGCCCTCCACATTAATGAAAGAGAATAA

Gene-Site-Direction: primer sequence
RsAAE-EcoRI-NcoI-F: ATAAGAATTCCATGGGTTTTGCTCCGCT
RsAAE-His-NotI-XmaI-R:
TGCGGCCGCCCGGGTTAATGGTGATGATGGTGATGTTCTCTTTCATTAATGTGGAGG

RsNNMT

ATGGCAGAGAAGCAGCAGGCAGTGACGGAGTTCTACAACAACACATCGCCAAGGGGAGCAT
GGGAGTTCCTCCTGGGAGACCATTTGCACGAAGGTTTTTATGACCCTGGAACAACCGCCACCA
TCTCCGGTAGCCAAGCCGCTGCAGCTCGAATGATCGATGAGGCTCTCCGTTTTGCCAACATTT
ACGATGATCCATCAAAGAAACCGAAAAACATGCTGGACATCGGATGTGGAGTAGGTGGGACT
TGTGTCCATGTAGCAAAGCAATATGGTATACAATGCAAAGGCATCACACTAAGCCCTGAGGA
AGTCAAATGTGCTCAAGGTATTGCAAAAGCCCAAGGACTAGAAGAAAAGGTCTCTTTCGATG
TGGGAGATGCCTTAAATCTGCCCTATAAAGATGGAACATTTGATCTGGTTCTCACCATTGAGT
GCATAGAACACGTTCAAGACAAAGAAAAGTTCATCCGCGAGATGATTCGGGTGGCAGCTCCT
GGTGCTCCTATAGTTATCCTGTCATACGCCCACCGGAACCTTTCTCCTTCGGCAGAATCCTTGA
AGCCAGATGAGAAGAAAGTACTGAAGAAGATATGTGATAACCTTGCTCTGTCATGTCTTTGTT
CTTCGGCTGATTTTGTGAGATGGTTGACACAACTTCCCGCTGAGGATATCAAGACTGCAGACT
GGACTCAAAACACCTCCCCATTTTTCCCTCTATTGATGAAAGAAACATTCACATGGAAGGGCT
TCACATCATTGCTCATGAAGGGTGGATGGACTGCTATCAAGGAGCTACTAGCACTGAGGATGA
TGTCTAAGGCAGCCGATGACGGTCTTCTTAAGTTCGTTGCAATTACATGCAGGAAATCAAAAT
AA

Gene-Site-Direction: primer sequence
RsNNMT-BglII-ApaI-F: ATAGATCTGGGCCCATGGCAGAGAAGCAGCA
RsNNMT-SacII-SphI-R: GGCCGCGGCATGCTTATTTTGATTTCCTGCATGTA

RsCAD1

ATGGCCGGAAAATCACCTGAAGAGGAGCACCCAGTGAAAGCCTACGGAGTGGCTGCTCGAGA
TTCGTCTGGGGTCCTTTCACCCTTCAAATTCTCCCGGAGGGCAACACTTGAGGATGATGTTAG
ACTCAAGGTGCTCTATTGTGGGTTATGTCATACTGACATTCATTTCCTCAAGAATGAGTGGGG
CTTTTCTACCTACCCTTTTGTACCGGGGCATGAAGTTGTAGGTGAAGTTATAGAGGTTGGTAGC
AAAGTTACAAAAGTCAAGGTTGGGGATAAAGTTGCTCATGGCGGCATTATCGGGTCATGCCGT
GCATGTGATAATTGTCATGCAGATATGGAGAGCTATTGTCCCAAAATGGTTATGGCCCATGGA
TCTCCAAATTTTGATGGAACCATTACTTATGGAGGCTTTTCCAATGAGATGGTCGTCAATGAG
CACTTTGTTATTCGTTACCCAGAGAACCTGCCACTTGCTGCTGGTGCACCATTGCTCTGTGCTG
GAATTACAGTGTACAGTCCAATGAAATACTATGGAATTGCAAAACCTGGAAACCACATAGGT
GTTAACGGTCTTGGTGGGCTGGGCCATATGGCTGTTAAATTCGCAAAGGCCTTGGGAGCAAAA
GTGACAGTCATCAGTTCATCCGAGAGCAAGAAAGACGATGCTATAAATCATCTGGGTGCAGA
TGCATTTTTACTGAGCAAAAATCCAGAAGAACTGCAGGCTGCAACAGGCACGTTGGATGGTAT
AGTCGATTGTGTTTCTGCTAAACACCCAATTATCCCATTGCTTGGTCTACTCAAGTCTCACGGA
AAGCTTGTTCTGGTTGGGGCACCTGGGGAGCCACTTGAGCTTCACTCTGCCCCTTTGCTTATGG
GAAGGAAGATGATCGGTGGAAGTGATGCTGGAGGAATGAAGGAGATTCAAGAAATGGTTGA
CCTTGCTGCAAAGCACAATATCACTGCAGATATCGAGCTTGTTTCCATGGACAACATCAACAC
AGTTGTGGAGCGCCTTGTCAAGGGTGATGTTAGATATCGCTTTGTTGTTGACGTTGCCAACAC
CTTGAAAGCTCCTTAA

