UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

Caractérisation des enzymes de la voie des phénylpropanoïdes chez la plante *Leucojum aestivum* produisant des alcaloïdes d'Amaryllidaceae et reconstruction de la voie centrale des alcaloïdes d'Amaryllidaceae dans *Saccharomyces cerevisiae*.

Characterization of phenylpropanoid pathway enzymes from Amaryllidaceae alkaloid-producing plant *Leucojum aestivum* and reconstruction of the Amaryllidaceae alkaloids core pathway in *Saccharomyces cerevisiae*.

THÈSE PRÉSENTÉE

COMME EXIGENCE PARTIELLE

DU DOCTORAT EN BIOLOGIE CELLULAIRE ET MOLÉCULAIRE

PAR

Vahid Karimzadegan

Novembre 2024

CONTENTS

ACKNOWLEDGEMENT ..... **ix**

RÉSUMÉ EN FRANÇAIS ..... **x**

Abstract ..... **xiii**

Chapter I **1**

1. Introduction **1**

1.1 Plant Specialized Metabolites 1

1.2 Biogenesis and Evolution of SMs Pathways 2

1.3 Classification of Plant SMs 3

1.3.1 Phenolics 3

1.3.2 Terpenoids 4

1.3.3 Alkaloids 6

1.4 Amaryllidaceae Alkaloids (AAs) 7

1.4.1 Biosynthesis of AAs 7

1.4.2 Phenylpropanoid pathway, contribution to AAs pathway 11

1.4.3 *Leucojum aestivum*, a source for AAs 22

1.5 Synthetic Biology, a Promising Sustainable Alternative 23

1.5.1 Multiplex CRISPER-Cas9 for rapid metabolic engineering of plant SMs 24

1.6 Problematic, Hypothesis, Objectives 26

1.6.1 Objective 1: Characterization of candidate hydroxylases in the phenylpropanoid pathway of AAs 27

1.6.2 Objective 2: establishment of AAs core pathway in baker's yeast 27

Chapter II **29**

2. Characterization of cinnamate 4-hydroxylase (CYP73A) and *p*-coumaroyl 3′-hydroxylase (CYP98A) from *Leucojum aestivum*, a source of Amaryllidaceae alkaloids **29**

Keywords 30

2.1 Abstract 30

2.2 Introduction 31
2.3 Materials and Methods 33
2.4 Results 40
2.4.1 Identification and sequence analysis of *LaeC4H* and *LaeC3'H* from *L. aestivum* 40
2.4.2 Prediction of substrate interaction for *LaeC4H* and *LaeC3'H* 43
2.4.3 Recombinant protein expression in yeast and bacterial host and in vitro enzymatic assays 45
2.4.4 Subcellular localization of *Lae*C4H and *Lae*C3′H in *N. benthamiana* 48
2.4.5 Metabolite profiling of *L. aestivum* plant organs 50
2.4.6 Differential Expression Analysis 51
2.5 Discussion 51
2.6 Conclusion 54
2.7 Credit Author Statement 55
2.8 Acknowledgments 55
2.9 Declaration of Interests 56
Chapter III 57
3. Establishment of the Amaryllidaceae alkaloid core pathway in yeast *saccharomyces cerevisiae* using CRISPR-Cas9 system 58
3.1 Abstract 58
3.2 Introduction 60
3.3 Materials and Methods 63
3.3.1 DNA construction and yeast stable transformation using the Multiplex CRISPR-Cas9 platform. 63
3.3.2 Plasmid-based platform construction 63
3.3.3 Yeast feeding assay and metabolite analysis 64
3.3.4 Metabolite analysis 64
3.4 Results and Discussion 65
Chapter IV 69
4. Conclusion 70
4.1 Final conclusion and future perspective 74
Appendix I 76
Supplementary Data of Chapter II 76

Appendix II **89**

Supplementary Data of Chapter III 89

Appendix III, Review Paper (Co-authorship) **92**

Biotechnological approaches to optimize the production of Amaryllidaceae alkaloids 92

Appendix IV; Research Paper (Co-authorship) **110**

Chemical synthesis and biological activities of Amaryllidaceae alkaloid norbelladine derivatives and precursors ...

Appendix V; Supplementary Data of A IV **132**

References **135**

List of the figures

Figure 1-1. Interconnection of plant specialized metabolism to the primary metabolism .....

Figure 1-2. The MVA and MEP pathways in plants…….

Figure 1-3. Proposed biosynthetic pathway of Amaryllidaceae alkaloids. .....

Figure 1-4. Simplified phenylpropanoid pathway and diversification to different branches of phenolics .....

Figure 1-5. The biosynthesis of lignin units.....

Figure 1-6. The schematic of the hydroxylation mechanism of plant CYP450 monooxygenases. .....

Figure 1-7. Distinct substrate preference of CYP98A P450s from different plant species .....

Figure 1-8. Proposed pathways for benzaldehyde biosynthesis in plants .....

Figure 2-1. Proposed pathway leading to 4′-O-methylnorbelladine, a common intermediate of AAs.....

Figure 2-2. The phylogenetic tree of L. aestivum enzymes, *Lae*C4H and *Lae*C3′H.....

Figure 2-3. Representation of *Lae*C4H and *Lae*C3′H structures.....

Figure 2-4. HPLC-MS/MS analysis of representative enzymatic assay for *Lae*C4H and *Lae*C3′H and their subcellular localization in *N. benthamiana* .....

Figure 2-5. *In vivo* enzymatic activity of *Lae*C4H and *Lae*C3′H.....

Figure 2-6. Metabolic profile and expression analyses.....

Figure 3-1. Schematic of Modular multiplex genome-edit (MMG)-CRISPR platform in yeast.....

Figure 3-2. Yeast engineering for 4′-O-methylnorbelladine production .....

Figure 3-4. Genotyping of the yeasts co-transformed with pESC-Ura:: *NBS*: NR & pESC-Leu:: *TYDC*: *OMT*. and yeasts co-transformed with pESC-Ura:: *NBS* & pESC-Leu:: *TYDC*: *OMT* (B). .....

Figure 3-5. LC-MS analysis of metabolites extracted from yeasts expressing AAs core pathway. .....

Figure 3-6. LC-MS analysis of metabolites extracted from SBY104 expressing AAs core pathway……...

List of the tables

Table 1-1. Phenylpropanoids, their subgroups, and functions in plants.....

Table 2-1. Kinetic parameters for *Lae*C4H and *Lae*C3′H.....

Table A1-1. List of primers used in Chapter Ⅱ .....

Table A1-2. Sequencing result of isolated C4H, C3′H, and APX/C3H from *L. aestivum*.....

Table A1-3. Sequence and structure comparison scores .....

Table A1-4. *Lae*C4H and *Lae*C3′H active site description and predicted interactions with docked substrates .....

Table A1-5. Summary of in vivo enzymatic assay *Lae*C4H and *Lae*C3′H .....

Table A2-1. List of primers used in ChapterⅢ .....

Table A3-1. Yields of uncommon AA of therapeutical interest in in vitro cultures. .....

Table A3-2. *In vitro* production of Amaryllidaceae alkaloids cultures following elicitor treatment.....

Table A4-1. EC50, CC50, and SI of norbelladine precursors and derivatives with antiviral effect.....

Table A4-2. Prediction of norbelladine derivatives interactions with butyrylcholinesterase. .....

Table A5-1. Inhibitory properties (IC50) of norbelladine precursors and derivatives towards enzymes implicated in Alzheimer's disease. .....

Avertissement

UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

BIOLOGIE CELLULAIRE ET MOLÉCULAIRE (DOCTORAT)

Direction de recherche :

Isabel Desgagné-Penix, PhD	Université du Québec à Trois-Rivières
Prénom et nom	Directeur de recherche

Co-Directeur de recherche :

Mehran Dastmalchi, PhD	McGill University
Prénom et nom	Directeur de recherche

Jury d'évaluation

Isabel Desgagné-Penix, PhD	Directeur de recherche
Prénom et nom	Fonction du membre de jury
Mehran Dastmalchi, PhD	Co-Directeur de recherche
Prénom et nom	Fonction du membre de jury
Hugo Germain, PhD	Président du jury
Prénom et nom	Fonction du membre de jury
Yang Qu, PhD	Évaluateur externe
Prénom et nom	Fonction du membre de jury
Dae-Kyun Ro, PhD	Évaluateur externe
Prénom et nom	Fonction du membre de jury

AKNOWLEDGEMENT

Doing and writing this PhD thesis has been a journey filled with challenges and triumphs, and I am deeply grateful to everyone who has supported and encouraged me along the way.

First and foremost, I want to extend my heartfelt thanks to my supervisor, Prof. Isabel Desgagné-Penix, at the Université du Québec à Trois-Rivières. Your unwavering support, insightful guidance, and constant encouragement have been invaluable. Your belief in my abilities kept me motivated even when things were tough.

I am also deeply thankful to my co-supervisor, Professor Mehran Dastmalchi, at McGill University. I appreciate your technical advice on the project's challenging steps and your significant guidance in improving my paper.

A big thank you to my thesis committee members, Prof. Yang Qu at the University of New Brunswick, Prof. Hugo Germain, and Prof. Prof. Dae-Kyun Ro at the University of Calgary, for your time dedicated to providing constructive suggestions and evaluations of my thesis.

To my friends and colleagues in the Department of Chemistry, Biochemistry, and Physics at the Université du Québec à Trois-Rivières and the Department of Biological Sciences at the University of Calgary, thank you for your support and the many stimulating discussions. You've made this journey not just bearable, but enjoyable. The laughs, shared frustrations, and collaborations have meant the world to me.

I am grateful to Prof. Dae-Kyun Ro at the University of Calgary for supporting my internship project. Working in your lab significantly advanced my knowledge of synthetic biology and opened a new avenue for the next chapter of my life.

I am also thankful for the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada Research Chair on Plant Specialized Metabolism, Canadian taxpayers, and the Canadian government. Your support made this research possible.

On a personal note, I want to express my deepest gratitude to my family. To my wife, Zahra Gozalzadeh, your endless support and belief in me have been my rock. Your encouragement and love have been crucial in helping me get to this point. Thank you to my parents and sisters for your patience, understanding, and love. Your support has been my anchor through this journey.

Lastly, I dedicate this thesis to my wife, Zahra Gozalzadeh, for being a strong, supportive, and wonderful companion in my life. Your patience and understanding over the past ten years made the challenging times bearable.

Thank you all for being part of this journey. I couldn't have done it without you.

RÉSUMÉ EN FRANÇAIS

Les thérapeutiques à base de plantes sont de plus en plus intégrées dans les traitements médicaux modernes, avec de nombreux médicaments anti-cancéreux actuels ayant des origines végétales. Des exemples notables incluent le Taxol, la vincristine et la vinblastine, pour n'en nommer que quelques-uns. De plus, de nombreux autres thérapeutiques dérivés des plantes présentent une large gamme d'activités biologiques. La famille des Amaryllidaceae synthétise un type unique d'alcaloïdes structurellement diversifiés avec diverses activités biologiques, notamment anti-Alzheimer, antivirales et anti-cancéreuses. Plus de 650 alcaloïdes structurellement différents ont été identifiés dans cette famille, et leurs applications médicinales ont été largement explorées, fournissant des informations précieuses sur leurs utilisations potentielles. Avec la galantamine étant un représentant approuvé par la FDA des alcaloïdes d'Amaryllidaceae (AAs), comprendre leurs voies biosynthétiques est un domaine crucial de la recherche actuelle en métabolisme spécialisé des plantes pour établir une plateforme plus durable afin de fournir des AAs.

Malgré leurs structures chimiques diverses, tous les AAs partagent un intermédiaire biosynthétique commun, la norbelladine. On pense que deux molécules précurseurs, la tyramine et le 3,4-dihydroxybenzaldéhyde (DHBA), subissent des réactions de condensation et de réduction pour produire la norbelladine. On suppose que le DHBA est dérivé de la phénylalanine par la voie des phénylpropanoïdes. La biosynthèse du DHBA nécessite deux étapes d'hydroxylation catalysées par les enzymes de la famille des cytochrome P450 monooxygénases dont CYP73A et CYP98A. Le CYP73A convertit l'acide *trans*-cinnamique en acide *p*-coumarique, tandis que le CYP98A hydroxyle les esters shikimate ou quinate du coumaroyl pour donner du shikimate ou du quinate de caféoyl. Le CYP98A peut potentiellement accepter l'acide *p*-coumarique comme substrat pour produire de l'acide caféique, bien que pour cette activité, il y a très peu d'études. L'acide caféique peut alors être converti en DHBA, bien que l'enzyme spécifique responsable de cette réaction soit encore inconnue. Alternativement, le DHBA pourrait être dérivé du 4-hydroxybenzaldéhyde (4-HBA), un dérivé proposé de l'acide *p*-coumarique. Des rapports récents suggèrent qu'une peroxydase bifonctionnelle de l'ascorbate peut catalyser l'hydroxylation en 3' de l'acide *p*-coumarique en acide caféique. Cependant, sa capacité à hydroxyler le 4-HBA pour produire du DHBA n'a pas été investiguée.

Nous avons utilisé une analyse bioinformatique pour identifier les enzymes candidates pour CYP73A, CYP98A et la peroxydase de l'ascorbate (APX) dans la plante Amaryllidaceae,

Leucojum aestivum. Des approches biochimiques ont révélé que CYP73A et CYP98A effectuent spécifiquement des hydroxylations en 4' et 3' sur l'acide *trans*-cinnamique et le *p*-coumaroyl shikimate, respectivement. Cependant, le CYP98A de *L. aestivum* n'a pas accepté l'acide *p*-coumarique ni le 4-HBA pour l'activité d'hydroxylation en 3'. L'activité *in vivo* de ces enzymes a été explorée en utilisant l'expression hétérologue dans *Nicotiana benthamiana*, démontrant que leur surexpression peut augmenter les niveaux des intermédiaires phénylpropanoïdes. Les essais enzymatiques *in vitro* de l'homologue d'APX de *L. aestivum* n'ont montré aucune activité d'hydroxylation sur l'acide *p*-coumarique ni le 4-HBA, ce qui soulève une controverse sur les études précédentes.

Lorsqu'ils sont exprimés dans *N. benthamiana*, le CYP73A et le CYP98A ont exhibé une localisation à la membrane du réticulum endoplasmique. Leurs profils d'expression génique n'ont pas montré de corrélation positive avec la teneur en alcaloïdes des Amaryllidaceae dans les tissus végétaux. Au lieu de cela, la teneur en alcaloïdes était plus positivement corrélée avec l'expression des enzymes de la voie en aval, y compris la *norbelladine synthase* (*NBS*), la *noroxomaritidine/norcraugsodine réductase* (*NR*) et la *norbelladine 4'-O-methyltransferase* (*N4'OMT*). Cette étude fournit une base pour une exploration plus approfondie des voies biosynthétiques des phénylpropanoïdes des Amaryllidaceae.

Dans la deuxième partie de cette thèse, la construction de la voie principale des AAs a été tentée en utilisant une plateforme CRISPR-Cas9 multiplex et un système d'expression plasmidique. La tyrosine décarboxylase (TYDC), NBS, NR et N4'OMT ont été intégrées de manière stable dans *S. cerevisiae* à 4 loci différents. Dans la plateforme d'expression plasmidique, ces gènes ont été exprimés à partir des plasmides pESC-LEU et pESC-URA. L'objectif était idéalement d'évaluer la production de la 4'-O-méthylnorbelladine en tant qu'intermédiaire central de la voie des AAs dans des levures supplémentées en DHBA. Les extraits métaboliques des levures ont été analysés par LC-MS après leur alimentation avec les substrats correspondants. Les souches de levures génétiquement modifiées ont pu produire de la norbelladine à une concentration de 0,2 mM après supplémentation avec de la tyramine et du DHBA. La tyramine produite de manière endogène par la TYDC surexprimée était probablement en dessous du seuil suffisant pour permettre la condensation. Bien que la 4'-O-méthylnorbelladine ait été détectable à des concentrations plus élevées de substrats, un signal de fond dans le contrôle négatif a été observé, ce qui reste à clarifier. Globalement, le test d'alimentation et l'analyse des métabolites nécessitent une optimisation

supplémentaire pour parvenir à une conclusion plus solide. Cette étude constitue une base pour une exploration plus approfondie des voies biosynthétiques des AAs et de leurs applications pharmaceutiques potentielles

Abstract

Plant-based therapeutics are increasingly integrated into modern medical treatments, with many current anti-cancer medicines having plant origins. Notable examples include Taxol, vincristine, and vinblastine, to name but a few. Additionally, numerous other plant-derived therapeutics exhibit a broad range of biological activities. The Amaryllidaceae family synthesizes a unique type of structurally diverse alkaloids with various biological activities, including anti-Alzheimer's, antiviral, and anti-cancer properties. Over 650 structurally different alkaloids have been identified in this family, and their medicinal applications have been extensively explored, providing valuable insights into their potential uses. With galanthamine being an FDA-approved representative of Amaryllidaceae alkaloids (AAs), understanding their biosynthetic pathways is a crucial area of current research in plant specialized metabolism to establish a more sustainable platform to supply AAs.

Despite their diverse chemical structures, all Amaryllidaceae alkaloids (AAs) share a common intermediate, norbelladine, in their biosynthesis. Two precursor molecules, tyramine and 3,4-dihydroxybenzaldehyde (DHBA), are believed to undergo condensation and reduction reactions to produce norbelladine. DHBA is hypothesized to be derived from phenylalanine through the phenylpropanoid pathway. This pathway converts the amino acid phenylalanine into various phenylpropanoids, including flavonoids, lignin, coumarins, *etc*. contributing to plants' structural integrity and environmental interaction. The biosynthesis of DHBA through the phenylpropanoid pathway requires two hydroxylation steps catalyzed by the Cytochrome P450 monooxygenase (CYP) enzymes CYP73A and CYP98A. CYP73A converts *trans*-cinnamic acid to *p*-coumaric acid, while CYP98A further hydroxylates shikimate or quinate esters of coumaroyl to yield caffeoyl shikimate or quinate. CYP98A can potentially accept *p*-coumaric acid as a substrate to produce caffeic acid, although there are limited reports of this activity. Caffeic acid may then be converted to DHBA, even though the specific enzyme responsible for this reaction is still unknown. Alternatively, DHBA could be derived from 4-hydroxybenzaldehyde (4-HBA), a proposed derivative of *p*-coumaric acid. Recent reports suggest that a bifunctional ascorbate peroxidase can catalyze the 3' hydroxylation of *p*-coumaric acid to caffeic acid. Still, its ability to hydroxylate 4-HBA to produce DHBA has not been investigated.

This study utilized bioinformatic analysis to identify candidate enzymes for CYP73A, CYP98A, and ascorbate peroxidase (APX) in the Amaryllidaceae plant *Leucojum aestivum*. Biochemical

approaches revealed that CYP73A and CYP98A specifically perform 4' and 3' hydroxylation on *trans*-cinnamic acid and *p*-coumaroyl shikimate, respectively. However, CYP98A from *L. aestivum* did not accept *p*-coumaric acid or 4-HBA for 3' hydroxylation activity. *In vivo* activity of these enzymes was explored using heterologous expression in *Nicotiana benthamiana*, demonstrating that overexpression of these enzymes can increase the levels of phenylpropanoid intermediates downstream of these enzymes. The *in vitro* enzymatic assays of the APX homolog from *L. aestivum* showed no hydroxylation activity on *p*-coumaric acid or 4-HBA, raising controversy for the previous report.

When expressed in *N. benthamiana*, CYP73A and CYP98A exhibited endoplasmic localization patterns with overlapping localization patterns to an endoplasmic fluorescent protein. Their gene expression patterns did not positively correlate with the Amaryllidaceae alkaloid content in plant tissues. Instead, the alkaloid content was more positively correlated with the expression of downstream pathway enzymes, including *norbelladine synthase* (*NBS*), *noroxomaritidine/norcraugsodine reductase* (*NR*), and norbelladine *4'-O-methyltransferase* (*N4'OMT*).

In the second part of this project, the construction of AAs core pathway was attempted using a Multiplex CRISPR-Cas9 platform and a plasmid expression system. Tyrosine decarboxylase (TYDC), NBS, NR, and N4′OMT were stably integrated in *S. cerevisiae* in 4 different loci. In the plasmid expression platform, these genes were expressed from pESC-LEU and pESC-URA plasmids. The goal ideally was to evaluate the production of 4'-*O*-methylnorbelladine as a central intermediate for the AAs pathway in yeasts supplemented with DHBA. Yeast metabolite extracts were analyzed by LC-MS after feeding them to the corresponding substrates. Genetically modified yeast strains could produce norbelladine at 0.2mM supplementation with tyramine and DHBA. Endogenously produced tyramine by the overexpressed TYDC probably was below the threshold sufficient for the condensation. Although 4'-*O*-methylnorbelladine was detectable at higher substrate concentrations, a background signal in negative control was observed that remains to be clarified. Overall, feeding assay and metabolite analysis need to be further optimized for a better conclusion. This study provides a foundation for further exploration of the biosynthetic pathways of AAs and their potential pharmaceutical applications.

Chapter I

1. Introduction

1.1 Plant specialized metabolites

Plants are sessile organisms, anchored in one place and cannot move. As a result, they have evolved various adaptations to respond to environmental changes and threats, such as producing specialized metabolites (SMs). Unlike primary metabolites, SMs or natural products fulfill diverse ecological functions rather than directly being involved in plant growth and reproduction. Plant SMs are multifaceted tools that facilitate environmental interactions, including defense against herbivores, communication with other organisms, and adaptation to various stresses. While primary metabolites are essential for basic cellular functions and are universally present in all organisms, SMs exhibit a remarkable degree of specialization tailored to specific ecological niches and environmental challenges [1, 11]. However, the differentiation between SMs and primary metabolites in function may not be easily distinguished. For instance, Phenylpropanoid-derived flavonoids, usually classified as SMs, play multiple roles involving the development of plant reproductive organs such as pollen tubes and seeds, seed dormancy, and maturation, essential for plant survival [12, 13]. These compounds have been reported to facilitate plant-microorganism communication to establish symbiosis in legume rhizobium during nodulation [13]. During water and oxidative stress, the upregulation of antioxidant phenylpropanoids prevents cellular damage by scavenging reactive oxygen species (ROS) [14]. A strong correlation between caffeic acid derivatives and flavonoid glycoside accumulation and salt and UV stress has been demonstrated, highlighting the role of SMs in abiotic stresses [15]. *Pseudomonas syringae* bacterial infection of tobacco (*Nicotiana tabacum* L.) was reported to significantly trigger the accumulation of phenylpropanoid derivatives, especially *p*-coumaric acid [16]. Phenylpropanoid derivatives were also shown to be involved in fungi resistance. For instance, a study showed a high feruloyl quinic acid content in resistant carrot genotypes (*Daucus carota* L.) compared to the susceptible genotypes when infected by *Alternaria dauci* [17]. In palm trees (*Phoenix dactylifera* L.), temporary upregulation of caffeoyl shikimic acid in certain physiological stages possibly contributes to resistance to *Fusarium oxysporum* [18]. The strong anti-fungal activity of

phenylpropanoid derivatives from the needles of *Picea neoveitchii*, as efficient as synthetic fungicide carbendazim, proposes a promising eco-friendly alternative [19]. As an anti-herbivory function, in chrysanthemums, chlorogenic acid and feruloylquinic acid are present in higher concentrations in thrips-resistant genotypes compared to susceptible ones, effectively differentiating their resistance to thrips [20].

In addition to their diverse role in native hosts, SMs possess a broad range of industrial applications, such as flavoring, cosmetics, dye, pharmaceuticals, nutritional supplements, natural preservatives, *etc*. [21]. It was reported that 40% of approved drugs between 1981 and 2014 were natural products [22] with plants as a leading source. Plant-based therapeutics are often expensive because of the challenges in chemical synthesis, low production rate in native hosts, purification from complex extracts, *etc*. [23]. A limited number of them have succeeded in commercial production using biotechnological approaches. Notable examples include Taxol, a potent anti-cancer agent; berberine, used for its anti-diabetic properties; shikonin, known for its anti-inflammatory benefits and applications in the cosmetic industry; and artemisinin, a key treatment for malaria [23].

The precursors of specialized metabolic pathways branch off from the basic metabolic processes found in all plants (Figure 1-1). Despite having structurally conserved precursors, SMs are unique to specific plant lineages. Their structural diversity in different plant lineages is a testament to the complexity and diversity of plant biology and contributes significantly to plant adaptation and survival in fluctuating habitats. [24, 25].

1.2 Biogenesis and evolution of SMs pathways

From a genetic perspective, despite the vast range of diversification and biological functions, specialized metabolite pathways utilize similar enzymes that catalyze oxidation, reduction, and conjugation processes, including glycosylation, methylation, and acylation. Through gene duplication, either by whole-genome duplication or tandem local duplication, and neofunctionalization of this common enzyme set throughout evolution, the diversification and species-specific repertoire of SMs likely emerged, which is considered divergent evolution of plant SMs [26-28]. As an example of divergent evolution, the specialized metabolic pathways responsible for glycoalkaloid production in tomatoes and potatoes belonging to the *Solanaceae* family have evolved divergently from a common triterpene pathway through the recruitment of

distinct enzymes. This exemplifies lineage-specific enzyme evolution, or divergent evolution, wherein unique biochemical pathways arise within related taxa [29].

In contrast to divergent evolution, convergent evolution is proposed as a possible explanation for the presence of structurally identical SMs in phylogenetically distant plant lineages [24]. For instance, sesamin, a specialized lignan that accumulates highly in dietary sesame seeds (*Sesamum* spp.), is observed not only in phylogenetically related species of Lamiales, such as *Paulownia* spp., but also in phylogenetically unrelated species of angiosperms and gymnosperms [30], highlighting the convergent evolution of this pathway in different lineages. Interestingly, CYP81Q1 from *Sesamum* spp. catalyzing two sequential methylenedioxy bridge formations is absent in phylogenetically distant sesamin-producing species. The other well-known example of SMs distributed in both related and distant species (*Theobroma cacao*, *Camellia sinensis*, *Citrus sinensis*) is caffeine from coffee trees belonging to purine alkaloids [31, 32]. Although the occurrence of specific SMs, like sesamin or caffeine, through convergent evolution is more reasonable due to their relatively simple biosynthetic pathways from core compounds like coniferyl alcohol, the presence of (deoxy)podophyllotoxin presents a more intricate case. This antitumor lignan is found in a variety of plant species across different phylogenetic lineages, suggesting a convergent evolution of a much more complex biosynthetic pathway [24].

1.3 Classification of plant SMs

SMs can be classified in various ways, including categorization based on their structural features, composition, the origin of their biosynthetic pathways, and solubility properties [1]. The major classification of SMs primarily relies on their chemical structure and biosynthetic pathways, which categorize them into three main classes: terpenoids, phenolics, and alkaloids belonging to the nitrogen-containing specialized metabolites (Figure 1-2).

1.3.1 Phenolics

As the most abundant SMs, phenolic compounds constitute approximately 40% of total organic carbon in plants [33] [34]. The structure of phenolic compounds consists of one or more phenolic rings containing one or more hydroxyl groups [34]. Phenolics are generally synthesized through the phenylpropanoid pathway and divided into subgroups, each with a distinct physiological function. There are some exceptions of phenolics that are synthesized directly through the shikimic

acid pathway (e.g., hydrolyzable tannins) [33]. The pathway and representative structures of phenolics and their physiological roles in plants are discussed in more detail in subsection 1.4.2.

1.3.2 Terpenoids

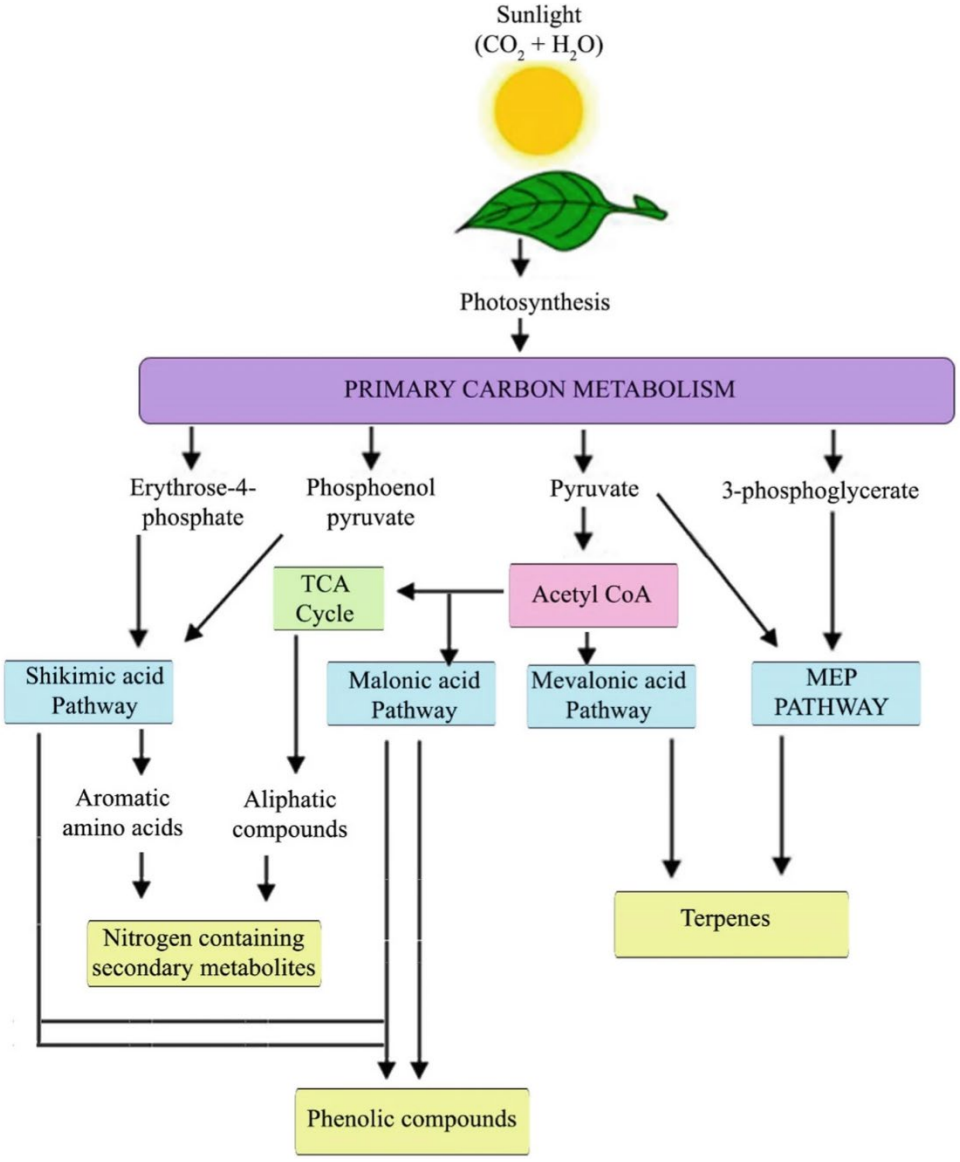


Fig 1-1. Interconnection of plant specialized metabolism to the primary metabolism. TCA, tricarboxylic acid; MEP, methyl erythritol phosphate. Source [1].

Terpenoids, the largest group (in terms of diversity) of SMs in the plant kingdom (in terms of diversity), encompass over 40,000 compounds with diverse physiological and ecological roles. They are crucial for plant growth and development, acting as essential components in various plant hormones (e.g., gibberellin phytohormones) and photosynthetic pigments (chlorophylls and

carotenoids) [35, 36]. This might contradict the early definition of SMs as non-essential components of plant growth and development components. Terpenoids are synthesized via the mevalonic acid (MVA) pathway in the cytosol and the methylerythritol phosphate (MEP) pathway in plastids, which both derive the product of primary metabolism into SMs. These pathways produce the 5-carbon isoprene units, isopentenyl diphosphate (IPP), and dimethylallyl diphosphate (DMAPP) (Figure 1-2). IPP and DMAPP combine to form geranyl diphosphate (GDP, C_{10}), farnesyl diphosphate (FDP, C_{15}), or geranylgeranyl diphosphate (GGDP, C_{20}), catalyzed by enzymes like geranyl diphosphate synthase (GDPS), geranylgeranyl diphosphate synthases (GGDPS), or farnesyl diphosphate synthase (FDPS). These diphosphates undergo rearrangement and cyclization, catalyzed by terpene synthases (TPS), to generate the basic carbon skeletons for terpenoids, including monoterpenes (C10), sesquiterpenes (C15), and diterpenes (C20), *etc*.

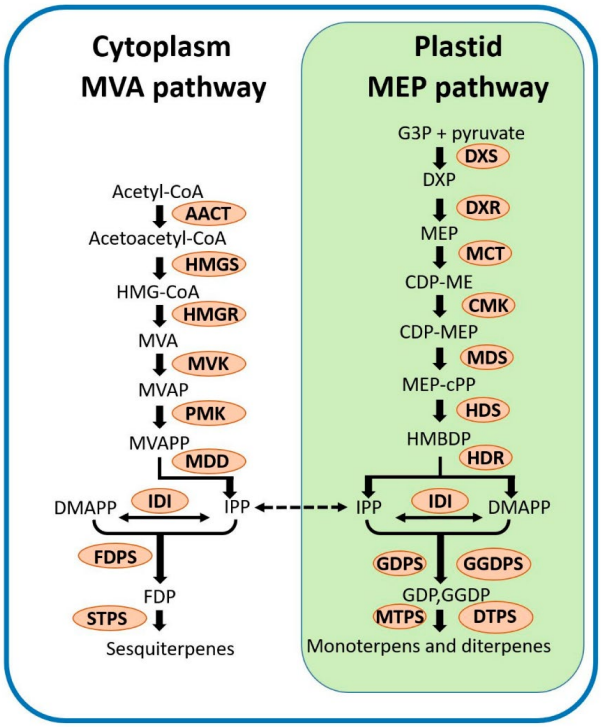


Fig 1-2. The MVA and MEP pathways in plants. AACT, acetoacetyl-CoA thiolase; CMK, 4-(cytidine 5′ -diphospho)-2-C-methyl-d-erythritol kinase; DMAPP, dimethylallyl diphosphate; DXR, 1-deoxy-d-xylulose 5-phosphate reductoisomerase; DXS, 1-deoxy- xylulose 5-phosphate synthase; FDP, farnesyl diphosphate; FPPS, farnesyl diphosphate synthase; G3P, d-glyceraldehyde 3-phosphate; GDPS, geranyl diphosphate synthase; GDP, geranyl diphosphate; HDR, (E)-4-hydroxy-3-methylbut-2-enyl diphosphate reductase; HDS, (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; HMGS, 3-hydroxy-3-methylglutaryl- CoA synthase; IDI, isopentenyl diphosphate isomerase; IPP, isopentenyl diphosphate; MCT, 2-C-methyl-d-erythritol 4-phosphate cytidylyltransferase; MDD, mevalonate diphosphate decarboxylase; MDS, 2-C-methyld-erythritol 2,4-cyclodiphosphate synthase; MVK, mevalonate kinase; MVAP, mevalonate 5-phosphate; MVAPP, mevalonate diphosphate; PMK, phosphomevalonate kinase; TPS, terpene synthase. Source [9].

Further modifications such as oxidation, reduction, and conjugation lead to the diverse structures and functions of terpenoids [35, 37].

1.3.3 Alkaloids

Alkaloids are naturally occurring SMs distinguished by the presence of nitrogen in their chemical structures. Alkaloids, which constitute about 20% of plant-based SMs, are characterized by their low-molecular-weight structures. Around 12,000 alkaloids have been isolated from various genera across the plant kingdom [38]. Alkaloids, mainly occurring as solids, are commonly found in higher plants. These compounds are notably abundant in a variety of botanical families, including Apocynaceae, Annonaceae, Amaryllidaceae, Berberidaceae, Boraginaceae, Gnetaceae, Liliaceae, Leguminosae, Lauraceae, Loganiaceae, Magnoliaceae, Menispermaceae, Papaveraceae, Piperaceae, Rutaceae, Rubiaceae, Ranunculaceae, Solanaceae, and many others [39]. Alkaloids are synthesized in various plant tissues, including leaves, roots, bark, and seeds, and are subsequently transported to specific tissues where they serve a defensive role. These compounds protect plants by deterring and combating abiotic and biotic stresses such as pests, fungi, bacteria, and insect larvae. Additionally, alkaloids exhibit allelopathic properties, inhibiting the growth of nearby competing plants. Through these mechanisms, alkaloids play a crucial role in plant survival and defense, contributing to the plant's ability to thrive in its environment [40, 41].

The diverse biological activities of alkaloids originate from the unique arrangements of atoms within their molecular frameworks. The structural variations in alkaloids lead to various pharmacological effects, making alkaloids valuable for medicinal and therapeutic applications. The intricate molecular architecture of alkaloids allows them to interact with multiple biological targets, resulting in properties such as analgesic, antimalarial, anticancer, and antibacterial activities [42]. The study of alkaloid structures and their bioactive potential continues to be a significant area of research in natural product chemistry and drug discovery.

1.3.3.1 Classification of alkaloids

Alkaloids are classified into three categories based on their biosynthetic origins: true alkaloids, protoalkaloids, and pseudoalkaloids. True alkaloids and protoalkaloids originate from amino acids such as tyrosine, tryptophan, lysine, and L-aspartate, but they differ in their chemical structures. True alkaloids incorporate a nitrogen atom within their heterocyclic ring (e.g., nicotine), while

proto alkaloids do not (e.g., colchicine). In contrast, pseudoalkaloids have a basic carbon skeleton not derived from amino acids [43]. For example, caffeine is a xanthine alkaloid originating from a nucleoside, adenosine [41].

In addition to their biosynthetic origin and chemical structure, alkaloids can be grouped based on their distribution in the same plant genera. For example, opium alkaloids generally refer to morphine, codeine, noscapine, thebaine, and papaverine alkaloids produced in *Papaver somniferum* L. [41]. Vinca alkaloids, including leurosine, vinblastine, and vincristine, are synthesized in *Catharanthus roseus*. Notably, vinblastine and vincristine are two prominent commercial anti-cancer chemotherapeutics [44].

1.4 Amaryllidaceae alkaloids (AAs)

AAs represent a diverse group of over 650 naturally occurring SMs found exclusively within the subfamily of Amaryllidoideae in the Amaryllidaceae plant family. Amaryllidaceae are characterized as perennial herbaceous species, predominantly thriving in tropical and subtropical regions worldwide [45]. Renowned for their ornamental beauty and medicinal properties, Amaryllidaceae plants have captured scientific interest due to their broad range of biological activities. These alkaloids exhibit significant therapeutic potential, making them subjects of active research in pharmacology and drug discovery. Their biological activities range from antitumor and antimicrobial effects to neurological and anti-viral properties, to name but a few [46, 47]. So far, galanthamine is the only FDA-approved AA that is prescribed for the treatment of mild symptoms of Alzheimer's disease due to its strong acetylcholinesterase inhibitory effect [48, 49]. Galanthamine is produced by many *Narcissus*, *Galanthus*, and *Leucojum* species, all belonging to the Amaryllidaceae family. AAs have been thoroughly explored from multiple angles, focusing on their biosynthetic origins, the discovery of associated enzymes, and their chemical structures [8, 45, 50]. The classification of AAs has changed frequently due to discoveries, group similarities, and varying scientific criteria, making a universal system challenging to establish (Figure 1-3) [50].

1.4.1 Biosynthesis of AAs

Amaryllidaceae alkaloids are isoquinoline alkaloids originating from the aromatic amino acid tyrosine [8, 50]. Despite structural diversity, the AA pathway shares a common precursor,

norbelladine (Figure 1-3). It has been proposed that DHBA and tyramine, derivatives of phenylalanine and tyrosine, undergo condensation and subsequent reduction to generate norbelladine [51, 52]. It was reported that norbelladine synthase (NBS) and norcraugsodine reductase (NR) could more effectively perform these two consecutive reactions together compared to individually [51]. Further protein-protein interaction assays confirmed the formation of a protein complex between these two enzymes, possibly allowing the intermediate to be channeled more effectively [51]. Norbelladine undergoes methylation to generate scaffold intermediate 4′-*O*-methylnorbelladine by the enzyme norbelladine 4′-*O*-methyltransferase (N4′OMT) [53]. Subsequently, 4′-O-methylnorbelladine can undergo different oxidative coupling reactions by CYP96T enzymes to initiate different branches of AAs [54, 55] (Figure 1-3).

In the Amaryllidaceae alkaloid pathway, tyramine and DHBA serve as precursor molecules, originating from the decarboxylation of tyrosine and the phenylpropanoid pathway, respectively (Figure 2-1). Tyrosine decarboxylase is a well-established enzyme that directly decarboxylates tyrosine to tyramine. This enzyme has been proposed to be a rate-limiting enzyme in the benzylisoquinoline alkaloids (BIA) biosynthetic pathway. Recombinant TYDC from opium poppy was able to decarboxylate L-dopa and L-tyrosine-tyrosine with comparative affinity to yield dopamine and tyramine, respectively [56]. Dopamine can be condensed with 4-hydroxyphenyl acetaldehyde (4-HPAA) to produce (*S*)-norcoclaurine, the first committed BIA molecule, which can be derivatized to several BIA such as berberine, sanguinarine, noscapine, papaverine, codeine, morphine, *etc.* [57]. A TYDC isoform from *A. thaliana* demonstrated a high level of specificity to

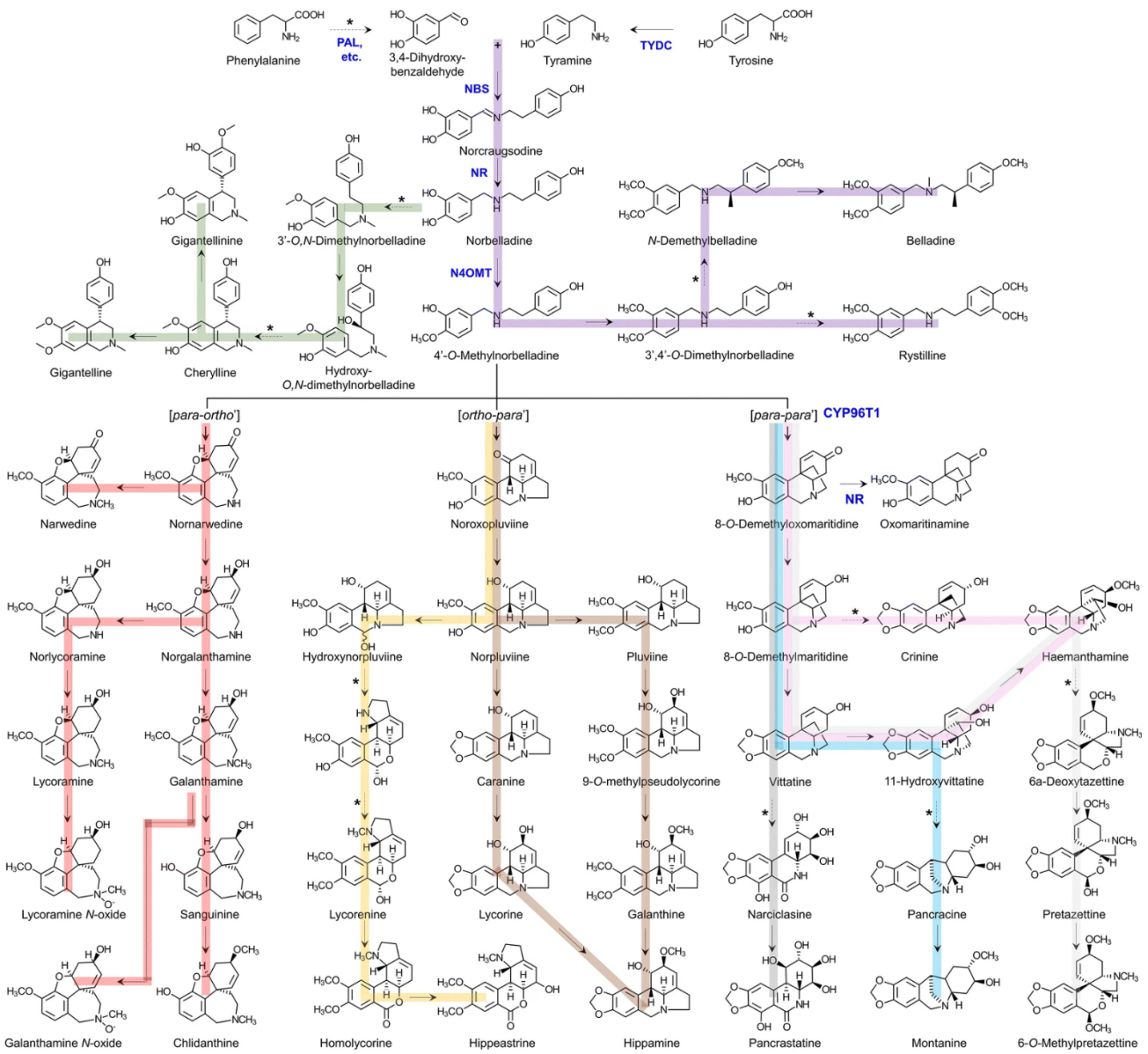


Fig 1-3. Proposed biosynthetic pathway of Amaryllidaceae alkaloids. Each class of AAs and corresponding branch are depicted in different colors: norbelladine (purple), cherylline (green), galanthamine (red), lycorenine (yellow), lycorine (brown), crinine/haemanthamie (pink), pancrastatin (dark gray), montanine (blue), and tazettine (light gray). Enzymes identified biochemically from Amaryllidaceae plants are shown in blue text. A solid arrow represents one enzymatic reaction, whereas a dashed arrow with an asterisk mark represents multiple enzymatic steps. Abbreviations are PAL, phenylalanine ammonia lyase; TYDC, tyrosine decarboxylase; NBS, norbelladine synthase; N4OMT, norbelladine 4'-O-methyltransferase; CYP96T1, cytochrome P450 monooxygenase 96T1; and NR, noroxomaritidine reductase. (Color figure online). Source [8].