Gene-Site-Direction: primer sequence
Rs90-BamHI-F: ATAGGATCCGATGGCCGGAAAATCACCTG
Rs90-SalI-tga-R: ATAGTCGACTTAAGGAGCTTTCAAGGTGTTGG

RsCAD3

ATGGGTAGCTTGGAAGCTGAGAGAAAGACTACAGGATGGGCAGCAGGAGACCCTTCTGGAGA
ACTCGCCCCCTACACCTACTCTCTCAGAAATACTGGGCCTGAAGATGTTTATCTCAAGGTGAT
ATGCTGCGGAGTCTGCCATACTGACATCCACCAGACAAAGAATCACCTCGGCATGTCCAATTA
TCCCATGGTTCCAGGGCATGAAGTAGTAGGTGAGGTGCTGGAGGTGGGATCCAACGTGACCA
AGTTCAGAGTTGGGGACCAAGTGGGAGTAGGTATAATCGTTGGATGCTGCAGAAACTGCCGC
CCGTGCCAAACAGATATTGAGCAATACTGCAACAAGAAGATTTGGACATACAATGATGTCTA
CACTGATGGCAACCCTACTCAAGGTGGATTTGCCAGTGCCATGGTCGTCGATCAGAAGTTTGT
GGTGAAAATCCCAGATGGTATGTCACCAGAGCAGGTAGCACCTCTGCTATGTGCTGGGGTAAC
GGTGTACAGTCCATTGAACCACTACGGGTTGAAGCAAAGTGGACTAAGAGGAGCCATATTAG
GACTTGGAGGGGTTGGGCATATGGGAGTGAAAATAGCCAAGGCCATGGGGCATCACGTAACC
GTCATAAGCTCTTCCGACAAGAAGAGGGAGGAGGCTTTGGACCACCTTGGTGCAGACCAGTA
CGTGGTGAGCTCGGACGAAGCCAAGATGCAGGAGGCTGCAGACTCACTAGACTACATTATTG
ACACGGTTCCCGTGTTTCACCCTCTCGAGCCATACCTATCGTTGTTGAAAGTTGATGGAAAGTT
GATTTTGATGGGCGTCATCAACCAACCCTTACAATTTGTCACCCCCATGGTTATGCTCGGAAG
GAAGTCGATCACAGGAAGCTTTATTGGTAGCATCAAAGAGACAGAGGAAGTTCTCGAGTTCT
GCAAGGAAAAAAACCTGACTTCCCAGATTGAAGTAGTGAAGATGGACTATATCAATAAGGCT
ATAGAAAGGCTGGAGAAGAATGATGTGAGATACAGATTCGTTGTGGACGTCGCCGGCAGCAA
TCTTGAGCAGTAG

Gene-Site-Direction: primer sequence
RsCAD3-EcoRI-F: ATAAGGATCCAATGGGTAGCTTGGAAGCTGAGA
RsCAD3-SalI-tga-R: GAAGTCGACCTTTCTACTGCTCAAGATTGCT