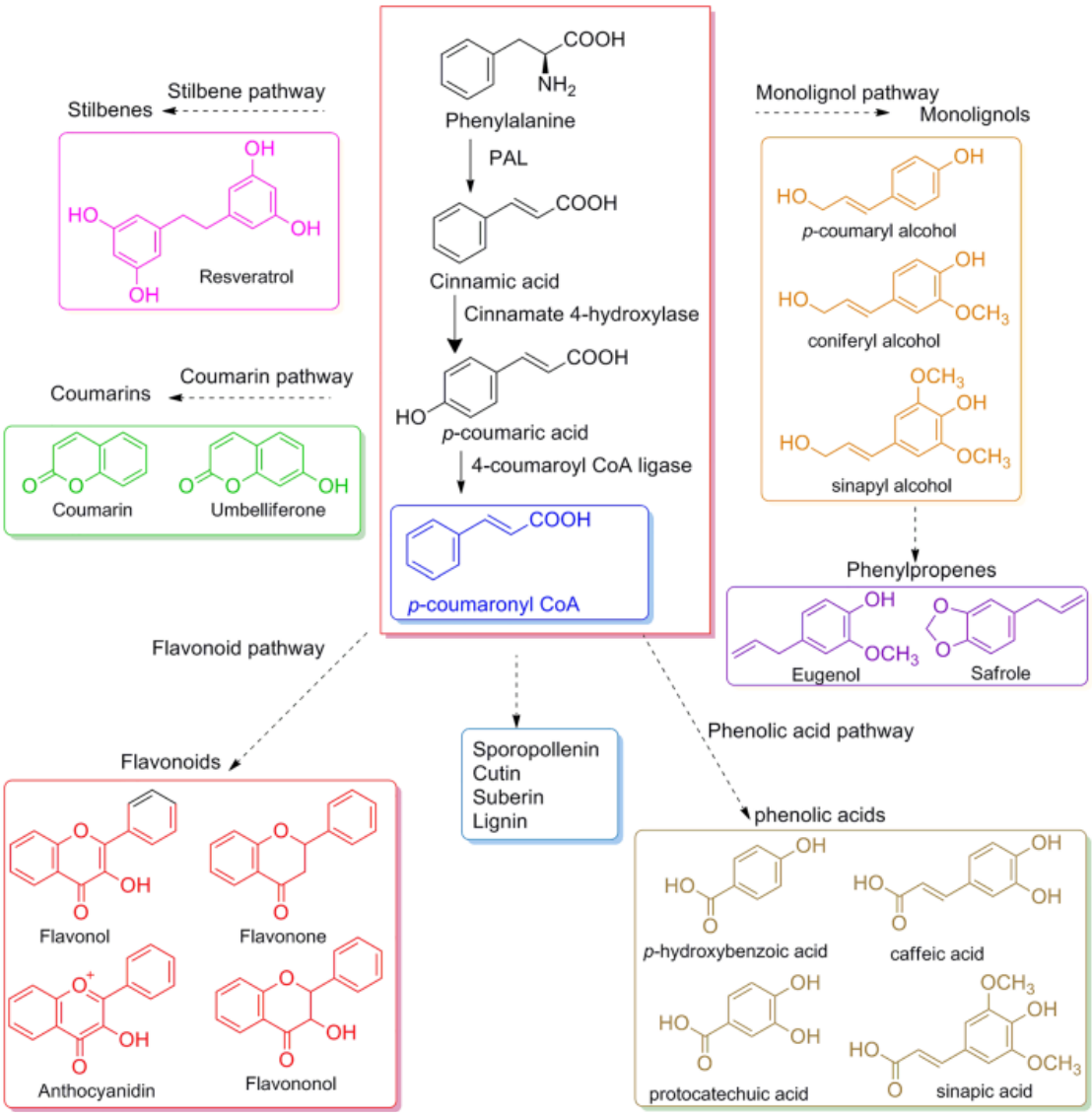


Figure 1-4. Simplified phenylpropanoid pathway and diversification to different branches of phenolics. Representatives of each subgroup are shown in the boxes. Source [2].

tyrosine rather than other candidate substrates. Expression of this gene was positively correlated with wounding, implying that tyramine might be exclusively involved in defense mechanisms in *A. thaliana* [58]. The TYDC form Amaryllidaceae plants has been reported from *Lycoris aurea* to catalyze the conversion of L-tyrosine more efficiently than L-dopa to corresponding products [59]. This enzyme was characterized from another Amaryllidaceae plant, *Lycoris radiata,* with the ability to catalyze the decarboxylation of L-tyrosine, and its transcriptional pattern was positively correlated with the galanthamine content of the plant tissues [60].

In contrast, the biosynthesis of DHBA, the other precursor of AAs via the phenylpropanoid pathway, is a multi-step process that has not been fully explored in Amaryllidaceae plants (Figure 2-1). DHBA indeed shares a common biosynthetic route with other major classes of SMs known as phenolics and phenylpropanoids. However, each of them may utilize a different intermediate as a branchpoint, leading to a distinct class of metabolites.

1.4.2 Phenylpropanoid pathway, contribution to AAs pathway

The phenylpropanoid pathway represents a crucial metabolic cascade involved in the biosynthesis of a wide variety of SMs utilizing the phenylalanine pathway or, in some cases, tyrosine as starting precursors. The pathway, dating back to early plant evolution, plays a fundamental role in carbon assimilation and the synthesis of diverse phenolic compounds essential for various facets of plant biology, including growth, defense, and adaptation to environmental stimuli (Figure 1-4) [2, 11, 61].

The phenylpropanoid pathway branches into multiple downstream pathways, producing a distinct class of phenolic compounds. Notable among these are flavonoids, lignin, coumarins, and phenolic acids (Figure 1-4). Flavonoids, known for their vivid colors, serve versatile roles such as UV protection, pigmentation, and attraction of pollinators [62]. Lignin, a complex polyphenolic polymer, gives structural integrity to plant cell walls, thereby contributing to mechanical support and water transport [63]. Coumarins, aromatic compounds widespread in various plant species, exhibit diverse pharmacological properties and are linked to defense mechanisms against biotic stresses [64]. Phenolic acids exhibit antioxidant and antimicrobial properties and play significant roles in plant defense responses against pathogens and herbivores.

Table1.1. Phenolics, their subgroups, and functions in plants. The information was mainly extracted from [2].

Phenolics	subgroups	Function & Application	Reference
Flavonoids	Flavones Flavanols Flavanonols Flavanones Aurones Auronols Anthocyanidins	plant growth & development, pigmentation, fertility, UV protection, scavenging reactive oxygen species (ROS)	[65], [66], [67], [68], [69]
Isoflavonoids	Isoflavones Isoflavones Isoflavones	chemoattractant, inducers of nitrogen-fixing root nodule formation	[70]
Stilbenes & stilbenoids	-	UV protection, defense against external stresses or infections, anti-aging properties, application in pharmaceutical, food, and cosmetic industries	[71], [72], [73]
Coumarins	Simple coumarins Furanocoumarins Pyranocoumarins Phenylcoumarins Dihydrofurocoumarins Biscoumarins	chemical defense, phytoalexins, iron mobilization, plant-microbe interactions, root-shoot communication, scavenging reactive oxygen species (ROS), application in the perfume industry	[2], [65], [74], [75], [76]
Monolignols	*p*- coniferyl alcohol sinapyl alcohol coumaroyl alcohol	building blocks of lignin and lignan, providing a mechanical barrier to pathogens, structural support, durability, longevity, and resistance of heartwood in tree species against pathogenic fungi (tree species), cytostatic, anti-inflammatory, antioxidant, antimicrobial, anti-obesity, and anti-cancerous effects	[2], [77], [65],
Phenolic acids	Hydroxycinnamic acid group Hydroxybenzoic acid group	plant growth and development, defense against pathogens, signaling molecules in plant-microbe interaction, antioxidant properties for humans, prevention of diabetes, cardiovascular diseases, and cancers	[78], [79], [80]
Phenylpropanoid polymers	Cutin suberin sporopollenin	Water Barrier (suberin, cutin), resistance to chemical and enzymatic degradation (sporopollenin)	[2], [81]

(Table 1-1). Furthermore, beyond their structural and defensive roles, these phenylpropanoid-derived metabolites serve as signaling molecules, mediating intercellular communication, and modulating physiological processes. Their involvement in signal transduction pathways underscores their importance in orchestrating adaptive responses to fluctuating environmental cues [2, 82].

The phenylpropanoid pathway in Amaryllidaceae plants is proposed to be linked to the AAs pathway via DHBA. The pathway is initiated by phenylalanine ammonia-lyase (PAL), or bifunctional phenylalanine/tyrosine ammonia-lyase (PTAL), and is interconnected with the shikimate pathway, channeling carbon into aromatic amino acids. PAL or PTAL activity is widespread in plants and some microorganisms but absent in animals [83]. The following sections discuss the candidate enzymes involved in the pathway potentially leading to DHBA.

1.4.2.1 PAL

PAL is a pivotal enzyme as the gateway enzyme in the general phenylpropanoid pathway. It is responsible for carbon flux from primary to specialized metabolism by catalyzing the conversion of the aromatic amino acid phenylalanine into *trans*-cinnamic acid by eliminating ammonia. PAL initiates the cascade of reactions that lead to synthesizing diverse phenolic compounds essential for various physiological functions in plants (Figure 1-5). This enzyme is part of a large superfamily of histidine ammonia-lyases (HALs) and PTALs found in bacteria. While PAL primarily acts on L-phenylalanine in dicotyledonous plants, PAL from certain monocots like maize can also catalyze the conversion of L-tyrosine to *p*-coumaric acid [84-86]. PAL arises from a multi-gene family in plants, resulting in various isoforms. These isoforms form tetrameric enzymes *in vivo*, with size and substrate affinity differences [61].

PAL expression is regulated transcriptionally in response to environmental and developmental cues. Different *PAL* genes exhibit unique spatial and temporal expression patterns, suggesting distinct functions in plant growth and ecological response. In *Populus* species, *PAL1* is predominantly expressed in non-lignifying tissues involved in condensed tannin synthesis, while *PAL2* isoform shows expression in both heavily lignified and non-lignified tissues, indicating its unique temporal and spatial expression pattern [87]. In *Arabidopsis thaliana*, *PAL1*, *PAL2*, and *PAL4* are closely linked to lignin biosynthesis, while only *PAL1* and *PAL2* are associated with

flavonoid biosynthesis [88]. PAL activity undergoes intricate regulation, spanning transcriptional and post-translational levels, encompassing processes such as ubiquitination, proteolysis, and

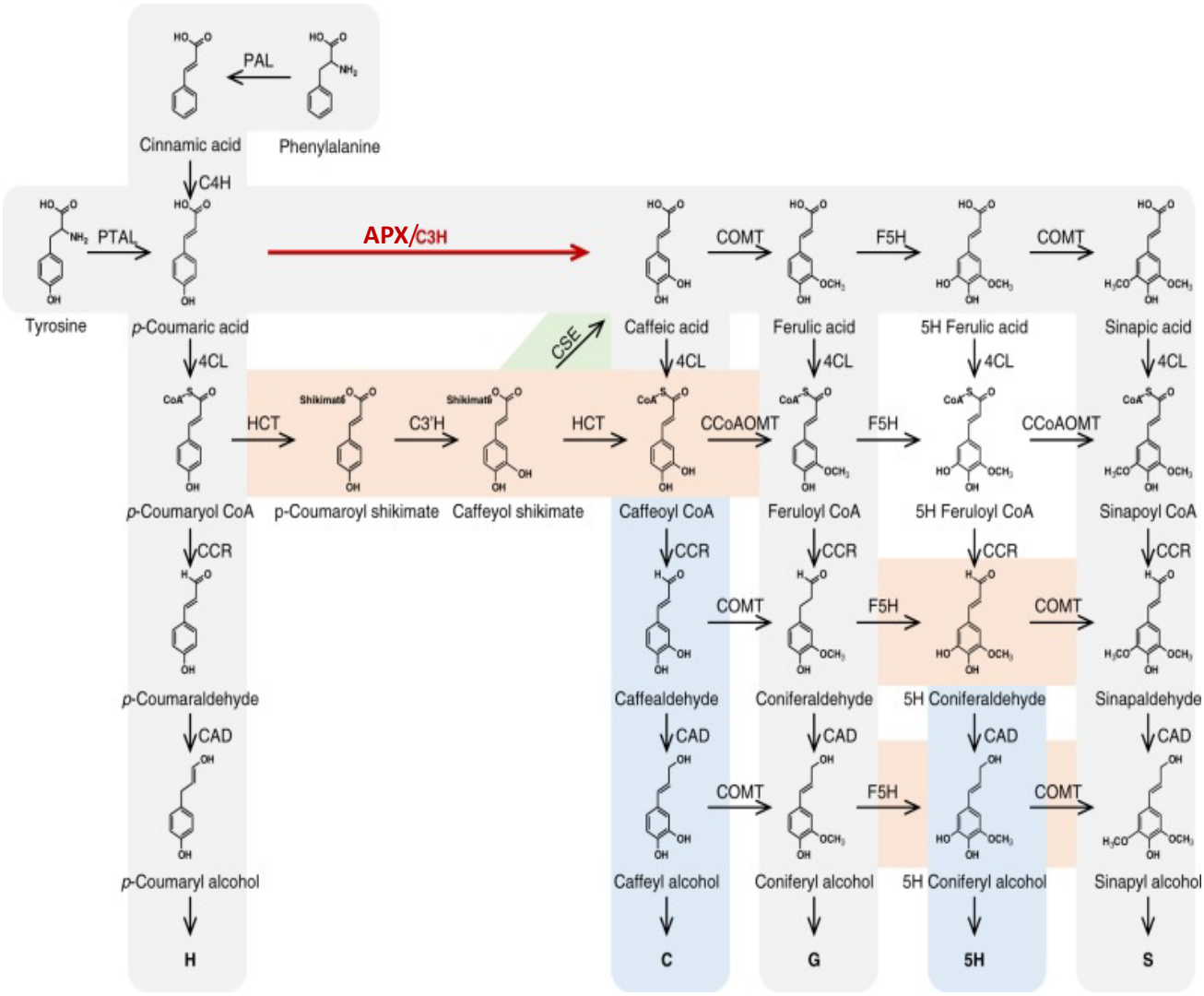


Figure 1-5. The biosynthesis of lignin units. The shaded areas represent major discoveries in the evolving understanding of lignin biosynthesis in plants. The gray shading illustrates the model of lignin biosynthesis as it was understood in the 1980s. The brown shading highlights the shikimate shunt and the substrate preferences of P450 enzymes, discoveries made in the 1990s. The blue shading indicates the identification of 5H- and C-lignin monomers that were subsequently discovered. The green shading marks the most recent steps identified in the lignin biosynthesis pathway. PAL, l-phenylalanine ammonia-lyase; PTAL bifunctional l-phenylalanine/l-tyrosine ammonia-lyase; C4H, cinnamate 4-hydroxylase; APX/C3H (in the original paper, it is labeled C3H), 4-coumarate 3-hydroxylase (bifunctional Ascorbate peroxidase/Coumarate 3-hydroxylase (APX/C3H) is proposed as a new candidate for 3' position hydroxylation of *p*-coumaric acid); COMT, caffeate/5-hydroxyferulate 3-O-methyltransferase; F5H, ferulate 5-hydroxylase/coniferaldehyde 5-hydroxylase; 4CL, 4-hydroxycinnamate:CoA ligase; HCT, 4-hydroxycinnamoyl CoA: shikimate/quinate hydroxycinnamoyl transferase; C3'H, 4-coumaroyl shikimate/quinate 3'-hydroxylase; CSE, caffeoyl shikimate esterase; CCoAOMT, caffeoyl CoA 3-O-methyltransferase; CCR, cinnamoyl CoA reductase, CAD, cinnamyl alcohol dehydrogenase. Source [10].

perhaps phosphorylation. Additionally, metabolic feedback mechanisms play a significant role in modulating its expression [61]. PAL's transcription and function are hindered by its product, *trans*-cinnamic acid [89], and other related intermediates like *p*-coumarate, caffeic acid [90], and flavonol aglycones [91]. Structural analyses indicate a tight binding of *trans*-cinnamic acid to the active site, leading to inhibition [92].

1.4.2.2 Cinnamate 4-hydroxylase (CYP73A, C4H)

Hydroxylation of the aromatic ring in cinnamates, catalyzed by cinnamate 4-hydroxylase (C4H), is a central process in the phenylpropanoid pathway (Figure 1-5), yielding a wide range of phenolic compounds crucial for plant structure, UV protection, antioxidant activity, antimicrobial defense, and flavor development [93]. C4H is a member of the extensive CYP450 enzyme family, which spans all life domains. Within the CYP450 enzymes, C4H is categorized under the CYP450 subfamily 73 (CYP73), which falls within the CYP71 clan. While C4H is traditionally considered exclusive to land plants, its product, *p*-coumarate, has been detected in diverse algae species [94]. This finding suggests the existence of an alternative pathway, which could involve either the direct hydroxylation of tyrosine by TAL enzymes or the activity of highly divergent C4H homologs. Notably, TALs have primarily been observed in monocots, indicating the potential involvement of a different CYP73 subfamily enzyme in algae [94, 95]. A phylogenetic analysis of C4H sequences from seven land plants, using a well-characterized *A. thaliana* C4H as a query, along with seven streptophyte algal sequences, revealed that all seven land plants grouped in a single clade, indicating their evolutionary relationship with C4H homologs from other plants. Interestingly, no algal sequences clustered with this group. Instead, four distinct clades of streptophyte algal CYP450 enzymes were identified. Some algal sequences showed closer genetic distances to the C4H-like clade than to sequences from land plants, suggesting potential candidates for para-hydroxylation activity that needs further investigation [25].

C4H belongs to the subfamily of class II monooxygenases within the CYP450 superfamily, which utilize molecular oxygen to hydroxylate substrates. This class of CYP450 enzymes uses NADPH as an electron donor, with CYP450 reductase, a flavoprotein, mediating the transfer of electrons between NADPH and the main CYP450 enzyme (Figure 1-6). These two enzymes are believed to function as a protein complex and colocalize on the exterior surface of the endoplasmic reticulum [96]. Most cytochrome P450 enzymes in eukaryotes, anchored to the endoplasmic

reticulum (ER) membrane, serve as nucleation sites for metabolon formation, increasing metabolic efficiency. For instance, PAL activity is associated with the ER membrane within the phenylpropanoid pathway, potentially interacting with cytochrome P450 enzyme C4H for metabolite channeling. Studies support this concept, showing co-localization of PAL and C4H, with PAL isoforms binding more strongly to C4H when co-overexpressed. These findings suggest a spatial organization of enzyme complexes influencing metabolic efficiency [97, 98]. In *Arabidopsis*, only one copy of the gene encodes C4H (*AtC4H*), expressed in all tissues and responds to various stimuli, indicating its diverse roles in phenylpropanoid metabolism. In contrast, other plants, like alfalfa, maize, orange, *etc*. have multiple *C4H* genes, resulting in multiple isoforms, each potentially playing distinct physiological roles in plants. However, the specific roles of individual C4H isoforms in SMs production are poorly understood [99].

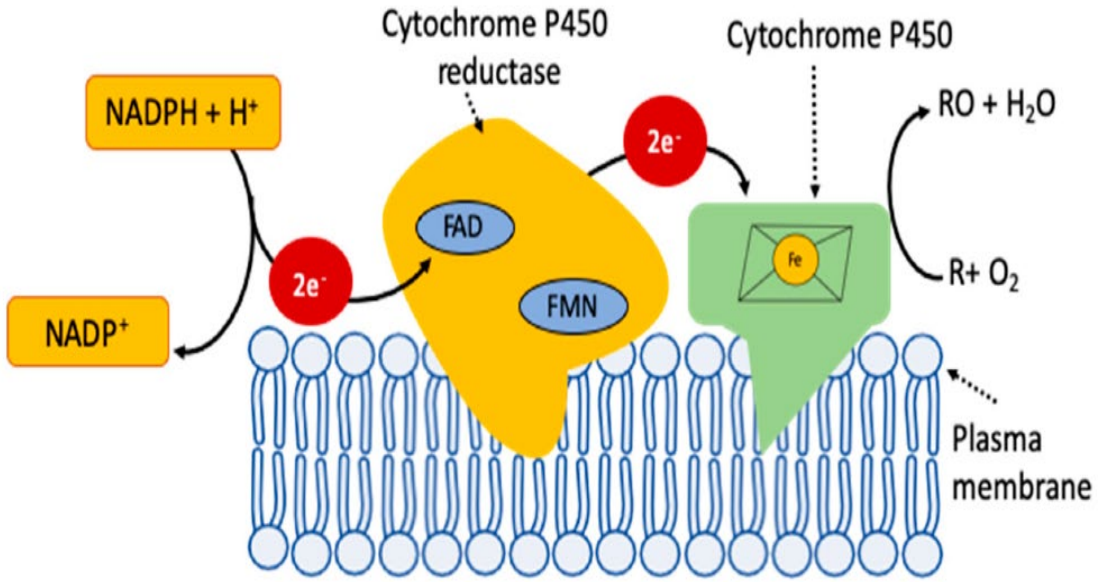


Figure 1-6. The schematic of the hydroxylation mechanism of plant CYP450 monooxygenases. Source [3].

1.4.2.3 4-coumarate Coenzyme A Ligase (4CL), main branch point of the phenylpropanoid pathway

p-coumaric acid, the product of C4H, is subsequently converted to p-coumaroyl-CoA by the 4-coumarate coenzyme A ligase (4CL) [100, 101]. *p*-coumaroyl-CoA is a key branch point metabolite that serves as a precursor for directing carbon flow toward the synthesis of a diverse array of metabolites, including (iso)flavonoids, monolignols, lignans, and sinapate esters [102]. 4CL often exists as multiple isoenzymes, each with distinct substrate affinities that align with specific metabolic functions (Figure 1-5) [101]. In *A. thaliana*, 4CL1 could activate *p*-coumaric acid and ferulic acid, but 4CL2 only accepts *p*-coumaric acid [103]. 4CL4 was later identified with a significantly higher affinity to sinapate [101].

1.4.2.4 Hydroxycinnamoyl CoA: shikimate/quinate Hydroxycinnamoyltransferase (HCT)

HCT catalyzes the transfer of hydroxycinnamoyl groups from *p*-coumaroyl-CoA to shikimate or quinate, forming hydroxycinnamoyl shikimate or quinate esters. This reaction is essential for producing intermediates, leading to the biosynthesis of monolignols, the building blocks of lignin. The activity of HCT helps regulate the flow of carbon through the phenylpropanoid pathway, influencing the composition and structure of lignin and other phenolic compounds (Figure 1-5) [82]. The dual function of this enzyme, transferring the caffeoyl from caffeoyl shikimate back to CoA moiety, generates caffeoyl-CoA that derives metabolic flux to other soluble phenolics and other subunits of lignin polymer [104]. Caffeoyl shikimate is generated by the second hydroxylation on *p*-coumaroyl shikimate catalyzed by the CYP98A enzyme, which is discussed in the following section.

1.4.2.5 The CYP98A enzyme family, 3′ hydroxylase of phenylpropanoids

The second hydroxylation in the phenylpropanoid pathway is catalyzed in the 3′ positions of coumaroyl esters by the CYP98A enzyme family. The substrate specificity of CYP98 enzymes varies across different plant families, a topic extensively explored by Alber et al. [4]. In angiosperms, the evolution of CYP98 enzymes involved a series of gene duplications and losses, resulting in multiple copies with diverse substrate preferences. Some CYP98 isoforms in angiosperms exhibit a narrow substrate preference, favoring *p*-coumaroyl-shikimate, an intermediate in lignin biosynthesis (Figure 1-5). This function directs upstream phenolic intermediates towards two monolignols: conifer alcohol, which is hydroxylated at the 4′ and 3′

positions of the phenolic ring, and sinapyl alcohol, which is hydroxylated at the 4′, 3′, and 5′ positions of the phenolic ring. These monolignols form guaiacyl (G) and syringyl (S) units in lignin (Figure 1-5). In angiosperms, G and S units are more abundant in the lignin structure [105].

Some CYP98A isoforms in angiosperms show distinct substrate acceptance, accommodating numerous *p*-coumaroyl-esters and -amines, producing a wide range of hydroxycinnamoyl conjugates with different functions. For instance, a CYP98 isoform in *A. thaliana* appears to accept *tri*-hydroxycinnamoyl-spermidines, playing a role in pollen coat biosynthesis rather than the lignin pathway [106]. In *Coleus blumei*, CYP98A14 is responsible for the 3′ hydroxylation in 4-coumaroyl-3′,4′-dihydroxyphenylactate and caffeoyl-4′-dihydroxyphenylactate, which serve as precursors for rosmarinic acid biosynthesis in this plant [107]. Similarly, CYP98A35 from the coffee tree can efficiently act on *p*-coumaroyl quinate, as on *p*-coumaroyl shikimate [108]. Two distinct isoforms of CYP98A from *Populus trichocarpa*, namely CYP98A23 and CYP98A27, exhibit contrasting substrate preferences. CYP98A27 demonstrates a more specific preference for *p*-coumaroyl shikimate/quinate, whereas CYP98A23 displays a broader range of substrate

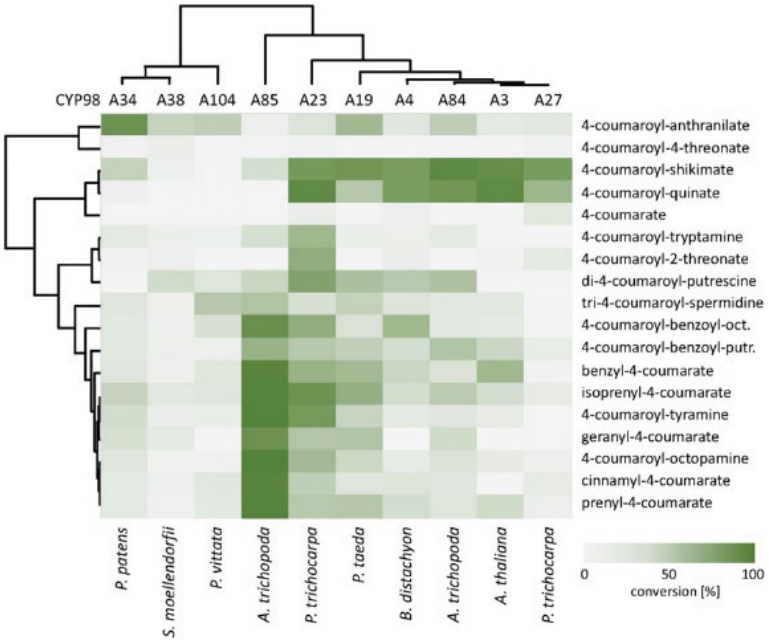


Figure 1-7. Distinct substrate preference of CYP98A P450s from different plant species. Source [4].

acceptance (Figure 1-7). CYP98A85 isoform of *Amborella trichopoda* showed a similar substrate range with *P. trichocarpa* CYP98A23, efficiently metabolizing 4-coumaroyltyramine, 4-coumaroyl-octopamine, and *tri*-coumaroylspermidine. Interestingly, CYP 98A85 preferred these substrates more than *p*-coumaroyl shikimate/quinate. In contrast, the CYP98A84 isoform of

Amborella trichopoda was more like CYP98A27 *P. trichocarpa* in terms of substrate range, with the highest preference for *p*-coumaroyl shikimate/quinate (Figure 1-7).

The pine (*Pinus taeda*, gymnosperm) enzyme CYP98A19 demonstrates versatility in substrate utilization. While it exhibits a preference for 4-coumaroyl-shikimate as its primary substrate, it also efficiently converts 4-coumaroyl-quinate and various other substrates, such as prenyl-4-coumarate, benzyl-4-coumarate, and 4-coumaroyl-anthranilate. Its substrate utilization profile aligns more closely with 4-coumaroyl-shikimate/quinate-specific angiosperm CYP98s than with angiosperm CYP98s with a broader substrate range (Figure 1-7) [109].

In the fern lineage, an early vascular plant, CYP98A104, showed the highest activity with 4-coumaroyl anthranilate and *tri*-4-coumaroyl-spermidine, with no activity observed for 4-coumaroyl shikimate/quinate [109]. CYP98A38 exhibited poor conversion rates for 4-coumaroyl anthranilate and no activity with 4-coumaroyl shikimate in lycopod plants, also an early vascular plant. In contrast, in moss plants (bryophytes), as representative of early non-vascular plants, CYP98A34 displayed the highest activity with 4-coumaroyl anthranilate and noticeable activity with 4-coumaroyl shikimate, isoprenyl 4-coumarate, and 4-coumaroyl tyramine (Figure 1-7) [4]. Overall, the substrate specificity of CYP98 enzymes varies across different plant families, reflecting the evolutionary diversification of phenylpropanoid metabolism in land plants. While CYP98A enzymes across various plant lineages exhibit promiscuity, a notable characteristic of this enzyme family is its distinct preference for *p*-coumaroyl shikimate and quinate esters in angiosperms, including both monocots and dicots, as well as gymnosperms [109]The hydroxylation activity of this enzyme on p-coumaroyl shikimate forms caffeoyl shikimate, which can subsequently be converted to caffeic acid by caffeoyl shikimate esterase (CSE) (Figure 1-5).

1.4.2.6 Caffeoyl shikimate esterase (CSE)

CSE in the phenylpropanoid pathway was identified first from *A. thaliana* through coexpression analysis and functional studies [110]. Genes coexpressed with known lignin biosynthetic pathway components were analyzed, and *CSE* was found to be consistently coexpressed with lignin pathway genes despite having no previously known role in lignin biosynthesis. CSE-reporter fusion proteins were created and detected in lignifying vascular tissues to validate their role, suggesting involvement in lignification. Two mutants of CSE (*cse-1* and *cse-2*) were studied, revealing reduced lignin content, altered lignin composition, and accumulation of the lignin

pathway intermediate caffeoyl shikimate, indicating that CSE converts caffeoyl shikimate into caffeic acid (Figure 1-5). Functional validation was further achieved as recombinant CSE hydrolyzed caffeoyl shikimate into caffeate, confirming CSE's critical role in the lignin biosynthetic pathway [110].

A chain-shortening enzymatic reaction can potentially yield DHBA directly from caffeic acid (Figure 2-1). While the enzyme responsible for DHBA synthesis remains uncharacterized in plant sources, *p*-hydroxycinnamoyl CoA hydratase/lyase (4-HCHL) from *Pseudomonas fluorescens*, and its close homolog in *Pseudomonas putida*, has been documented for catalyzing chain shortening of similar substrates, *p*-coumaroyl-CoA, caffeoyl-CoA, and feruloyl-CoA to yield respective aldehydes: 4-HBA, DHBA, and vanillin [111, 112].

1.4.2.7 Biosynthesis of 4-HBA from *p*-coumaric acid, a route to DHBA

The synthesis of 4-HBA, a precursor to SMs like vanillin (4-hydroxy-3-methoxybenzaldehyde) and AAs, involves multiple potential pathways, as evidenced by the diverse enzymatic activities in various plant systems. In *Vanilla planifolia* cell cultures, identifying 4-hydroxybenzaldehyde synthases (4-HBS) highlights one pathway [113]. This enzyme catalyzes the production of 4-HBA from *p*-coumaric acid in the presence of a thiol reagent without requiring a cofactor. The reaction proceeds non-oxidatively, as evidenced by the production of acetic acid and the absence of 4-hydroxybenzoic acid, a product of oxidative chain shortening. This process hydrates the 2,3 double bond of *p*-coumaric acid and cleaves the side chain, producing acetate and 4-HBA. Additionally, the specificity of the enzyme for *p*-coumaric acid suggests that this compound may be a major precursor in vanillin biosynthesis in *V. planifolia*. Other proposed pathways include converting *p*-coumaric acid to their Coenzyme A esters, chain-shortening analogous to fatty acid oxidation, or non-oxidative pathways directly converting *p*-coumaric acid to 4-HBA. These findings underscore the complexity of vanillin biosynthesis and suggest multiple routes by which plants may produce aromatic compounds like vanillin, contributing to the diversity of flavors and fragrances observed in nature [113].

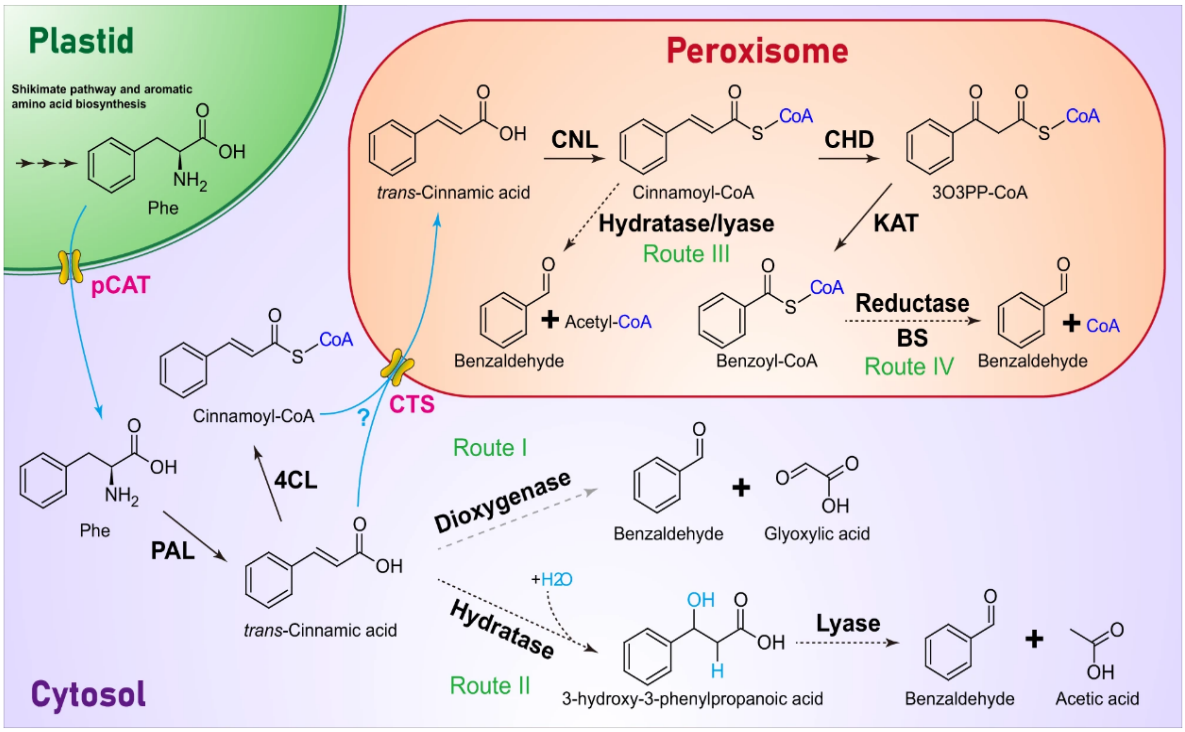


Figure 1-8. Proposed pathways for benzaldehyde biosynthesis in plants. Enzymes responsible for each biochemical reaction are shown in bold black. Solid black arrows show biochemical steps with their encoding genes discovered, dashed black arrows show steps established only at the enzyme level, and dashed gray arrows present putative steps. Transporters involved in metabolite transport between different subcellular locations are in bold magenta, with the flow of metabolites indicated by cyan arrows. CHD cinnamoyl-CoA hydratase/dehydrogenase, CNL cinnamate-CoA ligase, CTS peroxisomal cinnamic acid/cinnamoyl-CoA transporter COMATOSE. The substrate specificity of CTS is unclear, as indicated by a cyan question mark; KAT 3-ketoacyl thiolase, PAL phenylalanine ammonia lyase, pCAT plastidial cationic amino acid transporter, Phe phenylalanine, 3O3PP-CoA 3-oxo-3-phenylpropanoyl-CoA. Source [7].

For decades, the biosynthesis of benzaldehyde in plants has been a mystery. A recent study provides biochemical and genetic evidence showing that a heterodimeric enzyme consisting of two distinct subunits from the NAD(P)-binding Rossmann-fold superfamily, are responsible for plants' benzaldehyde formation. This enzyme converts benzoyl-CoA to benzaldehyde via the β-oxidative pathway using NADPH (Figure 1-8). It shows strict specificity for benzoyl-CoA and does not accept hydroxycinnamoyl-CoA thioesters. Benzaldehyde biosynthesis occurs in peroxisomes, where both the enzyme and its substrate, benzoyl-CoA, are found [7]. Hydroxylation of 4-HBA in the 3′ position can yield DHBA proposing a new route via phenylpropanoid pathway in Amaryllidaceae plants (Figure 2-1).

1.4.2.8 A cytosolic Ascorbate peroxidase, a novel candidate with 3′-hydroxylase activity in the phenylpropanoid pathway

It is well accepted that the three lignin polymer subunits, *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S), are derived from the phenylpropanoid pathway. Figure 1-5 illustrates the 3' position hydroxylation required for guaiacyl and syringyl subunits, which CYP98A facilitates through the ester route, involving enzymes such as CSE [10]. Notably, grasses lack CSE orthologues, and studies have shown that downregulating other enzymes in the ester route does not cause significant phenotypic changes [114]. This has led to speculation about an alternative 3' position hydroxylation pathway in these plants that could redirect flux to G and S units [10].

Barros' study reported hydroxylation activity of the bifunctional ascorbate peroxidase/Coumarate 3'-hydroxylase (APX/C3H) in plants such as *Brachypodium distachyon* and Arabidopsis (Figure 1-5) [10]. However, the orthologue enzyme from *Sorghum bicolor* did not exhibit significant coumarate hydroxylation activity compared to the negative control [115]. The phenotypic impairments observed in APX/C3H downregulated lines may not result from the lack of hydroxylation activity of this enzyme but could be an indirect effect, as APX has a broad range of physiological functions in plants [115]. Furthermore, The APX/C3H enzyme identified in *B. distachyon* in Barros' study was utilized for constructing the curcumin pathway in yeast, which requires the hydroxylation of *p*-coumaric acid to caffeic acid as an intermediate. The yeast expressing APX/C3H did not produce any detectable amounts of caffeic acid, unlike the yeast expressing CYP98A [6]. Overall, the reaction catalyzed by APX/C3H remains controversial and requires further investigation.

1.4.3 *Leucojum aestivum*, a source for AAs

Leucojum aestivum, or summer snowflake, is a Euro-Mediterranean plant belonging to the Amaryllidaceae family [116]. This plant is commercially grown for the extraction of galanthamine. Lycorine is another prominent alkaloid of this species with strong anti-viral [117] and antimitotic activities [118]. It was demonstrated that the quantitative and qualitative composition of alkaloids and phenolics of *L. aestivum* are influenced by different parameters such as genotype, temperature, light, soil type, growing season, harvesting time, *etc*. [119].

Despite significant advancements in the biochemical analysis of metabolites and their biological properties, pathway elucidation in this medically noteworthy species has been neglected. No enzymes from the upstream AA pathway leading to precursor molecules have been identified from

this plant. This study elucidates the crucial hydroxylation steps of the phenylpropanoid pathway in *L. aestivum* to explore routes toward DHBA (Figure 2-1). Although no unusual C4H activity was expected, CYP98A and the newly discovered 3' hydroxylase candidate APX/C3H suggest potentially different routes by acting on various substrates (Figure 2-1).

Understanding these pathways is crucial for future synthetic biology efforts to produce these valuable metabolites from simple carbon sources. This sustainable approach could reduce reliance on native plants and complex chemical synthesis, providing a stable supply of AAs for pharmaceutical and industrial applications.

1.5 Synthetic biology, a promising sustainable alternative

While the chemical synthesis of structurally complex alkaloids has been achieved, it often results in low final production yields due to their complexity. Synthetic biology offers a promising alternative to the challenges of the limited availability of natural products derived from native plants. This approach involves leveraging simple organisms like bacteria or yeast, which can be metabolically engineered to efficiently produce a wide array of desired natural compounds, offering a cost- and time-efficient method. Synthetic biology introduces genes for synthesizing specific metabolic pathways into microorganisms such as bacteria and yeast. This transformation effectively turns them into cell factories capable of producing plant-based therapeutics. The successful production of valuable compounds such as artemisinin, opioids, and vinblastine using microbial hosts is a testament to the impressive advancements in synthetic biology. For instance, artemisinin, a highly effective anti-malaria drug extracted primarily from *Artemisia annua* L., has been made via *de novo* semi-synthetic production of its precursor molecule, artemisinic acid, using yeast as a heterologous host system. This approach has led to more efficient and environmentally friendly production of this valuable compound [120]. Approximately one-third of the world's artemisinin supply now relies on microbial production. Additionally, overexpressing 21 and 23 enzymes from various sources in a yeast system successfully produced representative opioids, namely thebaine and hydrocodone, respectively [121]. More recently, the heterologous production of vinblastine and vincristine precursor molecules was accomplished by introducing 34 plant

enzymes, coupled with more than 22 genetic modifications in the yeast *Saccharomyces cerevisiae* genome [122].

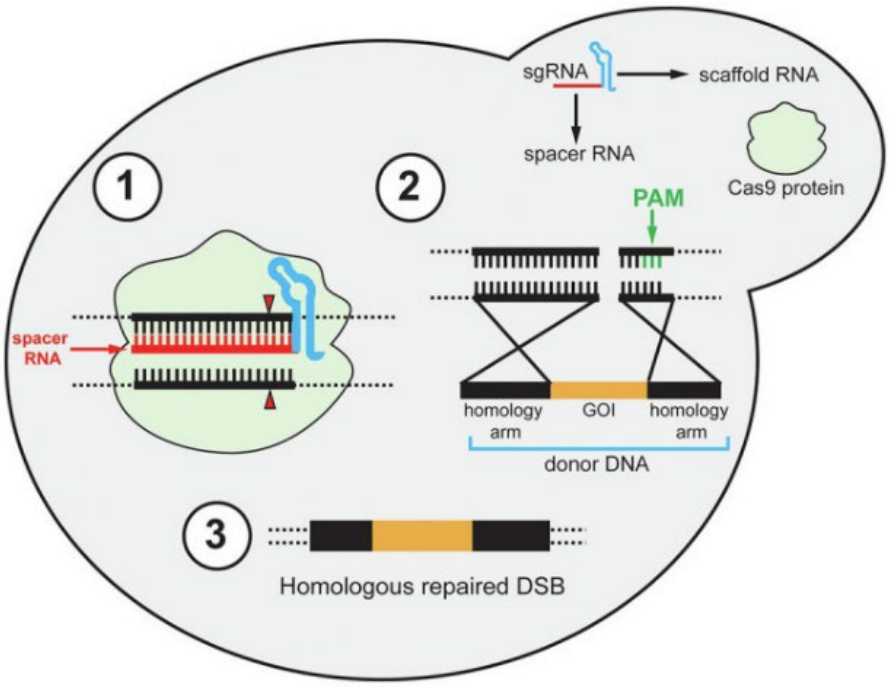


Figure 1-9. Schematic of CRISPR-Cas9 mediated gene integration in the yeast genome. A short guide RNA and Cas9 protein are two primary components of the system. Once complementation happens between gRNA and the target sequence in the genome, Cas 9 will make a DSB upstream of the PAM site. Yeast can repair DBS using donor DNA with 5' and 3' homology arms to the cut sites. Source [5].

1.5.1 Multiplex CRISPR-Cas9 for rapid metabolic engineering of plant SMs

Traditionally, modifying the yeast genome relies on the homology recombination approach, which necessitates selection markers and proves time-consuming, mainly when engineering multi-step pathways [5]. Subsequently, efforts were made to eliminate selection markers using alternative recombination methods like Cre-LoxP and Flp-FRT. However, a more time-efficient technique became imperative to enhance the metabolic engineering of intricate pathways [5].The Clustered Regularly Interspaced Palindromic Repeats (CRISPR) system, coupled with the Cas9 protein, is an efficient approach for rapid and precise genetic modification. Utilizing a single-strand guide RNA (gRNA), CRISPR-Cas9 creates double-strand breaks (DSBs) in specific regions, directing the Cas9-gRNA complex to sequences with complementary sequences to the gRNA [5, 6]. DSBs can be repaired through the non-homologous end-joining (NHEJ) mechanism or the homology-directed repair (HDR) system. However, it has been observed that the NHEJ mechanism is not an effective means in yeast for repairing DSBs, often leading to severe toxicity [123]. This presents

the opportunity for genome editing in yeast by employing donor DNA that shares homology with regions upstream and downstream of the DSBs (Figure 1-9). In the presence of a donor DNA that could contain the gene of interest, yeast exhibits efficient repair of the DSBs. However, any unrepaired break results in severe toxicity for yeast survival, essentially serving as a robust selection marker for genome editing facilitated by the CRISPR-Cas9 system [5].

The biosynthetic pathway of plant SMs often involves multiple enzymatic reactions that must be reconstituted in a heterologous host for synthetic biology applications. Integrating each enzyme using a single gRNA can be time-consuming, especially when several integrations are required for a given pathway. Although targeting multiple sites with multiple gRNAs has been achieved, the efficiency is often limited by the least effective gRNA [124]. The multiplexed CRISPR-Cas9 system offers a more time-efficient solution by simultaneously targeting multiple integration sites using a single gRNA, facilitating the construction of complex pathways more efficiently. Targeting multiple loci in the genome using a single gRNA requires multiple identical sites in the genome. Several strategies have been successfully implemented for simultaneously targeting multiple loci, which have been extensively discussed by Utomo et al. [124]. Some approaches have utilized pre-defined identical sequences scattered in the yeast genome, such as the delta (δ) site of Ty (transposons of yeast) elements [125] or ribosomal DNA (rDNA) repeats located on chromosome XII of yeast [126]. As reported, these approaches, despite having significant outcomes, have demonstrated unintended off-target results or severe toxicity caused by multiple cleavages in one chromosome, leading to genome instability [124, 127, 128]. Alternatively, some approaches utilizing synthetic gRNA target sites were developed for a more controlled multi-copy gene integration in the yeast genome. For example, Hou et al. created a wicket system that integrated a single target site in multiple intergenic regions with universal homology arms to construct the β-carotene biosynthetic pathway [129]. Bourgeois et al. developed a concept called landing pad, which consists of utilizing multiple pre-tested synthetic gRNA target sites in different copy numbers across different genomic loci to better control the copy number of a gene of interest. In this system, each synthetic target is targeted by a separate gRNA [130]. In the study done by Baek et al., a single gRNA landing pad was installed in up to 6 intergenic loci, each on different chromosomes, with unique homology arms allowing integration of each gene in a defined integration site. This system created multiple yeast strains capable of integrating 1 to 6 genes in a single transformation [5]. Shahsavarani et al. utilized this platform to engineer catharanthine and

vindoline, the intermediate molecules of vinblastine and vincristine [131]. This approach was later developed by Utomo et al. by employing additional synthetic gRNA target sites on another four chromosomes, allowing the integration of up to 8 genes in two sequential transformations [132].