RsCAD4

ATGGCCGGAAAATCACCAGAAGAGGAGCACCCAGTGAAGACCTACGGATGGGCTGCTCGTGA
TCCATCTGGGGCTCTTTCTCCCTTTAAATTCTCCAGGAGGGCAACACTTGATGATGATGTTAGA
TTCAAGGTACTCTATTGTGGGGTGTGCCATACTGACCTTCATTTCGTCAAGAATGACTGGGGCT
TTTCTACCTACCCTCTTGTACCGGGGCATGAAATCGTAGGAGAAGTTACAGAGGTTGGTAGTA
AAGTTACAAAAGTCAAGGTTGGAGATAAAGTTGGTGTTGGCTGCTTGGTTGGTTCATGCCGCA
CTTGTGATAATTGTAGTGCAGATCTTGAGAACTATTGTCCCAAAATGGTGGTAACCTATTCATT
TCCATATTTTGATGGAACCATTACATACGGAGGCTACTCCAATGAGATGGTCTGCAATGAGCA
CTTTATTATTCGTTTCCCAGAGAACCTGCCACTTGATGCTGGTGCACCATTGCTCTGTGCTGGA
ATTACAGTGTACAGTCCAATGAAATACTATGGCATTGCGAAACCTGGAAACCACATAGGCGTT
AACGGTCTTGGTGGGCTTGGCCATGTGGCTGTTAAGTTCGCAAAGGCCTTGGGAGCAAAAGTG
ACAGTCATCAGTACATCTGAGAGCAAGAAAGACGAAGCTATAAATCGTCTGGGTGCAGATGC
ATTTTTGCTGAGCAATAATCCAGAAGAACTAAAGGCTGCAACAGGCAAATTGGATGGTATAA
TCGACTGTGTTTCTGCTAAACACCAAATTATCCCATTGCTTGGTCTACTCAAATCTCATGGAAA
GCTTGTTCTAGTGGGGGCACCGGCAGAGCCTCTTGACCTTCATTCTGCGTCTTTGCTTATGGGG
AGGAAGATGATTGCTGGAAGTAACATTGGAGGATTGAGGGAGACTCAAGAGATGATTGATTT
TGCCGCAAAGCACAAAATCACTGCAGATATCGAACTTGTTTCCATGGACAGTATCAACACAGC
TTTGGAGCGCCTTGCCAAGGGTGACATTAGATATCGCTTTGTCATTGACGTTGCCAACACCTTG
AAATCTCCTTAA

Gene-Site-Direction: primer sequence
RsCAD4-BamHI-F: ATAGGATCCGATGGCCGGAAAATCACCAGAA

RsCAD4-SalI-tga-R: GAAGTCGACTTAAGGAGATTTCAAGGTGTTGGCA

RsCAD5

ATGGCAAAATCATTCGAGGAAGAACACCCTGTGAAGGCGTTTGGATGGGCAGCTAGAGACTC
ATCTGGGGTTCTTTCTCCCTTCAAATTCTCTAGAAGGGCTACTGGGGAGAAAGATGTGAGGTT
CAAGGTGTTGTTTACAGGGATATGTCACTCTGATCTTCACCATCTCAAGAATGAATGGGGCAC
CTCCAAGTACCCTTTAGTTCCTGGGCATGAGATCGTCGGTGTTGTGACAGAAGTGGGTAGCAA
GGTGGAGGGCTTCAAAGTAGGTGATAAAGTGGGTGTGGGATGCCTAGTTGGATCATGTCGCA
GCTGTGAAGATTGTACTGATGATCTTGAGAACTACTGCCCAAAGAAGATACTCACATATGATG
GCCTCCCATATCATGATGGAACCATGAACTTTGGAGGTTATTCCGATCACATGGTTGTCGATG
AACATTTTGCCCTTCAATGGCCGGAGAACCTCCCTCTTGACGCTGGTGCTCCGCTCCTTTGTGC
TGGGATTACAACCTATAGCCCAATGAAGTACTTTGGACTTGACAAGCCAGGTTTACACATTGG
TGTTGTAGGACTTGGTGGTTTAGGCCATGTGGCTGTGAAGTTTGTCAAGGCTTTTGGATCTAAG
GTAACTGTGATCAGTACTTCTCCTAGCAAGAAGGAGGAGGCACTCAAAAATCTTGGAGCTGAT
TCCTTTCTGGTTAGCCGCGATGCTGATGCGATGCAGGCTGCAGCAAGCACATTGCATGCTATT
ATTGACACAGTTTCTGCTGTTCACCCCCTTCTACCACTGATCAGCCTGTTGAAGAACCACGCTA
AGTACATAATGCTTGGTGCACCTGAAAAGCCACTTGAACTACCAGTTTTTCCTATGCTTATGG
GGAGAAAAATAGTGGCTGGGAGCAATATTGGAGGCTTGAAGGAGACACAAGAAATGCTCAAT
TTTGCCGCAGAACAGGGGATAACAGCAAATGTTGAAGTCGTTCCAATTGATAATGCGAATACT
GCCCTTGAGCGTCTTGCCAAAAATGATGTGAGGTATCGATTTGTCATCGATGTTGGGAACACA
TTGAAGGCTGCCTAG

Gene-Site-Direction: primer sequence
RsCAD5-BamHI-F: ATAGGATCCAATGGCAAAATCATTCGAGGAAG
RsCAD5-SalI-tga-R: GAAGTCGACTAGGCAGCCTTCAATGTGTTC

RsCAD6

ATGCTGATATCTCCCGTTTTATACACAAAACCACCCCTCCTCCCTACAGCTACGCCAGAATTAC
TTTACCAGACCGCCGGTAGGAAGTTGACCTCTATAATAATGGCCGGAAAATCACCGGAAGAG
CTGTTCCCAGTGAAGACTCATGGATGGGCTGCTCGTGACTCATCTGGGATTCTCTCCCTTTTCA
AATTCTCCAGGAGGGCGACACTTGAGGATGATATCAGATTCAAGGTTCTCTACTGTGGTATTT
GCCATACTGATCTTCACTTCATCAAGAATGAATGGGGCATATCGAGATATCCTCTTTTACCAG
GGCACGAGATTGTAGGTGAAGTTACAGAAGTTGGAAGCAAAGTTACAAAAGTAAAAGTTGGA
GACAAAGTGGGGGTTGGCTACTTGGTAGGATCATGCCGTAGTTGTGACAATTGTTCAGCGGAC
CTGGAGAACTATTGTCCTAAAATGGTTCTAACAGCTGGAGCTCTTTATTTTGATGGCACCCCTA
CATATGGTGGCTTTTCAAATGAAATGGTATGCAACGAGCACTTTGTGATTCGTTTCCCGGACA
ATTTGCCACTCGATGCTGGTGCTCCATTGCTTTGTGCTGGTGTCACTGTCTACAGTCCGATGAA
ATACTATGGCTTTGCCAAACCAGGAAACCATGTAGGAGTTAACGGGCTTGGAGGGCTTGGTCA
CGTGGCTGTTAAGTTTGCAAAGGCCTTTGGGGCAAAAGTCACGGTTATCAGTAGATCTTCTAA
AAAGAAGGAGGAAGCCATTGAGCATCTTGGTGCAGATGCATTCTTAGTGAGCCAAAATCCAG
AAGAAATGAAGGCTGCAATGGGCACCATGGATGGTATAATAGATTGTGTCTCAGCTAAGCAC
CAATTGGTGCCATTACTTGGTCTACTCAAGTATCATGGGAAGCTTGTTCTGGTTGGGGTACCA
GCAGAGCCACTTGACCTTCCTGTTCCTCCTTTGATTATGGGAAGGAAACTTGTTGGTGGAAGT
AATGTCGGAGGGCTAAAGGAGACACAAGAGATGATTGATTTTGCTGCAAAGCACAACATTAC
AGCAGATGTGGAGGTTATTTCTATGGACTATGTTAATACAGCTATGGAGCGTCTTGCCAAAGG
TGATATCAGATATCGCTTTGTCATCGACATCGGCAACACCTTGAAAGCTCCTTAA

Gene-Site-Direction: primer sequence
RsCAD6-BamHI-F: ATAGGATCCTTGCATGCTGATATCTCCCGTTTT
RsCAD6-SalI-tga-R: GAAGTCGACTTAAGGAGCTTTCAAGGTGTTGCC

RsCAD7

ATGGCAGCAGCAGAAACAGCAAAGACAATTGAAGCCTACGGATGGGCAGCCAGAGACGCAT
CTGGAGTTCTCTCTCCATTCAAGTTCCAAAGAAGGGCCACAGCAGAGCATGATGTCCAGCTTA
AAGTGTTGTATTGCGGGATGTGCGATTGGGATTCAATCGTAGTCAAGAATGGGTTTGGAACTA
CTAAATATCCCATTGTACCCGGGCATGAGGTGGTGGGTGTGGTGACTGAGATCGGTAGCAAG
GTGCAGAAATTCAAGGTTGGGGACATAGTAGGTGTCAGTAGCTATGTTCGAACATGTCGTAAA
TGTAAGAGATGTAAAGAAGGTCTTGACAGTTATTGCCCAAACTTGATAACTGGGGATGGTACT
TCGTTTAGTGATGGAAATGACTTATATTTCCACGATCCAAATGATACTGAAAGCAAGATATAC
GGTGGCTTTTCCAACATCACGGTTGTCGAGGAGTATTACGTTGTCCGTTGGCCTGAAAACTTTC
CTTTGGCCGCCGGTGTTCCTGTTTTATGTGCTGGTACCGTTCCCTACAGCCCCATGAGATGCTT
CGGATTTGATAAACCCGAAATTCATCTTGGTGTGGTCGGACTTGGTGGGATCGGCAGATTAAC
CGTGAAATTTGCTAAGGCTTTCGGAGCAAAAGTGACGGTAATCAGTACCTCCATTGACAAGAA
GCAAGAAGCTATTGAGAAATATGGTGCAGATAGTTTTTTACTCAGCAAAGAACCTGAGCAGCT
GCAGGCTGCGGTTGATACACTGGATGGCATTATTGACACAGTCCCTAGAATCCATCCTATTCT
TCCATTGATCAAATTGTTGAAATTCGACGGCACCCTTGTTTTGCTTGGAGCACCCCTGGAGCCA
TATGAGTTGCCGGTGTCTCCTTTGCTTATGGGGAGGAAGAGGGTGGTGGGGAGTGCCGGTGCG
AGCATGAAGGAAACGCAAGAGATGATGGATTTTGCAGCGAAGCATAACATAGTTGCAGATGT
AGAGATCATTCCAATGAATTATGCAAACACTGCAATTGAGCGGATAGAGAAGGGTGATTTCA
AGAACCGATTCGTGATTGATATAGAGAATACATTGAAATCTGCTTAG**CATAGATGCTGCATT**
TCAAGT