Problematic, hypothesis, objectives

So far, direct extraction from native plants has been the primary source for supplying AAs, notably galanthamine. However, this approach often presents significant challenges. The production rate of AAs in native plants is usually low. Meeting the pharmacological demand would lead to the over-exploitation of native plants in their natural habitats, endangering their existence. Furthermore, seasonal and environmental factors affect the production rate in natural habitats. Although it is possible to grow these plants commercially for medicinal and ornamental purposes, the expected demand for these valuable therapeutics in the future necessitates establishing more sustainable approaches to produce them cheaper and faster. Additionally, dedicating thousands of acres to cultivate these plants poses challenges such as land usage.

The key is to characterize the enzymes involved in precursor molecule production and those catalyzing downstream reactions and apply this knowledge to AAs pathway construction for synthetic biology. In recent years, biosynthetic enzymes involved in the AAs have been discovered continuously, including NBS, NR [85], N4′OMT [53], CYP96T [55], and many other novel enzymes [133]. However, more efforts are needed to elucidate the enzymes involved in precursor biosynthesis to produce these compounds from simple carbon sources via synthetic biology.

The phenylpropanoid pathway in *L. aestivum* remains largely unexplored, particularly in its enzymatic components. While partial characterization of enzymes such as phenylalanine PAL and truncated C4H has been reported in other Amaryllidaceae species [126, 134], the roles of CYP98A and a novel hydroxylase candidate APX/C3H have not been investigated in this plant family. Exploring the activities of CYP98A and APX/C3H and their potential substrates could provide essential biochemical insights and potentially reveal a shorter route toward DHBA, a precursor for the heterologous construction of the AAs pathway in suitable host systems. Such insights are pivotal for advancing synthetic biology approaches aimed at the sustainable production of these valuable metabolites, thereby reducing reliance on native plant sources and complex chemical synthesis.

1.5.2 Objective 1: Characterization of candidate hydroxylases in the phenylpropanoid pathway of the Amaryllidaceae plant, *L. aestivum*

The main objective of this study is to characterize candidate hydroxylases and their substrate preferences in the phenylpropanoid pathway (Figure 2-1) of *L. aestivum,* discussed in Chapter II. This involves several specific aims: (i) conducting an *in silico* analysis of the *L. aestivum* transcriptome [52] to identify candidate sequences for cinnamate 4-hydroxylase (CYP73A), CYP98A, and APX/C3H using homologs from other plant families; (ii) biochemically characterizing these enzymes using appropriate heterologous systems, such as yeast for CYP450 enzymes, *N. benthamiana* for *in vivo* evidence, and *E. coli* for *in vitro* enzymatic tests of APX/C3H; (iii) performing subcellular localization studies of candidate enzymes in *N. benthamiana* using a fluorescent protein tagging system to provide initial evidence for potential protein-protein interactions and the existence of a metabolon in the phenylpropanoid pathway; and (iv) conducting differential expression analysis of genes in the phenylpropanoid and core Amaryllidaceae pathways, correlating these data with metabolite profiling of phenolics and alkaloids in various tissues. This approach aims to identify the most relevant steps associated with AA content.

1.5.3 Objective 2: Establishment of AAs core pathway in baker's yeast

The AAs core pathway has not yet been reconstructed in microbial systems, except for some studies that individually reported the *in vivo* activity of a few AAs pathway enzymes in yeast and bacterial systems [51, 135]. Establishing a platform containing the necessary phenylpropanoid pathway genes and early genes for Amaryllidaceae alkaloids is essential. This platform could serve as a chassis to further produce any desired Amaryllidaceae alkaloid. This will allow us to link the primary metabolism of microbial systems to the specialized metabolic pathways of plant natural metabolites, opening new possibilities for sustainable and efficient production. Therefore, it was first attempted to reconstitute the AAs central pathway consisting of enzymes TYDC, NBS, NR, and N4′OMT, sourced from Amaryllidaceae species, in yeast using multiplex CRISPR-Cas9 system as previously described [5, 132]. Future studies would further engineer the strain producing 4-O-methylnorbelladine for the upstream phenylpropanoid pathway to *de novo* produce this compound. As 4-*O*-methylnorbelladine is a core intermediate for the most biologically important AAs, including galanthamine, this objective would provide the initial perception of the pathway

bottlenecks for future research. The preliminary outcome of this attempt is discussed in Chapter III.

Chapter II

This chapter explores the characterization of potential hydroxylase enzymes within the phenylpropanoid pathway of Amaryllidaceae plants. This pathway is believed to synthesize DHBA, a precursor molecule for AAs, which requires two hydroxylation steps. The chapter was written in article format and published in the **Plant Physiology and Biochemistry journal (doi.org/10.1016/j.plaphy.2024.108612)**, and the content unavoidably overlaps with other chapters.

Contribution: Vahid Karimzadegan performed bioinformatic analysis, yeast heterologous expression, subcellular localization, metabolite, and differential expression experiments. Manoj Koirala contributed to metabolite and gene expression analysis and data visualization. Natacha Merindol mainly did protein modeling and molecular docking and contributed to writing the draft and revision. Sajjad Sobhanverdi helped with the *in vivo* enzymatic assay. Bharat Bhusan Majhi provided technical mentorship during the project. Sarah-Eve Gélinas performed LC-MS analysis. Vitaliy I. Timokhin and John Ralph provided the substrate for the CYP98A enzyme assay and assisted with the paper's revision. Mehran Dastmalchi co-supervised the project and assisted with the revision of the paper. Isabel Desgagné-Penix supervised the project, provided funding and material for the project, and assisted with the writing and revising the paper.

2. Characterization of cinnamate 4-hydroxylase (CYP73A) and *p*-coumaroyl 3′-hydroxylase (CYP98A) from *Leucojum aestivum*, a source of Amaryllidaceae alkaloids

Vahid Karimzadegan[a], Manoj Koirala[a], Sajjad Sobhanverdi[a], Natacha Merindol[a], Bharat Bhusan Majhi[a,] Sarah-Eve Gélinas[a], Vitaliy I. Timokhin[b], John Ralph[b,c], Mehran Dastmalchi[d], and Isabel Desgagné-Penix[a,*]

[a]Department of Chemistry, Biochemistry and Physics, Université du Québec à Trois-Rivières, Trois-Rivières, Québec, Canada.

[b]Department of Energy's Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, Madison, WI 53726, USA,

[c]Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, USA
[d]Department of Plant Science, McGill University, Montréal, Québec, Canada

*Corresponding author: Isabel Desgagne-Penix, Department of Chemistry, Biochemistry and Physics, Université du Québec à Trois-Rivières, 3351 boul. des Forges, Trois-Rivières, QC, G9A 5H7, Canada. (Tel: 819-376-5011, Fax: 819-376-5014).

E-mail: Isabel.Desgagne-Penix@uqtr.ca

Keywords

Phenylpropanoids, specialized metabolism, Ascorbate peroxidase/4-coumarate 3-hydroxylase, endoplasmic reticulum membrane localization, coumaric acid, metabolic profile

2.1 Abstract

The biosynthesis of AAs starts with the condensation of tyramine with DHBA. The latter derives from the phenylpropanoid pathway that involves modifications of *trans*-cinnamic acid, *p*-coumaric acid, caffeic acid, and possibly 4-HBA, potentially catalyzed by hydroxylase enzymes. Leveraging bioinformatics, molecular biology techniques, and cell biology tools, this research identifies and characterizes key enzymes from the phenylpropanoid pathway in *Leucojum aestivum*. Notably, we identified and characterized *trans*-cinnamate 4-hydroxylase (*Lae*C4H) and 3′-hydroxylase (*Lae*C3′H), two essential cytochrome P450 enzymes, and the ascorbate peroxidase/4-coumarate 3-hydroxylase (*Lae*APX/C3H). Although *Lae*APX/C3H consumed *p*-coumaric acid, it did not result in the production of caffeic acid. Yeasts expressing *Lae*C4H converted *trans*-cinnamate to *p*-coumaric acid, whereas *Lae*C3′H catalyzed a 3-hydroxylation specifically of *p*-coumaroyl shikimate rather than of free *p*-coumaric acid or 4-HBA. *In vivo* assays conducted *in planta* in this study provided further evidence for the contribution of these enzymes to the phenylpropanoid pathway. Both enzymes demonstrated typical endoplasmic reticulum membrane localization in *Nicotiana benthamiana,* adding spatial context to their functions. Tissue-specific gene expression analysis reveals roots as hotspots for phenylpropanoid-related transcripts and bulbs as hubs for AA biosynthetic genes, aligning with the highest AA concentration. This investigation contributes to adding valuable insights into the phenylpropanoid pathway within Amaryllidaceae, laying the foundation for the development of sustainable production platforms for AA and other bioactive compounds with diverse applications.

2.2 Introduction

The Amaryllidaceae are pharmacologically potent plant species owing to the production of phyla-restricted SMs, namely the Amaryllidaceae alkaloids [136-139]. The structurally diverse alkaloids from this family are associated with many pharmacological applications. *Leucojum aestivum* L. (Summer snowflake) of the Amaryllidaceae family is a reported source of galanthamine, an AA approved as a treatment for mild symptoms of Alzheimer's disease [140]. The plant additionally accumulates anticancer and antiviral lycorine- and haemanthamine-type AA [141-143]. However, due to their variable-to-low abundance in nature, extraction from plants would pose a threat to their habitats and the growth of native plants [140]. Microbial platforms, such as bacteria or yeast, offer promising alternatives for producing valuable plant-derived compounds through a synthetic biology approach [144, 145]. For example, the reconstitution of vinblastine biosynthesis, an essential anti-cancer drug, was successfully achieved in *Saccharomyces cerevisiae* [146]. Such an approach requires complete resolution of pathways before reassembly in heterologous hosts.

The biosynthesis of AA starts with the condensation of DHBA, resulting from the phenylpropanoid pathway, and tyramine that is formed through the decarboxylation of tyrosine [8]. The phenylpropanoid core pathway starts with synthesizing *trans*-cinnamic acid from phenylalanine by the well-characterized cytosolic enzyme phenylalanine ammonia-lyase (PAL) (Figure 2-1). The hydroxylation of *trans*-cinnamic acid to *p*-coumaric acid in the early phenylpropanoid pathway is a key step, catalyzed by a cytochrome P450 (CYP450) enzyme, *trans*-cinnamate 4-hydroxylase (C4H), that belongs to the CYP73A subfamily [102]. Considering the pivotal role of *p*-coumaric acid derivatives in plant metabolism, structure, development, and defense [147], the catalytic role of C4H is of great importance.

The hydroxylation of *trans*-cinnamic acid into *p*-coumaric acid is expected to be catalyzed by a C4H enzyme in *L. aestivum* as evidenced by similar enzymatic characterizations in various plant species such as liverworts [148], Madagascar periwinkles [149], poplars [150], and other plants [151, 152]. To date, two studies have isolated and partially characterized C4H from Amaryllidaceae species, reporting that the overexpression in bacteria of truncated C4H from *Lycoris radiata* and *L. aurea* yielded *p*-coumaric acid [134, 153].

Two routes have been suggested, from *p*-coumaric acid to caffeic acid (Figure 2-1). It was first proposed that direct hydroxylation at the acid level could be catalyzed by a CYP98A family enzyme, as described following the heterologous expression of *Arabidopsis thaliana At*CYP98A (referred to as C3H in the presence of *p*-coumaric acid) in the cyanobacterium *Synechocystis* spp. [154]. A later report suggested that hydroxylation by *At*CYP98A instead occurred via *p*-coumarate esters, including *p*-coumaroyl shikimate and *p*-coumaroyl quinate (referred to as C3′H in the presence of coumarate esters) [155]. Recent research indicates that substrate specificity of angiosperm C3′H varies between species and between isoforms of the same species, some displaying promiscuity and favoring *p*-coumaroyl shikimate and *p*-coumaroyl quinate as substrates over various natural and non-natural compounds, including prenyl-, isoprenyl-, benzyl-, and threonyl-coumarates [109]. On the other hand, a bifunctional ascorbate peroxidase/4-coumarate 3-hydroxylase (APX/C3H), a soluble enzyme purified from multiple plant species, including *A. thaliana*, maize, and *Brachypodium*, was reported to catalyze the direct hydroxylation of free *p*-coumaric acid to caffeic acid [10]. However, a recent study challenged this hypothesis when testing the activity of APX/C3H from *Sorghum bicolor* on *p*-coumaric acid [156]. Caffeic acid can potentially yield DHBA to be incorporated into AAs, but the enzyme catalyzing the reaction has not yet been identified.

Here, we report the isolation and functional characterization of C4H, APX/C3H, and C3′H from the AA-producing *L. aestivum* (*Lae*). Heterologous expression of soluble *Lae*APX/C3H was conducted in *Escherichia coli.* Recombinant *Lae*C4H and *Lae*C3′H were expressed in a yeast expression system. The subcellular localization of *Lae*C4H and *Lae*C3′H fusion proteins and *in planta* enzymatic activity was investigated in *Agrobacterium*-infiltrated *Nicotiana benthamiana*. *L. aestivum* tissue-specific expression patterns of *Lae*C4H and *Lae*C3′H and other transcripts involved in the phenylpropanoid and AA pathways were investigated and compared with targeted metabolite profiles. This study reports on C4H (CYP73A), APX/C3H, and C3′H (CYP98A) from the Amaryllidaceae family to clarify their role in the phenylpropanoid pathway upstream of the AA biosynthetic pathway.

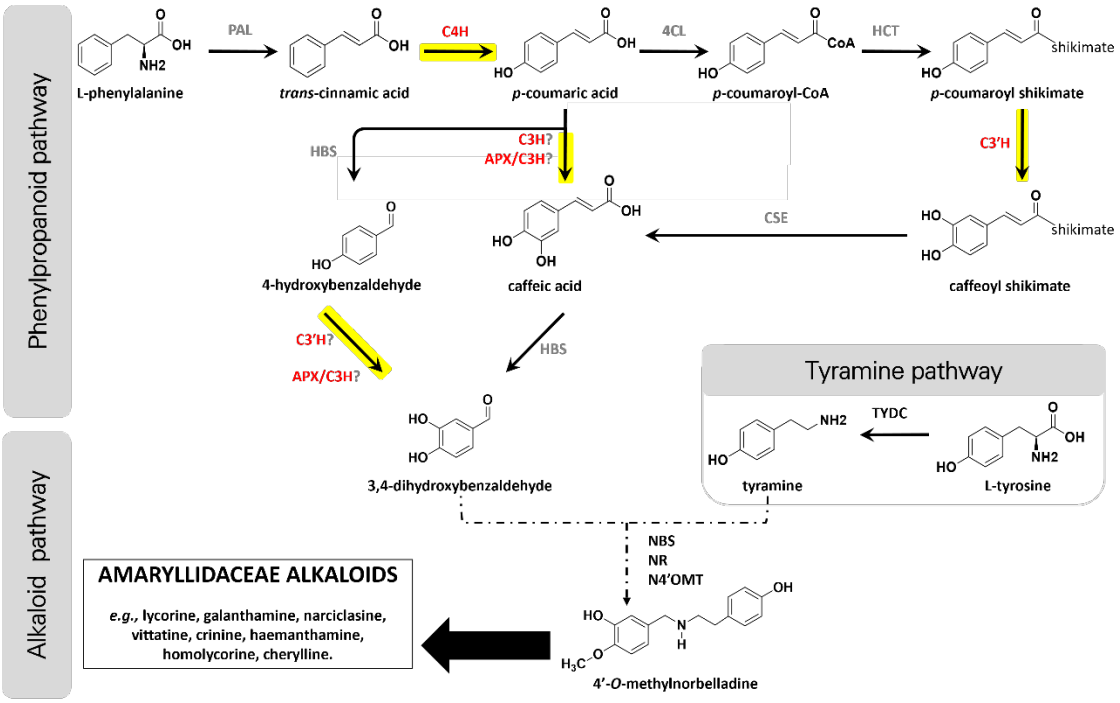


Figure 2-1. Proposed pathway leading to 4′-O-methylnorbelladine, a common intermediate of AAs. Enzymes shown in red were the primary enzymes studied in this work. Enzymes shown in black have been characterized in Amaryllidaceae, whereas enzymes in grey have not been characterized in Amaryllidaceae. Highlighted in yellow are the reactions investigated in this study. Abbreviations: PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase (CYP73A); 4CL, 4-coumarate:CoA ligase; HCT, hydroxycinnamoyl-CoA: shikimate hydroxycinnamoyl transferase; C3′H, *p*-coumaroyl shikimate hydroxylase (CYP98A); C3H, coumarate 3-hydroxylase (CYP98A); CSE, caffeoyl shikimate esterase; APX/C3H, ascorbate peroxidase/coumarate 3-hydroxylase; HBS, hydroxybenzaldehyde synthase; TYDC, tyrosine decarboxylase; NBS, norbelladine synthase; NR, noroxomaritidine/norcraugsodine reductase; N4′OMT , norbelladine 4′-*O*-methyltransferase.

2.3 Materials and Methods

2.3.1 Plant materials and growth condition

The bulbs of *L. aestivum*, commonly known as summer snowflake, were bought from Vesey's (York, PE, Canada). The bulbs were grown outdoors in Trois-Rivières (Québec, Canada) until flowering. *N. benthamiana* seeds were germinated and grown indoors in autoclaved AGRO MIX G6 potting soil (Fafard, Saint-Bonaventure, QC, Canada) with a long photoperiod (16 h light/8 h dark) at 22 °C.

2.3.2 Bioinformatic analysis, molecular homology modeling and docking

The open reading frames (ORF) and the accession numbers were obtained from NCBI (ncbi.nlm.nih.gov). Molecular weight (MW) and isoelectric point (IP) were determined using Expasy [157]. Sequence alignment utilized MegAlign Pro (DNASTAR, MegAlign Pro, version 17.4.1.17, Madison, WI: DNASTAR, Inc) and the Clustal Omega algorithm. MEGA11 was employed for Phylogenetic analysis with 1000 bootstrap replicates [158]. GraphPad Prism 8.0.1 (GraphPad Software, San Diego, California, USA) was used to visualize data. Amino acid sequences corresponding to *Lae*C3′H and *Lae*C4H were uploaded onto the Protein Homology/analogY Recognition Engine V 2.0 (Phyre2) [159] website, RosettaFold [160], and MOE 2020.09 software (Chemical Computing Group) for modeling the proteins. By comparing with crystalized orthologs (1PQ2 for human drug-metabolizing CYP450 2C8, 6VBY for cinnamate 4-hydroxylase (C4H) from *Sorghum*. *bicolor*) and Alphafold-predicted C4H from *Petunia hybrida* (AF-F1B282-F1-model_v4), the most consistent models were selected from Phyre predictions.

MOE was used to analyze the resulting homology model conformations and to prepare receptors for docking, as described previously [161]. The structure preparation consisted of correcting issues, capping, charging termini, selecting appropriate alternatives, and calculating optimal hydrogen position and charges using Protonate 3D. Energy minimization was performed for each fixed receptor with a tethered active site that included positioned template substrates inside. Ready-to-dock substrates were uploaded from ZINC15 [162], or built from smiles codes with MOE builder. All protomers predicted at the enzymatic reaction pH were included as possible substrates. The MMFF94× force field was used. Each receptor's active site was predicted using MOE Site Finder and used as a docking site to place substrate using Triangle Matcher as placement method for 200 poses and tethered induced fit as a refinement to perform flexible docking. Then, the resulting poses were analyzed, and the best pose was presented for each substrate according to a comparison with a template's active site and docking scores. The protein-ligand interaction profiler (PLIP) was used to analyze the interactions between substrates and receptors' residues [163], and the images were further processed using PyMOL (Shrödinger).

2.3.3 Yeast and bacteria growth conditions

E. coli DH5α (Invitrogen, Carlsbad, CA, USA) was cultured in Luria-Bertani (LB) medium with 50 μg.mL^{-1} ampicillin at 180 rpm for 16 h at 37 °C. *Agrobacterium tumefaciens* GV3101 [164] was grown in LB medium with kanamycin, rifampicin, and gentamicin (Fisher Scientific, ON, Canada) at 50, 50, and 30 μg.mL^{-1} final concentrations, respectively, with overnight incubation at 28 °C and 200 rpm stirring. *Saccharomyces cerevisiae*-derived yeast strain INVSc-1 (Invitrogen, Fisher Scientific) was grown in YPD complete media.

2.3.4 Chemicals

Caffeic acid (98%), ferulic acid (99%), and papaverine (98%) reference standards were purchased from Millipore Sigma (Massachusetts, USA). *p*-coumaric acid (98%), *trans*-cinnamic acid (98%), 4-HBA, and DHBA reference standards were bought from Fisher Scientific (Ontario, Canada). *p*-coumaroyl shikimate and caffeoyl shikimate were chemically synthesized [165]. Analytical LC-MS grade methanol (99.9%) and formic acid (99%) were purchased from Fisher Scientific. Standard stock solutions of each reference standard were prepared at 100 mg. mL^{-1} in methanol and stored in the dark at -20 °C.

2.3.5 RNA extraction, cDNA synthesis, and differential expression analysis

One hundred fifty mg of plant tissues, including bulb, root, stem, leaf, and flower of L. aestivum, was ground in liquid nitrogen using a mortar and pestle and immediately transferred to 1.5 mL Eppendorf tubes. Total RNA was extracted using TRIzol reagent (Invitrogen, Fisher Scientific) according to the manufacturer's instructions. After treatment with DNAse, quantification was done using a Nanophotometer (Implen, Munich, Germany), and 1 μg of total RNA was used for cDNA synthesis using SensiFAST's cDNA synthesis kit (Bioline, London, England, United Kingdom) according to the company's instructions, using both oligo(d)T and random hexamers.

Real-time quantitative PCR (RT-qPCR) was performed to investigate the expression pattern of the genes involved in Amaryllidaceae alkaloid and precursor pathway (Figure 2-1 and Table A1) with Luna Universal qPCR Master Mix (New England Biolabs). The cycle program was set as 95 °C for 2 min (1 cycle), [95 °C for 15 s, 60 °C for 30 s] (45 cycles)] followed by dissociation step 95 °C for 10 s, 50 °C for 5 s and 95 °C for 5 s. *Histone3* from *L. aestivum* was used as the internal control. The threshold cycle (Ct) value of each gene was normalized against the Ct value of the

reference gene. The relative gene expression levels were determined using the comparative ΔΔCt method [166] by utilizing the average Ct values obtained from the technical triplicates. The obtained results were analyzed and visualized via CFX Maestro software (Bio-Rad).

2.3.6 Cloning, transformation procedure, and protein expression

Gene-specific forward and reverse primers (Table A1-1) were designed to isolate the coding sequences of *C4H*, *C3'H*, and APX/C3H from the cDNA of *L. aestivum*. The open reading frames (ORFs) of *LaeC4H* and *LaeC3'H*, each in-frame with a Myc tag, were sequenced (Table A1-2) and cloned into the yeast expression vector pESC-LEU-*CroCPR*. Specifically, *LaeC4H* was inserted between ApaI and SalI restriction sites, while *LaeC3'H* was inserted between BamHI and SalI restriction sites. This vector already harbors the coding sequence of CYP450 reductase (CPR) from *Catharanthus roseus* (provided by Prof. Yang Qu at the University of New Brunswick). The ORF of *LaeAPX*/*C3H* was cloned in the bacterial expression vector pMAL-c2X (New England Biolabs) using BamHI and SalI restriction enzymes in frame with maltose-binding protein (MBP). Restriction enzymes were purchased from New England Biolabs. Using a heat-shock transformation method, all the constructs were delivered to chemically competent *E. coli* DH5α, and selection was carried out on LB agar medium with ampicillin (50 $\mu g.mL^{-1}$). A colony PCR was performed to isolate positive colonies using gene-specific primers and Taq DNA polymerase. The resulting plasmids were subjected to DNA sequencing to confirm the integrity of the target sequences.

The *S. cerevisiae* strain INVSc-1 was used to express recombinant *Lae*C4H and *Lae*C3′H proteins. Transformation with constructed vectors was performed using the Yeastmaker™ Yeast Transformation System 2 kit (Takara Bio). Positive yeast transformants were selected on yeast minimal leucine drop-out media. A single colony was precultured in 2 mL minimal leucine drop-out medium overnight at 30 °C, 200 rpm. This preculture was then inoculated into a 500 mL synthetic nitrogen base minimal medium lacking leucine with 5% glucose or dextrose (w/v) and incubated in the same condition. Yeast cells were collected, washed, and transferred to an induction medium (synthetic nitrogen base leucine drop-out) with 10% galactose for 24 hours in the same condition. Microsome preparation, adapted from [27], involved resuspending cells in TES buffer (20 mM Tris-HCl [pH 7.5], 1 mM EDTA, 0.6 M sorbitol) and dividing into pre-chilled 1.5 ml tubes. The TES-yeast solution, containing 1mm glass beads, was chilled, and cells were

mechanically lysed using a TissueLyser (Qiagen, ON, Canada) at 30 Hz for 5 min (3 repetitions with cooling intervals). Lysed cells were centrifuged at 5,000 rpm for 20 min at 4 °C, and the supernatant was subjected to ultracentrifugation (Optima l-90k, Beckman Coulter, ON, Canada) at 24,000 rpm for 60 min at 4 °C. The microsome pellet was resuspended in TEG buffer (50 mM Tris-HCl [pH 7.5], 1 mM EDTA, 20% v/v glycerol) and stored at -80 °C. The expression of recombinant proteins was verified using Western blot analysis.

To produce recombinant *Lae*APX/C3H, purified plasmids were transformed using heat shock transformation to chemically competent *E. coli* Rosetta (DE3) pLysS (Novage) strain. Colony PCR was conducted to screen for positive transformants. Expression and purification steps of *Lae*APX/C3H followed as described in [10, 161]. SDS-PAGE was performed to validate protein expression and purification.

2.3.7 *In vitro* enzymatic assay for *Lae*C4H, *Lae*C3′H and *Lae*APX/C3H

The *in vitro* enzymatic assay for *Lae*C4H was performed according to [167], with a minor change in terms of incubation. Briefly, the reaction was adjusted to 600 µL total volume containing 50 µg microsomal fractions, NADPH (0.5 mM), sodium phosphate buffer (100 mM, [pH 7.4]), *trans*-cinnamic acid (0.1 mM), and incubated at 30 °C for 10, 30, 60, and 120 min, and overnight. The negative controls for the assays were: the reaction mixture with no NADPH, no substrate, no microsomal fractions, and microsomal fractions extracted from yeast harboring an empty vector. At the end of each time point, the reaction was terminated by adding 40 µL 6 M HCl. All reactions were performed in triplicate, and each determination was repeated at least twice.

The activity of *Lae*C3′H was investigated as described previously [155], with minor modifications. Three different substrates, *p*-coumaroyl shikimate, free *p*-coumaric acid, and 4-HBA (all at 0.1 mM final concentration) were tested in a 200 µL reaction containing 50 µg microsomal fractions, NADPH (0.6 mM), and sodium phosphate buffer (100 mM, [pH 7.4]). The samples were incubated at 28 °C for 30, 60, and 120 min, and overnight. The negative controls were the same as those described above. The reaction was terminated by adding 20 µL of acetic acid. All reactions were performed in triplicate, and each determination was repeated at least twice.

The *in vitro* hydroxylase activity of *Lae*APX/C3H toward *p*-coumaric acid and 4-HBA was tested using the purified recombinant protein as previously described [10].

Following the reaction termination, papaverine (10 mg. L^{-1} final concentration) was added to all the reactions, serving as an internal standard for the relative quantification of detected compounds. The reaction samples were mixed using a vortex, centrifuged for 10 min at 12,000 rpm, and diluted 10-fold in the mobile phase (formic acid 0.1% v/v in milli-Q water and formic acid 0.1% v/v in methanol (90:10)). Analyses were conducted using HPLC-MS/MS as described below.

To determine the kinetic parameters of *Lae*C4H and *Lae*C3′H, enzymatic assays using a constant concentration of microsomal fractions while varying the concentration of respective substrates (in the case of *Lae*C3′H, only *p*-coumaroyl shikimate was used) ranging from 100 nM to 30 µM were carried out under the same conditions described above. The initial rate of reaction was measured for each substrate concentration. The enzymes' maximum reaction rate (Vmax) and Km were determined using the Michaelis-Menten equation and non-linear regression analysis in GraphPad Prism 8.0.1.

2.3.8 Instrumentation and chromatographic conditions for enzymatic assays

A high-performance liquid chromatography (HPLC) system coupled with a tandem mass spectrometer (MS/MS) (Agilent Technologies, Santa Clara, California, USA) equipped with an Agilent Jet Stream ionization source, a binary pump, an autosampler, and a column compartment were used for the analysis. Compound separation was achieved using a Kinetex EVO C18 column (150 × 4.6 mm, 5 µm, 100 Å; Phenomenex, Torrance, USA). Five microliters of each sample were injected onto the column that was set at 30 °C. A gradient method made of (A) formic acid 0.1% v/v in milli-Q water and (B) formic acid 0.1% v/v in methanol with a flow rate of 0.4 mL/min was used to achieve chromatographic separation. The HPLC elution program is described as follows: 0 min, 35% B; 10.0 min, 50% B; 16.0 min, 60% B; 16.2 min, 35% B. The total run time was 20 min per sample to allow column reconditioning before the next injection. The parameters used in the MS/MS source were set as follows: gas flow rate 10 L. min^{-1}, gas temperature 300 °C, nebulizer 45 psi, sheath gas flow 11 L. min^{-1}, sheath gas temperature 300 °C, capillary voltage 4000 V in ESI^+ and 3500 V in ESI^- and nozzle voltage 500 V. Agilent MassHunter Data Acquisition (version 1.2) was used to control the HPLC-MS/MS, and MassHunter Qualitative Analysis (version 10.0) was used for data processing.

2.3.9 Fusion fluorescent protein constructions, transient protein expression in *N. benthamiana*, and confocal microscopy

The open reading frame of *LaeC4H* and *LaeC3'H* genes were cloned into the *KpnI* and *XbaI* sites of the pBTEX binary vector in a frame with a yellow fluorescent protein (YFP) under the CaMV 35S promoter [168]. The resulting vectors, pBTEX-*LaeC4H-YFP* and pBTEX-*LaeC3'H-YFP*, were transformed into the *E. coli* DH5α chemical competent cells using heat-shock. The transformation mixtures were grown on LB agar medium with 50 $\mu g.mL^{-1}$ kanamycin, and colony PCR was done to isolate positive colonies as described above. Each of the constructed vectors was transferred separately to *A. tumefaciens* strain GV3101 by electroporation and grown on LB agar containing kanamycin, rifampicin, and gentamicin, at the final concentration of 50, 50, and 30 $\mu g.mL^{-1}$ respectively. Colony PCR was performed to select the positive transformants.

The transient expression in *N. benthamiana*, using *A. tumefaciens* bacteria, was conducted as described previously [161]. To explore the subcellular location of the *Lae*C4H and *Lae*C3′H, the leaves of 4-week-old *N. benthamiana* plants were co-infiltrated with *A. tumefaciens* harboring pBTEX-*LaeC4H-YFP* and *A. tumefaciens* containing either an ER marker (ER-mCherry) or a nuclear marker (Nls-CFP) in a 1:1 mixture. Forty-eight hours after co-infiltration, the fluorescent signals were visualized using a Leica TCS SP8 confocal laser scanning microscope (Leica, Wetzlar, Hesse, Germany). The excitation and emission wavelengths for YFP visualization were 488 and 500-525 nm, 405 and 420-490 nm for CFP, and 587 and 610 nm for mCherry, while excitation and emission wavelengths to detect chlorophyll were set on 550 and 630-670 nm, respectively. The images were merged in Las X software (Leica 720 Microsystems).

2.3.10 Metabolite extraction and analysis

One gram of *L. aestivum* sample corresponding to the different plant parts, such as flower, stem, leaf, bulb, and root, was ground under the liquid nitrogen using a mortar and pestle. A portion of the homogenized powder was kept for RNA extraction. Crude metabolites extraction was performed by using 1 mL of methanol for 150 mg of plant tissue fresh weight. Extraction was carried out for 24 h at room temperature, followed by centrifuging at 10,000 x g to remove plant debris. Samples were dried using a speed vac concentrator, and crude metabolite extract was reconstituted in methanol to have a final concentration of 1000 mg/ml and was filtered (0.2 µm Acrodisc® syringe filter, Pall Corporation, NY). Target metabolite analysis was performed for

phenylpropanoids and AA. The LC-MS/MS method used is described in subsection 2.11 with some modification in the HPLC elution program, which was set as follows: 0-10 min, 10% B; 20-25 min, 100% B; 26 min, 10% B. The total run time was 30 min per sample to allow the reconditioning of the column before the next injection. Agilent MassHunter Data Acquisition (version 1.2) was used to control the HPLC-MS/MS, MassHunter Qualitative Analysis (version 10.0), and MassHunter Quantitative QQQ Analysis (version 10.0) were used for data processing. During extraction and metabolites analysis, papaverine was used as an internal standard. All the relative quantities of different metabolites were normalized with the highest amount in each tissue type and visualized by using Graph pad Prism 8.0.1

2.3.11 *In vivo* enzymatic assay of *LaeC4H* and *LaeC3'H*

*Lae*C4H and *Lae*C3′H were transiently expressed in *N. benthamiana* leaves, as described in subsection 2.8, to evaluate the impact of the overexpressed enzymes on phenolic compound levels *in planta*. A vector expressing yellow fluorescent protein (YFP) was used as a negative control. Additional appropriate controls are listed in Table A1-3. Subsequently, 2 days post-agroinfiltration, leaves were treated with 100 µM of substrates and then incubated under the same conditions. After 24 hours, four leaves from each plant were harvested, pooled, and ground with liquid nitrogen for storage at -80°C until analysis. Metabolite extraction and analysis were conducted using 100 mg of each sample as described in subsection 2.9 with a minor modification. A second analysis was performed in duplicate following a 10-fold metabolite concentration before injecting in HPLC-MS/MS.

2.4 Results

2.4.1 Identification and sequence analysis of *LaeC4H* and *LaeC3'H* from *L. aestivum*

The full-length open reading frame (ORF) sequences of predicted *LaeC4H* (505 amino acids (aa)) and *LaeC3'H* (509 aa), and *LaeAPX*/*C3H* (248 aa) were extracted from *L. aestivum* transcriptome [52]. The NCBI accession numbers are as follows: UIP35210 (*LaeC4H*), UIP35212 (*LaeC3'H*), and MW971972.1 (*LaeAPX*/*C3H*) (Table A1-2). Predicted hypothetical molecular weights (MW) and isoelectric points (IP) for these proteins are as follows: *Lae*C4H (57.97 kDa, 9.12), *Lae*C3′H (57.8 kDa, 8.6), and *Lae*APX/C3H (5.84 kDa, 27.2). Multiple sequence alignments were performed for *Lae*C4H and *Lae*C3′H, with homologous sequences selected from monocot, dicot, gymnosperm, pteridophyte, and bryophyte species. *Lae*C4H shared a high amino acid sequence

identity of 94.5% with Amaryllidaceae *Lycoris radiata*. Substantial identity is also observed with counterparts from various plant groups, including monocots (78% to *S. bicolor*), dicots (84% to *A. thaliana* and 85% to *C. roseus*), gymnosperms (80% to *Ginkgo biloba*), and a fern, *Selaginella moellendorffii*, with 80% identity (Figure 2-2 A and Figure A1-1). Lower but still high levels of identity were noted between *Lae*C4H and *Marchantia paleacea* (65%) and *Physcomitrella patens* (70%) as representatives of bryophytes. *Lae*C3′H shared a high level of identity (95%) with the predicted homolog in *Narcissus papyraceus*. Compared to *Lae*C4H, *Lae*C3′H was more distantly related to other homologs, such as monocot *Sbi*C3′H (71%), dicot *Ath*C3′H (73%), gymnosperm *Pta*C3′H (73%), and bryophyte *Ppa*C3′H (64%) sequences (Figure 2-2 B and A1-2). *Lae*C4H and *Lae*C3′H and their Amaryllidaceae paralogs formed a clade with monocot species adjacent to the dicot clade but further from a gymnosperm, pteridophyte, and bryophyte homologs (Figure 2-2).

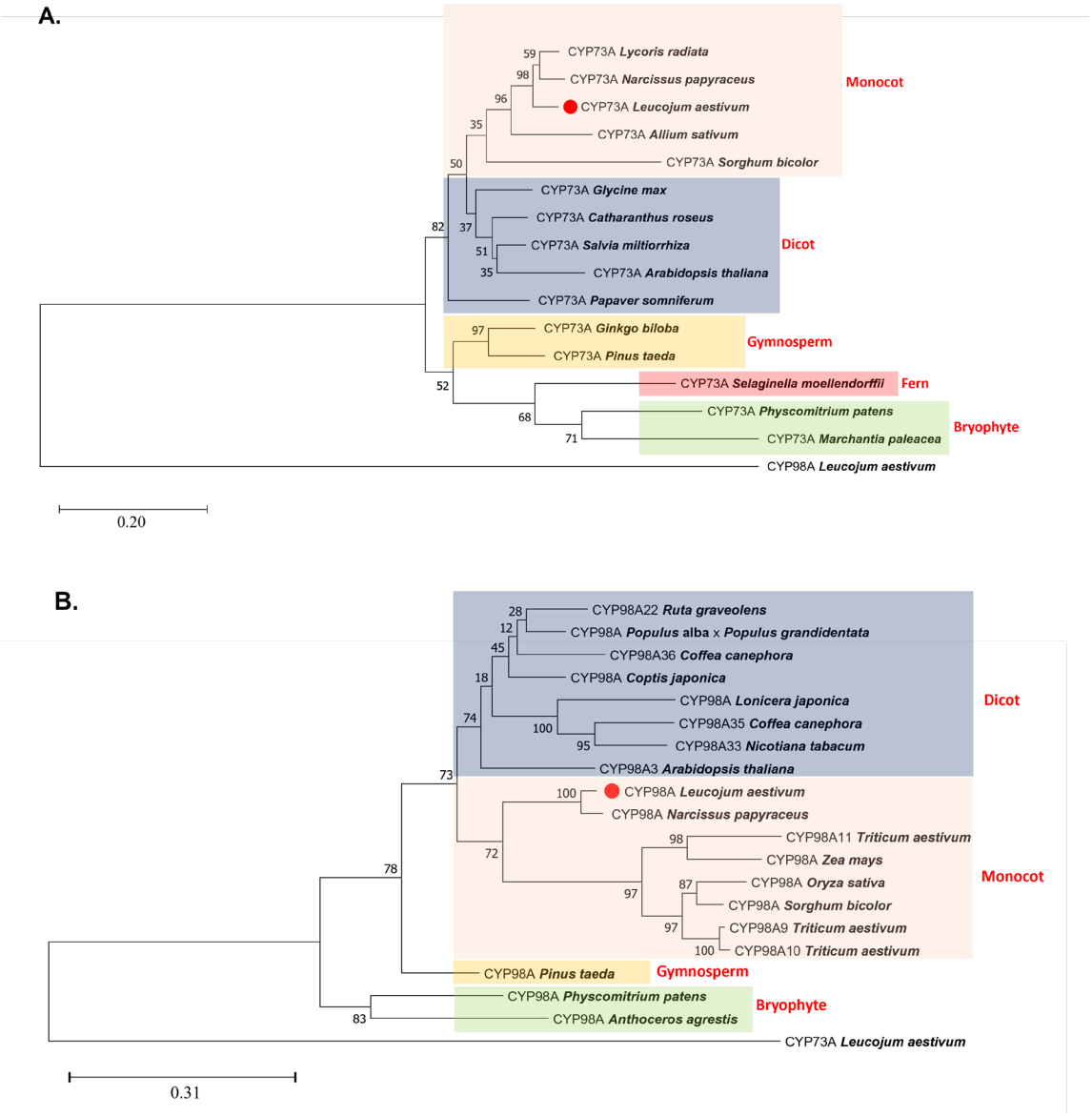


Fig 2-2. The phylogenetic tree of *L. aestivum* enzymes, *Lae*C4H and *Lae*C3′H, indicated with circles (●). The analysis was performed using the Maximum Likelihood method with 1,000 bootstrap replicates with the Mega11 software for **A.** *Lae*C4H of species *L. radiata* (AWW24970), *N. papyraceus* (AXU39895), *L. aestivum* (UIP35210), *Allium sativum* (ADO24190), *S. bicolor* (AAK54447), *Glycine max* (ACR44227), *C. roseus* (CAA83552), *Salvia miltiorrhiza* (ABC75596), *A. thaliana* (AAC99993), *Papaver somniferum* (XP_026426522), *G. biloba* (AAW70021), *Pinus taeda* (AAD23378), *S. moellendorffii* (EFJ22128), *P. patens* (ADF28535), *M. paleacea* (ASA39648). The CYP98A (UIP35212) from *L. aestivum* was used as an outgroup; **B.** *Lae*C3′H with species *Ruta graveolens* (AEG19446), *Populus alba* × *Populus grandidentata* (ABY85195), *Coffea canephora* (ABB83676, ABB83677), *Coptis japonica* var. dissecta (BAF98473), *Lonicera japonica* (AGQ48118), *Nicotiana tabacum* (ABC69384), *A. thaliana* (NP 850337), *L. aestivum* (UIP35212), *N. papyraceus* (AXU39897), *Triticum aestivum* (CAE47491, CAE47489, CAE47490), *Zea mays* (PWZ32976), *Oryza sativa* (AAU44038), *S. bicolor* (XP_002440001), *P. taeda* (AAL47685), *P. patens* (XP_024360823), *Anthocero agrestis* (QPI70542), CYP73A (UIP35210) from *L. aestivum* was used as an outgroup. The bootstrap values are indicated at the branch points.

*Lae*C4H and *Lae*C3′H belong to the CYP73A and CYP98A enzyme families, respectively. As CYP450, they share multiple domains, such as a common N-terminal membrane-anchoring domain that enables binding to the cytoplasmic surface of the endoplasmic reticulum membrane [169, 170]. Using multiple sequence alignment, we identified the characteristic membrane-anchoring domain in *Lae*C4H (Figure A1-1), which appeared to be highly conserved among the C4Hs from monocot and dicot plants but less so within other plant groups such as bryophytes. By contrast, for *Lae*C3′H, no conserved pattern was observed in the membrane binding domain region of C3′H homologs (Figure A1-2). The proline-rich motif (PPGPLPV) is another conserved signature of CYP450 that is hypothesized to be involved in the folding and proper integration of heme and the stability of microsomal proteins that was observed in both enzymes (Figure A1-1 and A1-2) [170-172]. The helix I motif (AAIET including conserved Ala306 and Ala307), speculated to play a role in proton transfer and oxygen activation, was conserved in all aligned C4H sequences, including *Lae*C4H (Figure A1-1) [173]. Similarly, a proposed oxygen binding motif (AGMDT) [174] was found to be conserved among C3′H homologs. The heme-binding domain (FGVGRRSCPG in C4H and FGAGRRVCPG in C3′H) is a key feature of the CYP450 superfamily, which was indeed preserved in both *Lae*C4H and *Lae*C3′H along with other plant species (Figure A1-1 and A1-2). This segment encompasses the cysteine pocket enclosing the heme and interaction with Cys447-S in a hydrophobic environment.

2.4.2 Prediction of substrate interaction for *LaeC4H* and *LaeC3'H*

*Lae*C4H and *Lae*C3′H structures and active sites were modeled to gain further understanding of their activity. *Lae*C4H and *Lae*C3′H were structurally similar, visible as a α-helix-rich triangular shape, like other CYP450s (Figure 3A and B, Figure A1-3, Table A1-4). The active site in both enzymes forms a large canal crossing the enzyme. The heme pocket is deeply buried within the protein large cavity, and two lobes extend from this stem on both sides (Figure 2-3A and B, Table A5). The left side lobe lined by the F and G α-helices is speculated to be the substrate entry port (Figure A1-3) [147]. The disparity between *Lae*C3′H and *Lae*C4H in the sequences of the A, F, and G helices and the connecting loop with B′ α-helix likely confers unique specificity onto each CYP450 enzyme towards various substrates. The right extension is longer in the *Lae*C3′H predicted structure, possibly allowing the binding of larger substrates. Molecular docking was used to study substrates' orientation, affinity, and interaction with *Lae*C4H and *Lae*C3′H. The

interactions between *trans*-cinnamic acid and *Lae*C4H were mapped based on comparing similar

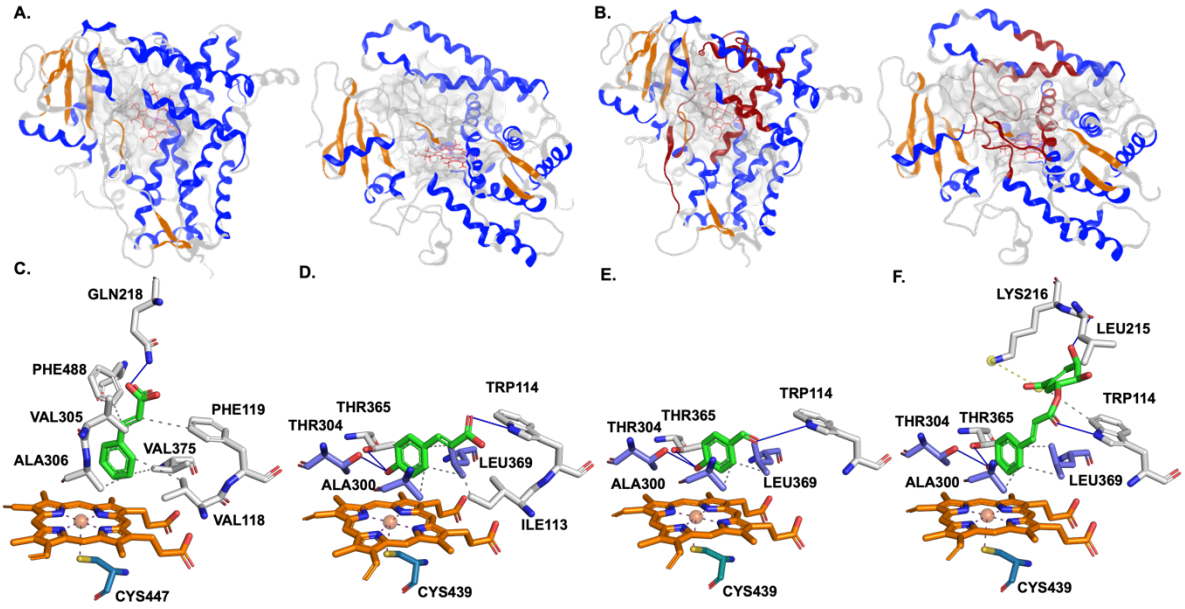


Fig 2-3. Representation of LaeC4H and LaeC3′H structures. A. Cartoon representation of *Lae*C4H with grey transparent surface-active site, top of the heme view (left) and perpendicular side of the heme plan (right). **B.** Cartoon representation of *Lae*C3′H with grey transparent surface-active site, top of the heme (left), and perpendicular side of the heme plan (right). Heme group is shown as red sticks. Conserved α-helix (blue) and β-strands (orange) secondary structures are shown. Conserved CYP98A substrate recognition is displayed in red. **C.** *trans*-cinnamic acid (green) interactions with *Lae*C4H residues at its active site; **D**, **E,** and **F.** interactions of *p*-coumaric acid (green), 4-HBA (green), and *p*-coumaroyl shikimate (green) with *Lae*C3′H at its active site. Residues are shown as sticks. The heme group is orange; grey residues interact with substrates only; purple residues interact with both heme and substrate. H-bonds are shown as blue lines, hydrophobic interactions are shown as dashed green lines, and salt bridges are shown as dashed yellow lines.

enzymes in other plant species (Figure 2-3C, Table A1-5). HEPES was first docked into *Lae*C4H to validate our model by comparison with previously crystallized structures. It docked in an axial position on the *Lae*C4H active site with a score of –7.32 kcal.mol^{-1} in a position analogous to HEPES in the crystalized structure of *Sb*C4H (PDB: 6VQY), reflecting the affinity of the enzyme for this substrate [147] (Figure A1-4). *trans*-cinnamic acid docked with a score of –5.12 kcal.mol^{-1} (Figure 2-3C). The smaller atomic size of this ligand compared to HEPES could be reflected in a smaller score [175]. Nonetheless, it positioned itself similarly at the stem, on the distal side of the heme, opposite to the axial thiolate, stabilized by hydrophobic interactions with Val118, Phe119, Val305, Ala306, Val375, and Phe488 with its phenylpropene portion, as well as an H-bond between Gln218 located in the F helix and the O at the carboxyl group of the substrate (Figure 2.3C). Positioning and detected interactions were similar to those in a previous study on *Sbi*C4H

[147]. The substrate phenol ring C4 is at 4.2 Å of the heme, a positioning consistent with the ferrous ion being ligated to the molecular oxygen, yielding hydroxylation of this atom to obtain *p*-coumaric acid.