Gene-Site-Direction: primer sequence
RsCAD7-BamHI-F: ATAGGATCCGATGGCAGCAGCAGAAAC
RsCAD7-SalI-tga-R: GAAGTCG**ACTTGAAATGCAGCATCTATG**C

RsCAD8

ATGTCCACTGTAAGTGTCGACGAGGATTGTCTCAGCTGGGCCGCAAGAGATTCATCTGGAGTT
CTGTCACCCTACAAATTTAGCCGAAGAGTGATCGGGGCTGATGATGTTGATATAAAAATTGCA
TTTTGTGGAGTTTGTTATGCTGATGTTGTTTGGAGCAGGAATATCCTGGGAACTACAAAGTATC
CTTTGGTGCCTGGACACGAAATTGTTGGGATTGTAAGAGAAGTTGGCCCCAATGTTCAGCGTT
TTAAAGTTGGTGACCATGTAGGAGTTGGAACTTACGTTGGTTCTTGCAGACAATGTGAATACT
GTGACGATGGATTAGAAGTCCATTGCTCAGAAGTAGTCCTCACTTTCGATGGTATTGATGTGG
ATGGTACAGTCACTAAAGGAGGATATTCTAGTCATATTGTTGTTCACGAGAGGTACTGCTTTA
AAATACCCGACAATTACCCACTTGCATTGGCAGCGCCTTTGCTTTGTGCTGGGATTACTGTCTA
CACGCCCATGATGCGTCACAACATGAACCAACCTGGCAAATCTTTGGGTGTGATTGGGCTAGG
TGGTCTTGGTCACTTAGCAGTTAAGTTTGGAAAGGCTCTTGGACTGAAAGTAACAGTTTTCAG
CACAAGTACATCAAAAAGGGATGACGCACTGAATCTTCTAGGAGCAGACAATTTTGTAGTCTC
ATCTGACGAACAGCAGATGATGAGGCTGGCTAAATCACTTGACTTCATAATCAACTCAGCTTC
AGCAGAATTTCCTTTTGATCCATACCTATCTCTGTTGAAGACTGCGGGCATTCTTGTGCTGGTG
GGTTTTCCACGTGAAGTCAAATTCAGCCCGGGAAGCCTAATTATGGGTATGAAGACCATATCT
GGCAGCGCAACTGGTGGAACGAAACAGACGCAGGAAATGTTGGAGTTCTGTGCTTCACACAA
AATTTATCCAGAAATTGAAATAATTCCAATTCAACAGTCAAATGAGGCTCTTGAGAGGATGAT
CAAGAAGGATGTGAAATATCGTTTCGTGATAGATGTTGCAAATTCGCTCAAGTGA

Gene-Site-Direction: primer sequence
RsCAD8-BamHI-F: ATAGGATCCAATGTCCACTGTAAGTGTCGACGA
RsCAD8-xhoI-tga-R: GCTCGAGTCACTTGAGCGAATTTGCAACATC

Curriculum Vitae

Candidate's full name: Jun Guo

Universities attended: University of New Brunswick, BSc (First Class Hons.) Biology-Chemistry, 2018-2022

Publications:

Kim, K., Shahsavarani, M., Garza-Garćia, J. J. O., Carlisle, J.E., **Guo, J.**, De Luca, V., Qu, Y. (2022). *Biosynthesis of Kratom Opioids*. bioRxiv https://doi.org/10.1101/2022.12.25.521902