The interactions between *p*-coumaric acid, 4-HBA, or *p*-coumaroyl shikimate substrates and the active site of *Lae*C3′H were also studied (Figure 2.3D, E, and F). These bulkier ligands were included as possible substrates to clarify the controversy surrounding the catalytic activity of *Lae*C3′H, probing the compatibility with its active site. *p*-coumaric acid, 4-HBA, and *p*-coumaroyl shikimate docked with a score of -5.52, -4.67, and -7.04 kcal/mol, respectively (Table A1-5). Upon docking, substrates oriented analogously to that of *p*-coumaroyl shikimate with wheat CYP98A10 and CYP98A11 [176], *i.e.,* their phenylpropanoid moiety close to the heme center in a bent conformation, aromatic carbon C3 oriented toward the heme at a distance of 3.7 (*p*-coumaroyl shikimate), 3.9 (*p*-coumaric acid) and 4.1 Å (4-HBA) (Figure 2.3D, E and F). For all substrates, the propylbenzene group was stabilized by: 1) hydrophobic interactions with Ala300 and Leu369 and 2) hydrogen bonds with Thr304 and Thr365. Trp114 formed a hydrogen bond with the carbonyl group of all three substrates. *p*-coumaric acid interacted with the active site through additional hydrophobic bonds with Ile113 and Thr365, Ala300, Thr304, and Leu369 participated in both substrate interaction stabilization through hydrophobic bonding and stabilization of the heme. A hydrogen bond with Leu215 and a salt bridge with Lys216 stabilized the shikimate portion of *p*-coumaroyl shikimate (Figure 2-3F). Potential electron donors and acceptors were present in the site interacting with all substrates, and the orientation of each docked substrate appeared to be favorable for its *m*-hydroxylation. *p*-coumaric acid and 4-HBA docking scores were lower, with fewer H-bonds and no salt bridge between the active site residues and the substrates (Table A1-5), and a higher distance to the heme, compared to *p*-coumaroyl shikimate. These results suggest that *p*-coumaroyl shikimate is the most suitable substrate for *Lae*C3′H.

2.4.3 Recombinant protein expression in yeast and bacterial host and *in vitro* enzymatic assays

CYP450 enzymes are commonly known to be localized to the ER membrane by an N-terminal hydrophobic anchor, which is necessary for optimal functionality. We, therefore, chose a eukaryotic yeast expression system to provide canonical localization. CYP450s were cloned into vectors containing the requisite redox partner, a CYP450 reductase from *C. roseus* (*Cro*CPR).

Microsomal fractions were isolated from strains expressing either *Lae*C4H-Myc or *Lae*C3′H-Myc recombinant proteins. Expression was confirmed by SDS-PAGE and Western blot analysis, which revealed the expected bands. (Figure A1-5).

The microsomal fractions were subjected to *in vitro* enzymatic assay. *Lae*C4H microsomal fractions were supplemented with *trans*-cinnamic acid, and the reaction mixture was incubated for 10 and 30 min, 1 and 2 h, and overnight at 30 °C. Representative results corresponding to 30 min incubation time are shown (Figure 2-4A). The reaction product, *p*-coumaric acid, was detected at all incubation periods above 10 min (Figure A1-6). Thus, the results showed that C4H from *L. aestivum* can efficiently hydroxylate *trans*-cinnamic acid to produce *p*-coumaric acid.

Similarly, recombinant *Lae*C3′H-Myc was tested with various substrates, including *p*-coumaroyl shikimate, identified as the most probable substrate, free *p*-coumaric acid, and 4-HBA. The reaction mixtures were subjected to HPLC-MS/MS analysis. A signal corresponding to caffeoyl shikimate in the test tube supplemented with *p*-coumaroyl shikimate was detected at all time points (Figure 2-4B and Figure A1-7). In contrast, no reaction product, including no signal for caffeic acid or DHBA, was detected when *p*-coumaric acid or 4-HBA was used as substrates.

The kinetic parameters of *Lae*C4H and *Lae*C3′H were determined using enzymatic assays. Microsome fractions were maintained at a constant concentration, while the concentration of the respective substrates varied from 100 nM to 30 μM (Figure A1-8). The maximum rate of reaction (Vmax) and Km of the enzymes were determined using the Michaelis-Menten equation and non-linear regression analysis (Table 2-1). *Lae*C4H (Km 3.32 μM; Vmax 0.17 μM/min) shows higher affinity and speed of catalysis toward its substrate than *Lae*C3′H (Km 22.16 μM; Vmax 5.89 $\mu M.min^{-1}$).

Table 2.1. Kinetic parameters for *Lae*C4H and *Lae*C3′H.

***Lae*C4H**	0.12	1.21
***Lae*C3′H**	5.89	22.16

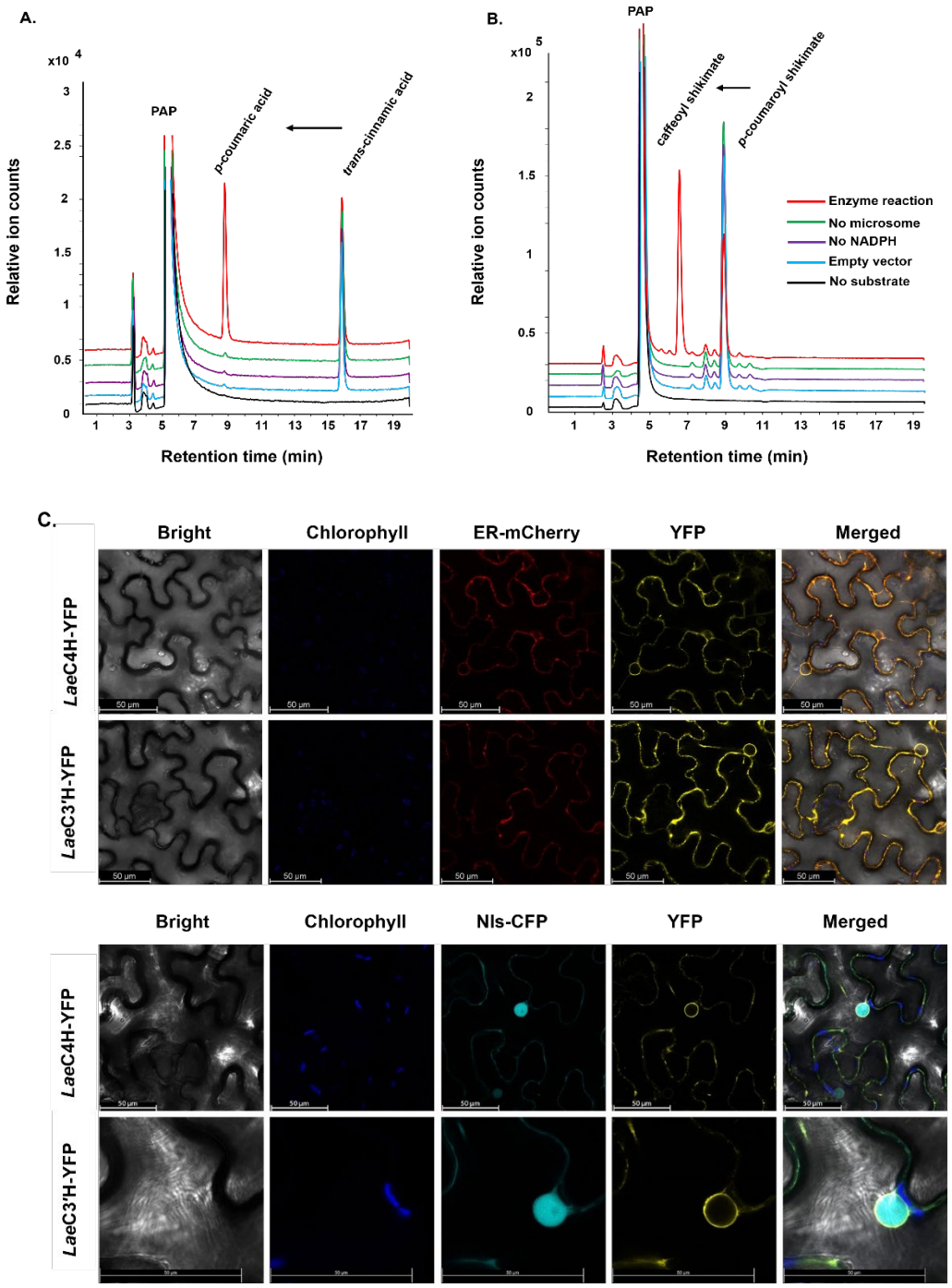


Figure 2-4. HPLC-MS/MS analysis of representative enzymatic assay for *Lae*C4H (A) and *Lae*C3′H (B), and their subcellular localization in *N. benthamiana* (C). A. *trans*-Cinnamic acid was tested on microsome fractions extracted from yeast harboring the *Lae*C4H-Myc construct, and *p*-coumaric acid (red) formation was detected; **B.** Assays with microsomes containing *Lae*C3′H-Myc using *p*-coumaroyl shikimate, and the samples were subjected to the detection of caffeoyl shikimate. Red peaks resulting from the sample containing microsomal recombinant proteins (*Lae*C4H-Myc and *Lae*C3′H-Myc); Purple peaks from the sample without NADPH; Blue

peaks from the sample containing empty vector microsomal fractions; Green peaks from the sample without recombinant protein; and black represents the peak from the sample without substrate. Papaverine (PAP) was used as a standard control; **C.** Subcellular localization of *Lae*C4H and *Lae*C3′H in *N. benthamiana* epidermal leaf. Fusion fluorescent proteins were co-expressed with ER marker (ER-mCherry) (upper panel) and nucleus marker (Nls-CFP) (bottom panel) in N. benthamiana leaves by the Agrobacterium-mediated transient expression method, and subsequently, the fluorescent signals were detected by confocal microscopy 48 h after co-infiltration. The scale bars in the pictures indicate 50 µm.

*Lae*APX/C3H was expressed in a bacterial system and validated using SDS-PAGE (Figure A1-9). Enzymatic activity of *Lae*APX/C3H was performed using *p*-coumaric acid and 4-HBA to assess 3-hydroxylation. A reaction mixture without *Lae*APX/C3H was used as a negative control. Caffeic acid was detected in the enzymatic reaction supplemented with *p*-coumaric acid; however, caffeic acid was higher in the negative control (Figure A1-10 and A1-11). Using an increasing amount of enzyme in the reaction, p-coumaric acid progressively disappeared, but the caffeic acid level did not rise above the level of the negative control (Figure A1-11). This suggests that the enzyme consumed *p*-coumaric acid, but caffeic acid was not the corresponding product. This result is consistent with the study of [115] reporting that the caffeic acid level in the reaction without and with a homologous APX/C3H from monocot *S. bicolor* was not significantly different. DHBA was detected in both enzymatic reactions and in the negative control with 4-HBA as a substrate. No significant changes were observed in 4-HBA consumption by increasing the enzyme concentration, suggesting that the enzyme did not accept 4-HBA as a substrate (Figure A1-11).

2.4.4 Subcellular localization of *Lae*C4H and *Lae*C3′H in *N. benthamiana*

C-terminal YFP-fusion *Lae*C4H and *Lae*C3′H were co-expressed with an Endoplasmic reticulum marker (ER-mCherry) and a nucleus marker (Nls-CFP) in *N. benthamina* leaf using agroinfiltration to explore their subcellular localization. As expected, *Lae*C4H and *Lae*C3′H colocalized to the ER (Figure 2-4C, upper panel), which is consistent with previously reported cellular localizations for both enzymes from other species. Co-expression with a nucleus marker indicated that the target proteins are not localized inside the nucleus. Yellow fluorescence was observed in the vicinity of the nuclear membrane, which would suggest the localization of both enzymes in the membrane region closely associated with the nuclear envelope (Figure 2-4C, bottom panel). The integrity of the expressed fusion proteins was validated by western blot analysis using an anti-YFP primary antibody (Figure A1-12).

*Lae*C4H and *Lae*C3′H were transiently expressed, and their respective substrates were fed to *N. benthamiana* leaves to confirm their functionality *in planta.* Unexpectedly, upon expressing

*Lae*C4H and feeding with *trans*-cinnamic acid, *p*-coumaric acid was not detected, but *p*-coumaroyl shikimate was observed (Figure 2-5 and Table A1-3). This suggested that *Lae*C4H expression yielded *p*-coumaric acid from *trans*-cinnamic acid, but the latter was rapidly converted to *p*-coumaroyl shikimate by endogenous enzymes (Figure 2-1). Interestingly, caffeoyl shikimate, produced by the endogenous C3′H, was also detectable in these plants. To gain further precision, samples were concentrated 10 times, revealing that plants infiltrated with *Lae*C4H and *trans*-cinnamic acid had increased production of three downstream products, *i.e., p*-coumaroyl shikimate, caffeoyl shikimate, and caffeic acid. In corresponding controls, these compounds were not detected. In plants overexpressing *Lae*C3′H and fed with *p*-coumaroyl shikimate, caffeoyl shikimate, the direct product of this enzyme, was detectable as well, whereas plants fed with *p*-coumaroyl shikimate without *Lae*C3′H demonstrated no detectable level of caffeoyl shikimate. Additionally, caffeic acid, a downstream product of these enzymes, also accumulated to detectable levels in plants overexpressing the enzymes, providing indirect evidence for the activity of *LaeC4H* and *LaeC3'H* (Figure 2-5 and Table A1-3) *in planta*.

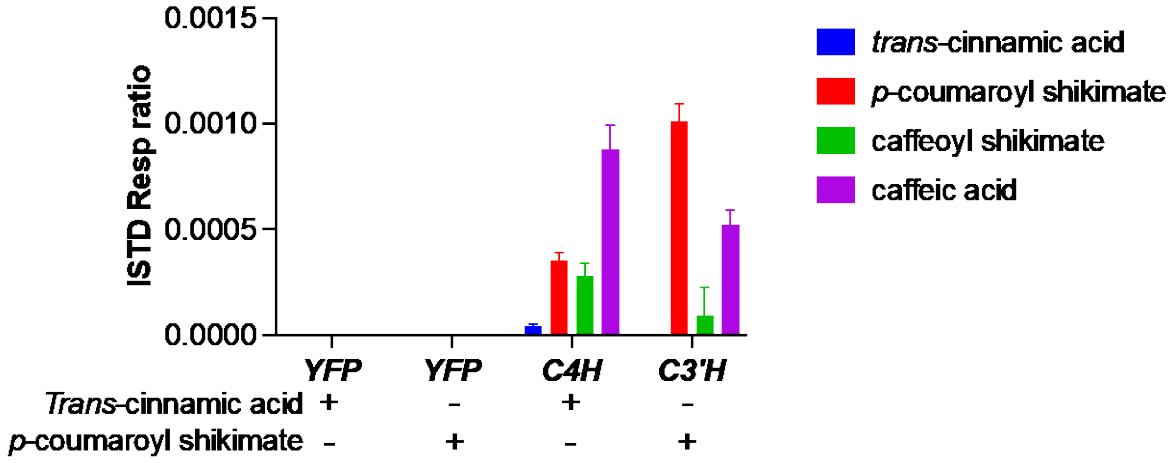


Figure 2-5. *In vivo* enzymatic activity of *Lae*C4H and *Lae*C3′H. Metabolite extracts from samples listed in Table A1-3 were concentrated and analyzed using HPLC-MS/MS. Control samples included: 1) plants without infiltration and feeding to distinguish the level of endogenous phenolic compounds from those altered by experimental conditions, 2) plants expressing the YFP, and 3) YFP-expressed plants fed with the substrates to assess the effect of target genes versus potential effects of agrobacterium infection on phenolic compounds. NA: not applicable. This experiment was performed in biological duplicates of 5 leaf pooled samples.

2.4.5 Metabolite profiling of *L. aestivum* plant organs

Targeted LC-MS/MS analyses of metabolite extracts from different parts of *L. aestivum* were performed to identify and measure relative metabolite levels for intermediates of AA biosynthesis (Figure 2-6). Tyrosine and phenylalanine were detected at higher levels in the bulb and flower tissues (Figure 2-6A). Tyramine was detected in the bulb and flower tissues but primarily accumulated in the roots (a 2-fold increase compared to the bulb). *p*-coumaric acid, caffeic acid, and ferulic acid were detected only in the flowers. In general, AA was detected in several tissues and higher quantities in the bulbs and flowers compared to other tissues. Specifically, lycorine accumulated almost 4- and 2-fold more in bulbs and flowers as compared to leaves and roots respectively. Norgalanthamine, vittatine and 11-hydoxyvittatine were more abundant in the flower as compared to other parts of *L. aestivum*. Interestingly, flowers were the only tissues to concurrently accumulate all the tested precursor metabolites and intermediates of AA biosynthesis.

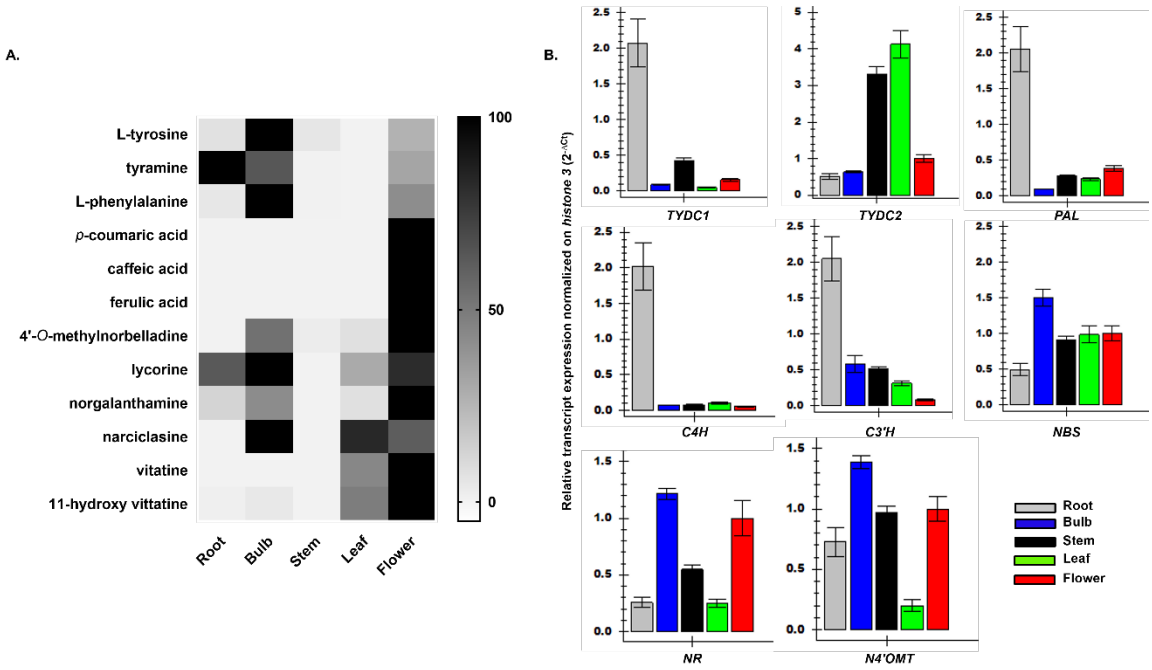


Figure 2-6. Metabolic profile and expression analyses. A. Targeted metabolite profiling from different tissues of L. *aestivum.* Heat Map showing the relative abundance of AAs (precursors, intermediates, and end products) from different parts of *L. aestivum*. Relative abundance corresponds to the mean value of three independent replicates. Values were normalized to the sample with the highest level for each compound in different tissues; **B.** qRT-PCR analyses of precursor and AAs specific genes from different plant tissues of *L. aestivum,* including bulb, flower, leaf, root, and stem. The genes include *phenylalanine ammonia lyase* (*PAL*), *tyrosine decarboxylase* (*TYDC1*, *TYDC2*), *C4H*, *C3'H*, and several AA-specific genes such as *norbelladine synthase* (*NBS*), *noroxomaritidine/norcraugsodine reductase* (*NR*) and *norbelladine 4-O-methyltransferase* (*N4'OMT*). Three biological and two technical replicates were performed for each gene.

2.4.6 Differential Expression Analysis

Both C3′H and C4H are key enzymes of the phenylpropanoid pathway. Their differential expression and transcript levels of genes encoding enzymes involved in the biosynthesis of AA precursors and end-products were explored in different plant tissues of *L. aestivum* (Figure 2-6B). RT-qPCR analysis showed that transcripts involved in the phenylpropanoid pathway were expressed higher in roots than in other parts of the plant. *PAL*, *C4H*, *C3'H*, and *TYDC1* were all expressed at significantly higher levels in roots (Figure 2-1). Relatively higher expression of *TYDC2* and C3′H along with other AA-specific genes, including *NBS*, *NR*, and *N4'OMT* (Figure 2-6B) in alkaloid-containing tissues could highlight their potential role in providing the intermediates to AA biosynthesis. The high abundance of tyramine in roots was consistent with increased transcript levels of *TYDC1* compared to other tissues. Unlike *TYDC1*, TYDC2 was mostly expressed in leaves and stems unrelated to the tyramine content of those tissues.

2.5 Discussion

Amaryllidaceae plants accumulate SMs, such as specific alkaloids known for their therapeutic potential. Their production is controlled by the enzymes involved in their biosynthesis and the formation of phenylpropanoids through a complex process. The phenylpropanoid pathway also catalyzes the formation of the precursors of lignin, lignans, and flavonoids. The biosynthesis of the AA precursor DHBA starts with phenylalanine and requires two hydroxylation reactions [8]. C4H was shown to introduce the first hydroxyl group at the 4-position of the aromatic ring of *trans*-cinnamic acid in several plant species [177] (Figure 2-1). The enzyme responsible for the 3-hydroxylation remained uncharacterized for a long time, but in recent years, C3′H (CYP98) was proposed as the major 3-hydroxylase in the phenylpropanoid pathway in *A. thaliana* [155]. Other studies suggested that APX/C3H rather catalyzed it in *Brachypodium distachyon* [10] but not in *Sorghum bicolor* [156]. In the current study, a survey of the transcriptome data of *L. aestivum* [52] allowed the identification of putative *Lae*C4H, *Lae*APX/C3H, and *Lae*C3′H genes.

The *Lae*C4H sequence was highly similar to other monocot C4Hs with distinctive conserved domains and active residues. This suggests a very low level of evolutionary pressure during the diversification of the land plants. In the predicted model of *Lae*C4H, *trans*-cinnamic acid was docked in a position consistent with its 4-hydroxylation, yielding *p*-coumaric acid, as observed in

previous reports [134, 153, 178, 179]. We further confirmed that microsomes extracted from *Lae*C4H-transformed yeast catalyze the 4-hydroxylation of *trans*-cinnamic acid *in vitro*. The Km value for *Lae*C4H towards cinnamic acid was consistent with values reported elsewhere [148], supporting *trans*-cinnamic acid as the C4H substrate. Previous work on the heterologous expression of truncated Amaryllidaceae C4H from *L. aurea* and *L. radiata* in bacteria showed that the enzyme was functional without its N-terminal membrane domain [134, 153]. However, the latter mediates the anchoring to the exterior surface of the ER [180] and ensures the correct sub-cellular localization. This localization facilitates interactions with the electron donor reductase partner and other pathway enzymes [28], potentially promoting the formation of a metabolon for efficient metabolite channeling [181].

The *Lae*C3′H sequence was also similar to those of monocots, although a little more distant than *Lae*C4H. Alber et al. have discussed the evolutionary analysis of the C3′H family. [109], reporting that they had a single origin in a common ancestor of all land plants. The study highlighted multiple gene duplications within angiosperms, leading to small clusters of gene families in most species.

Previous studies revealed a wide range of possible substrates for C3′H. In dicot and monocot angiosperms, *p*-coumaroyl shikimate has been reported to be the preferred substrate for the CYP98A enzyme family [109]. In *A. thaliana*, C3′H (CYP98A3) also catalyzed 3′-hydroxylation of *p*-coumaroyl quinate [155], whereas C3′H (CYP98A8 and CYP98A9) showed 3′-hydroxylase activity toward spermidine-conjugated phenolics [182], and naringenin (in the case of CYP98A9) [183]. In wheat, a CYP98A12 isoform could hydroxylate *p*-coumaroyl tyramine to caffeoyl tyramine that was subsequently methylated to feruloyl tyramine as part of a pathogen-induced defense response [184]. Considering the substrate promiscuity of C3′H in different plant groups, we hypothesized that *L. aestivum* C3′H was capable of hydroxylating 4-HBA along with previously reported substrates like *p*-coumaroyl shikimate and free *p*-coumaric acid. Modeled structure and docking analysis suggested that the *Lae*C3′H pocket was more consistent with 3-hydroxylation of the large substrates such as *p*-coumaroyl shikimate stabilized by key additional hydrophobic interactions, H-bonds, and a salt bridge with Lys216 with the shikimate portion, compared to HBA or *p*-coumaric acid. This was confirmed by *in vitro* enzymatic assay yielding caffeoyl shikimate from *p*-coumaroyl shikimate, whereas C3′H did not hydroxylate other substrates. Thus, *Lae*C3′H does not catalyze the synthesis of caffeic acid or DHBA.

Recently it has been found that direct hydroxylation of free *p*-coumaric acid is instead catalyzed by a bifunctional ascorbate peroxidase/4-coumarate 3-hydroxylase, a non-membrane enzyme purified from different plant species [10]. We identified and cloned a homologous *Lae*APX/C3H and successfully expressed it in a prokaryotic system. *Lae*APX/C3H did not catalyze the 3-hydroxylation of HBA or *p*-coumaric acid. Although caffeic acid was detected when *p*-coumaric acid was used as a substrate, it was more abundant in the reaction mixture lacking *Lae*APX/C3H, showing that its production was not enzyme-dependent. A progressive consumption of *p*-coumaric acid was detected with increased concentration of enzymes, implying that this enzyme consumes *p*-coumaric acid, possibly for polymerization, as evidenced in *sorghum bicolor* [115].

The enzymes were expressed to verify the activity of *Lae*C4H and *Lae*C3′H in *planta*, and their substrates were fed in *N. benthamiana* leaves. Infiltration with *Lae*C4H and feeding with *trans*-cinnamic acid led to the specific detection of *p*-coumaroyl shikimate, caffeoyl shikimate, and caffeic acid. In contrast, *Lae*C3′H combined with *p*-coumaroyl yielded caffeoyl shikimate and caffeic acid. Unexpectedly, *p*-coumaric acid, the direct product of *Lae*C4H, was not detected. We hypothesized that *p*-coumaric acid was converted to the detected downstream products by endogenous enzymes such as 4CL and HCT (Figure 2-1). In various plant species, *p*-coumaric acid can rapidly be metabolized to *p*-coumaroyl CoA, serving as a pivotal entry point for different metabolic pathways, including the biosynthesis of flavonoids, stilbenes, coumarins, monolignols of lignin, among others. Many of these compounds trigger defense mechanisms against wounds and external infections [185, 186]. Overall, these results provide *in planta* evidence of the functionality of *Lae*C4H and *Lae*C3′H.

Overall, we showed that C4H catalyzes the production of *p*-coumaric acid, which can be converted to *p*-coumaroyl-CoA and *p*-coumaroyl shikimate by 4CL and HCT, respectively. Then, C3′H catalyzes the production of caffeoyl shikimate, which can produce caffeic acid by caffeoyl shikimate esterase (CSE) [110]. The enzymes were detected in the ER. *A. thaliana* C3′H (CYP98A3) and C4H (CYP73A5) were shown to colocalize and dimerize in the ER of the transfected *N. benthamiana* leaf and to associate with HCT and 4CL1 as soluble partners to form an enzyme complex of the phenylpropanoid pathway [187]. This suggests their ER localization is pivotal to efficiently channeling metabolites between the branches. Future studies are required to show whether these enzymes are part of a metabolon in Amaryllidaceae plants.

At expression levels, transcripts also shared similar tissue repartition. *PAL*, *C4H*, *C3'H*, and *TYDC1* in *L. aestivum* showed the highest relative expression levels in roots, in which the precursors L-tyrosine, tyramine, and L-phenylalanine were detected. Their co-expression pattern is consistent with the hypothesis that they could complex together to form a metabolon. In *L. radiata*, gene expression analysis revealed higher *PAL* and *C4H* expression in root tissues than in bulb and leaf. Interestingly, the products of phenylpropanoid enzymes, such as *p*-coumaric acid, caffeic acid, and ferulic acid, were predominantly detected in flowers. AA biosynthetic genes (*NBS*, *NR*, and N4′OMT) were expressed in all tissues, and N4′OMT) were expressed in all tissues, and N4′OMT) was expressed in all tissues but particularly abundant in bulbs and flowers. In this latter tissue, in addition to the phenylpropanoid precursors necessary for AA synthesis, 4'-*O*-methylnorbelladine, lycorine, haemanthamine, narciclasine, norgalanthamine, vittatine, and 11-hydroxyvittatine were all detected. Our finding supports an inverse correlation of the upstream phenylpropanoid genes expression pattern with the phenylpropanoid and alkaloid metabolites content in the same tissue [134]. Inter- and intra-cellular transport of phenylpropanoid compounds between cell compartments and also different cell types could occur between different tissues of the Amaryllidaceae plant [188]. To our knowledge, nothing has been reported on the transport or single-cell multi-omics in Amaryllidaceae which could help elucidate further alkaloid metabolism in this plant family.

The phenylpropanoid pathway also catalyzes the formation of the precursors of lignin, lignans, flavonoids, and alkaloids in Amaryllidaceae. Screening for and identifying the *Lae*C4H and *Lae*C3′H genes is crucial for elucidating AA metabolism in *L. aestivum*.

2.6 Conclusion

In the present study, we identified C4H, APX/C3H, and C3′H in the *L. aestivum* transcriptome. The sequences of *Lae*C4H and *Lae*C3′H shared a high identity with *bona fide* C4Hs and C3′H enzymes from several species and harbored typical CYP450 domains. The recombinant *Lae*C4H and *Lae*C3′H proteins were successfully expressed in yeast. *Lae*C4H catalyzed the expected 4-hydroxylation of *trans*-cinnamic to *p*-coumaric acid, whereas *Lae*C3′H catalyzed the 3-hydroxylation of *p*-coumaroyl shikimate to caffeoyl shikimate. *Lae*APX/C3H used *p*-coumaric acid as substrate but did not form caffeic acid. This work unravels the reactions involved in the first key steps of the biosynthesis of AA. Deciphering AA precursor biosynthesis will facilitate the

development of the biosynthetic tools required to produce AA *in vitro* and help produce important pharmaceuticals, such as galanthamine, heterologous hosts to treat the symptoms of Alzheimer's disease.

2.7 Credit Author Statement

Vahid Karimzadegan: Conceptualization, Methodology, Investigation, Formal analysis, Validation, Visualization, Writing- Original draft; **Manoj Koirala:** Methodology, Resources, Writing- Reviewing and Editing; **Sajjad Sobhanverdi:** Methodology, Investigation, Formal analysis, Validation, Visualization; **Natacha Merindol:** Conceptualization, Methodology, Project administration, Resources, Writing- Original draft, Writing- Reviewing and Editing; **Bharat Bhusan Majhi:** Methodology, Resources, Writing- Reviewing and Editing; **Sarah-Eve Gélinas:** Methodology, Resources, and Editing; **Vitaliy I. Timokhin:** Methodology, Resources, Chemical synthesis and Editing; **John Ralph:**, Methodology, Resources, Chemical synthesis, supervision and Editing; **Mehran Dastmalchi:** Supervision, Writing- Reviewing and Editing; **Isabel Desgagné-Penix:** Conceptualization, Methodology, Resources, Funding acquisition, Supervision, Writing- Original draft, Writing- Reviewing and Editing

2.8 Acknowledgments

The authors thank Professor Yang Qu (University of New Brunswick, Fredericton, NB, Canada) for generously providing yeast vector pESC-LEU-*Cro*CPR for protein expression. We also wish to thank Dr. Elisa Fantino for precious technical advice during this study, and Mélodie B. Plourde and Snehi Gazal for their kind technical assistance and patience with confocal microscopy acquisition and analyses. We thank the reviewers and editors for their careful reading and helpful comments on this manuscript. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) award number RGPIN-2021-03218 to I.D-P and the Canada Research Chair on Plant Specialized Metabolism Award No CRC-2018-03218 to I.D-P. Many thanks are extended to the Canadian taxpayers and government for supporting the Canada Research Chairs Program. V.I.T. and J.R. were supported by the DOE Great Lakes Bioenergy Research Center (DOE Office of Science BER DE-SC0018409).

2.9 Declaration of Interests

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Chapter III

This chapter discusses the establishment of the AAs core pathway in baker's yeast using the CRISPR-Cas9 multiplex platform. It was conducted as a collaborative internship project at the University of Calgary in Professor Dae-Kyun Ro's lab. The chapter was written in article format, and its content might overlap with other chapters. Due to time constraints, the results are not finalized at the time of writing and may be updated in the final deposit of the thesis.

Contributions: Vahid Karimzadegan carried out pathway construction, feeding assays, and writing the manuscript draft. Rahul Kumar performed the LC-MS analysis. Sajjad Sobhanverdi performed the second feeding assay, metabolite extraction. Sarah-Eve Gélinas performed the metabolite analysis of the second metabolite extracts. Basanta Lamichhane provided vectors harboring NBS, NR, and N4′OMT coding sequences. Dae-Kyun Ro supervised the internship project and provided space and materials for it. Isabel Desgagné-Penix supervised the project and assisted with revising the manuscript draft. Mehran Dastmalchi co-supervised the project and assisted with the revision.

3. Establishment of the Amaryllidaceae alkaloid core pathway in yeast saccharomyces cerevisiae using CRISPR-Cas9 system

Vahid Karimzadegan[a], Rahul Kumar[b], Sajjad Sobhanverdi[a], Basanta Lamichhane[a], Dae-Kyun Ro[b], Mehran Dastmalchi[c], and Isabel Desgagné-Penix[a, *]

[a] Department of Chemistry, Biochemistry and Physics, Université du Québec à Trois-Rivières, Trois-Rivières, Québec, Canada.

[b] Department of Biological Sciences, University of Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada.

[c] Department of Plant Science, McGill University, Montréal, Québec, Canada

*Corresponding author: Isabel Desgagne-Penix, Department of Chemistry, Biochemistry and Physics, Université du Québec à Trois-Rivières, 3351 boul. des Forges, Trois-Rivières, QC, G9A 5H7, Canada. (Tel: 819-376-5011, Fax: 819-376-5014).

E-mail: Isabel.Desgagne-Penix@uqtr.ca

3.1 Abstract

Amaryllidaceae alkaloids (AAs) are a large group of plant-derived specialized metabolites, many of which have valuable pharmaceutical or biological activity. There are ~650 AAs produced by thousands of plant species of the Amaryllidaceae family. The diverse chemical structures in this metabolite class originate from 4′-*O*-methylnorbelladine, the last common biosynthetic intermediate for all AAs enzymatic pathways. Reconstitution of biosynthetic pathways in a heterologous host is a promising strategy for the rapid and inexpensive production of complex molecules found in plants. Here, we demonstrate how 4′-*O*-methylnorbelladine can be produced in a *Saccharomyces cerevisiae* host from known AAs pathway genes. Tyrosine decarboxylase (TYDC), norbelladine synthase (NBS), noroxomaritidine/norcraugsodine reductase (NR), and norbelladine 4′-*O*-methyltransferase (N4′OMT) were either integrated in the yeast genome or introduced by plasmid, to produce 4′-*O*-methylnorbelladine from dihdroxybenzaldehyde (DHBA).

Our first experiment showed that 4′-*O*-methylnorbeladine was produced in the genome-integrated strain in particular. However, the low production rate and the detection of that in one of the negative control replicates in high substrate concentration (1mM) made it inconclusive. Ultimately, based on our second experiment, it can be said that genome-integrated strain produces norbelladine, and the absence of 4′-*O*-methylnorbeladine remains to be explored. This work is the first attempt to develop the production of complex AAs in a microbial system, and future works can advance the *de novo* AAs by assembling enzymes from different organisms when plugged into the core pathway

3.2 Introduction

The Amaryllidaceae plant family synthesizes a unique type of specialized metabolite known as Amaryllidaceae alkaloids (AAs), which are restricted to these plants. AAs' broad biological activities hold immense potential for future therapeutic applications. This family is considered among the most therapeutically important plant species [8]. For instance, galanthamine is the most well-known AA, displaying an acetylcholine esterase inhibitory effect. It is one of the few clinically important anti-Alzheimer drugs available today. Many other AAs, such as lycorine and haemanthaine, have been found to exhibit strong anti-cancer properties, while compounds like cherylline possess anti-viral activities [141, 189, 190]. They accumulate trace amounts in native plants, causing challenges for clinical purposes and research applications.

In recent years, CRISPR-Cas9 multiplex genome editing has been developed for yeast, allowing the targeting of multiple loci within the yeast genome. This advancement facilitates the metabolic engineering of complex, multi-enzymatic pathways in a time-efficient manner by employing multiple gRNAs that target different loci. A recent breakthrough in this field is the single gRNA-mediated (SGM)-CRISPR system, as reported by Baek et al. [5]. This method utilizes a single pre-installed synthetic gRNA binding site across multiple loci, enabling targeting these loci with just one gRNA. Baek et al. integrated six flunisolide biosynthetic pathway genes in the yeast genome using this system. Further advancements include the development of a modular multiplex genome-edit (MMG)-CRISPR platform and an extension of the SGM-CRISPR system. This platform incorporates an additional pre-integrated gRNA-binding site, allowing the integration of up to eight genes in two sequential transformation steps. Utomo et al. utilized this system to reconstitute the curcumin pathway in *S. cerevisiae* by introducing up to eight genes (Figure 3-1) [6].

Using the MMG-CRISPR platform, we aimed to produce intermediate 4′-*O*-methylnorbelladine in yeast *S. cerevisiae* (Figure 3-2). The biosynthesis of AA starts with the condensation of DHBA, resulting from the phenylpropanoid pathway, and tyramine that is formed through the decarboxylation of tyrosine [8]. The enzymes norbelladine Synthase (NBS) and noroxomaritidine/norcraugsodine reductase (NR) were reported to catalyze this condensation co-operatively [51]. Further *O*-methylation by norbelladine 4′-*O*-methyltransferase (N4′OMT) results in 4′-*O*-methylnorbelladine [53], which is a central intermediate for the majority of AAs. A new study identified a substantial number of biosynthetic enzymes in the Amaryllidaceae alkaloids pathway, including enzymes involved in the biosynthetic pathway of galantamine and haemanthamine, which are derived from 4′-*O*-methylnorbelladine through para-ortho' and para-

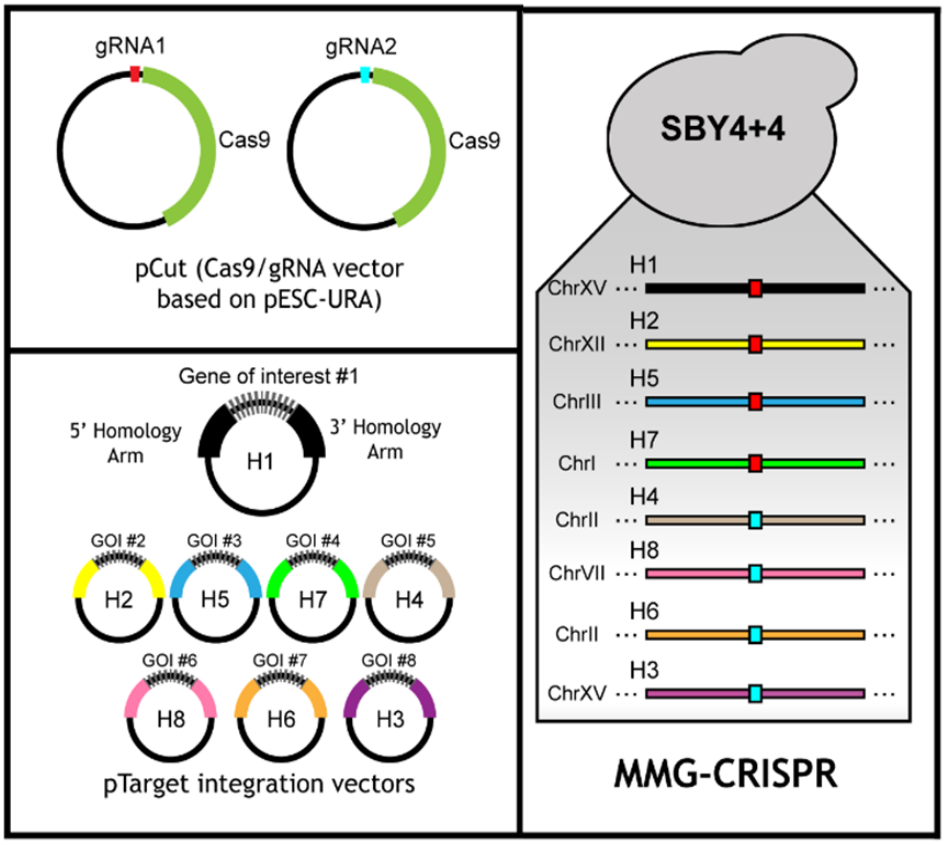


Figure 3-1. Schematic of Modular multiplex genome-edit (MMG)-CRISPR platform in yeast. As represented, for each gRNA, four identical synthetic landing pads have been inserted in four different loci in the yeast genome. This platform can integrate 8 genes in two sequential transformations. Combination of pCut plasmid encoding gRNA1 or gRNA2 and Cas9, with donor DNAs having homology arms corresponding to H1, H2, H5, H7 (target sites for gRNA1), and H3, H4, H6, H8 (target sites for gRNA2) will facilitate the integration of 8 genes in 8 different loci. For more details, refer to the source [6].

para phenol coupling reactions, respectively [133]. Therefore, future studies would plug the proposed central 4′-O-methylnorbelladine producing block into different downstream pathways, advancing microbial production of therapeutic Amaryllidaceae alkaloids from cheaper intermediates or simple sugar sources.

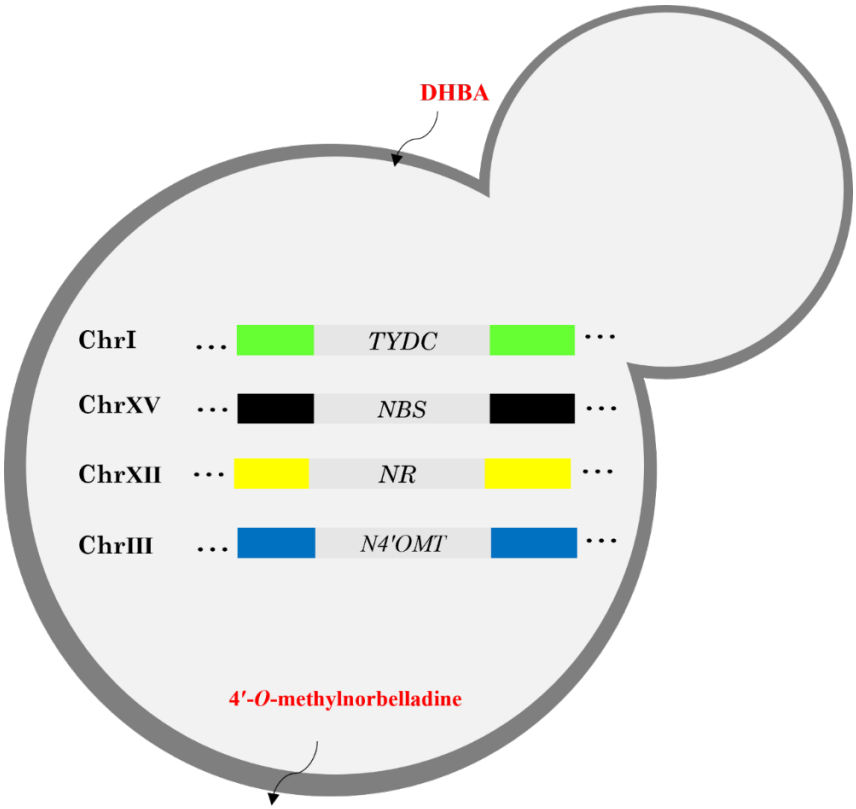


Figure 3-2. Yeast engineering for 4′-O-methylnorbelladine production. Genes in the core pathway are shown in gray boxes on different chromosomes. The different colors flanking the genes of interest represent unique homology arms matching with homology arms in the corresponding pTarget plasmids. DHBA is the substrate and 4′-*O*-methylnorbelladine is the expected metabolite.

3.3 Materials and Methods

3.3.1 DNA construction and yeast stable transformation using the Multiplex CRISPR-Cas9 platform.

The *S. cerevisiae* yeast strain SBY104 (MATα, HIS3Δ1, LEU2Δ0, LYS2Δ0, URA3Δ0) was used for CRISPR-Cas9-based genomic gene integration as described by Baek et al. (2021) [5]. The open reading frames of *NBS* (GenBank: MW971977.1), *NR* (GenBank: MW971982.1), *N4'OMT* (GenBank: MW971978.1), and *TYDC* (GenBank: MT468182.1) were cloned under galactose-inducible promoter (Gal1) in donor DNA plasmids containing left and right homology arms corresponding to the H1, H2, H5, and H7 sites in yeast, respectively. The resulting plasmids were transformed into *E. coli* TOP10 chemically competent cells.

Using appropriate primers listed in Table A2-1, linear donor DNA fragments were amplified from four donor plasmids using Q5 DNA polymerase (Figure A2-1). Yeast transformation was performed using the lithium acetate method outlined by Baek et al. (2021). Overnight yeast cultures were diluted 1:100 into 30 ml YPD media and grown until the OD_{600} reached 0.4–0.8. The cells were then collected by centrifugation and washed with sterile H_2O. Approximately 2×10^7 cells were mixed with 200 µl of 60% (w/v) PEG, 1 µg pESC-URA: Cas9-gRNA plasmid, and 2 µg of each linear donor DNA incubated for 15 minutes. Subsequently, 50 µl of 2 mg/ml boiled salmon sperm DNA and 18 µl of 2 M lithium acetate were added. The mixtures were incubated at 42 °C for 40 minutes, cooled on ice, centrifuged, and then incubated in synthetic complex (SC)-Ura with glucose media for 2 hours at 30 °C before being plated on SC-Ura plates. Plates were incubated overnight at 37 °C, followed by 2–3 days at 30 °C. Genomic gene integrations were confirmed by colony PCR using Phire Plant Direct PCR Master Mix (Thermo Fisher) using primers listed in Table A2-1. Each colony was genotyped at the integration sites using intergenomic primers, resulting in an 800 bp product for negative colonies and a 2–3 kb product for positive colonies.

3.3.2 Plasmid-based platform construction

The ORFs of *NBS* and *NR* were cloned under the GAL1,10 promoter in frame with FLAG and MYC tags, respectively, in the pESC-URA vector. The ORFs of the *N4'OMT and TYDC* were

cloned in a pESC-LEU plasmid in the same order. The cloning was performed using the Gibson assembly method [191]. Primers are listed in table A2-1. The resulting plasmids were transformed into *E. coli* TOP10 and two plasmids were co-transformed into the yeast BY4742 using the lithium acetate method described above. The yeast transformants were selected in the SC-Ura-Leu agar plates, and colony PCR was conducted using Phire Plant Direct PCR Master (Thermo Fisher Scientific) Mix to identify positive colonies for the feeding assay.

3.3.3 Yeast feeding assay and metabolite analysis

Yeast strains constructed in this study were tested to produce 4′-*O*-methylnorbelladine. The CRISPR-Cas 9 integrated strain was grown in 5 ml synthetic complete nitrogen base (SC) media supplemented with 2% glucose overnight at 30 ◦C with shaking at 200 rpm. The plasmid-based yeast strains harboring either four genes or three genes (NR excluded) were grown in 5 ml SC-Ura-Leu supplemented with 2% glucose in the same condition. Overnight cultures were diluted 1/100 in 40 ml matching fresh media supplemented with 0.2% glucose and 2% galactose and incubated in the same conditions. After 24 h, the cultures were supplemented with 0.2, 0.4, and 1mM DHBA and tyramine. To test if TYDC can produce tyramine endogenously, yeasts were also solely supplemented with DHBA with the above-described concentrations. After supplementation, yeast cultures were incubated in the same culture conditions for 72 hours.

Metabolite extraction was performed using ethyl acetate. Six milliliters of ethyl acetate were added to the culture, vortexed for 1 min, and centrifuged at 3500 rpm for 15 min. The upper phase was collected and dried entirely under nitrogen [6]. The extracts were reconstituted in 0.5 ml mobile phase (50% acetonitrile with 0.1 % formic acid) before LC-MS analysis.

3.3.4 Metabolite analysis

A Vanquish UHPLC system (ThermoFisher) with an Accucore Polar Premium (150 mm × 2.1 mm i. d.) C18 column and TSQ Fortis MS were used for sample analysis. The mobile phase consisted of water (A) and acetonitrile (B) mixed with 0.1% formic acid. The method used for separation consisted of a 7 min hold at 90:10 (A/B), a 3.5 min increase to 0:100, a 4.5 min hold at 0:100, a 1-minute increase to 90:10, and a 4 min hold at 90:10 to equilibrate the column. The solvents were passed through the column at 400 µL/min, and the temperature was held at 30°C. Metabolites were ionized via H-ESI (heated electrospray ionization). The MS was set to the

following parameters: spray voltage of 3700 V, sweep gas flow rate of 0 respective arbitrary units, sheath gas flow rate of 5 respective arbitrary units, auxiliary gas flow rate of 5 respective arbitrary units, ion transfer tube temperature of 325 °C, and vaporizer temperature of 325 °C. Full scans in positive and negative modes from 100-300 m/z were performed for substrate, intermediate, and product detection. For precise detection of 4′-*O*-methylnorbelladine, selective reaction monitoring (SRM) parameters were developed using a 4′-*O*-methylnorbelladine standard. Positive ion mode was used for SRM, with a selected precursor ion [M + H+] of 274.1 and product ions of 93.883, 121.883, and 136.967.

3.4 Results and Discussion

This study aimed to produce 4′-*O*-methylnorbelladine in *S. cerevisiae* following the introduction of the genes encoding the enzymes NBS, NR, N4′OMT, and TYDC and feeding with DHBA (Figure 3-2). Previous studies suggest that NBS and NR cooperatively catalyze the condensation of DHBA and tyramine to form norcraugsodine, which is then reduced to norbelladine. This observation indicates that the intermediates might efficiently channel between these two enzymes through a proposed protein complex, resulting in a higher production rate [192, 193]. Norbelladine is methylated to for central 4′-O-methylnorbelladine. The core pathway was successfully constructed using genome integration via a multiplex CRISPR-Cas9 platform (Figure 3-3) and a plasmid-based expression system (Figure 3-4).

Yeast strains in this study were fed with different concentrations of tyramine and DHBA or DHBA solely 24 hours after transferring to the induction media. Our preliminary metabolite analysis of yeast extracts was not conclusive. A weak signal corresponding to 4′-O-methylnorbelladine seems to be detected in metabolite extracts of yeast SBY104 strain engineered

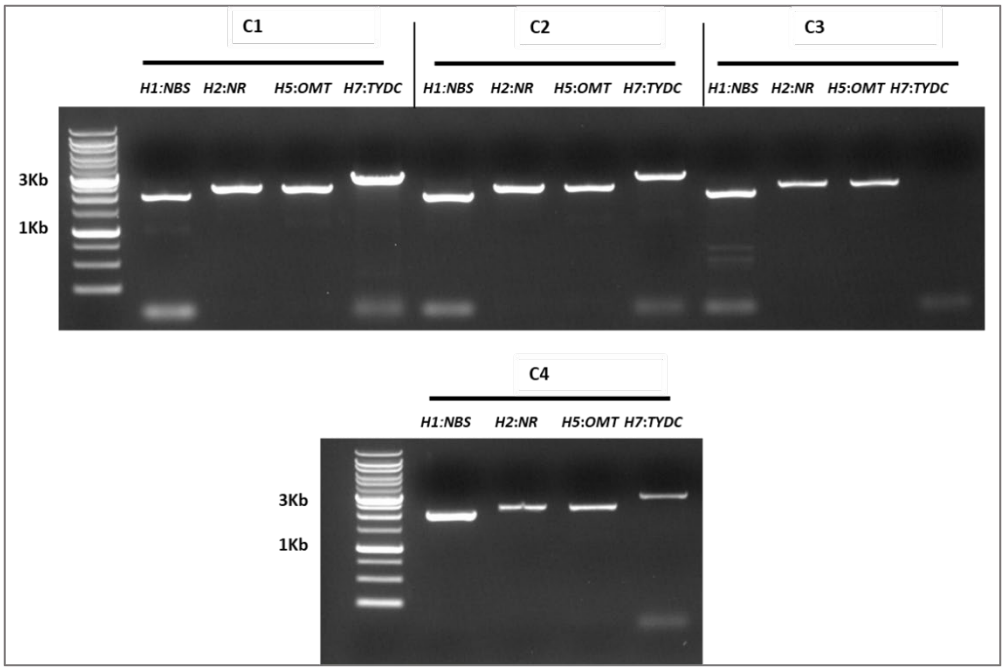


Figure 3-3. The genotyping of yeast (SBY104) transformants for 4′-O-methylnorbelladine pathway genes. The genotyping was conducted using two primers listed in Table A2-1. Four colonies have been screened for integration (C; Colony). Bands between 2-3kb approximately indicate the integration, while bands near 1kb could represent no integration. In C1, C2, and C4, successful integration is shown for all the sites. In C3, no positive or negative band was amplified for the H7 site, possibly due to a pipetting issue. *OMT*; *N4OMT*

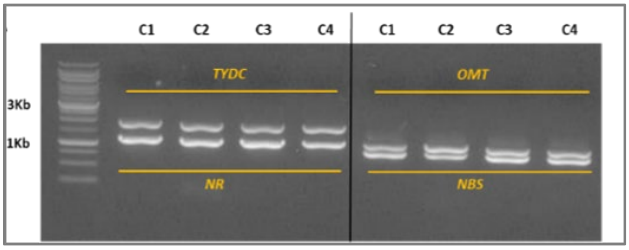


Figure 3-3. Genotyping of the yeasts co-transformed with pESC-Ura::NBS: NR & pESC-Leu::TYDC: OMT. and yeasts co-transformed with pESC-Ura::NBS & pESC-Leu::TYDC: OMT (B). Four colonies from each transformant were subjected to colony PCR using two sets of primes from the vector backbone, one amplifying *TYDC* and *NR* (same binding sites in pESC-Ura and pESC-Leu) and the other amplifying *NBS* and *OMT* (*N4'OMT*) (same binding sites in pESC-Ura and pESC-Leu). The vectors without inserts could show a band below 250 that was absent in the screened clones. The primers are listed in Table A2-1.

for four genes (Figure 3-5) when supplemented with 0.2mM of both substrates. Although a 4′-O-methylnorbelladine peak was detected for all yeast strains expressing AAs core pathway genes in higher concentrations of substrates, a weak signal detected from negative control makes it inconclusive. It might propose that in higher concentrations of substrates, the spontaneous condensation reaction of tyramine and DHBA produces a detectable level of norbelladine, followed by methylation via yeast endogenous methyltransferases despite low possibility. Norcoclaurine, a counterpart of norbelladine, is a central intermediate in the biosynthesis of benzylisoquinoline alkaloids. It forms through the condensation of dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA). This reaction can proceed spontaneously, especially under higher concentrations of the substrates [194, 195].

The feeding experiment was repeated with the lowest concentration of each substrate tested before. The metabolite analysis using HPLC-MS/MS indicated that the yeast with genome-integrated core pathway genes produces norbelladine (Figure 3-6), which was absent in the negative control. 4′-O-methylnorbelladine was not detected in the sample showing norbelladine. The absence of 4′-O-methylnorbelladine might indicate that methylation requires a threshold norbelladine concentration or that the methylation process is otherwise constrained. Another possibility is that a recent study reported that norbelladine methyltransferase exhibits substrate promiscuity and can methylate other compounds such as DHBA as well as 4′-O-methylnorbelladine or even catalyze 3′-methylation of norbelladine. This could provide an alternative explanation for the absence of detectable 4′-O-methylnorbelladine in the metabolite extracts and should be closely investigated.

Furthermore, metabolite analysis of samples fed with only DHBA did not show 4′-O-methylnorbelladine. Although introducing TYDC was ideally aimed at producing tyramine endogenously, it might not be sufficient for the detectable production of norbelladine. A study on substrate preference of NBS and NR showed that tyramine and DHBA are the best candidate substrates for these enzymes; however, the best *in vitro* enzymatic assay condition could have a 10% conversion rate [193]. Therefore, a concentration threshold is probably needed for the condensation reaction, and the tyramine produced by overexpressed TYDC might be insufficient.

On the other hand, although it was not investigated meticulously in this study, one possibility is that tyramine can be consumed with competing pathways in yeast or undergo catabolic pathways.

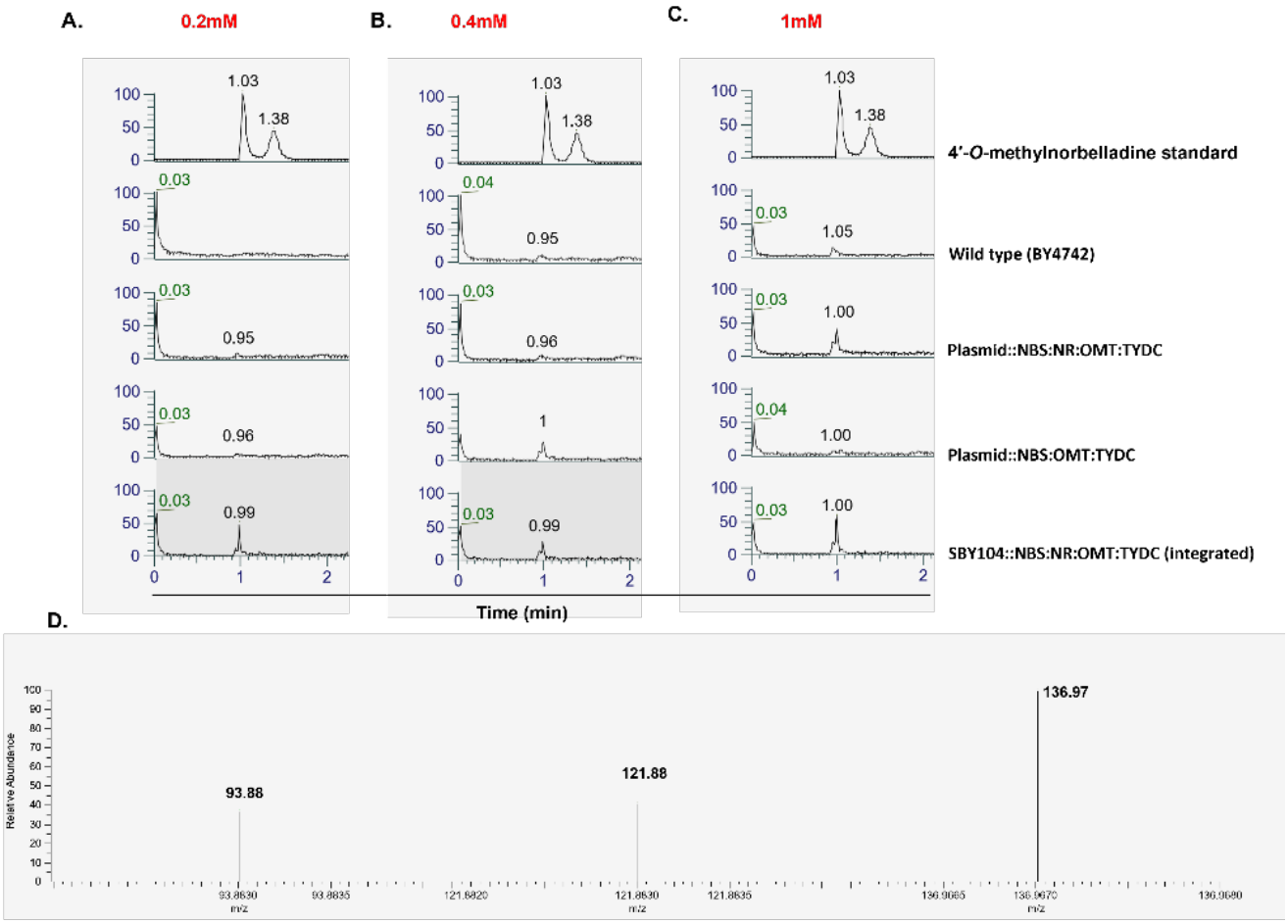


Figure 3-5. LC-MS analysis of metabolites extracted from yeasts expressing AAs core pathway. A; supplemented with 0.2mM (A), 0.4mM (B), 1mM (C) of tyramine & DHBA. BY4742 harboring empty pESC-Leucine and pESC-Uracil was considered as a control, D; The corresponding mass spectra fragmentation for 4′ -*O*-methylnorbelladine $[M+H]^+$= 274.147.

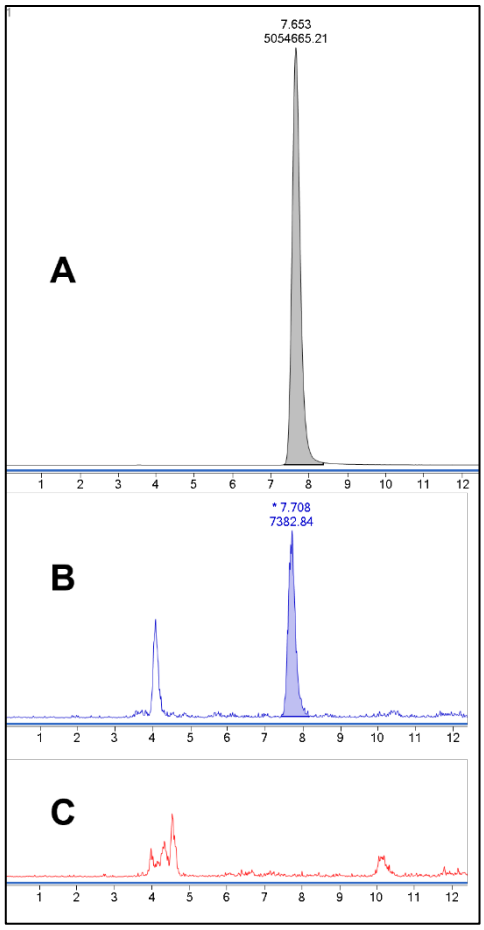


Figure 4-6. LC-MS analysis of metabolites extracted from SBY104 expressing AAs core pathway. A; Norbelladine standard, B; extracts of SBY104 expressing AAs core pathway supplemented with 0.2mM tyramine & DHBA (the extraction was performed from three independent cultures. A peak corresponding to the norbelladine was detected in all replicates but only the chromatogram of the first replicate is shown), C; BY4742 harboring empty pESC-Leucine and pESC-Uracil was considered as a control.

4. Conclusion

In conclusion, this study demonstrates the potential of *S. cerevisiae* as a platform for producing norbelladine and related AAs through the integration of a CRISPR-Cas9-based multiplex genome editing system. While the production of 4′-O-methylnorbelladine remains inconclusive due to potential limitations in substrate availability, methylation constraints, and competing pathways, the findings highlight the importance of optimizing metabolic engineering strategies and experimental conditions to achieve efficient biosynthesis of target metabolites. Future studies should focus on refining fermentation protocols and addressing enzyme-specific challenges to enhance pathway efficiency and product yield.

Chapter IV

Conclusion

Despite being a source for AAs, predominantly FDA-approved galanthamine [196] and lycorine with promising anti-cancer and anti-viral properties [45, 197], the biosynthetic pathway of AAs' precursors remains unexplored, mainly in *L. aestivum*. The AAs biosynthetic pathway utilizes two precursor molecules, DHBA, and tyramine, which are phenylalanine and tyrosine derivatives [51]. Tyramine is synthesized by a single enzymatic reaction catalyzed by TYDC. DHBA is believed to be synthesized through the phenylpropanoid pathway, and different possible routes have been proposed to yield it using this pathway [8]. Regardless of all possible routes, biosynthesis of DHBA from phenylalanine requires two hydroxylation steps. This study primarily aimed to characterize candidate enzymes responsible for the hydroxylation steps of phenylpropanoid intermediates through different routes (Figure 2-1).

As reported by other studies, C4H (CYP73A), C3′H (CYP98A), and, recently, bifunctional APX enzymes are the candidates for the hydroxylation steps. Specifically, C4H has been reported from many plant families to catalyze the first hydroxylation of *trans*-cinnamic acid in the phenylpropanoid pathway in the 4′ position [11, 198, 199]. The best candidate for the second hydroxylation is C3′H (CYP98A), which catalyzes 3′ hydroxylation on *p*-coumaroyl shikimate/quinate with less possibility on free *p*-coumaric acid. This enzyme demonstrated activity on various substrates across different plant families [4]. Recently, an isoform of APX has been proposed to play a dual role in some plants, such as *A. thaliana* and *B. distachyon,* capable of hydroxylating *p*-coumaric acid and scavenging oxygen free radicals [10].

This study has advanced our understanding of the phenylpropanoid and Amaryllidaceae alkaloid (AA) biosynthetic pathways by employing a multifaceted approach combining bioinformatics, heterologous expression systems, and structural analysis. Through the identification and functional characterization of key enzymes such as *Lae*C4H and *Lae*C3′H, we have provided insights into their roles and mechanisms within the phenylpropanoid pathway in Amaryllidaceae.

In silico analyses of the full-length ORF sequences of *Lae*C4H, *Lae*C3′H, and *Lae*APX/C3H from *L. aestivum* showed high sequence identity with their homologs in other plant species,

especially with monocots. CYP450 enzymes displayed conserved motifs typical of the CYP450 superfamily, such as the N-terminal membrane anchoring domain, proline-rich, and hem-binding signature motifs. However, compared to *Lae*C4H, the N-terminal membrane domain in *Lae*C3′H displayed relatively less identity than phylogenetically distant species.

*Lae*C4H exhibited efficient *in vitro* catalytic activity when expressed in the yeast *S. cerevisiae*, a common heterologous host for CYP450 enzyme expression. The kinetic parameters (Km and Vmax) observed for *Lae*C4H were consistent with those of previously characterized isoforms from other plant species. However, due to a lack of expertise, we were unable to quantify the enzyme's net content in total microsomal fractions to determine the Kcat value despite several attempts.

Structural modeling and molecular docking further suggested that *Lae*C4H effectively interacts with *trans*-cinnamic acid, leading to hydroxylation. The sequence of the C4H enzyme is highly conserved across different plant lineages, reflecting the critical and universal role this enzyme plays in plant metabolism. This conservation potentially indicates intense evolutionary pressure to maintain C4H's structure and function, as any significant variation could disrupt its essential role in the phenylpropanoid pathway. The high sequence conservation and strict substrate preference of C4H may be necessary to ensure the consistent production of *p*-coumaric acid, thereby supporting the diverse and essential metabolic processes that rely on this compound.

*Lae*C3′H, while typically favoring *p*-coumaroyl shikimate or quinate for the second hydroxylation step in the phenylpropanoid pathway of angiosperms, has been reported in some studies to hydroxylate other substrates such as spermidine-conjugated phenolics [106], naringenin, *p*-coumaroyl tyramine [183], and various natural and non-natural substrates [4]. *Lae*C3′H was tested with three potential substrates: *p*-coumaroyl shikimate, *p*-coumaric acid, and 4-HBA to evaluate its activity on different phenylpropanoid intermediates, each potentially yielding DHBA. Only *p*-coumaroyl shikimate was accepted for hydroxylation activity when tested in yeast microsomal fractions expressing this enzyme. Like *Lae*C4H, this enzyme efficiently could hydroxylate the corresponding substrate to caffeoyl shikimate with kinetic parameters consistent with other reports.

For *Lae*C3′H, as expected, molecular docking indicated that *p*-coumaroyl shikimate is the most suitable substrate, interacting through hydrophobic interactions and hydrogen bonds, with its phenylpropanoid moiety positioned favorably for hydroxylation. In contrast, *p*-coumaric acid and 4-HBA exhibited lower docking scores, fewer hydrogen bonds, and a lack of salt-bridge

interactions with the active site residues. It is hypothesized that the large catalytic pocket of *Lae*C3′H is more suitable for accommodating larger substrates like *p*-coumaroyl shikimate rather than smaller ligands such as 4-HBA or *p*-coumaric acid. It remains to be examined whether substituting bulkier amino acids in the active site of *Lae*C3′H or other CYP98A homologs can stabilize the smaller substrates such as *p*-coumaric acid or 4-HBA. This could be interesting from a synthetic biology perspective as it can shorten the pathway to yield DHA by acting on 4-HBA or *p*-coumarric acid.

The study successfully confirmed the functionality of *Lae*C4H and *Lae*C3′H enzymes *in planta* through transient expression in *N. benthamiana* leaves, followed by substrate feeding experiments. The observed conversion of *trans*-cinnamic acid and *p*-coumaroyl shikimate into downstream products provides compelling evidence for the enzymatic activity of *Lae*C4H and *Lae*C3′H within a living plant system. Although direct detection of *p*-coumaric acid, the product of *Lae*C4H, was not achieved, it is likely due to its rapid turnover by endogenous enzymes such as 4CL and HCT in the phenylpropanoid pathway. This rapid turnover highlights the dynamic nature of metabolic processes *in planta*. The increased levels of downstream phenolic intermediates relative to controls indirectly support the functionality of *Lae*C4H. Similarly, plants overexpressing *Lae*C3′H exhibited elevated levels of its direct product, caffeoyl shikimate, and subsequent intermediates like caffeic acid, further validating the catalytic role of *Lae*C3′H *in vivo*. Future research should focus on expressing these enzymes in mutant lines of model plants such as Arabidopsis, where the genetic background can be more precisely controlled and the metabolic fluxes more accurately monitored.

The *E. coli*-expressed *Lae*APX/C3H candidate enzyme did not hydroxylate *p*-coumaric acid or 4-HBA. However, it was observed that the enzyme consumed *p*-coumaric acid, though the level of caffeic acid remained unchanged compared to the negative control, suggesting that the detected caffeic acid might result from non-enzymatic conversion. This finding aligns with reports on the APX/C3H homolog from *S. bicolor*, which reported polymerization of *p*-coumaric acid rather than hydroxylation [115], and challenges studies reporting the hydroxylation activity of this enzyme in plants like *A. thaliana* and *B. distachyon* [10]. It is proposed that the hydroxylation activity of this enzyme may be species-specific and not ubiquitous across all plant families [115]. However, the non-functionality of APX/C3H from *B. distachyon* expressed in yeast further questions the enzyme's ability to hydroxylate *p*-coumaric acid [132]. The latest study isolated the APX/C3H

homolog from the Amaryllidaceae plant, *L. aurea*, using the APX/C3H sequence from *B. distachyon* as a reference. It was claimed recently that this enzyme could hydroxylate 4-HBA to DHBA. However, despite being a homolog of the enzyme from *B. distachyon*, it remains unclear why *p*-coumaric acid was not tested as a substrate, given that it was primarily reported to be a substrate for the *B. distachyon* enzyme. Interestingly, the crucial catalytic residues identified in the study are conserved between the APX/C3H homologs from *B. distachyon* and *L. aestivum*, suggesting a potential for similar catalytic activity.

The subcellular localization of *Lae*C4H and *Lae*C3′H was investigated by tagging these enzymes with fluorescent proteins and using Agrobacterium-mediated transient expression in *N. benthamiana*. Both enzymes exhibited endoplasmic reticulum (ER) localization when co-expressed with an ER-localizing fluorescent protein. This supports the hypothesis that these enzymes are part of a larger metabolon within the phenylpropanoid pathway, requiring co-localization on the ER's exterior surface [98, 200]. Such an arrangement likely facilitates efficient channeling of intermediates in response to biotic and abiotic stresses. Further protein-protein interaction studies, such as bimolecular fluorescent complementation (BiFC) assays, could provide additional evidence for interactions between these enzymes with upstream and downstream phenylpropanoid enzymes in Amaryllidaceae plants.

Differential expression analysis of phenylpropanoid-related genes and AA-specific transcripts was conducted using qRT-PCR in various tissues of *L. aestivum*. The genes encoding PAL, C4H, C3′H, and TYDC1enzymes showed the highest transcript levels in the roots. In contrast, *NBS*, *NR*, and *N4'OMT* transcripts were most abundant in the bulbs, which accumulate the highest quantity of AAs. This suggests that the expression patterns of phenylpropanoid pathway enzymes do not directly correlate with alkaloid content in a tissue. This is unsurprising, as these enzymes are involved in different branches of the phenylpropanoid pathway, which produce various phenylpropanoid compounds for downstream branches. In other words, the Amaryllidaceae alkaloid pathway is one of several pathways utilizing phenylpropanoid products as precursors.

By understanding the tissue-specific expression patterns of key enzymes, we can genetically engineer plants to overexpress or suppress specific genes in targeted tissues, leading to increased production of desired compounds. For example, enhancing the expression of NBS, NR, and N4′OMT in bulb tissues could potentially increase the yield of AAs, although the transformation

techniques for genetic engineering of Amaryllidaceae plants need to be well established. In addition, understanding bottleneck enzymes would be useful for synthetic biology purposes.

Synthetic biology presents a promising alternative for producing plant natural products in a timely and cost-effective manner. Introducing the enzymes involved in specialized metabolite pathways into microbial systems makes it possible to link microbes' primary metabolism with plants' specialized metabolism to produce plant-derived therapeutics [146].

As an internship project, it aimed to establish the AAs core pathway for producing 4′-*O*-methylnorbelladine by incorporating enzymes such as TYDC, NBS, NR, and N4′OMT into baker's yeast. This initiative was intended to demonstrate the effectiveness of these enzymes within microbial systems and to advance the production of these metabolites through synthetic biology. Successful integration of the four genes was achieved using a multiplex CRISPR-Cas9 platform [5, 6]. For the plasmid expression system, we utilized two yeast vectors, pESC-URA and pESC-LEU, which harbored the four (all the genes). Preliminary yeast feeding assay and metabolite analysis showed a signal corresponding to 4′-*O*-methylnorbelladine. A background signal was observed in the negative control at higher substrate concentrations, rendering the results inconclusive. It is unlikely that yeast produce these metabolites endogenously, suggesting the background signal may be due to a technical error or cross-contamination. However, when the experiment was repeated using lower substrate concentrations, the CRISPR-based platform demonstrated clear production of norbelladine, while 4'-O-methylnorbelladine was not detected. Further experiments will be needed to optimize fermentation conditions, metabolite extraction, and analyses. Strains producing 4′-*O*-methylnorbelladine can be further engineered for *de novo* production of AAs in microbial systems. This requires complete elucidation of enzymes involved in DHBA synthesis. The first part of this project is such an attempt at DHBA pathway characterization. It can be coupled with the second objective of this work to connect AAs specific pathway to yeast primary metabolism.

4.1 Final conclusion and future perspective

This study contributes to elucidating the biosynthetic pathway of AAs precursor, DHBA, by characterizing key enzymes involved in phenylpropanoid hydroxylation steps in *L. aestivum*. *Lae*C4H and *Lae*C3′H, when expressed in yeast and *N. benthamiana*, demonstrated catalytic activity on *trans*-cinnamic acid and *p*-coumaroyl shikimate, respectively, confirming their roles in

the biosynthesis of intermediates of the phenylpropanoid pathway. Furthermore, the study challenges previous reports on the hydroxylation activity of *Lae*APX on *p*-coumaric acid to yield caffeic acid. Future research might focus on expressing these enzymes in mutant lines of model plants to provide definitive evidence of their *in planta* activity and further elucidate the biosynthesis of DHBA.

Similar subcellular localization of *Lae*C4H and *Lae*C3′H in ER might provide evidence for metabolon existence in the phenylpropanoid pathway of Amaryllidaceae plants. Additional protein-protein interaction studies will be crucial in understanding the dynamic assembly of metabolons in this pathway.

The differential gene expression analysis further indicates that the phenylpropanoid pathway's regulation is complex and tissue-specific, with distinct expression patterns of enzymes involved in Amaryllidaceae alkaloid biosynthesis.

The core pathway of Amaryllidaceae alkaloids (AAs) has been successfully engineered in a yeast system to demonstrate the proof of concept to produce 4′-O-methylnorbelladine. This compound is the latest intermediate in the biosynthesis of all downstream AAs, including galanthamine, within a microbial system. These engineered yeast strains provide a foundation for future research to further advance synthetic biology approaches for AAs. By integrating plant-derived enzymes into yeast, this method offers a promising alternative to the over-exploitation of natural plant resources for producing valuable plant secondary metabolites.

Appendix I

Supplementary Data of Chapter II

Table A1-1. List of primers used in this study.

Primer name	Sequence	Purpose
***LaeC4H*-F**	ATCGGGGCCCATGGATCTCATCCTACTAGAGAAGTCCC	Cloning
***LaeC4H*-R**	GATCGTCGACGAACGCTCTAGGTTTGGCCAC	Cloning
***LaeC3'H* -F**	AGTCGGATCCGTAATACGACTCACTATAGGGCCCATGGATCTCTCTCCTCTCACAGTCT	Cloning
***LaeC3'H* -R**	AGCCGTCGACCATATCTACGGCCATGCGCTTG	Cloning
***LaeAPX/C3H* -F**	GGATCCATGACGAAATCGTACCCCAAGGT	Cloning
***LaeAPX/C3H* -R**	GTCGACTTAGGCAAACCCGAGCTCAGA	Cloning
***PAL*-F**	GAAAGAGAGATCAATTCGGTG AA	RT-qPCR
***PAL*-R**	AGGAAGCCCGTTGTTGTAGA	RT-qPCR
***TYDC1*-F**	GTAGCTCCGCATCACCATCC	RT-qPCR
***TYDC1*-R**	GACTGCTGCTGCCTCTGG	RT-qPCR
***TYDC2*-F**	CTTTGCCAATGGGTTACGCC	RT-qPCR
***TYDC2*-R**	GCCCCTGACCATCCTGAAAG	RT-qPCR
***C4H*-F**	GACGACCTCAACCACCGCA	RT-qPCR
***C4H*-R**	CGTCCGCGACCCGAACTC	RT-qPCR
***C3'H* -F**	GATCGCCCGAGGAACCG	RT-qPCR
***C3'H* -R**	GTAGTGGGGCCCGTAATCTG	RT-qPCR
***NBS*-F**	AGGAAGCATTTGCTGCATGT	RT-qPCR
***NBS*-R**	GAGGGAGGGCACTTGG	RT-qPCR
***NR*-F**	GCCGATTCGAAATTGGTGGC	RT-qPCR
***NR*-F**	AACAATGCAGGGACAGCCAT	RT-qPCR
***OMT*-F**	GGGATATCTCGACACGGGGA	RT-qPCR
***OMT*-R**	CCACGAGCGATTAGTGAAGC	RT-qPCR
***Histone3*-F**	ACAAGCCTTTGGAAGGGCAG	RT-qPCR
***Histone3*-R**	GGAGTGAAGAAGCCCCACC	RT-qPCR

Table A1-2. Sequencing result of isolated *C4H*, *C3'H*, and *APX*/*C3H* from *L. aestivum*. There is 100% identity with predicted *C4H* from RNA-seq data obtained from *L. aestivum*. A variation in nucleotide number 1080 was observed that shifts amino acid codes from tryptophan to isoleucine; however, analyzing the sequences of *C4H* homologs in other species, we found the isoleucine is conserved in almost all the species. Therefore, we speculated that this variation originated from RNA sequencing errors rather than mutation resulting in PCR amplification. A 100% and 99% identity were observed in amino acid levels for *LaeC3'H* and *LaeAPX*/*C3H*.

Name of the gene	Sequenced obtained in this study
LaeC4H	ATGGATCTCATCCTACTAGAGAAGTCCCTCCTCGCCGTATTCTTCGCCATCATCCTCAGCATCGTAATCTCCAAACTCCGCGGGAAGAAGCTCAGGCTCCCCCCTGGGCCGCT CCCGGTCCCCGTATTCGGGAACTGGCTCCAGGTCGGCGACGACCTCAACCACCGCAACCTCACCTCCCTCGCCAAGAGATTCGGCGACTTATTGATGCTGAGAATGGGCCAG AGGAATCTCGTGGTCGTGTCGTCACCCGACCTCGCCCGCGACGTCCTGCACACTCAAGGGGTCGAGTTCGGGTCGCGGACGAGGAACGTGGTTTTCGATATATTTACCGGG AAAGGGCAGGACATGGTTTTCACAGTGTATGGTGAGCACTGGCGCAAGATGCGCCGGATCATGACCGTGCCGTTTTTCACCAACAAAGTTGTGAATCAGTACAGGTTTGGT TGGGAAGACGAAGCTGGTAAAGTGGTGGAGGATGTGAAGAAGGAGCCCACGGCCGCGACCACGGGGATTGTTCTGAGGAAGAGGTTGCAGTTGATGATGTACAACAAC ATGTATAGGATTATGTTCGACAGGAGATTCGAGAGCGAGGAGGATCCCTTGTTCTTGAGGCTGAAGGCCTTGAATGGAGAAAGGAGCAGGCTCGCGCAGAGTTTCGAGTA CAATTACGGAGATTTTATCCCCATTTTGAGGCCTTTCTTGAGAGGATATTTGAAGATTTGCAAGCAGGTCAAGGAGAGGAGGTTGCAATTGTTCAAGGATTATTTTCTGGAG GAGAGGAAGAAGTTGGCAAGCACAAATCCAGCAGATAATGCCGGACTAAAGTGTGCAATTGATCACATCTTGGATGCTGAAAAGAAAGGGGAGATCAATGAGGACAATG TCCTCTACATTGTAGAGAACATCAATGTTGCAGCCATCGAGACAACCCTCTGGTCGATCGAATGGGGAATAGCAGAGCTAGTGAACCACCCCACCATCCAACAAAAGCTCCG CGATGAGCTCGACACCCTCCTCGGCCCCGGCGTCCAAATCACCGAGCCCGACACCTACCGCCTCCCGTACCTCCAAGCCGTGATCAAAGAGACCCTCCGCTTACGCATGGCC ATCCCGCTCCTCGTCCCCCACATGAACCTCCACGACGCCAAGCTCGGTGGCTACGACACCCCCGCCGAAAGCAAGATCCTCGTCAACGCGTGGTGGCTCGCCAACAACCCTG CCCACTGGAAAGACCCCCAAGAGTTCCGCCCCGAGAGGTTCCTGGAGGAGGAGGCCAAGGTGGAAGCCAATGGCAACGATTTCAGATACATTCCGTTTGGAGTGGGGAG GAGGAGTTGCCCAGGGATCATACTTGCATTGCCGATCTTGGGGATCACCTTTGGAAGGCTTGTTCAGAACTTTGAATTGCTGCCTCCACCTGGCCAAGACAAGCTTGATACC ACCGAGAAGGGAGGACAATTTAGCCTGCACATTTTGAAGCATTCCACCATCGTGGCCAAACCTAGAGCGTTCTAA
LaeC3'H	ATGGATCTCTCTCCTCTCACAGTCTCCATAGCCATCTTCTCTCTCCTTATCTCCTTTCTCCTCTTCAAGAAGCTGACCAGCAAACTGCCCCCGGGCCCAAAGCCATGGCCGGTTG TAGGCAACCTCTACGACATAAAGCCGATCCGGTTCAGGTGCTTCGCCGAGTGGGCCCAAACCTACGGCCCCATCATGTCGGTCTGGTTCGGGACCATCCTCAACATCGTCGT CTCAAGCTCCGAGCTGGCGAAGGAGGTGCTCAAGGAGAAGGACCAGAGCCTGGCCGATCGCCCGAGGAACCGGGCAGCTGCCCGGTTCAGCAGGGATGGGAAGGATTT GATCTGGGCAGATTACGGGCCCCACTACGTGAAGGTGAGGAAGGTCTGCAATCTTGAGCTCTTCTCACCGAAGAGGTTGGAGGGGTTGCGGCCGATTCGGGAGGATGAG GTCACCGCCATGGTCGAGTCCATTTTCAGGGAATGCACAAAGCAAGATAAGATCGGGAAGAGTTTGGTTGTAAAAAATCACCTCTCTGGTGTCGCCTTCAACAACATTACGA GGCTTGCATTTGGAAAGCGCTTTGTGAACTCAGAAGGCGTGATAGACGAACAAGGCTTGGAATTCAAGGCGATTGTGGCCAATGGATTGAAATTCGGCGCATCTCTCTCAC TTGGAGAATACGTGACATGGCTTAAATGGATGTTCCCTTTGGATGAAGAGGCCTTTGCCAAGCACGGAGCTCGTAGAGATCGACTCACTAGAGCCATCATGGAAGAGCACA CACGTACTCGCAAGGAGAGTGGTGCAAGACAACACTTTGTTGATGCATTGCTTACCCTTCAAGAAAAGTACGACCTTAGTGAAGATACCATCATTGGCCTTCTTTGGGACAT GATCACCGCAGGAATGGACACGACCGTGATTACTGTGGAATGGGCCATGGCGGAGCTAGTAAAGCACCCAAGAGTACAAGAAAAGGTACAAGAAGAACTAGACCGAGTC ATTGGAATCGACCGCATCATGACCGAATCCGATTTTCCCAACCTCCCATACCTCCATTGCGTTGTAAAAGAAGCCCTTCGCCTCCACCCTCCAACCCCCCTCATGCTCCCCCAC AAAGCCTCTGCCAACACAAAGCTTGGTGGATATGACATCCCAAAGGGAGCGGTTGTGCACGTCAATGTGTGGGCCATCGCACGCGACCCGAATTTGTGGAAAAATCCATTA GAATTTCGCCCGGAAAGATTTCAAGAAGAGAATATTGATATCAAAGGGAACGATTTTCGCGTGCTTCCATTTGGAGCGGGTAGAAGGGTTTGTCCTGGGGCACAATTGGGG ATCAATTTGGTGCAATCCATGTTAGGGCATCTGCTGCATCATTTTAAGTGGGAGTTGCCTGAAGGGATGAAACCCGAGGAAATCGACATGGCGGAGAATCCTGGGATGGTT ACTTTCATGCACACGCCTTTGCAGGCTGTTGCGATTCCTAGGCTGCCATCGCATTTGTACAAGCGCATGGCCGTAGATATGTAA
LaeAPX/C3H	ATGACGAAATCGTACCCCAAGGTTGGCGACGAGTACTTGAAAGCAGTCGAGAAATGCAAGCGAAAGCTTCGTGGATTCATTGCCGAGAAGAATTGCGCTCCTTTGATGCTT CGATTGGCGTGGCATTCGGCGGGGACTTACGATGTGAAGACGAAGACCGGTGGTCCGTTTGGGACGATGAGGTACAAGGCTGAGCTCGGTCATGGCGCGAATAATGGAC TTGATATTGCGGTTAGGCTTTTGGAGCCGATCAAGGAGCAGTTTCCGATCTTGTCTTATGGAGATTTTTATCAGCTTGCTGGAGTTGTCGCTGTGGAGGTAACTGGAGGACC TGAGGTTCCATTCCACCCTGGCAGAGAGGACAAAGCTGAGCCCCCTCTTGAAGGCCGCCTGCCCGATGCCACAAAAGGTTCTGATCATCTTAGAACTGTCTTTGGTCAACAA ATGGGATTAAGTGACCAGGACATTGTTGCTCTGTCTGGAGGTCACACCCTGGGTAGGTGCCACAAGGAGAGATCAGGCTTTGAGGGTGCTTGGACTTCAAATCCCCTCATC TTTGACAACTCATATTTCAAGGAGCTTTTGAGTGGGGAGAAGGAAGGCCTTCTTCAACTGCCATCAGACAAGACCCTTCTGAGCGATCCAGTCTTCCGCCCACTCGTGGATA AATATGCAGCGGATGAGGATGCCTTCTTTGCCGACTATGCTGAGGCTCACCTGAAGCTTTCTGAGCTCGGGTTTGCCTAA

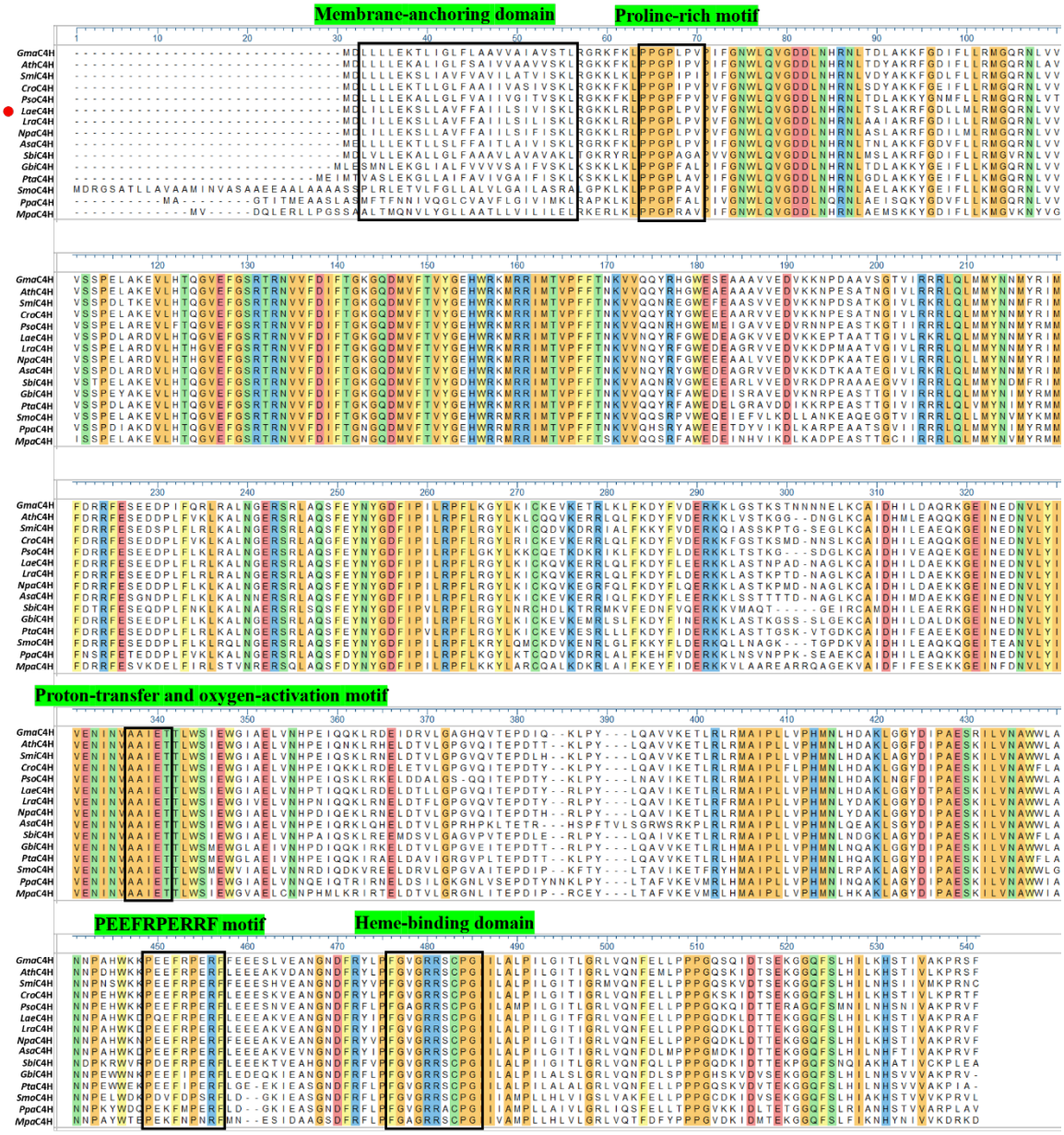


Figure A1-1. Multiple sequence alignment of C4H amino acid residues from different plant species using DNASTAR MegAlign Pro software and the clustal omega algorithm. *Lae*C4H (●), the amino acid sequence, was used as a query. The most functionally important domains conserved among all sequences are shown in black boxes. *Gma*, *Glycine max* (ACR44227); *Smi*, *Salvia miltiorrhiza* (ABC75596); *Ath*, *Arabidopsis thaliana* (AAC99993); *Cro*, *Catharanthus roseus* (CAA83552); *Pso*, *Papaver somniferum* (XP_026426522); *Lae*, *Leucojum aestivum* (UIP35210); *Lra*, *Lycoris radiata* (AWW24970); *Npa*, *Narcissus papyraceus* (AXU39895); *Asa*, *Allium sativum* (ADO24190); *Sbi*, *Sorghum bicolor* (AAK54447); *Gbi*, *Ginkgo biloba* (AAW70021); *Pta*, *Pinus taeda* (AAD23378); *Smo*, *Selaginella moellendorffii* (EFJ22128); *Ppa*, *Physcomitrium patens* (ADF28535); *Mpa*, *Marchantia paleacea* (ASA39648).

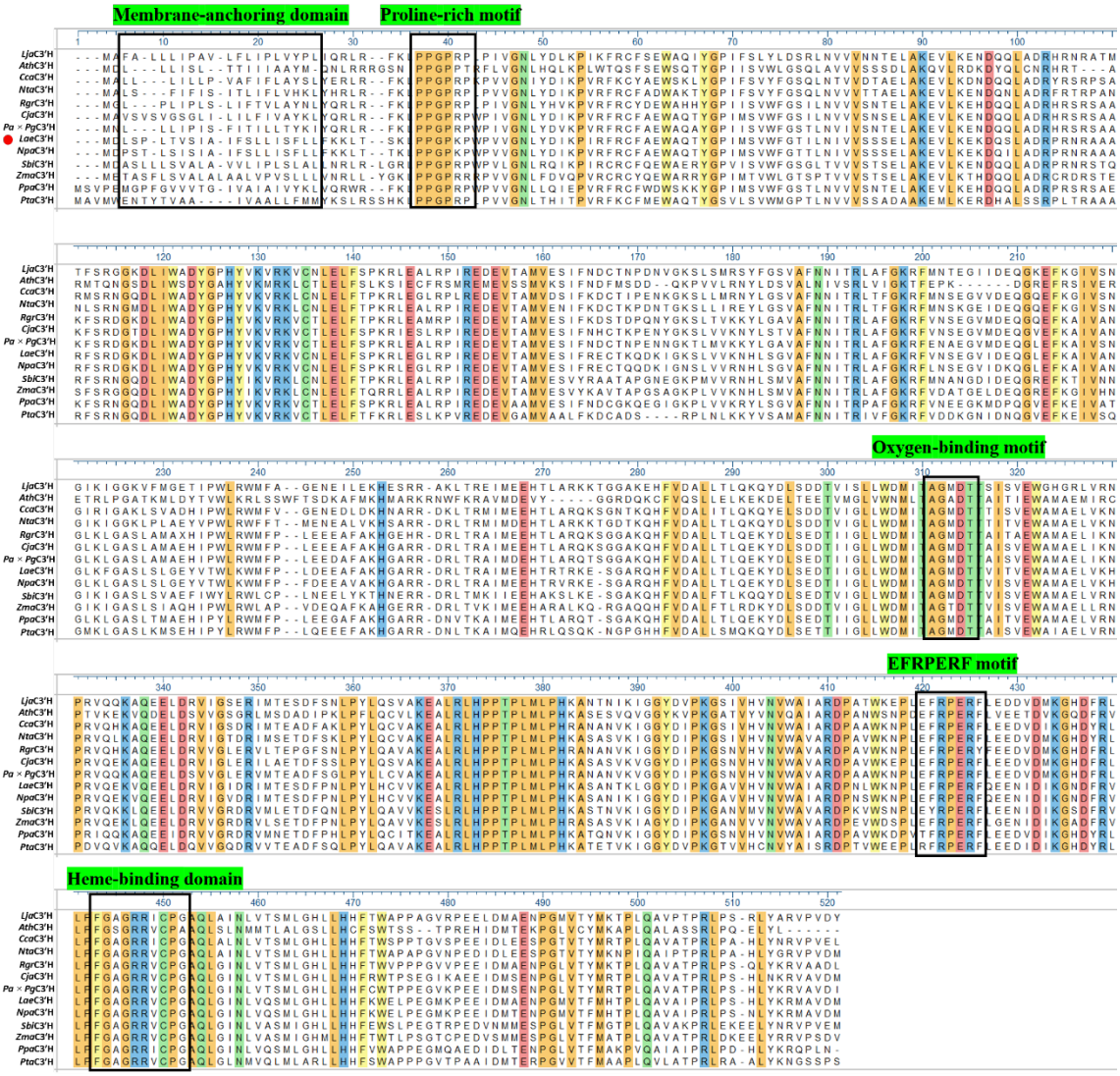


Figure A1-2. Multiple sequence alignment of C3′H amino acid residues from different plant species using DNASTAR Megalign Pro software and the clustal omega algorithm. *Lae*C3′H (●), the amino acid sequence was used as a query. The most functionally important domains conserved among all sequences were shown in black boxes. *Lja*, *Lonicera japonica* (AGQ48118); *Ath*, *Arabidopsis thaliana* (NP_177595); *Cca*, *Coffea canephora* (ABB83676); *Nta*, *Nicotiana tabacum* (ABC69384); *Rgr*, *Ruta graveolens* (AEG19446); *Cja*, *Coptis japonica* var. dissecta (BAF98473); *Pa×Pg*, *Populus alba* × *Populus grandidentata* (ABY85195); *Lae*, *Leucojum aestivum* (UIP35212); *Npa*, *Narcissus papyraceus* (AXU39897); *Sbi*, *Sorghum bicolor* (XP_002440001); *Zma*, *Zea mays* (PWZ32976); *Pta*, *Pinus taeda* (AAL47685); *Ppa*, *Physcomitrium patens* (XP_024360823).

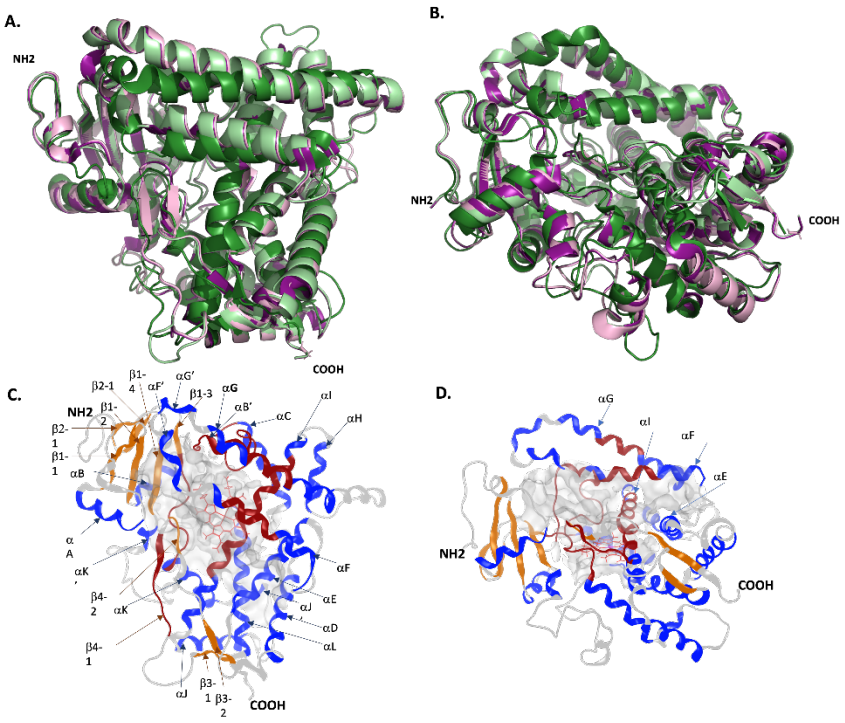


Figure A1-3. A, Superimposed models of *Lae*C4H (light pink), *Lae*C3′H (pale green), *Hs*C3′H (Forest green), *Sb*C4H1 (deep purple), top view compared to the heme; **B,** Superimposed models of *Lae*C4H (light pink), *Lae*C3′H (pale green), *Hs*C3′H (Forest green), *Sb*C4H1 (deep purple) side view of the heme; **C**, Cartoon representation of *Lae*C3′H with annotated secondary structures according to [201], and grey transparent surface-active site on top of the heme (shown as red stick) view; **D**, Cartoon representation of *Lae*C3′H with grey transparent surface-active site, perpendicular side of the heme plan. Conserved α-helix and β-strands of secondary structures are shown in blue and orange, respectively. Conserved CYP98A domains substrate recognition is displayed in red. Heme group is shown as red sticks. The left side lobe lined by the F and G α-helices is speculated to be the substrate entry port [147]. In other CYP98, such as *Hs*2C8, the active site cavity includes the α-helices A (N-terminus) and I [184, 202] [201].

Table A1-3. Sequence and structure comparison scores. Id: identity, Simi: similarity, computed by MOE, RMSD: root mean square deviation computed by Pymol, na: not applicable.

	2C2			*Sb*C4H1			C4H1		
	Id (%)	Sim (%)	RMSD (Å)	Id (%)	Sim (%)	RMSD (Å)	Id (%)	Simi (%)	RMSD (Å)
*Lae*C3′H	24.1	42.4	1.879	31.1	51.7	0.110	30.7	52.8	0.015
*Lae*C4H	22.6	41.3	1.795	77.5	87.9	0.127	na	na	na

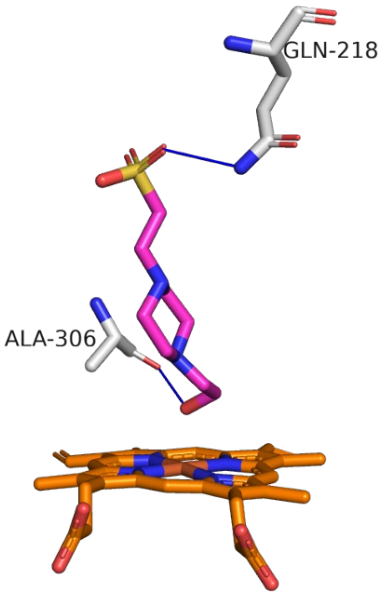


Figure A1-4. HEPES (pink sticks) with *Lae*C4H at its active site, respectively. Residues are shown as sticks. H-bonds are shown as blue lines, hydrophobic bonds as dashed green, and salt bridges as dashed yellow.

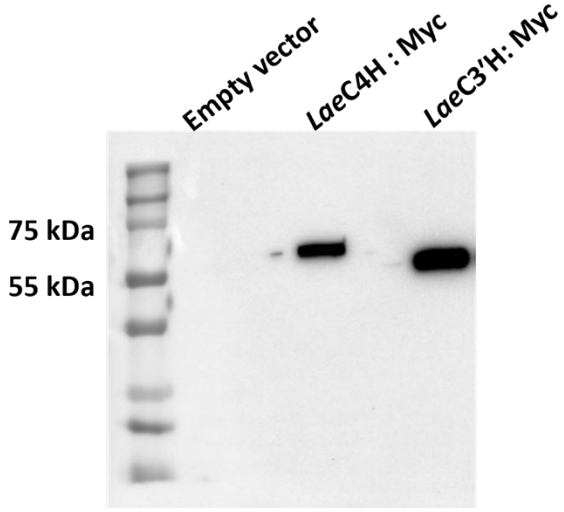


Figure A1-5. Western blot analysis of microsomal proteins isolated from yeast transformed with empty vector pESC-LEU-*CroCPR* (negative control), *LaeC4H*-Myc, or *LaeC3'H*-Myc constructs. Microsomal fractions (10 µg) were probed with anti-Myc (mouse) and anti-mouse antibodies, conjugated with peroxidase. An empty well separates each microsomal fraction.

Table A1-4. *Lae*C4H and *Lae*C3′H active site description and predicted interactions with docked substrates. NA; Not applicable.

Enzyme	Active-site residues	Ligand	Score (kcal.mol^{-1})	Interactions		
				Hydrophobic	H-bonds	Other
***Lae*C4H**	1: Arg101, Met117, Val118, Trp126, Arg130, Thr134, Phe137, Phe138, Met186, Ile189, Met190, Asn302, Ile303, Val305, Ala306, Ala307, Thr310, Thr311, Ser314, Leu365, Met369, Ala370, Ile371, Leu374, Val375, His377, Pro439, Phe440, Gly441, Gly443, Arg444, Arg445, Ser446, Cys447, Pro448, Gly449, Leu452, Ala453, Ile456, Leu457, Phe488 2: Gly50, Asp51, Asp52, Leu53, Met73, Arg76, Leu78, Thr102, Arg103, Asn104, Phe107, Phe119, Ala217, Gln218, Ser219, Phe220, Tyr222, Asn223, Asp226, Phe227, Pro376, Lys395, Leu397, Lys484, Gly485, Gly486, Gln487, Phe488, Ser489, Leu490, His491	Heme	NA	Val118, Phe138, Ala307, Thr310, Leu365, Ile371, Leu374, Val375, Phe440, Pro448, Leu452, Leu457	Trp126, Val375	Salt bridges: Arg101, His377, Arg445; Metal complex (n=4), Cys447
		trans-cinnamic acid	-5.1207	Val118, Phe119, Val305, Ala306, Val375, Phe488	Gln218	
***Lae*C3′H**	Pro47, Ile48, Arg49, Phe53, Phe68, Thr70, Ile71, Leu72, Asn73, Arg96, Pro97, Arg98, Asn99, Arg100, Ala101, Ala102, Ala103, Asp108, Leu112, Ile113, Trp114, Tyr121, Val122, Arg125, Lys126, Asn129, Leu132, Phe133, Arg145, Phe181, Ile184, Leu205, Lys208, Ala209, Ala212, Gly214, Leu215, Lys216, Phe217, Gly218, Ser220, Leu221, Leu223, Gly224, Glu225, Val227, Met233, Trp295, Asp296, Met297, Ile298, Thr299, Ala300, Asp303, Val306, Ile307, Thr365, Met368, Leu369, Pro370, Val389, His391, Val394, Leu430, Pro431, Phe432, Gly433, Arg437, Val438, Cys439, Gly445, Leu446, Leu448, Val449, Gly480, Met481, Val482	Heme	NA	Leu112, Ile113 (n=2), Leu132, Ala300, Thr304, Pro364, Met368, Leu369 (n=2), Phe432, Pro440, Leu444	Tyr121, Leu369	Salt bridge: Arg96, His371, Arg 437. Metal complexes: Cys439 (n=4)
		p-coumaroyl quinate	-7.2687	Ile113, Ala300, Thr365, Leu369 (n=2)	Trp114, Lys216, Thr304, Thr365	Salt bridge : Lys216
		p-coumaroyl shikimate	-7.0370	Trp114, Ala300, Leu369 (n=2)	Trp114, Leu115, Thr304, Thr365	Salt bridge: Lys216
		p-coumaric acid	-5.5228	Ile113, Ala300, Thr365, Leu369	Trp114, Thr304, Thr365	none
		4-HBA	-4.668	Ala300, Leu369	Trp114, Thr304, Thr365	none detected

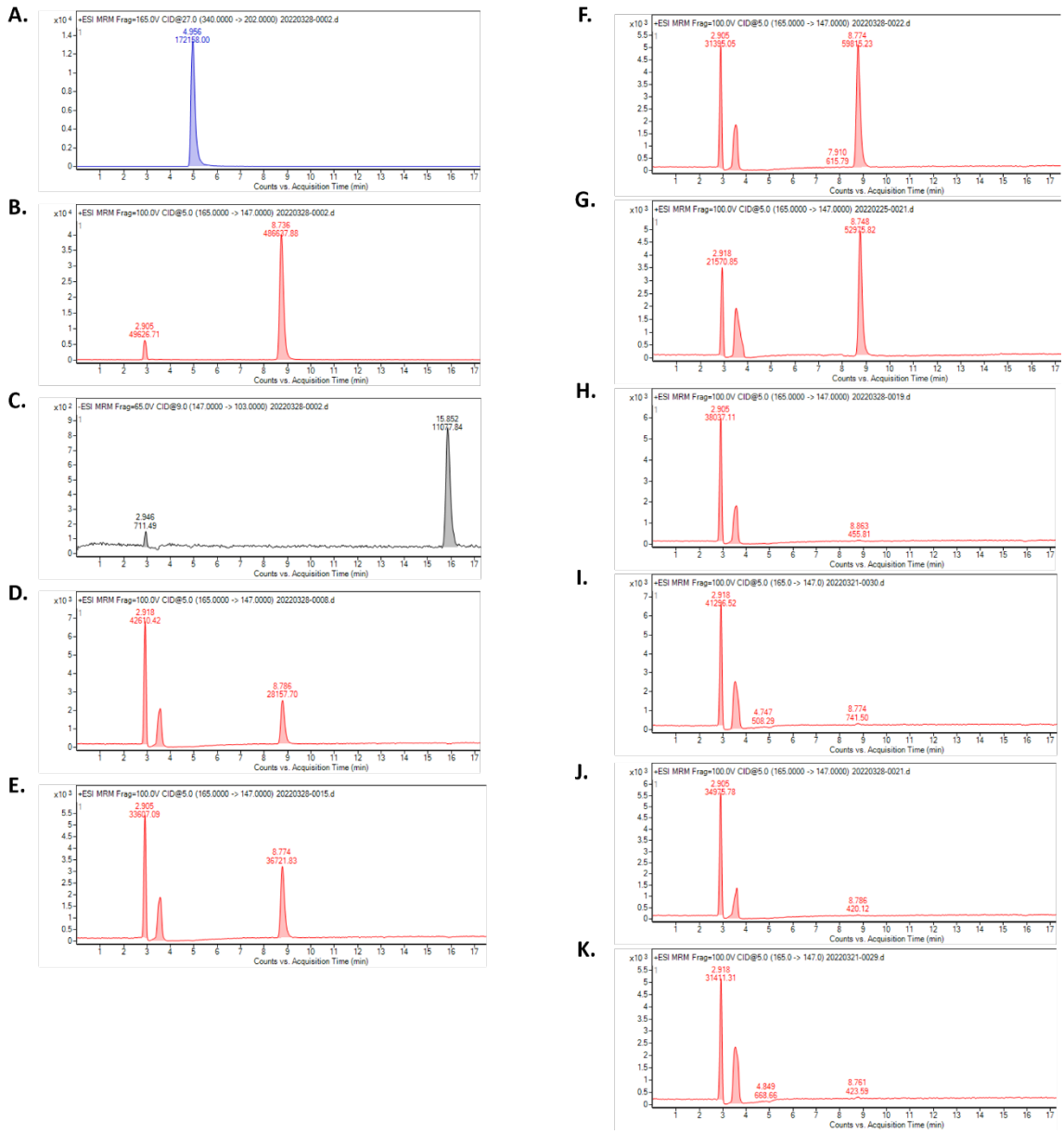


Figure A1-6. HPLC-MS/MS analyses of the enzymatic assay performed with recombinant enzymes *Lae*C4H-Myc. A-C, injection of a mixture of standards (papaverine, *p*-coumaric acid, and *trans*-cinnamic acid); D-G, injection of reaction mixture performed with *Lae*C4H enzyme in 30 min, 1 and 2 h, and overnight incubation; H-K, injection of reaction mixture performed with microsomal proteins extracted from yeast harboring empty vector, No NADPH, No microsomal proteins, No substrate.

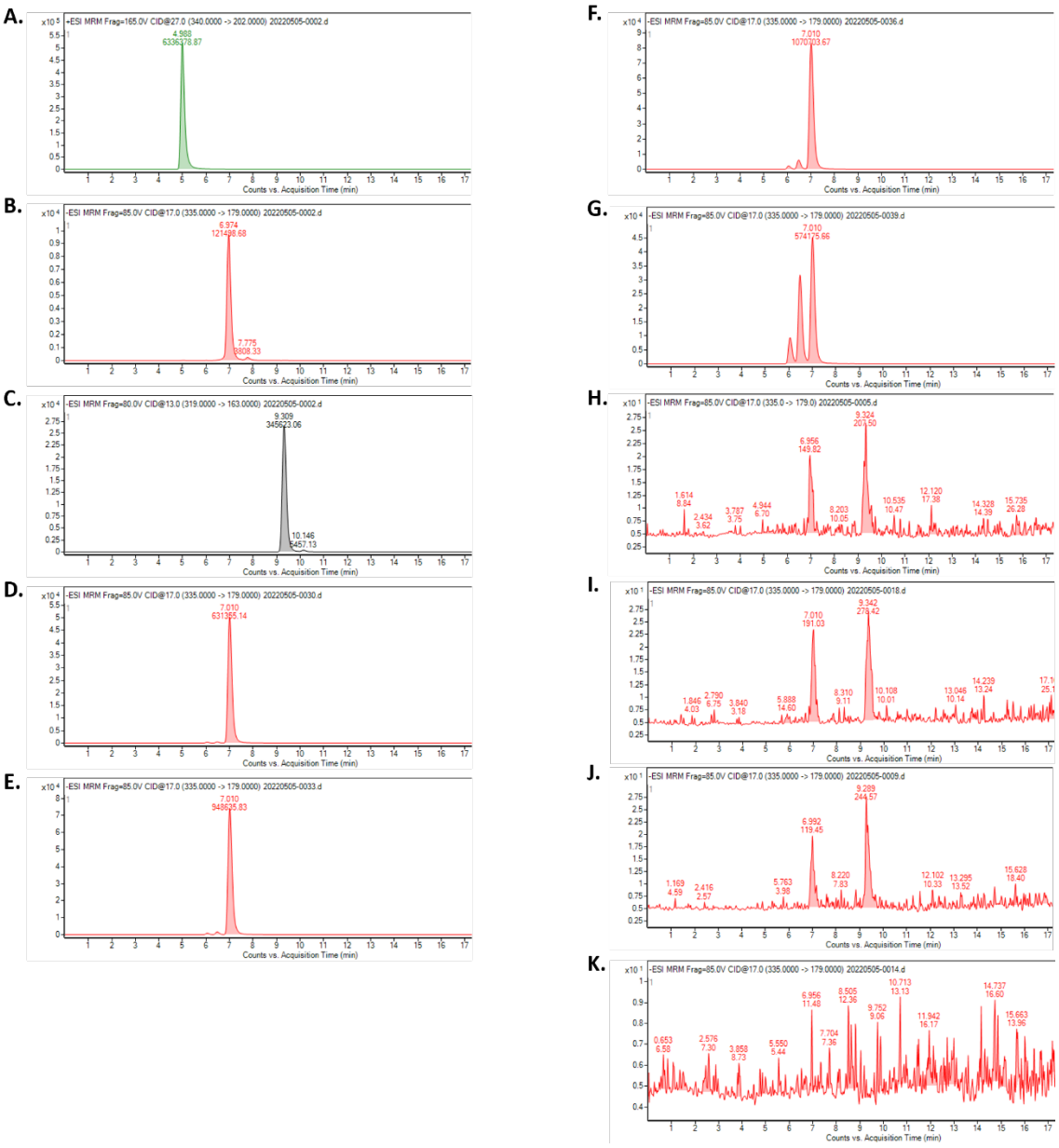


Figure A1-7. HPLC-MS/MS analyses of the enzymatic assay performed with recombinant enzymes *Lae*C3′H -Myc. A-C, injection of a mixture of standards (papaverine, caffeoyl shikimate, and *p*-coumaroyl shikimate); D-G, injection of reaction mixture performed with *Lae*C3′H enzyme in 30 min, 1 and 2 h, and overnight incubation; H-K, injection of reaction mixture performed with microsomal proteins extracted from yeast harboring empty vector, No NADPH, No microsomal proteins, No substrate.

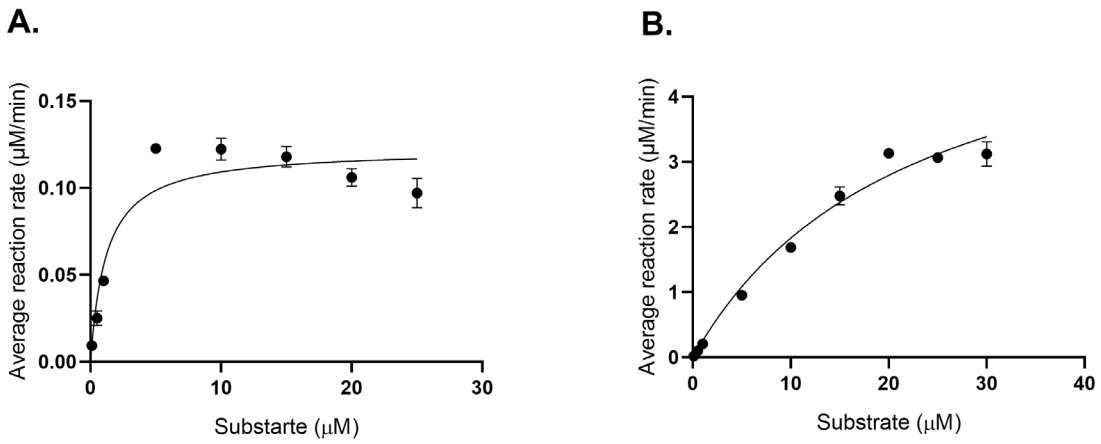


Figure A1-8. Michaelis–Menten plot of initial velocity versus substrate concentration for *Lae*C4H (A) and *Lae*C3′H (B). The production of *p*-coumaric acid and caffeoyl shikimate was monitored for *Lae*C4H and *Lae*C3′H, respectively. The reactions utilized a range of substrates from 0 to 30 μM while keeping the enzyme concentration constant.

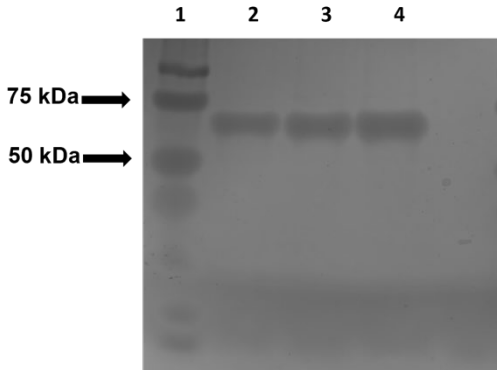


Figure A1-9. SDS-PAGE (10%) analysis of purified APX/C3H fused to MBP tag. Lane 1, protein ladder; Lane 2-4, 2, 4, and 6 μg of purified protein, respectively.

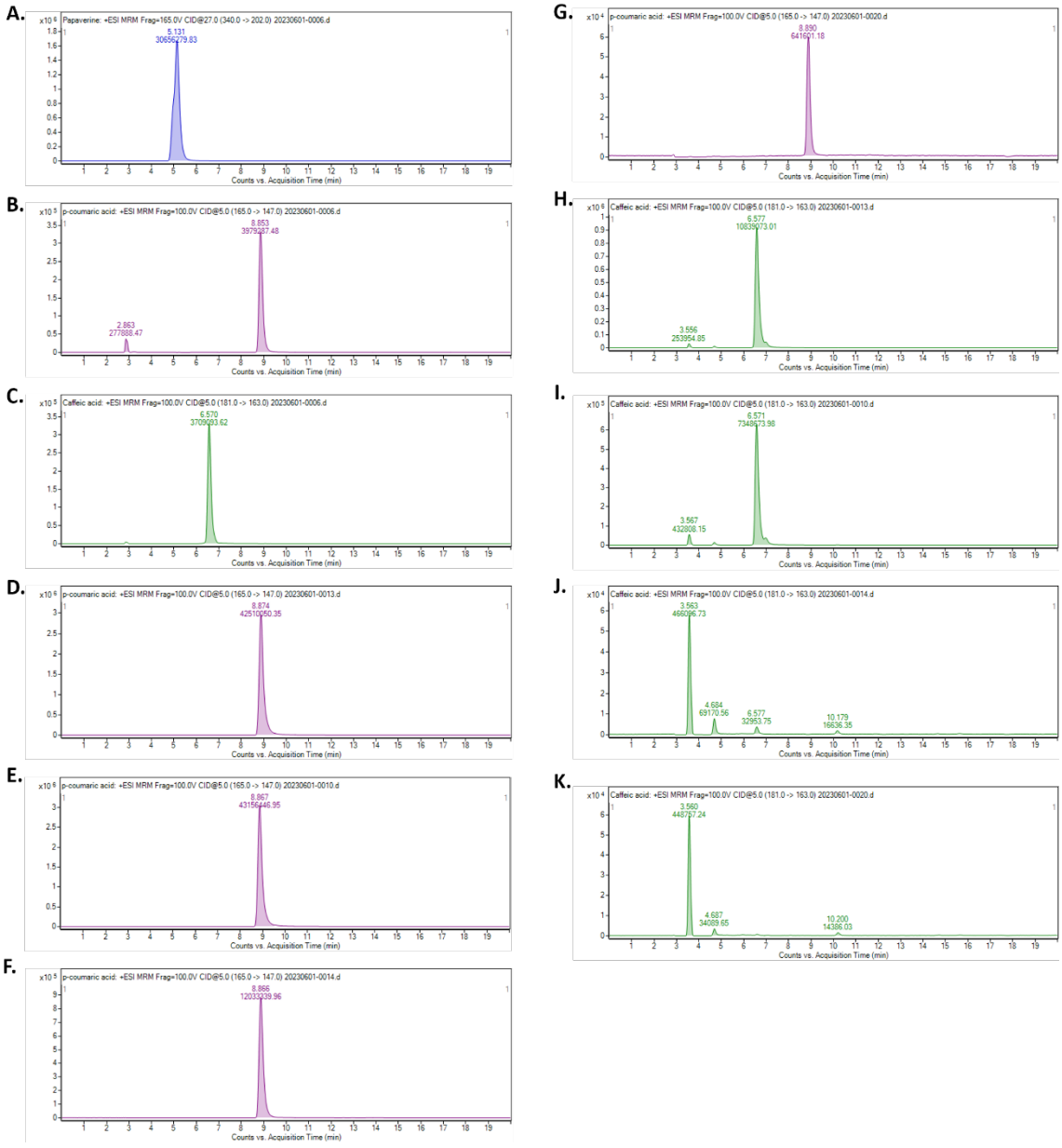


Figure A1-10. HPLC-MS/MS analyses of the enzymatic assay performed with recombinant enzymes *Lae*APX/C3H using *p*-coumaric acid as a substrate. A-C, injection of a mixture of standards (papaverine, *p*-coumaric acid, and caffeic acid); D-G, the peak corresponding to *p*- coumaric acid level in the reaction mixtures performed with 0, 500, 2500, and 5000ng of *Lae*C3′H; H-K, the peak corresponding to caffeic acid level in the reaction mixtures performed with 0, 500, 2500, and 5000ng of *Lae*C3′H.

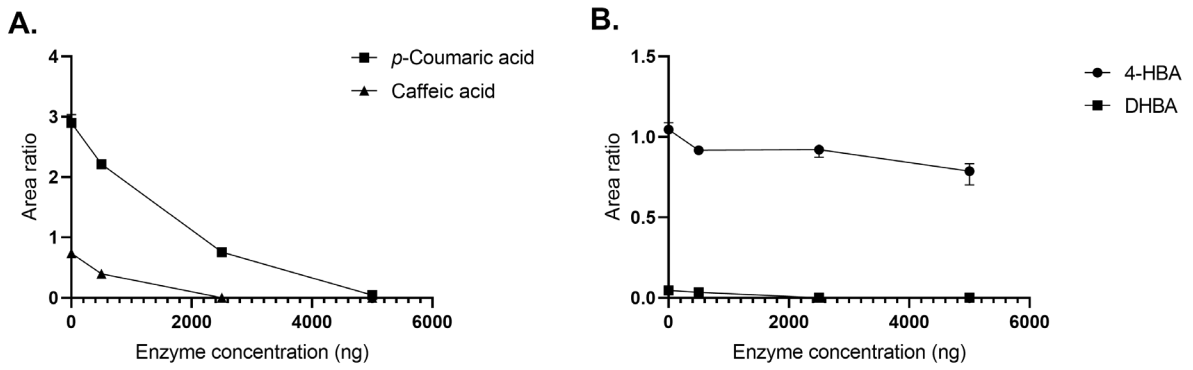


Figure A1-11. Graph representing the area ratio of various compounds corresponding to the *Lae*APX/C3H enzyme assay. A represents *p*-coumaric acid consumption and caffeic acid production in the reactions with enzyme concentrations ranging from 0 to 5000 ng; B represents 4-HBA consumption and DHBA production in the reactions with enzyme concentrations ranging from 0 to 5000 ng. The Y-axis represents the relative area under the curve, calculated as the ratio of each compound's area under the curve to the area under the internal standard (papaverine) curve obtained from HPLC-MS analysis.

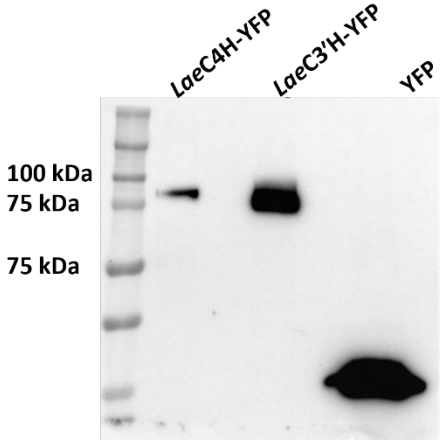


Figure A1-12. Western blot analysis of protein samples extracted from **N. benthamiana** expressing *Lae*C3′H-*YFP* and *LaeC4H-YFP* constructs, respectively. A vector containing the *YFP* sequence alone was used as a negative control. Crude protein extract (50µg) first were separated using SDS-PAGE and then were transferred to a PVDF membrane. Transferred proteins were probed with mouse polyclonal anti-GFP primary antibody and incubated overnight. The membrane was subjected to a secondary antibody conjugated with peroxidase enzyme. Immediately after adding the substrate, the signals were detected by chemiluminescence.

Table A1-5. Summary of *in vivo* enzymatic assay *Lae*C4H and *Lae*C3′H. Three biological replicates were considered for each infiltration condition. The control samples were as follows: a plant without infiltration and feeding to differentiate between endogenous phenolic compounds and altered levels resulting from experimental conditions; plants expressing YFP constructs to differentiate the effect of target genes and the possible effect of *Agrobacterium* infection and consequent changes in phenolic compounds. NA; not applicable.

Infiltrated construct	**Feed**	**Detected compounds**			
		p-coumaroyl shikimate	Caffeoyl shikimate	Caffeic acid	Ferulic acid
NA	NA	-	-	-	-
YFP	NA	-	-	-	+
***Lae*C4H**	NA	-	-	-	-
***Lae*C4H**	*trans*-cinnamic acid	+	+	+	-
NA	*trans*-cinnamic acid	-	-	-	-
***Lae*C3′H**	NA	-	-	-	-
***Lae*C3′H**	*p*-coumaroyl shikimate	+	+	+	+*
NA	*p*-coumaroyl shikimate	+	-	-	-
***Lae*C4H+*Lae*C3′H**	NA	-	-	+	-

Appendix II

Supplementary Data of Chapter III

Table A2-1. Primers were used in the second chapter of the project.

Primer name	Sequence (5'……3')	Purpose of use
NBS-F	GTCAGGATCCAAAATGAAGGGAAGTCTCTCCC	Cloning of *NBS* in pTarget (H1 site) plasmid
NBS-R	TCACTGCTAGCTTACAGATCCTCTTCAGAGATGAGTTTCTGCTCCGCTACAATAGCTTTTTGCT	
NR-F	AGTCAGAATTCAAAATGTTTACAGGAAGAGAAGAGGAA	Cloning of *NR* in pTarget (H2) plasmid
NR-R	ACTGTTAATTAATTACTTATCGTCGTCATCCTTGTAATCACCGTTTATGCCCCG	
OMT-F	ATATACCTCTATACTTTAACGTCAAGGAGAAAAAACCCCGGATCCAAAATGGGTGCTAGCCAAGATGAT	Cloning of *N4'OMT* in pTarget (H5) plasmid
OMT-R	TTACTCGAGGTCTTCTTCGGAAATCAACTTCTGTTCCATGTCGACATAAAGACGTCGACAAATAGTCACTCCA	
TYDC-F	CAACCCTCACTAAAGGGCGGCCGCACTAGTAAAATGGGCAGCCTTGG	Cloning of *TYDC* in pTarget (H7) plasmid
TYDC-R	TCGTCATCCTTGTAATCCATCGATACTAGTTTTTGTAGTCTTTTGTGTTCCCC	
H1-F	GAAGGTGAGACGCGCATAAC	To amplify donor DNAs from corresponding pTarget plasmids
H1-R	GTTTAGCCCATTATGTCTTGTC	
H2-F	GTAGCCTCCCCATAACATAAAC	
H2-R	CACACGAAAAGTCAGAAGAG	
H5-F	GATGACTTCCCATACTGTAATTG	
H5-R	GAACCTTCGATTGCTTGTTAC	
H7-F	CTGTTCATGGCAACGTCAC	
H7-R	GAGCCGAAATTGTGGAAG	

H1-G-F	GTAGTGTGCGTGAATGAAGG	Genotyping of H1 site
H1-G-R	GAAGGTGAGACGCGCATAAC	
H2-G-F	GATTGCAAGGAGAGTGAAAGAG	Genotyping of H2 site
H2-G-F	GTAGCCTCCCCATAACATAAAC	
H5-G-F	AAATCTTGGACAGACAACTTTGAAG	Genotyping of H5 site
H5-G-R	GATGACTTCCCATACTGTAATTG	
H7-G-F	GGTACCTAGCATCATATGGGAAG	Genotyping of H7 site
H7-G-R	CTGTTCATGGCAACGTCAC	
NBS-pESC-U-F	TCGAATTCAACCCTCACTAAAGGGCATGAAGGGAAGTCTCTCCCATG	To clone *NBS* in MSC1 in pESC-URA under Gal10
NBS-pESC-U-R	TCTTATCGTCGTCATCCTTGTAATCCGCTACAATAGCTTTTTG	
NR-pESC-U-F	TCAAGGAGAAAAAACCCCGGATCCGATGTTTACAGGAAGAGAAGAGG	To clone *NR* in MSC2 in pESC-URA under Gal1
NR-pESC-U-F	CTTCTTCGGAAATCAACTTCTGTTCACCGTTTATGCCCC	
OMT-pESC-L-F	TCTTATCGTCGTCATCCTTGTAATCATAAAGACGTCGACAAATAGTCAC	To clone *N4'OMT* in MSC1 in pESC-LEU under Gal10
OMT-pESC-L-R	TCAACCCTCACTAAAGGGCGGCCGCATGGGTGCTAGCCAAG	
TYDC-pESC-L-F	GTCAAGGAGAAAAAACCCCGGATCCATGGGCAGCCTTGG	To clone *TYDC* in MSC2 in pESC-LEU under Gal1
TYDC-pESC-L-R	CTTCTTCGGAAATCAACTTCTGTTCCCCAGCAAGGC	
pESC-URA-1	GCCCTTTAGTGAGGGTTG	To amplify the pESC-URA backbone for the assembly of *NBS* in MCS1
pESC-URA-2	GATTACAAGGATGACGACGATAAG	
pESC-URA-3	GAACAGAAGTTGATTTCCGAAGAAGACCTC	To amplify the pESC-URA-*NBS* backbone for the assembly of *NR* in MCS2
pESC-URA-4	GATTACAAGGATGACGACGATAAGATCTGAGC	
pESC-LEU-1	GCGGCCGCCCTTTAGTGAGGGTTGAATTCG	To amplify the pESC-LEU backbone for the assembly of *N4'OMT* in MCS1
pESC-LEU-2	GGATCCGGGGTTTTTTCTCCTTGACG	
pESC-LEU-3	GAACAGAAGTTGATTTCCGAAGAAGACCTC	To amplify the pESC-LEU -*N4'OMT* backbone for the assembly of *TYDC* in MCS2
pESC-LEU-4	GAACAGAAGTTGATTTCCGAAGAAGACCTC	
GAL10-F	ATGTATATGGTGGTAATGCCATGTAATATG	To verify the insertion of *NBS* or *N4'OMT* in MCS1 of pESC-URA or pESC-LEU plasmids
tADH1-R	GAGCGACCTCATGCTATACCTG	
GAL1-F-Mod	TTCAACATTTTCGGTTTGTATTACTTCTTATTC	

Cyc1TerCasseteRev	CTTCGAGCGTCCCAAAACC	To verify the insertion of NR or TYDC in MCS2 of pESC-URA or pESC-LEU plasmids

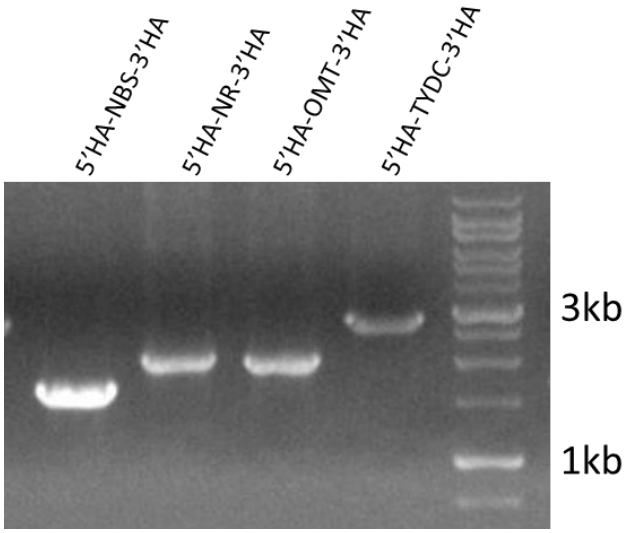


Figure A2-1. PCR amplified donor DNAs on 1% agarose gel. The donor DNAs were amplified from pTarget plasmids, H1, H2, H5, and H7, plasmids harboring *NBS*, *NR*, *OMT*, and *TYDC* sequences using primers listed in Table A2.1. HA; Homology arm.

Appendix III, Review Paper (Co-authorship)

Review

Biotechnological approaches to optimize the production of Amaryllidaceae alkaloids

Manoj Koirala[1], Vahid Karimzadegan[1], Nuwan Sameera Liyanage[1], Natacha Mérindol[1], and Isabel Desgagné-Penix [1,2,*]

[1] Department of Chemistry, Biochemistry and Physics, Université du Québec à Trois-Rivières, 3351, boul. des Forges, C.P. 500, Trois-Rivières, Québec, Canada G9A 5H7; Manoj.Koirala@uqtr.ca (M.K.); Natacha.Merindol@uqtr.ca (N.M.); Isabel.Desgagne-Penix@uqtr.ca (I.D-P.)

[2] Groupe de Recherche en Biologie Végétale, Université du Québec à Trois-Rivières, 3351, boul. des Forges, C.P. 500, Trois-Rivières, Québec, Canada G9A 5H7; Isabel.Desgagne-Penix@uqtr.ca (I.D-P.)

* Correspondence: Isabel.Desgagne-Penix@uqtr.ca

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Biomolecules* **2022**, *12*, x. https://doi.org/10.3390/xxxxx

Academic Editor: Firstname Lastname

Received: date
Accepted: date
Published: date

Abstract: Amaryllidaceae alkaloids are plant specialized metabolites with therapeutic properties exclusively produced by the Amaryllidaceae plant family. The two most studied representatives of the family are galanthamine, an acetylcholinesterase inhibitor used as a treatment of Alzheimer's disease, and lycorine displaying potent *in vitro* and *in vivo* cytotoxic and antiviral properties. Unfortunately, the variable level of AA' production *in planta* restricts most of the pharmaceutical applications. Several biotechnological alternatives, such as *in vitro* culture or synthetic biology, are being developed to enhance the production and fulfil the increasing demand for these AA plant-derived drugs. In this review, current biotechnological approaches to produce different types of bioactive AA are discussed.

Keywords: Amaryllidaceae alkaloids; bioactive molecules; biotechnological approach; biosynthesis; *in vitro* cultures; synthetic biology

1. Current challenges in the production of Amaryllidaceae alkaloids

Amaryllidaceae alkaloids are isoquinoline alkaloids exclusively isolated from the Amaryllidaceae plant family. AA are structurally diverse biomolecules classified into different types (or groups) based on their structure, biogenetic origin, or chemical nature [203-206]. They can be divided into nine groups: norbelladine, cherylline, galanthamine, lycorine, homolycorine, crinine, pancratistatin, pretazettine and montanine, according to their ring type and biosynthetic origin [207]. Among all the AA, the reversible acetylcholinesterase inhibitor galanthamine is yet the only one to be approved for medicinal purposes to treat early symptoms of Alzheimer's disease

in humans [208]. Although the mechanism of action is still not fully understood, two main speculations have been proposed. Galanthamine reversibly, competitively, and selectively inhibits acetylcholinesterase, an enzyme known for acetylcholine degradation, so that the neurotransmitter associated with memory formation and learning will be available for a longer time in the synaptic cleft of cholinergic neurons to transfer neuro signals [209, 210]. In addition, galanthamine allosterically binds to nicotinic acetylcholine receptors of the central nervous system that control the release of different neurotransmitters, altering their conformation and increasing neurotransmitter secretion [209]. AChE inhibitory action of galanthamine also decreases the level of reactive oxygens [166], oxidative stress being a common adverse effect of many human diseases such as Alzheimer's, Parkinson's, Down syndrome, cancer, *etc*. this hints towards a neuroprotective effect.

Galanthamine production has mainly relied on natural resource exploitation from species such as *Galanthus*, *Leucojum*, *Narcissus*, *etc*. Providing galanthamine to the 55 million people living with dementia cannot solely rely on the plant source, and in the case of some species like *Leucojum*, it has already endangered the biodiversity of the wild population in the past years [211]. As an alternative strategy, the chemical synthesis of galanthamine has been attempted [212, 213]. However, multi-step synthesis of structurally complex compounds such as galanthamine is not economically competitive compared to extraction from native plants due to the low final yield [214].

Lycorine, another prominent AA, exhibits a broad spectrum of biological activities, including anti-viral, anti-bacterial, anti-parasitic, and anti-inflammatory properties, and it has been particularly studied for its anticancer activity [215]. Lycorine's antitumor potency involves several pathways, such as induction of apoptosis and necrotic cell death, inhibition of cell cycle, autophagy, and metastasis, probably aiming at multiple molecular targets [215]. Its high cytotoxic potency at low concentrations makes lycorine's structure an interesting leading molecule for designing new anticancer drugs. Recently, the less abundant AA cherylline was also shown to possess anti-flaviviral potential, inhibiting both dengue and Zika viruses at the viral RNA replication step, with EC_{50} of 8.8µM and 20.3µM respectively [190]. In fact, novel AA with anti-acetylcholinesterase, anti-viral, cytotoxic, anticonvulsant, antitumor, hypotensive, and anti-inflammatory properties are continuously discovered [190, 205, 207, 216]. Their pharmacological potential depends on their complex chemical structure, including their region-specific

functionalization and chirality [217]. Consequently, it is often challenging and not always cost-effective nor ecological to chemically synthesize intact structures of AA. Although there are some reports on the successful chemical synthesis of AA, such as galanthamine, lycorine, or cherylline, the multiple steps involved lead to a low overall yield, while the end products may contain a mixture of *R* and *S* enantiomers [213, 218-223].

Currently, plants are the main source of AA. Even though many AA with interesting pharmacological potentials were identified, clinical application and further research are restricted mainly because of the variable and low production levels *in planta*. For example, cherylline-type AA are rare because they are specifically recovered from few species of *Crinum* (0.004% crude alkaline solution in *C. powelli*) and in about twenty ornamental cultivars of *Narcissus* (extraction from the leaves of jonquilla and apodanthus daffodil cultivar "sundial") [224, 225]. Furthermore, the synthesis and accumulation of AA in plants vary with environmental and seasonal changes throughout the year. For example, the alkaloid content of *Cyrtanthus contractus* changes from 667.4 to 1020.6 µg/g between months of the same year [226]. In addition, massive harvesting for the extraction of alkaloids decreases the number of Amaryllidaceae in nature and has led some species to become endangered, such as *Narcissus asturiensis* [227].

Therefore, intensive research on sustainable alternative techniques is carried out to achieve economical and eco-responsible production of pharmacologically active AA [228-231]. In recent years, a book chapter by Laurrain-Mattar *et al.,* and two reviews well describing such techniques were published, emphasizing the growing interest in this matter [232-234]. As an alternative to native source harvesting or chemical synthesis, biotechnological strategies offer many advantages to sustainably producing AA (Figure A3-1). This includes cultivating plants and plant parts in artificial systems or synthetic biology (metabolic engineering) of heterologous hosts for AA production (Figure A3-1). In addition, different approaches have been reported to optimize AA yield in plants and *in vitro* cultures. In this review, we discuss recent progress on the biotechnological approaches and overall factors affecting their efficiency, together with future perspectives to boost the synthesis and accumulation of AA.

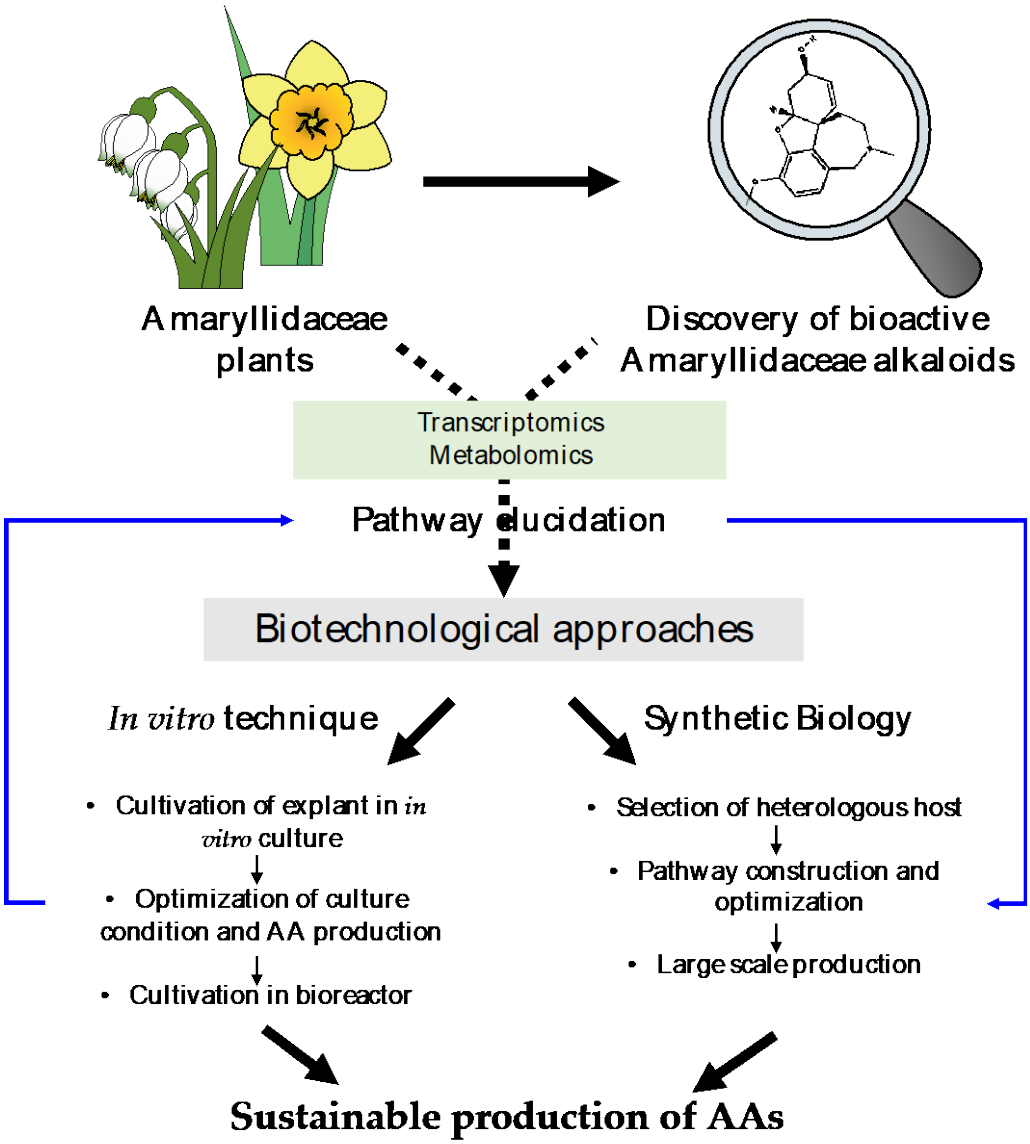


Figure A3-1: Current and future biotechnological approaches to produce AAs.

2. *In vitro* techniques to produce AAs

In vitro systems hold the beneficial ability to produce plant SMs in a sustainable way continuously. They also enable the control of environmental factors, a complicated task in nature, providing the opportunity to analyze the effect of different variables in the production of SMs [235]. Several therapeutic and marketed metabolites have been produced using *in vitro* cultures, such as the anti-bacterial and anti-inflammatory naphtoquinone shikonin from *Lithospermum erythrorhizon*, the chemotherapeutic agent paclitaxel from *Taxus baccata, the antioxidant* saponins from *Panax ginseng* cells, the bioactive alkaloids berberine and sanguinarine from *Coptis japonica* and *Papaver somniferum cultures, respectively, as reviewed in* [236] and [237].

Thus, reports on many other plant families have shown that *in vitro* cell culture techniques can be a fantastic platform for producing specialized molecules and understanding their biosynthesis [139, 238]. The specific interest in Amaryllidaceae *in vitro* cultures to produce alkaloids was first reported in 1963 by Fales et al. [196] and has been intensively and continuously studied.

In vitro techniques involve the transfer of healthy sterile explants into artificial conditions using suitable growth media. It can be applied to grow whole plants, parts, or undifferentiated tissues. Micropropagation, a technique that enables rapid vegetative/clonal multiplication of plants from limited or small-size plants, has been successfully applied with Amaryllidaceae species such as *Rhodophiala pratensis*, *Lapiedra martinezii*, *Eucrosia stricklandii*, and *Lycoris sprengeri* leading to plant development with similar morphometric traits [227, 239-242]. Other cultivation methods of plant material *in vitro* (bulblets, seedlings, plantlets, shoots, roots, shoot-clump, callus) also provide an interesting opportunity to produce AA, effective for conservation, long-term growth, and industrial purposes. Specifically, callus induction is defined as the growth of undifferentiated tissues from any plant parts. Because different plant parts produce different amounts and types of alkaloids, the obtained type of callus and its metabolite content may be related to the type of tissue used as a starting material [243]. The production of uncommon but interesting AA such as cherylline, tazettine, haemanthamine, and mesembrenone was reported in various studies of *in vitro* propagation of Amaryllidaceae species (Table A3-1). For example, up to 6.9 mg/100 g DW of anti-acetylcholinesterase and anti-viral cherylline [190, 244] was observed in bulblet cultures of *Crinum moorei* cultivated in the presence of charcoal [238].

Table A3-1. Yields of uncommon AA of therapeutical interest in *in vitro* cultures.

Target metabolites	Species	Tissue type	Maximum yield	Ref
Cherylline	*Crinum moorei*	Bulblets	6.9 mg/100 g DW	[245]
Haemanthamine	*Rhodophiala pratensis*	Callus	6.9 μg/mg Ext	[239]
	Narcissus cv. Hawera	Plants	25.5 μg/100 mg Ext	[246]
Powelline	*Crinum moorei*	Bulblets	46.84 mg/100 g DW	[245]
Tazettine	*Rhodophiala pratensis*	Callus	2.68 μg/mg Ext	[239]
	Narcissus confuses	Shoot-clump culture	0.043 % DW	[247]
Mesembrenone	*Narcissus pallidulus*	Plants	337.6 μg/100 mg Ext	[246]
	N. cv. Hawera	Plants	214.8 μg/100 mg Ext	[246]

DW = Dry weight, Ext: Extract

Additionally, the production of well-known bioactive AA has been investigated in *in vitro* cultures (Table A3-2). Although undifferentiated tissues do not always yield high amounts of alkaloids, elicitation helps increase the yield of various AA in *vitro* cultures (Table A3-2) [248, 249] Still, this technique is advantageous because it can maintain growth for long periods of time, and callus can be used as a gateway for micropropagation, plant cell suspension cultures, or other *in vitro* systems to produce alkaloids. Indeed, because of somaclonal variations, shoots grown from callus displayed galanthamine production in some studies [250].

Table A3-2: *In vitro* production of AAs cultures following elicitor treatment.

Species	Culture condition (tissue type)	Amaryllidaceae alkaloid	Yield and type of condition	Elicitor and yield	Ref
Narcissus confuses	Liquid-shake culture (shoot clumps)	Galanthamine	ut. 2-2.5 mg / culture	MeJa (3.8 X)	[251]
***N. pseudonarcissus* cv. Carlton**	Callus	Galanthamine	ut. 7.88 µg/g FW	MeJa (5.6 X) Chitosan (3 X)	[252]
Lycoris longituba	Liquid medium (seedling)	Galanthamine	ut. n.a.	MeJa (2.71 X)	[253]
		Lycorine	ut. n.a.	MeJa (2.01 X)	
		Lycoramine	ut.n.a.	MeJa (2.85 X)	
	Seedling (culture in tray)	Galanthamine	white light n.a.	Blue light (2.45X)	[254]
		Lycorine	white light n.a.	Blue light (1.74 X)	
		Lycoramine	white light n.a.	Blue light (1.92 X)	
Lycoris chinensis	seedling	Galanthamine	ut. n.a.	MeJa (1.49 X) YE (1.62 X) SNP (1.72 X)	[255]
		Lycorine	ut. n.a.	MeJA (1.37 X) YE (1.38 X)	
L. aestivum	*In vitro* plants	Galanthamine	ut. n.a.	Melatonin (58.6X)	[256]
		Lycorine	ut. n.a.	Melatonin (1.5 X)	
	Liquid shoot culture	Galanthamine	ut. n.a.	JA (1.36 X)	[257]
		Lycorine	ut. n.a.	JA (1.40-1.67 X) MeJa (1.3 X)	
		Norgalanthamine	ut. n.a.	JA (2X) MeJa (2X)	
	temporary immersion system (bulblets, leaves)	Galanthamine	ut. 372.2-1719.6 µg/g DW	MeJa (468.6-2202.5 µg/g DW)	[258]
L. aestivum L.	RITA Bioreactor	Galanthamine	ut. n.a.	MeJa (0.1 mg /g DW) ACC (0.10 mg/ g DW)	[259]
		Lycorine	ut. 0.2-0.25 mg /g DW	MeJa (0.6 mg /g DW) SA (1 mg /g DW) Ethephon (0.46 mg /g DW)	
***L. aestivum* Gravety Giant**	RITA Bioreactor	Galanthamine	ut. 0.08-0.1 mg/g DW	MeJa (0.4 mg/g 5X DW) SA (8x) ACC (0.60 mg/g)	[259]
		Lycorine	ut. 0.15 -0.62 mg/g DW	MeJa (1.15 mg/g DW 1.85X) SA (5X) ACC (0.54mg/g DW, 3.6X)	

Abbreviations; n.a.: not available; ut: untreated, basal condition, MJ: methyl jasmonate, JA: jasmonic acid; SA: salicylic acid; ACC: 1-aminocyclopropane-1-carboxylic acid; FW: Fresh weight; DW : Dry weight; X: fold change; YE: yeast elicitor; SNP: sodium nitroprusside; Ref: reference.

In addition to a proper selection of tissue and *in vitro* propagation technique, many other factors affect the growth and the efficiency of alkaloid production. In natural conditions, the growth of Amaryllidaceae plants and their ability to synthesize different types of alkaloids vary throughout the year and are influenced by biotic and abiotic environmental factors that affect the synthesis and accumulation of AA [226, 260, 261]. Therefore, understanding and controlling both the effects of growth conditions and the plant defense response mechanism would help increase biomass production with higher yields of SMs such as AA. Remarkably, these factors can be monitored and optimized using *in vitro* methods. Studies have shown that Amaryllidaceae plants grow differently under distinctive artificial conditions. The overall goal of all culture techniques is to provide optimal growth conditions and boost the production of alkaloids. The focus of most of the published research was galanthamine production, while there is also abundant data on lycorine production optimization. Table A3-2 illustrates the impact of different elicitation conditions on AA production in *in vitro* systems. Bergoñón *et al.* achieved a total production of 2.50 mg of galanthamine per culture by cultivating shoot clumps in shaking-liquid media which is the highest amount of *in vitro* production of galanthamine ever reported to date [262].

In the following section, we briefly discuss the effect of the different physical and chemical parameters used in the studies summarized here. These parameters may be applied to the in vitro system and eventually affect the growth and development of Amaryllidaceae plant tissue culture, as well as the synthesis and accumulation of alkaloids.

2.1 Physical parameters

Different cultivation methods have been tested to optimize the *in vitro* growth of Amaryllidaceae plants; this includes solid media culture, shaken-flask submerged condition, and temporary immersion or fully-subversive techniques in RITA® bioreactor (Table A3-1 and 2) [263-265]. Pavlov *et al. grew Leucojum aestivum 80 shoot culture in shaken flasks following induction from callus and monitored their growth index*. They observed that the maximum biomass was obtained

at day 35 and that AA biosynthesis intensified at late exponential to early stationary growth phases [228]. In the following study, to optimize the production of target metabolites (mostly galanthamine), L. aestivum 80 shoots were cultivated in a temporary immersion RITA system with a higher growth index (i.e., 2.98) compared to shaken-flasks culture. The main advantage of the temporary immersion system was that cultivated shoots increased significantly, while shoots generated from meristematic cell [265]. The system was further improved for *L. aestivum* 80 shoot culture using an advanced modified gas column bioreactor with a 1.7 mg/L maximum production of galanthamine [266]. *L. aestivum* shoot cultures show balanced growth at all tested regimes in this bubble-column bioreactor. Similar techniques using twin scale from bulbs as starting materials have shown that *Narcissus confuses* shoot-clump culture in a liquid-shake medium leading to an efficient micropropagation system to produce galanthamine (2.50 mg per culture in a day-long photoperiod) [262].

In addition to the type of *in vitro* cultivation system, temperature is a key factor that modulates both growth and alkaloid production in Amaryllidaceae species. Some studies have shown that among different culture temperatures (*i.e.,* 18°C, 22°C, 26°C, and 30°C), the maximum yield of galanthamine was achieved at 26°C whereas the best combination of highest amount of dry mass (20.8 g/L) and galanthamine content (1.7 g/L) was achieved at 22°C when shoot culture were grown under 18 L/(L·h) flow rate of inlet air [265-267]. Others have been more successful at lower temperatures. Ivanov et al. obtained 18 different AA from the shoot culture of *L. aestivum* 80 and reported that lower temperature (18°C) favored galanthamine production while inhibiting the production of lycorine- and galanthamine-types of alkaloids. They concluded that temperature possibly alters the activity of the enzymes catalyzing phenol-oxidative coupling reaction of 4'-*O* methylnorbelladine [268].

Light is another important factor that can boost the production of AA in in vitro cultures. In general, studies suggest that light has a positive impact on alkaloid production in Amaryllidaceae tissues [228, 261, 269, 270]. In shoot cultures of N. confuses light affected both morphology and alkaloid content [262]. In N. tazetta L., bulblets and leaves proliferation per explant were higher in light condition light/dark photoperiod (16/8h), as compared to a 24h dark condition. Also, regenerated bulblets contained 40 µg/g dry weight of galanthamine under exposed photoperiod compared to 20 µg/g dry weight for 24 h dark condition [270]. The light condition and quality have been studied in relation to the production of AA. For example, it was observed that increases

in the production of galanthamine (2.45 times), lycorine (1.74 times), and lycoramine (1.92 times) were generated by blue light condition compared to white light in in vitro plantlets of Lycoris longituba [254]. These studies demonstrate that a complex combination of physical parameters impacts alkaloid production, and that cultivation system, light, and temperature should be optimized in each system, which may vary between species.

2.2. Chemical factors

Plants require diverse types of micro- and macro-elements for their growth and development. Generally, the type of media and the addition of growth factors, hormones, or other regulators, such as charcoal, affect both the growth and the production of metabolites [271]. Activated charcoal is known to promote plant cells and callus differentiation, possibly through the induction of genes from the phenylpropanoid biosynthesis pathway, which could lead to an increase in alkaloid production[271-273]. The quantity of carbon and nitrogen in the media affects both growth and the production of SMs. Studies have shown that the type and concentration of different carbon sources play an important role in plant tissue culture. In plant tissue culture research, sucrose is widely used as a carbon source. The effect of its concentration on N. confuses culture was measured following growth in a liquid-shake medium by Sellés et al. (1997) [247]. They showed that sucrose concentration affects biomass production and galanthamine synthesis. Assessing the effects of concentrations ranging between 3% - 18% of sucrose, an optimal combination of growth and galanthamine production was achieved with 9% [247]. In 2020, Ptak et al. showed that both the concentration and the type of carbohydrate are critical for the synthesis of AA. Similar to Sellés study, they demonstrate that the highest amount of L. aestivum biomass was obtained with 9% sucrose when cultivated in RITA® bioreactor. However, they also show that using other types of sugar can increase the success of AA synthesis, as the highest amount of galanthamine was recorded with 3% fructose [274].

The type of phytohormones and their concentration play a key role in tissue survival and differentiation, including in organogenesis and eventually in the synthesis of AA. Auxins, abscisic acid, cytokinin, ethylene, and gibberellins are commonly recognized as the five main classes of naturally occurring plant hormones. The isolated or combined effect of auxin or cytokinin on *in vitro* cultures' nutrient uptake and metabolism was demonstrated [275]. In 2011, Tahchy *et al.* observed that an increment of auxin (2,4-dicholophenoxyacetic acid or 2,4-D) in growth media

reduced the survival of *in vitro* cultured tissues of *N. pseudonarcissus, Galanthus elwesii,* and *L. aestivum*, whereas an increment of cytokinin (6-benzylaminopurine, BAP) increased it [276]. The concentration and the type of phytohormones also influence the type of tissue that will develop, eventually affecting the alkaloid profile. Hence, the alkaloid profile of *L. aestivum* and *N. pseudonarcissus cv. Carlton* obtained from *an in vitro* system revealed that differentiated cells are more suitable for producing AA than undifferentiated cells (*callus),* in which alkaloid contents were lower [228, 248]. A study completed on different varieties of *Narcissus* showed that undifferentiated calli development was induced following treatment with auxin concentrations of 25 µM of 1-naphthalene acetic acid (NAA), 50 µM of 2,4-D, and picloram. In contrast, organogenesis only happened when calli were treated with NAA, or when using a higher concentration of picloram. Interestingly, AA demethylmaritidine and tazettine were only detected on the differentiated callus [249]. Most studies have confirmed that the amount and the type of auxin correlated with the alkaloid profile and tissue differentiation during *in vitro* culture [248, 249]. However, no consensual combination can be clearly defined at this point.

Although the eco-physiological role of many plants' SMs is not clear, studies demonstrated that different biotic and abiotic factors or signaling agents (elicitors) can boost the production of AA [255, 261, 277]. Different elicitors, such as fungal elicitors, methyl jasmonate, jasmonic acid, salicylic acid, and melatonin, have been used to enhance the synthesis of AA in *in vitro* culture (Table A3-2). The induction of AA using methyl jasmonate treatment on seedlings of *Lycoris aurea* was well-studied by Wang *et al.* (2017) [278]. Others have deciphered that methyl jasmonate and jasmonic acid increased the production of AA in *L. aestivum.* Shoot cultures cultivated in submerged conditions by stimulating two enzymes involved in the formation of AA precursors [257]. Melatonin addition (10 µM) during *in vitro* culture of *L. aestivum L.* reduced the negative effect of NaCl (*i.e.,* salt stress) and enhanced the biomass production together with an increased accumulation of galanthamine and lycorine by 58.6 and 1.5 folds, respectively (Table A3-2) [256]. In conclusion, optimizing media components and elicitor type in *in vitro* culture of Amaryllidaceae provides an alternative and sustainable source of AA.

3. Genetic engineering of heterologous host for alternative production of AA

Bioengineered microbial hosts that grow rapidly can produce plant-target SMs faster than whole plant systems. In addition, producing plant metabolites in heterologous hosts can reduce the

downstream extraction process, eventually becoming more economically sustainable. For the successful synthesis of plant metabolites such as AA, heterologous hosts require the introduction of a reconstructed biosynthetic pathway, requiring key enzymes. This requires comprehensive knowledge of the enzymatic reactions involved in the biosynthesis of the compound of interest in the native host organism (i.e., plants).

3.1 Molecular understanding of AAs biosynthesis

Even though AA's pharmacological aspects have been extensively explored, fully understanding the AA biosynthetic pathway and the characterization of enzymes responsible for catalyzing the different biosynthetic reactions demands more effort. This knowledge would enable establishing improved systems or sustainable platforms to produce these valuable biologically active compounds. Combined application of early labeling study followed by the latest omics strategies have accelerated the discovery of AA biosynthetic enzymes [8, 279]. After the proposition of the biosynthetic route of different intermediates, several biosynthetic enzymes were predicted based on the nature of the biochemical reaction and by homology with enzymes involved in alkaloid biosynthesis of other plant families. Databases generated from transcriptomic and metabolic analysis of different species of Amaryllidaceae support the presence of different enzyme families involved in the AA pathway [280-282].

The AA biosynthetic pathway utilizes two common amino acids, L-tyrosine and L-phenylalanine, building blocks to produce a vast range of alkaloids with diverse biological activities. AA biosynthesis's first reactions involve forming the ‘precursors’ from the phenylpropanoid and tyramine pathways (Figure A3-2). As such, L-tyrosine is decarboxylated by the enzyme tyrosine decarboxylase (TYDC) to yield tyramine while the production of the second building block, DHBA, is achieved via the phenylpropanoid pathway by the action of enzymes such as phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H), p-coumarate 3-hydroxylase (C3H), to name but a few. TYDC was characterized by *Lycoris radiata*, a galanthamine-producing Amaryllidaceae plant [283]. The functional characterization of PAL and C4H in *L. radiata* was reported using heterologous expression in bacteria [134].

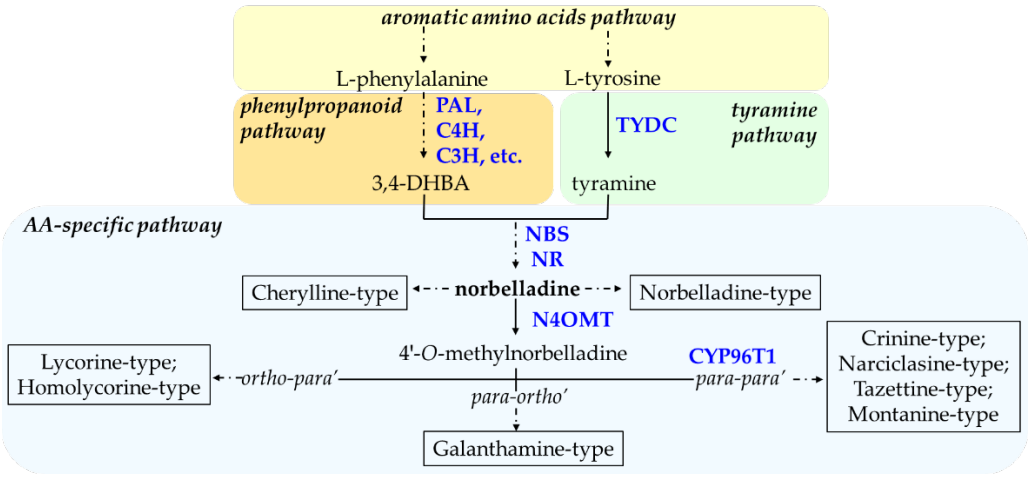


Figure A3-2. Biosynthetic routes to main types (boxed) of Amaryllidaceae alkaloid (AA). Arrows without labeling reflect chemical reactions where no enzyme was characterized. Enzymes that have been identified are labeled in blue. A solid arrow shows one enzymatic step, whereas a broken arrow symbolizes multiple enzymatic reactions. Following 4'*O*-methylnorbelladine, the regioselective phenol-phenol coupling reaction is indicated in the broken arrow leading to various AA-types. Enzyme abbreviations: DHBA, 3,4-dihydroxybenzaldehyde; PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; C3H, coumarate 3-hydroxylase; TYDC, tyrosine decarboxylase; NBS, norbelladine synthase; NR, noroxomaritidine/norcraugsodine reductase; N4′OMT, norbelladine 4'-*O*-methyltransferase; CYP96T1, cytochrome P450 monooxygenase 96T1.

Despite having a remarkable structure and biological activity diversity, all AAs are derived from a common intermediate, norbelladine. The condensation of tyramine and DHBA yields norbelladine and was shown to be catalyzed either by norbelladine synthase (NBS) or by noroxomaritidine/norcraugsodine reductase (NR), in both cases with low yield [280, 284, 285]. NBS was characterized by *N. pseudonarcisus* king Alfred and *L. aestivum* [280, 285]. GFP-tagged *La*NBS and CFP-tagged NR showed that both enzymes are localized to the cytosol, which suggests that the first committed step of AA biosynthesis probably occurs in the cytosol. [51, 52].

Norbelladine can either be utilized directly to generate norbelladine- and cherylline-type AA or be further methylated by norbelladine 4'-*O*-methyltransferase (N4′OMT) to give 4'-*O*-methylnorbelladine (Figure A3-2). The structural feature of cherylline-type AA suggests the occurrence of 3'-*O*-methylation during the biosynthesis of these types of AA, although it remains to be proven. In addition to a norbelladine 3'-O-methyltransferase (N3OMT), cherylline biosynthesis would require steps catalyzed by 3'-O-methylnorbelladine N-methyltransferase (MNB-NMT) and 3'-O, N-dimethylnorbelladine 2-hydroxylase (DMNB-2H) or a cherylline synthase (CherySyn.) (Figure A3-3). The specific synthesis of both 3'-*O*-methylated and 4'-*O*-

methylated AA suggests that regioselective methylation is essential to determine the types of the product of the AA biosynthesis route. The characterization of norbelladine OMT from *Narcissus sp. aff. pseudonarsissus* suggests that methylation by *Np*N4′OMT happens specifically at the 4'-*O* position of norbelladine [286]. However, later studies on *L. radiata* OMT (*Lr*OMT) propose that methylation can occur either in the 3'-*O* or 4'-*O* position of norbelladine, DHBA, or caffeic specifically at the 4'-*O* position of norbelladine [286]. However, later studies on *L. radiata* OMT (*Lr*OMT) acid. A kinetic study of *Lr*OMT indicates a higher affinity for DHBA as a substrate than norbelladine. The methylated forms of DHBA (*i.e.,* vanillin and isovanillin) could also be condensed with tyramine to generate 3' or 4'-*O*-methylnorbelladine. However, until now, none of the possible methylated forms of 3, 4-DHBA substrate have been tested as substrates for NBS.

One step deeper in the AA pathway and depending on the type of phenol-coupling reaction, the 4'-*O*-methylnorbelladine can be directed to 1) galanthamine-type through *para-ortho'*, 2) lycorine-type AA by *ortho-para'* and 3) crinine-type of AA by *para-para'* phenol coupling reactions (Figure A3-2). These C-C phenol-coupling reactions are putatively catalyzed by members of the cytochrome P450 enzyme family. For example, *Np*CYP96T1 was shown to catalyze the *para-para* oxidative reaction of 4'-*O*-methylnorbelladine into noroxomaritidine and was also shown to catalyze the formation of the *para-ortho'* phenol coupled product, *N*-demethylnarwedine as less than 1% of the total product [55]. Aside from CYP96T1 and NR, no other steps (genes or enzymes) have been identified in the formation of phenol-coupled AA-types to date (Figure A3-2). In a nutshell, galanthamine production could be catalyzed by para-ortho' phenol oxidative coupling of 4'-O-methylnorbelladine to yield nornarwedine by nornarwedine synthase (NNS); then a nornarwedine reductase (NNR) and finally a norgalanthamine N-methyltransferase (NG-NMT) (Figure A3-3).

Plants synthesize SMs using complex biosynthetic routes derived from primary metabolic pathways. AA biosynthesis is a multistep process that involves different regulatory elements and gene functions. Certain genes' expression in plant metabolism also changes with different climatic and environmental factors [226]. Furthermore, it also varies within different developmental stages of plants [243]. It remains challenging to correlate gene expression and metabolite accumulation *in planta*, as the site of metabolite synthesis may differ from the site of accumulation. For example, nicotine biosynthesis occurs in the root of tobacco but accumulates in the aerial part of the plant [287] whereas morphine biosynthesis starts in sieve elements of the phloem but accumulates in

adjacent laticifer cells in opium poppy [288]. As such, *in vitro* cultures have been essential to decipher the alkaloid biosynthesis pathway. In 2011, Tahchy *et al.* used deuterium-labeled precursors fed to *in vitro* cultures of *L. aestivum*. In this study, the authors followed the transfer of labeled precursor 4'-*O*-methyl-d_3-norbelladine from media into the shoot and its metabolization into lycorine and galanthamine. This study demonstrated that 4'-*O*-methylated-norbelladine was a key intermediate AA [289]. Until now, AA-specific genes such as NBS, *N4'OMT*, *CYP96T1*, and *NR* have been characterized and confirmed from Leucojum sp., Narcissus sp., and Lycoris sp. cultures [44-47]. However, our molecular understanding regarding this complex biosynthesis route of AA and its regulation is still unclear. Furthermore, the relative expression pattern of putative AA biosynthetic genes (in fields versus *in vitro* and in differentiated *versus* undifferentiated tissues of *Narcissus* development) added some clear knowledge regarding their role in alkaloid biosynthesis [290]. A study performed on callus culture of *L. radiata* showed how different factors, such as temperature (cold treatment), osmotic pressure (PEG treatment), or elicitor treatment (methyl jasmonate), can influence *Lr*OMT gene expression pattern [291]. Thus, in vitro system cultures are a powerful tool for uncovering AA biosynthesis and gene regulation that should be thoroughly exploited.

3.2. Synthetic biology for AA biosynthesis

Although the complete biosynthetic pathway of AA is not resolved, and up to now, the AA demand has not been sufficiently fulfilled by plant sources, a synthetic biological approach could be a powerful approach to produce AA. Recent achievements in synthetic biological approaches include the production of complex biomolecules such as noscapine (a benzylisoquinoline alkaloid from opium poppy) and its halogenated derivatives (anticancer) in *Saccharomyces cerevisiae*, assembling 30 biosynthetic enzymes from plant, bacteria, and mammal and yeast itself including seven plant endoplasmic reticulum localized genes [292]. This success raises the hope of producing complex biomolecules such as AA using a synthetic biological approach.

Proper selection of the host organism is the starting point of the synthetic biological approach. The chosen organism should be producing (or easily modified to produce) enough core metabolites, such as aromatic amino acids L-phenylalanine and L-tyrosine, precursors needed for the biosynthesis of target SMs such as AA. Selection of host species will also rely on prior knowledge of their ease of engineering, established cloning tools, culture techniques, and the

possibility of scaling up to industrial requirements. Due to rapid growth and easy handling, microbial hosts such as yeast (*Saccharomyces cerevisiae*) and, to a lesser extent, *Escherichia coli,* were used to produce plant-derived high-value alkaloids, like morphinan alkaloids [293-295]. Furthermore, the production of aromatic amino acid (precursor for AA) and associated upstream genes/enzymes were well studied in these hosts [296]. Precursors such as L-tyrosine and *p*-coumaric have already been produced in *E. coli* [297, 298]. Recently, unicellular photosynthetic organisms, such as microalgae and cyanobacteria, became engaging research platforms because of their unicellular physiology and photosynthetic, heterotrophic, and mixotrophic lifestyles. Moreover, plant-based genetic engineering techniques are also emerging in model plants such as *Nicotiana tabacum* and *N. benthamiana* [299].

Once a host organism is selected, the availability of precursor molecules can be enhanced by modifications to its metabolic pathway, such as gene deletions, swapping of endogenous enzymes with more active homologs, or overexpression of endogenous metabolic genes. Then, a route to the desired SMs can be planned and implemented. A candidate pathway is first outlined through a selection of stepwise chemical intermediates leading from host metabolism to the target compound, followed by the selection of enzymes to carry out each specified reaction [300, 301] .Even though the lack of knowledge in the AA biosynthetic pathway hinders this approach as of yet, it could be partially overcome by creating libraries of gap-filling gene candidates generated from a huge plant transcriptomic database, as available for thousands of plants or as part of the PhytoMetaSyn project [301]. In addition, the decrease in the cost of DNA synthesis helps accelerate gene characterization from its native source and ultimately facilitates the production of complex biomolecules like AA [301, 302]. Such work was done to produce polyketides. Soon, the synthetic approach platform will not only provide techniques to produce AA but also help in the biosynthesis of novel AA derivatives with improved biological and physiological properties. For example, once we have completed the identification of methyltransferases, reductases, and other enzymes required for the biosynthesis of galanthamine, we can add one more gene in the transgenic construct that will increase glycosylation, shifting the polarity of the parent molecule and eventually improve drug uptake by the human body.

4. Conclusions

Studies on the production of AA using *in vitro* systems have mainly focused on the commercially available galanthamine and the abundant lycorine. The demand for natural therapeutic metabolites obtained from plants is growing fast. However, overexploitation of the native plants to meet this demand will be insufficient and endanger the biodiversity of wild populations. Alternative chemical synthesis requires a multi-step process to produce intact complex compounds. Fortunately, biotechnological approaches, including in vitro platforms or synthetic biology, are promising strategies to establish a more reliable, economical, and environmentally friendly system to produce plant-derived metabolites. Currently, the production of AA from *in vitro* systems does not achieve the levels produced by wild-type plants. However, they have several advantages, i.e., they enable the production of target metabolites independently from environmental factors affecting the production yield, biodiversity concerns, and land usage; they facilitate the discovery of biosynthetic pathways and the understanding of its regulation in a short period. Currently established platforms of *in vitro* systems can be used to determine the effect of different variables on plants in a controlled environment with stable chemical and physical parameters. Notable effects of biotic and abiotic stresses on AA biosynthesis and accumulation in *in vitro* systems can be used as a basis platform for transcriptomic and metabolomic level studies, which generate a huge amount of data regarding the production of AA and their regulation *in planta*. This generated data can serve as fundamental units for the synthetic biological approach. It can be utilized to: 1) establish an *in vitro* production system with optimized parameters economically comparable to extraction from natural sources yet sustainable, decreasing the need for native plant harvesting; 2) be linked to other branches of science such as bioinformatics, cell biology, biochemistry, 3) to produce metabolites in a fast-growing heterologous organism such as yeast, bacteria and new emerging platforms like microalgae.

In this way, research combining biologists, biochemists, bioengineers, physicists, and computer scientists will enhance a deep understanding of AA metabolism, enabling their (re)design in selected heterologous hosts such as bacteria or yeast systems.

Author Contributions: Conceptualization, M.K., N.M., and I.D-P.; writing—original draft preparation, M.K., V.K. and N.S.L.; writing—review and editing, M.K., V.K., N.S.L., N.M. and I.D-P.; visualization, M.K., N.M., and I.D-P.; supervision, N.M. and I.D-P.; project administration, N.M. and I.D-P.; funding acquisition, N.M. and I.D-P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Natural Sciences and Engineering Research Council of Canada – Discovery Grants Program Award No RGPIN-2021-03218 to I.D-P and by the Canada Research Chair on Plant Specialized Metabolism Award No 950-232164 to Isabel Desgagné-Penix. Many thanks are extended to the Canadian taxpayers and the Canadian government for supporting the Discovery and the Canada Research Chairs Program.

Acknowledgments: The authors wish to acknowledge and dedicate this review to Professor Antonio Evidente. His achievements in natural product research, covering AAs' chemistry, biology, and medicinal potential, are inspiring.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix IV; Research Paper (Co-authorship)

Article

Chemical synthesis and biological activities of Amaryllidaceae alkaloid norbelladine derivatives and precursors

Marie-Pierre Girard [1,†], Vahid Karimzadegan [1,†], Marianne Héneault [1], Francis Cloutier [1], Gervais Bérubé [1], Lionel Berthoux [2], Natacha Merindol [1], Isabel Desgagné-Penix [1,*]

[1] Département de Chimie, biochimie et physique, Université du Québec à Trois-Rivières, 3351, Boul. des Forges, C.P. 500, Trois-Rivières, Québec, G8Z 4M3, Canada
[2] Département de Biologie médicale, Université du Québec à Trois-Rivières, 3351, Boul. des Forges, C.P. 500, Trois-Rivières, Québec, G8Z 4M3, Canada
* Correspondence: Isabel.Desgagne-Penix@uqtr.ca
† These authors contributed equally to this work.

Abstract: Amaryllidaceae alkaloids (AAs) are a structurally diverse family of alkaloids recognized for their diverse therapeutic properties, such as antiviral, anti-cholinesterase, and anticancer. Norbelladine and its derivatives, whose biological properties are poorly studied, are key intermediates required for the biosynthesis of all reported ~650 AAs. To gain insight into their therapeutic potential, we synthesized a series of *O*-methylated norbelladine-type alkaloids and evaluated their cytotoxic effect on two types of cancer cell lines, their antiviral effect against the dengue virus (DENV) and the human immunodeficiency virus -1 (HIV-1), as well as their anti-Alzheimer's disease (anti-cholinesterase and -prolyl oligopeptidase) properties. In monocytic leukemia cells, norcraugsodine was highly cytotoxic (CC_{50}=27.0 μM), while norbelladine was the most cytotoxic to hepatocarcinoma cells (CC_{50}=72.6 μM). HIV-1 infection was impaired only at cytotoxic concentrations of compounds. The DHBA (Selectivity Index (SI)=7.2), 3',4'-*O*-dimethylnorbelladine (SI=4.8), 4'-*O*-methylnorbelladine (SI>4.9), 3'-*O*-methylnorbelladine (SI>4.5) and norcraugsodine (SI=3.2), reduced DENV-infected cells with EC_{50} ranging from 24.1 to 44.9 μM. *O*-methylation of norcraugsodine abolished its anti-DENV potential. Norbelladine and its *O*-methylated forms also displayed butyrylcholinesterase inhibitory properties (IC_{50} ranging from 26.1 to 91.6 μM). Altogether, the results provide hints of the structure-activity relationship of norbelladine-type alkaloids, an important knowledge for the development of new inhibitors of DENV and butyrylcholinesterase.

Keywords: Amaryllidaceae alkaloid; norbelladine; dengue virus; anti-cholinesterase; specialized metabolism; *O*-methylation; galanthamine; Alzheimer's disease; antiviral.

1. Introduction

Amaryllidaceae are a family of monocotyledonous plants of the order Asparagales, composed of 1100 species and 75 genera, found in tropical, subtropical and warm regions of the world [303]. For centuries, Amaryllidaceae plant extracts have been recognized worldwide for their varied therapeutic properties, including anti-acetylcholinesterase (AChE), anti-microbial, and anti-tumor [304-307]. Their medicinal potency is mainly attributed to the presence of SMs of the alkaloid group, named the AAs [8, 304]. To date, more than 650 AAs have been reported and classified into norbelladine-, cherylline-, galanthamine-, lycorine-, lycorenine-, crinine-, narciclasine-, tazettine-, and montanine-type based on the proposed biosynthetic origin of the ring structure and their carbon skeleton [8, 207, 308]. All AAs are derived from norbelladine, a common metabolic intermediate formed through the condensation of tyramine and DHBA [285, 309].

AAs are known to have various pharmacological properties [310]. For example, several AAs, such as sanguinine and galanthamine (Figure A4-1), are strong anti-acetylcholinesterase inhibitors, the latter being currently used as a treatment for symptoms of Alzheimer's disease (AD) [49, 311]. Lycorine, like many other AAs, exhibits anticancer (*i.e.*, cytotoxic activity) [312-316], but also exerts an inhibitory effect against flaviviruses such as DENV (dengue virus), and against viruses belonging to other families [46, 317-319]. Recently, we demonstrated that AAs cherylline, paracrine, haemanthamine, and haemanthidine display antiviral effects against DENV, and except for cherylline, also against human immunodeficiency virus (HIV-1) [190, 320]. Developing antiviral therapeutics based on AAs may provide decisive medical solutions to catastrophic pandemics. Since many of these molecules exert their antiviral action by targeting host factors, they present opportunities to develop broad-spectrum treatments that are less susceptible to the emergence of drug resistance [321].

Few studies have been performed on the biological potential of norbelladine-type AAs. One study demonstrated that norbelladine has slight *vitro* anti-inflammatory and anti-oxidant properties [322]. In another study, synthetically designed complex alkaloid derivatives of carltonin A and B of the norbelladine-type were shown to exhibit anti-butyrylcholinesterase (BuChE) and -prolyl

oligopeptidase (POP) properties, both considered attractive targets for AD [40, 323, 324]. However, the pharmaceutical properties of norbelladine-type alkaloids concerning AD, viral infections, and cytotoxicity remain largely unknown.

In this study, we investigated the biological activities of norbelladine-type molecules. Since Amaryllidaceae plants do not accumulate a high level of such metabolites, we report on the chemical synthesis of norbelladine, norcraugsodine, and their *O*-methylated derivatives (*i.e.,* 3'-*O*-methylnorbelladine, 3'-*O*-methylnorcraugsodine, 4'-*O*-methylnorbelladine, 4'-*O*-methylnorcraugsodine, 3',4'-*O*-dimethylnorbelladine and 3',4'-*O*-dimethylnorcraugsodine) (Figure A4-1). We assessed their antiviral potential *in* cellulo and that of their precursors, DHBA and tyramine, using a propagative DENVGFP vector and a non-propagative HIV-1GFP vector. We analyzed their cytotoxicity against acute monocytic leukemia THP-1 cells; hepatocytic cellular carcinoma-derived Huh7 cells. We also measured their anti-AD potential by assessing anti-AChE and -BuChE, as well as -POP activity. We report for the first time that DHBA and O-methylated norbelladine derivatives inhibit DENV infection and that norbelladine displays anti-butyrylcholinesterase activity.

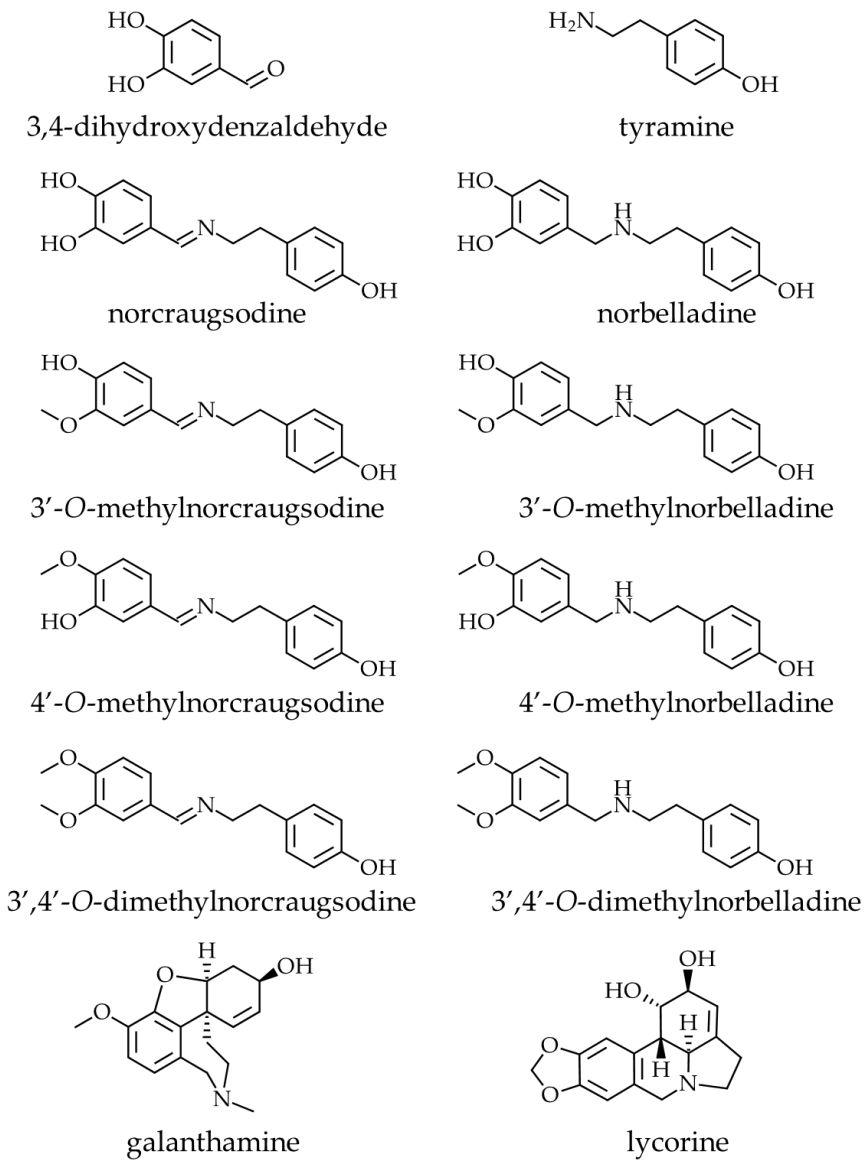


Figure A4-1. Chemical structure of molecules used in this study. Amaryllidaceae alkaloid precursors (DHBA and tyramine); intermediates (norcraugsodine and norbelladine); their corresponding *O*-methylated derivatives (3'-*O*-methylnorcraugsodine, 3'-*O*-methylnorbelladine, 4'-*O*-methylnorcraugsodine, 4'-*O*-methylnorbelladine, 3',4-'*O*-dimethylnorcraugsodine and 3',4'-*O*-dimethylnorbelladine); and well-known AA galanthamine and lycorine.

2. Materials and Methods

2.1. Chemical synthesis and purification of alkaloids

The starting material, reactant, and solvents were obtained commercially and used as such or purified and dried by standard methods [325]. The infrared spectra were recorded on a Nicolet Impact 420 FT-IR spectrophotometer. Nuclear magnetic resonance (NMR) spectra were recorded

on a Varian 200 MHz NMR apparatus. Samples were dissolved in dimethyl sulfoxide (DMSO)-d6 for data acquisition (δ 2.49 ppm for ^{1}H NMR and 39.95 ppm for ^{13}C NMR) using the residual solvent signal as internal standard (δ 0.00 ppm). Chemical shifts (δ) are expressed in parts per million (ppm), whereas the coupling constants (J) are expressed in hertz (Hz). Multiplicities are described by the following abbreviations: s for singlet, d for doublet, t for triplet, m for multiplet, and bs for broad singlet.

.

2.1.1. Two-step synthesis of Norbelladine and methylated analogs:

Norcraugsodine, norbelladine, 3'-*O*-methylnorcraugsodine, 3'-*O*-methylnorbelladine, 4'-*O*-methylnorcraugsodine, 4'-*O*-methylnorbelladine, 3',4'-*O*-dimethylnorcraugsodine and 3',4'-*O*-dimethylnorbelladine were obtained by organic synthesis following a two-step reaction sequence as described below. The products were characterized by infrared (IR) spectroscopy, proton (^{1}H NMR), and carbon nuclear magnetic resonance (^{13}C NMR) spectroscopy.

Step 1: General procedure for the preparation of the imine intermediates

An equimolar quantity of the relevant benzaldehyde and tyramine were added as powders to a flask containing dichloromethane (20 mL). The solution was stirred gently overnight (about 12 hours) at room temperature to yield the imine intermediate. The solvent was evaporated under reduced pressure using a rotatory evaporator followed by mechanical pumping to remove the residual solvent and water. The resulting imines were obtained quantitatively (100%) and were sufficiently pure to be used as such in the hydrogenation step.

Norcraugsodine. Step 1 with DHBA (537 mg, 3.88 mmol), tyramine (533 mg, 3.88 mmol), and dichloromethane (20 mL). Norcraugsodine (0.99 g, 99%). IR (cm^{-1}): 3345 (OH), 3038 (aromatic) and 1648 (C=N); ^{1}H NMR (200 MHz, DMSO-d6) δ: 7.98 (1H, s, CH imine), 7.15 (1H, d, J = 2.0 Hz, CH-Ar), 6.99 (2 H, d, J = 8.6 Hz, CH-Ar), 6.90 (1 H, dd, J_1= 2 Hz and J_2 = 8.2 Hz, CH-Ar), 6.65 (1 H, d, J = 8.6 Hz, CH-Ar), 6.63 (2 H, d, J = 8.6 Hz, CH-Ar), 3.63 (2H, t, J = 7.2 Hz, CH=NC$\underline{H_2}$CH$_2$), 2.73 (2H, t, J = 7.4 Hz, CH=NCH$_2$C$\underline{H_2}$); ^{13}C NMR (200 MHz, DMSO-d6) δ: 160.9, 155.9, 149.5, 146.1, 130.5, 130.1, 127.8, 121.9, 115.8, 115.4, 113.9, 62.4, 36.8.

3'-O-methylnorcraugsodine. Step 1 with 4-hydroxy-3-methoxybenzaldehyde (vanillin) (561 mg, 3.68 mmol), tyramine (506 mg, 3.68 mmol), dichloromethane (20 mL). 3'O-Methylnorcraugsodine (0.98 g, 98%). IR (cm^{-1}): 3008 (OH), 1639 (C=N); ^{1}H NMR (200 MHz, DMSO-d6) δ: 8.37 (1H, s, CH imine), 6.99 (4H, m, CH-Ar), 6.69 (3 H, m, CH-Ar), 3.73 (2H, m, CH=NC$\underline{H_2}CH_2$ and 3H, s, OMe), 2.79 (2H, t, J = 7.03 Hz, CH=NCH$_2$C$\underline{H_2}$); ^{13}C NMR (200 MHz, DMSO-d6) δ: 166.4, 156.1, 152.6, 148.6, 130.1, 129.7, 123.5, 118.5, 117.8, 115.5, 114.9, 59.9, 56.1, 36.3.

4'-O-methylnorcraugsodine. Step 1 with 3-hydroxy-4-methoxybenzaldehyde (isovanillin) (561 mg, 3.68 mmol), tyramine (506 mg, 3.68 mmol), dichloromethane (20 mL). 4'-O-methylnorcraugsodine (1.00 g, 100%). IR (cm^{-1}): 3508, 2900 (OH), 1638 (C=N); ^{1}H NMR (200 MHz, DMSO-d6) δ: 8.8 (1H, OH), 8.04 (1H, s, CH imine), 7.19 (1H, d, J = 1.56 Hz, CH-Ar), 6.97 (4H, m, CH=Ar), 6.65 (2H, d, J = 8.2 Hz, CH-Ar), 3.77 (3H, s, OMe), 3.64 (2H, t, J = 7.03Hz, CH=NC$\underline{H_2}CH_2$), 2.74 (2H, t, J = 7.03 Hz, CH=NCH$_2$C$\underline{H_2}$); ^{13}C NMR (200 MHz, DMSO-d6) δ: 160.8, 155.9, 150.4, 147.1, 130.5, 130.1, 129.8, 121.2, 115.4, 113.6, 111.9, 62.8, 55.9, 36.7.

3',4'-O-dimethylnorcraugsodine. Step 1 with 3,4-dimethoxybenzaldehyde (582 mg, 3.50 mmol), tyramine (481 mg, 3.50 mmol), and dichloromethane (20 mL). 3',4'-O-dimethylnorcraugsodine (1.00 g, 100%); IR (cm^{-1}): 2938 (OH), 1638 (C=N); ^{1}H NMR (200 MHz, DMSO-d6) δ: 9.05 (1H, OH), 8.13 (1H, s, CH imine), 7.32 (1H, s, CH-Ar), 7.15 (1H, d, J = 7.81 Hz, CH-Ar), 6.99 (3H, m, CH-Ar), 6.66 (2H, d, J = 8.2 Hz, 3.77 (6H, s, 2 x OMe), 3.67 (2H, t, J = 7.03 Hz, CH=NC$\underline{H_2}CH_2$), 2.77 (2H, t, J = 7.03 Hz, CH=NCH$_2$C$\underline{H_2}$); ^{13}C NMR (200 MHz, DMSO-d6) δ: 160.7, 155.9, 151.4, 149.4, 130.4, 130.1, 129.6, 122.8, 115.4, 111.6, 109.4, 62.9, 55.9, 55.8, 36.7.

Step 2: General procedure for the preparation of the final amine products

The relevant imine was dissolved in a mixture of ethylacetate/methanol (9:1, 10 mL) and hydrogenated to the amine using 30 mol% palladium on carbon (Pd/C 10 %) under a H_2 atmosphere using a balloon. The hydrogen was bubbled three times (t = 0, 30 and 60 minutes) during the hydrogenation process. The mixture was agitated for 2 to 3 hours (or until disappearance

of the starting material by TLC) and then filtered on a silica gel using ethylacetate/methanol (4:1 mixture) to remove the Pd/C. The solvent was evaporated under reduced pressure using a rotatory evaporator followed by mechanical pumping to yield the desired amine. The final amines were obtained from 43% to 98% yields.

Norbelladine. Step 2 with norcraugsodine (150 mg, 0.58 mmol), under H_2 atmosphere, 10% Pd/C (40 mg), ethyl acetate/methanol (10 mL). Norbelladine (130 mg, 86%). IR (cm^{-1}): 3021 (-OH and NH); ^{1}H NMR (200 MHz, DMSO-d6) δ: 6.94 (2 H, d, J = 8.6 Hz, CH-Ar), 6.63 (4 H, m, CH-Ar), 6.52 (1H, dd, J_1 = 1.7 Hz and J_2 = 7.7 Hz, CH-Ar), 3.49 (2H, s, Ar-C$\underline{H}_2$-NH), 2.58 (4 H, m, NH-C$\underline{H}_2$C$\underline{H}_2$-Ar); ^{13}C NMR (200 MHz, DMSO-d_6) δ: 155.8, 145.4, 144.3, 132.1, 130.9, 129.8, 119.2, 116.0, 115.6, 115.5, 53.1, 51.1, 35.4.

3'O-methylnorbelladine. Step 2 with 3'-O-methylnorcraugsodine (182.5 mg, 0.67 mmol), under H_2 atmosphere, 10% Pd/C (40 mg), ethyl acetate/methanol (10 mL). 3'O-methylnorbelladine (182.5 mg, 99%). IR (cm^{-1}): 2934 (OH and NH); ^{1}H NMR (200 MHz, DMSO-d6) δ: 6.98-6.65 (7H, m, CH-Ar), 3.81 (2H, s, Ar-C$\underline{H}_2$-NH), 3.71 (3H, s, OMe), 2.64 (4H, m, NH-C$\underline{H}_2$C$\underline{H}_2$-Ar); ^{13}C NMR (200 MHz, DMSO-d6) δ: 160.8, 152.7, 151.8, 134.8, 134.6, 129.1, 126.0, 123.2, 120.3, 116.4, 60.8, 55.2, 55.1, 39.4.

4'-O-methylnorbelladine. Step 2 with 4'-*O*-methylnorcraugsodine (266 mg), under H_2 atmosphere, 10% Pd/C (40 mg), ethyl acetate/methanol (10 mL). 4'-O-methylnorbelladine (114 mg, 43%). IR (cm^{-1}): 2989 (OH and NH); ^{1}H NMR (200 MHz, DMSO-d6) δ: 6.96-6.62 (7H, m, CH-Ar), 3.71 (3H, s, OMe), 3.55 (2H, s, Ar-C$\underline{H}_2$-NH), 2.60 (4H, br s, NH-C$\underline{H}_2$C$\underline{H}_2$-Ar); ^{13}C NMR (200 MHz, DMSO-d6) δ: 155.9, 146.8, 146.7, 133.6, 130.7, 129.8, 119.0, 115.9, 115.5, 112.4, 56.1, 52.9, 51.0, 35.3.

3',4'-O-dimethylnorbelladine. Step 2 with 3',4'-*O*-dimethylnorcraugsodine (404 mg, 1.41), under H_2 atmosphere, 10% Pd/C (70 mg), ethyl acetate/methanol (10 mL). 3',4'-O-dimethylnorbelladine (400 mg, 98%). IR (cm^{-1}): 3261 (OH), 2953 (OH and NH)); ^{1}H NMR (200 MHz, DMSO-d6) δ: 6.97-6.55 (7H, m, CH-Ar), 3.69 (6H, s, OMe), 3.60 (2H, br s, Ar-C$\underline{H}_2$-NH),

2.61 (4H, br s, NH-C$\underline{H}_2$C$\underline{H}_2$-Ar); ^{13}C NMR (200 MHz, DMSO-d6) δ: 155.9, 149.0, 147.9, 133.9, 130.9, 129.8, 120.2, 115.4, 112.1, 111.9, 55.9, 55.8, 53.0, 51.1, 35.4.

2.2 Preparation of lycorine, DHBA, tyramine and other commercial inhibitors stocks

Lycorine was isolated from *Crinum jagus* according to the method reported in [47] and provided by Antonio Evidente (Universita di Napoli Federico II, Italy). DHBA (purity: 97%, formula: $C_7H_6O_3$, molar mass: 138.12, CAS number: 139-85-5) was obtained from Acros Organics. Tyramine (purity: ≥98%, formula: $C_8H_{11}NO$, molar mass: 137.18, CAS number: 51-67-2) was obtained from Sigma Aldrich. Raltegravir (purity: 99.85%, formula: $C_{20}H_{20}FN_6O_5K$, molar mass: 482.52, CAS number: 871038-72-1) was obtained through the NIH HIV Reagent Program, Division of AIDS, NIAID (Isentress; MK-0518). Each compound was dissolved in dimethylsulfoxide (DMSO) at a final concentration of 100 mM and stored at -20ºC until subsequent use. Rivastigmine and galanthamine hydrobromide were purchased from Millipore Sigma (Sigma-Aldrich, Canada) and solubilized in water at 20 mM before experiments.

2.3. Cell lines and culture

The human hepatocarcinoma Huh7 cell line was kindly provided by Hugo Soudeyns (University of Montréal, QC, Canada). Huh7, Crandell-Rees Feline Kidney Cell (CRFK), and Vero cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin solution (all from Wisent, Inc., Canada). The human leukemia monocytic THP-1 cell line was maintained in Roswell Park Memorial Institute (RPMI) medium supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin solution (all from Wisent, Inc., Canada). All cell lines were incubated at 37ºC and 5% CO_2.

2.4. Cytotoxicity assay

Cytotoxicity assays of norbelladine and its derivatives were performed on Huh7 and THP-1 cells by measuring ATP levels using the Cell-Titer GLO assay kit (Promega, Madison, WI, USA). Briefly, 7.5x10^3 Huh7 cells/well or 2x10^4 THP-1 cells/well were plated in black 96-well plates and incubated at 37C° for 24 hours. The next day, precursors, norbelladine, its derivatives, DMSO, and lycorine were serially diluted by a factor of 2 in DMEM complete medium (for Huh7) or RPMI complete medium (for THP-1) at room temperature. Each dilution was added to the cell plates to

obtain final concentrations of 6.25 µM to 200 µM for alkaloids and DMSO and 0.39 µM to 50 µM for lycorine. DMSO was used as a negative control, and lycorine as a positive control since its cytotoxic effect was previously demonstrated [326]. The plates were incubated at 37ºC and 5% CO_2 for 72 hours. Cell-Titer GLO reagent was added to the plates previously equilibrated to room temperature. The plates were then mixed on an orbital shaker for 2 minutes and incubated for 10 minutes at room temperature. The luminescence signal was measured with a microplate spectrophotometer (Synergy H1, Biotek, QC, Canada). Viability percentages were obtained by calculating the ratio of the signal corresponding to each alkaloid concentration to the signal of the equimolar DMSO control. All experiments were performed at least twice. Median cytotoxic concentrations (CC_{50}) were calculated using QuestGraph IC50 calculator software (MLA Quest Graph™ IC_{50} Calculator. AAT Bioquest, Inc).

2.5. Viral vectors

To investigate the antiviral effect of norbelladine and its derivatives, we used a dengue virus propagative vector ($DENV_{GFP}$) and a non-propagative human immunodeficiency virus (HIV)-1 pseudotyped VSV-G vector (HIV-1_{GFP}), both encoding green fluorescent protein (GFP). The plasmid used to obtain the $DENV_{GFP}$ vector (pFK-DVs-G2A) was provided by Ralf Bartenschlager (Heidelberg University, Germany) and Laurent Chatel-Chaix (Institut national de la recherche scientifique, Québec, Canada) [47, 327]. The two plasmids used to obtain the HIV-1_{GFP} vector were PMD2.G and pNL4-3-GFPΔEnvΔNef [328]. For $DENV_{GFP}$, viral titer was measured by plaque assay in Vero cells, as described by [329]. For HIV-1_{GFP}, the viral titer was obtained by measuring the infectivity of serially diluted vectors in CRFK cells, as described by [330].

2.6. Antiviral assays

Briefly, $7.5x10^3$ Huh7 cells/well or $2x10^4$ THP-1 cells/well were plated in 96-well plates at 37C° for 24 hours. The next day, norbelladine and its derivatives, as well as DMSO-dissolved lycorine (Huh7) or DMSO-dissolved raltegravir (THP-1), were serially diluted by a factor of 2 in DMEM or RPMI medium, respectively. Each dilution was added to the cell plates to obtain final concentrations of 1.56 µM to 200 µM for alkaloids and matched concentrations of DMSO and 0.05 µM to 6.4 µM for lycorine or 0.078 µM to 10 µM for raltegravir. Lycorine and raltegravir were used as DENV and HIV-1 inhibitor controls, respectively [317, 331]. $DENV_{GFP}$ and HIV-

1_{GFP} were then added at a multiplicity of infection (MOI) of 0.025 and 0.1, respectively. Plates were incubated at 37°C and 5% CO_2 for 72 hours. Afterward, the cells were fixed in 3.7% formaldehyde, and the percentage of infection was measured by flow cytometry with an FC500 MPL cytometer (Beckman Coulter, Inc., Brea, CA, USA). Data analysis was performed using the Flowjo software (BD, FlowJo LLC, Ashland, OR, USA). All experiments were performed at least twice. EC_{50} was calculated using QuestGraph IC50 calculator software (MLA Quest Graph™ IC50 Calculator. AAT Bioquest, Inc).

2.7. Anti-acetylcholinesterase (AChE) and -butyrylcholinesterase (BuChE) activity

Pharmacological properties specific to AD were tested first on AChEs according to the kit (ab138871, Abcam) that provides a colorimetric measure of enzyme activity and inhibition. Briefly, the reaction was performed in a final volume of 100 µL in 96-well microplates. A preliminary screen to identify the most potent AChEs inhibitors was performed. DMSO-dissolved test compounds were added to a final concentration of 1 mM (1% DMSO) in duplicates. Next, a five µL reaction mixture containing equal amounts of acetylthiocholine (20x) and DTNB (20x) was added to each well. Ultimately, the enzyme solution was added to a final concentration of 0.25 U/mL, and absorbance was measured at 412 nm in kinetic mode for 10 minutes using a microplate reader (Synergy H1, Biotek, QC, Canada). The same procedure was utilized for the BuChEs (equine, Sigma-Aldrich) activity test, except that the enzyme concentration was 2 U/mL per reaction. DTNB (Bis(3-carboxy-4-nitrophenyl) disulfide, Ellman's Reagent), acetylthiocholine iodide, and butyrylthiocholine iodide were purchased from Sigma-Aldrich (Canada). Galanthamine (10 µM) and rivastigamine (2 mM) were used as a positive control for AChE and BuChE assays, respectively. Molecules showing inhibition during preliminary screenings were selected for further assessment of IC_{50} using serially diluted concentrations. Experiments were performed at least twice. Inhibition was calculated as follows:

$$I = 100x(1-\Delta i/\Delta e) \quad [40]$$

where Δi is the difference of absorbance between two time points in the presence of an inhibitor, and Δe is the difference of absorbance using two time points in the presence of DMSO or appropriate solvent.

2.8. Prolyloligopeptidase (POP) inhibition assay

POP enzyme activity was measured using the Fluorogenic POP Assay Kit (BPS Bioscience Inc, USA). The reaction was carried out in a final volume of 50 µL in a low-binding NUNC microtiter plate. All compounds were tested at a final concentration of 250 µM in duplicate. DPP substrate was added to a final concentration of 2.5 µM, and the POP enzyme solution was subsequently added to a final concentration of 1 ng/µL. Optical densities at 440 nm were recorded using a microplate reader (Synergy H1, Biotek, QC, Canada).

2.9. Docking of norbelladine derivatives with BuChE

Docking was performed using the crystal structure of human BuChE in a complex with tacrine (PDB: 4BDS) using MOE 2020.09 software (Chemical Computing Group). Tacrine was removed. Structure issues were corrected using the structure preparation tool and amino acids were protonated using the protonate3D tool. Ligands were protonated at pH=7 using the protomer option. Ligands were protonated at pH=7 using the protomers tool. The active site (Asn68, Ile69, Asp70, Gln71, Ser72, Gly78, Ser79, Trp82, Tyr114, Gly115, Gly116, Gly117, Gln119, Thr120, Gly121, Thr122, Leu125, Tyr128, Glu197, Ser198, Ala199, Trp231, Glu276, Ala277, Val280, Gly283, Thr284, Pro285, Leu286, Ser287, Val288, Asn289, Phe290, Ala328, Phe329, Tyr332, Phe398, Trp430, Met437, His438, Gly439, Tyr440, Ile442) was predicted by the site finder tool of the MOE software and validated with literature data. Dummy atoms across the active site were created and used as docking sites. Water and solvent molecules were removed, residues further than 8 Å from dummy atoms were fixed, and active site residues were tethered using the QuickPrep default parameters. The triangle matcher method was used to place ligands in the active site using the London dG score, with 200 poses, and the induced fit was used as a refinement option with 10 poses and GBVI/WSA score. The first pose of the most abundant configuration was chosen to be conserved for ligand-protein interaction analysis. This corresponded to the 1st pose for all ligands except 3'4'-*O*-methylnorbelladine, for which the 1st pose displayed a flipped structure; hence the 2nd pose was selected. Protein-Ligand Interaction Profiler (PLIP) was used to analyze the interaction of the ligands with the binding site following the docking procedure [332]. Pymol (Schrödinger) was used to visualize and present PLIP results.

2.10. Statistical analysis

All the analysis and related graphs were performed using GraphPad Prism version 8.0.0 (GraphPad Software, San Diego, California USA).

3. Results

3.0. Chemical synthesis

Synthesis of norbelladine as well as the different methylated analogs, was performed following a two-step reaction sequence reported in the literature for norbelladine [322, 333]. The condensation of relevant aldehyde with tyramine initially yielded the imide intermediates quantitatively. Next, a simple catalytic hydrogenation allowed us to generate the final derivatives with yields ranging from 43% to 99%.

3.1. Cytotoxic assay

Several alkaloids of the Amaryllidaceae family have been reported to be cytotoxic [334]. Therefore, we evaluated the cytotoxic activity of the AA precursors DHBA, tyramine, and the eight norbelladine-derived molecules on two types of cancer cell lines, including human monocytic leukemia cells (THP-1) and human hepatocarcinoma cells (Huh7). Tyramine is a natural by-product of the breakdown of the amino acid tyrosine and is found in plants and animals. As expected for both cell lines, tyramine was not cytotoxic and did not affect cell viability, whereas lycorine was cytotoxic [335] at concentrations below 50 µM (Figure A4-2, Table A4-1). For THP-1 cells, norcraugsodine and DHBA at 100 µM and norbelladine and 3'-O-methylnorbelladine at 200 µM were highly cytotoxic, killing the majority of the cells (Figure A4-2a). For Huh7 cells, norbelladine, norcraugsodine, 3',4'-*O*-dimethylnorbelladine, and DHBA at 200 µM were cytotoxic, representing 50% cell death (Figure A4-2b).

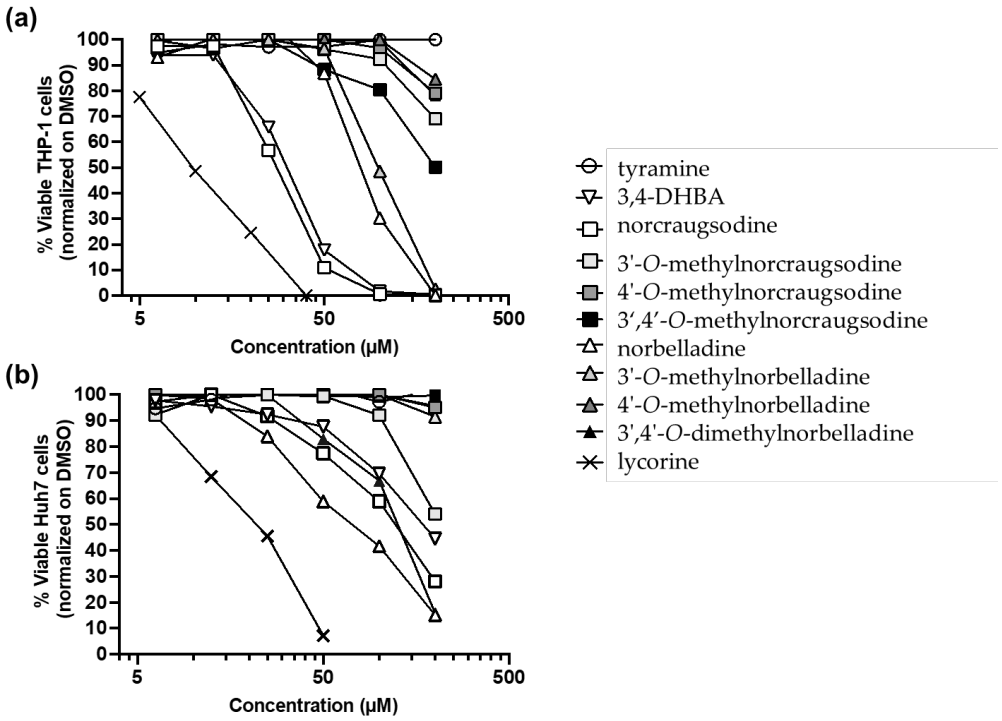


Figure A4-2. Cytotoxic effect of norbelladine precursors and its derivatives on Huh7 and THP-1 cells. To assess cell viability, the cellular ATP levels were measured on **(a)** THP-1 and **(b)** Huh7 cells 72 h after alkaloid addition at concentrations of 6.25 µM to 200 µM. Lycorine was utilized as a positive control at concentrations of 0.3 µM to 40 µM. Results were normalized to equivalent concentrations of DMSO, and the x-axis is displayed in $\log_{10}$. DMSO: dimethyl sulfoxide; ATP: adenosine triphosphate.

The median cytotoxic concentrations (CC_{50}) causing a 50% decrease in cell viability were estimated for all alkaloids reaching this value (Figure A4-2; Table A4-1). Norcraugsodine, DHBA, norbelladine, and 3'-*O*-methylnorbelladine were highly cytotoxic on THP-1 cells, with CC_{50} ranging from 27.0 µM to 99.0 µM (Table 1). The 3'- and 4'-*O*-methylnorcraugsodine, 3'-*O*-methylnorbelladine, 3',4'-*O*-dimethylnorcraugsodine, and 4'-*O*-methylnorbelladine exhibited only moderate cytotoxicity (<50% of cell death) at 200 µM on THP-1 cells (Figure A4-2a; Table A4-1). The norbelladine, norcraugsodine, 3',4'-*O*-dimethylnorbelladine, and DHBA were cytotoxic to Huh7 cells (Figure A4-2b), with CC_{50} ranging from 72.6 µM to 173.1 µM (Table A4-1). The 3'-*O*-methylnorcraugsodine was weakly cytotoxic on Huh7 cells at tested concentrations (Figure A4-2b). All other tested molecules were not cytotoxic (Figure A4-2b; Table A4-1).

In summary, DHBA, norcraugsodine, and norbelladine were the most cytotoxic to both types of cell lines, whereas 3'-*O*-methylnorbelladine displayed cytotoxic specificity to THP-1, and 3',4'-*O*-dimethylnorbelladine to Huh7 cells (Figure A4-2; Table A4-1).

Table A4-1. EC_{50}, CC_{50}, and SI of norbelladine precursors and derivatives with antiviral effect. EC_{50}: median effector concentration; CC_{50}: median cytotoxic concentration, SI= Selectivity index.

Alkaloids	EC_{50} HIV-1 (μM)	CC_{50} THP-1 (μM)	SI HIV-1	EC_{50} DENV (μM)	CC_{50} Huh7 (μM)	SI DENV
Tyramine	>200-	>200-	<1.0	>200-	>200-	<1.0
DHBA	51.5*	31.65	0.6	24.1*	173.1	7.2
Norbelladine	50.5	82.2	1.6	50.4	72.6	1.4
Norcraugsodine	55.5	27.0	0.5	37.7	121.8	3.2
3'-*O*-methylnorcraugsodine	107.2*	>200	>1.9	176.3*	>200	>1.1
3'-*O*-methylnorbelladine	134.7*	99.01	0.73	44.9*	>200	>4.5
4'-*O*-methylnorbelladine	108.3*	>200	>1.8	40.5*	>200	>4.9
4'-*O*-methylnorcraugsodine	>200-	>200-	<1.0	>200-	>200-	<1.0
3',4'-*O*-dimethylnorbelladine	98.3*	>200	>2.0	27.5	131.4	4.8
3',4'-*O*-dimethylnorcraugsodine	>200-	>200-	<1.0	>200-	>200-	<1.0
Raltegravir	0.098	>10-	>102	u.i.	u.i.	u.i.
Lycorine	u.i.	10,7	u.i.	0,090	17,4	193,3

EC_{50} and CC_{50} of antiviral compounds (norbelladine, norcraugsodine, 3'-*O*-methylnorcraugsodine, 3'-*O*-methylnorbelladine, 4'-*O*-methylnorbelladine, 3',4'-*O*-dimethylnorbelladine, and DHBA) were calculated using QuestGraph IC50 calculator (MLA Quest Graph™ IC50 Calculator. AAT Bioquest, Inc). SI = CC_{50}/EC_{50}. "-": EC_{50} or CC_{50} was not achieved. "*": compound addition does not yield complete viral inhibition. "u.i." is unidentified.

3.2. Antiviral assay

Several studies have shed light on the outstanding antiviral properties of alkaloids extracted from Amaryllidaceae, such as lycorine, cherylline, haemanthamine, haemanthidine, and pancracine [17, 21, 22]. Hence, we measured the antiviral activity of DHBA, tyramine, and the eight norbelladine-derived molecules towards HIV-1_{GFP} and $DENV_{GFP}$ in THP-1 and Huh7 cells, respectively. Infection levels were measured 72 hours post-infection, where a dose-dependent inhibition of HIV-1_{GFP} (Figure A4-3; Supplementary File A5-1) and $DENV_{GFP}$ (Figure A4-4; Supplementary File A4-2) was generated. In addition, the effective concentration inhibiting infection by 50% (EC_{50}) was calculated along with the selectivity index (SI) that was determined by the ratio of CC_{50} and EC_{50} (Table A4-1).

3.2.1. Inhibition of HIV-1$_{GFP}$

Norcraugsodine, DHBA, and norbelladine prevented HIV-1 infection in most cells at 100 µM, while their methylated forms and tyramine were mostly inactive. At 200 µM, all three compounds and 3'-*O*-methylnorcraugsodine impeded HIV-1 infection (from 94% to 99% inhibition) (Figure A4-3; Supplementary File A5-1), with EC_{50} ranging from 50.5 µM to 107.2 µM, and selectivity indices (SI) of 0.5 to 1.6 (Table 1). Hence, at these concentrations, AAs were also associated with significant toxicity, raising concerns about the specificity of their antiretroviral properties.

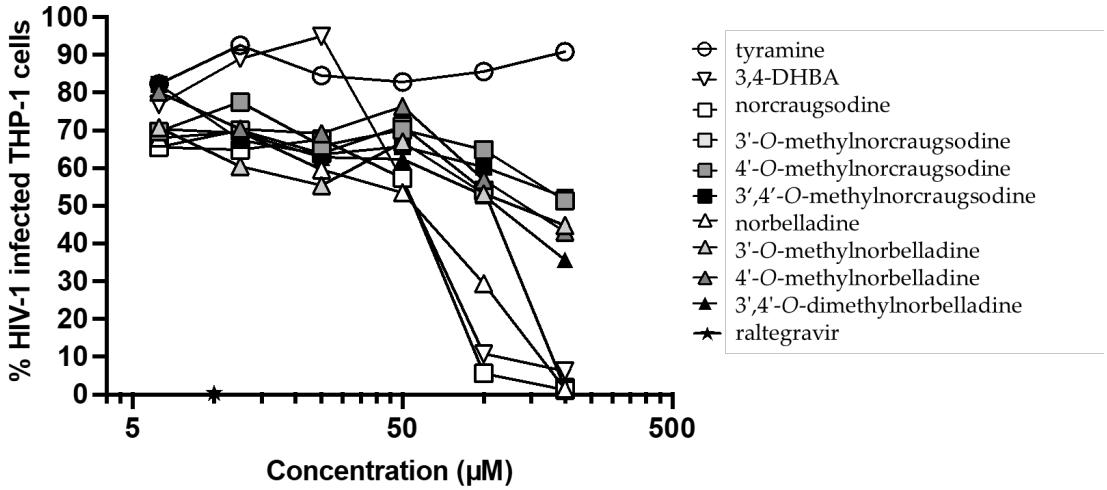


Figure A4-3. Antiretroviral effect of norbelladine precursors and derivatives on HIV-1$_{GFP}$. The antiviral activity against HIV-1$_{GFP}$ of norbelladine precursors and derivatives was evaluated 72 hours post-infection using THP-1 cells by flow cytometry at concentrations ranging from 6.25 µM to 200 µM. Infections were performed with the non-propagative HIV-1$_{GFP}$ virus at a multiplicity of infection (MOI) of 0.1. Raltegravir served as a positive control and DMSO as a negative control at concentrations equivalent to tested alkaloids. Results were normalized to the value with HIV-1$_{GFP}$ infection without treatment, and the x-axis is displayed as $\log_{10}$.

3.2.2. Inhibition of DENV$_{GFP}$

DHBA, 3',4'-*O*-dimethylnorbelladine, norcraugsodine, 4'-*O*-methylnorbelladine 3'-*O*-methylnorbelladine, and norbelladine show potent inhibition of DENV$_{GFP}$ infection at 50, 100 and 200 µM. Huh7 cells treated with 200 µM of these compounds resulted in 89% to 100% decrease of infection, with EC_{50} ranging from 24.1 µM to 50.4 µM, and SI from 1.5 to 6.2 (Table A4-1,

Figure A4-4; Supplementary File A5-2). DHBA, 4'-*O*-methylnorbelladine, 3',4'-*O*-dimethylnorbelladine and 3'-*O*-methylnorbelladine are the most selective, with SI > 4.5. The other alkaloids tested showed little or no specific antiviral effect against $DENV_{GFP}$ infections.

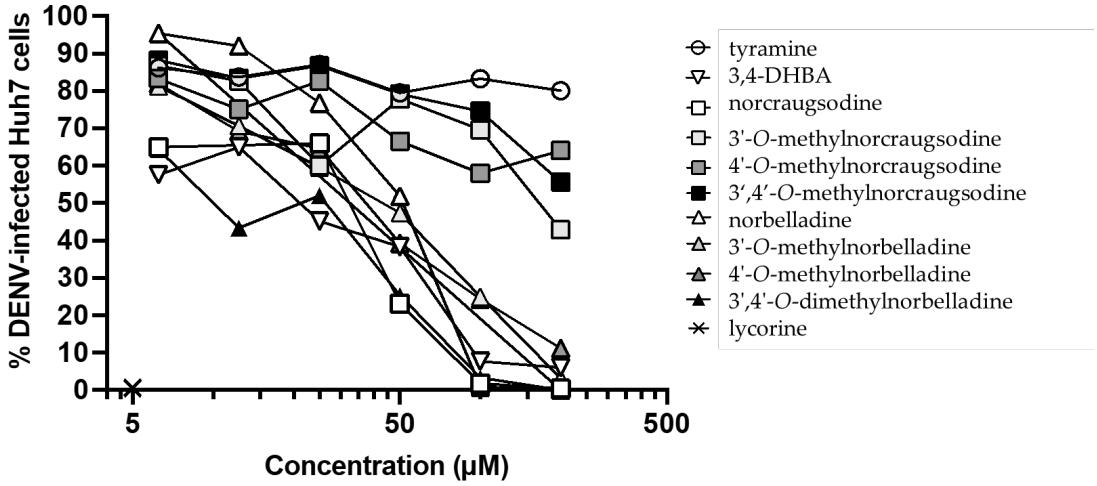


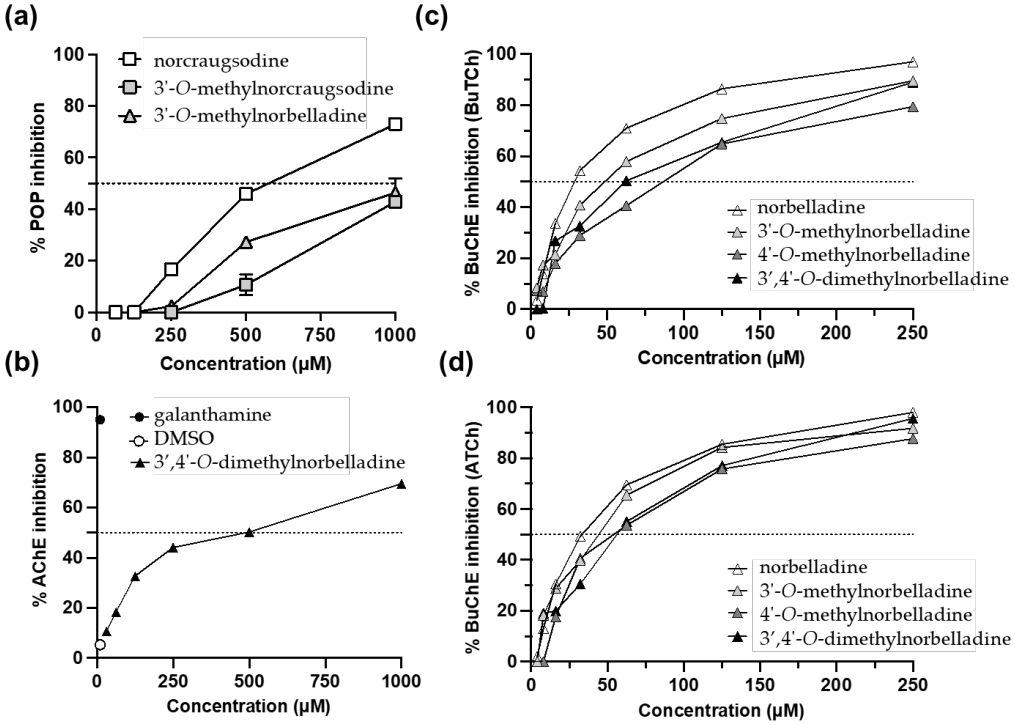
Fig A4-4. Antiflaviviral effect of norbelladine precursors and derivatives on $DENV_{GFP}$. The antiviral activity against $DENV_{GFP}$ of norbelladine precursors and derivatives was evaluated 72 hours post-infection using Huh7 cells using flow cytometry at concentrations ranging from 6.25 µM to 200 µM. Infections were performed with propagative $DENV_{GFP}$ virus at an MOI of 0.025. Lycorine was used as a positive control, and DMSO was used as a negative control at concentrations equivalent to tested alkaloids. Results were normalized to the value of $DENV_{GFP}$ infection without treatment, and the x-axis is displayed as $\log_{10}$.

3.3. Choline Esterase and Prolyloligopeptidase inhibitory effect

Enzymatic inhibitions were first trialed using one mM of compounds in duplicates (Table A5-1). Further experiments were carried out only on selected active molecules. Of all the compounds tested, only norcraugsodine, 3'-O-methylnorcraugsodine, and 3'-O-methylnorbelladine inhibited POP activity, norcraugsodine being the most potent with IC_{50} = 463.8 µM (Figure A4-5a). Among all AAs and precursors, only 3',4'-O-dimethylnorbelladine exhibited moderate inhibition of AChE, with IC_{50} = 319.6 µM (Figure A4-5b). Among the 10 molecules tested, only norbelladine and its methylated forms significantly blocked BuChE activity using both acetylthiocholine and butyrylthiocholine as substrates (Figure A4-5c and 5d). Norbelladine was the most potent with an

IC_{50} of 33.26 μM and 26.13 μM, respectively (Table A5-1). The addition of two methyl groups in 3'- and 4'-O positions leads to a ~3-fold decrease in inhibition.

Figure A4-5. Anti-Alzheimer's disease properties. (a) Prolyl oligopeptidase inhibition by norcraugsodine, 3'O-methylnorcraugsodine, and 3'O-methylnorbelladine. **(b)** Acetylcholinesterase inhibition of 3'4'O-methylnorbelladine



using acetylthiocholine as a substrate. **(c)** Butyrylcholinesterase (BuChE) inhibition of norbelladine and derivatives using butyrylthiocholine (BuTCh) as substrate. **(d)** Butyrylcholinesterase inhibition of norbelladine and derivatives using acetylthiocholine (ATCh) as substrate. Galanthamine (10 μM) was used as a positive control for AchE assays, while rivastigmine (2 mM) was used for BuChE assays (100% of inhibition not shown on the graph).

3.4. Molecular docking of norbelladine derivatives with BuChE.

To better understand the interactions between norbelladine derivatives and BuChE, we performed docking using the crystal structure of human BuChE (PDB: 4BDS) (Table 2, Figure A4-6). The active site of BuChE is located at the bottom of a profound gorge (20 Å) comprising 6 conserved aromatic residues and 6 aliphatic (Leu286 and Val288) and polar residues [324]. It includes a catalytic triad (Ser198, Glu325, and His438) mediating the choline esters hydrolysis, an anionic site (Trp82, Tyr128, Phe329), essential for the reaction, an oxyanion hole (Gly116, Gly117, Ala199) stabilizing the transition state through hydrogen-bond interactions, while the acyl pocket (Ala199, Leu286, and Val288) is responsible for substrate specificity [336], and the peripheral anionic site (PAS or P-site) (Asp70, Tyr332) at the entry of the active gorge, interacting with

cationic substrate guiding them down the gorge to the catalytic triad, is implicated in interactions with beta-amyloid and in the binding with many inhibitors [337, 338](Figure A4-6). Docking scores were very similar for norbelladine and its derivatives, ranging from -6.68 to -7.03 kCal/mol (Table 2). All the molecules interacted with the key residue of the anionic site Trp82. The hydrophobic and H-bonds interaction with Trp82 were stabilized by aromatic π-π-stacking in the case of norbelladine and 3'-O-methylnorbelladine, consistently with other inhibitors such as tacrine [339], while the docking conformation of the 4'-O-methylated compounds did not allow for such interaction. In the case of norbelladine and 3'-O-methylnorbelladine, the interaction with BuChE was also supported with H-bonds with PAS amino acids (Tyr332 and Asp70, respectively), and with other binding site residues (Tyr440 for norbelladine, Thr120 and Trp430 for 3'-O-methylnorbelladine). In addition, 3'-O-methylnorbelladine was H-bonded to the catalytic residue His438. Thus, the docking results are consistent with inhibition mechanisms like other previously reported inhibitor molecules such as tacrine and possibly reflect a stronger and more stable inhibition potential for norbelladine and 3'-O-methylnorbelladine compared to 4'O-methylated norbelladine derivatives (Figures A4-5 and A4-6).

Table A4-2. Prediction of norbelladine derivatives interactions with butyrylcholinesterase.

	Score (kCal/mol)	Interaction		
		H-bonds	hydrophobic	π -stack
norbelladine	-6,6829	Trp82, Tyr332, Tyr440	Trp82, Tyr440	Trp82
3'-*O*-methylnorbelladine	-6,8174	Asp70, Trp82, Thr120, Trp430, His438	Trp82	Trp82
4'-*O*-methylnorbelladine	-6,8119	Trp82	Trp82	nd
3',4'-*O*-dimethylnorbelladine	-7,0368	Thr120	Trp82, Tyr440	nd

Nd: None detected.

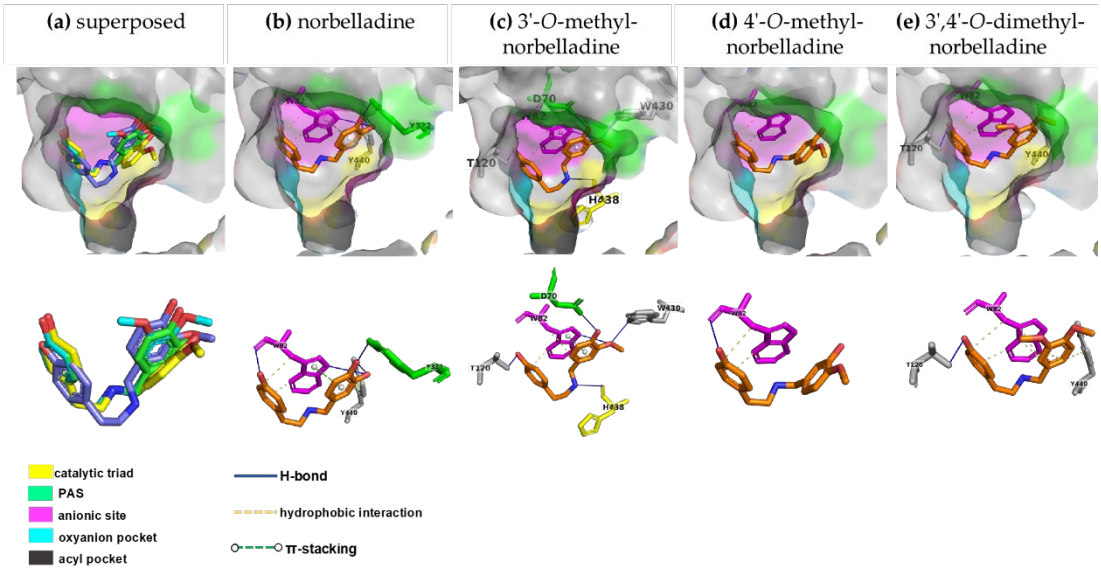


Figure A4-6. Prediction of norbelladine derivatives interaction with butyrylcholinesterase. Grey surface representation of BuChE (4BDS) active site with key subsites highlighted in different colors (catalytic triad is a yellow, pre-anionic site (PAS) is green, the anionic site is pink, oxyanion pocket is turquoise, and acyl pocket is dark grey. **(a)** Superimposition of docked ligands in the BuChE active site, norbelladine is green, 3'-*O*-methylnorbelladine purple, 4'-*O*-methylnorbelladine yellow, 3',4'-*O*-dimethylnorbelladine turquoise. **(b)** Non-covalent (H-bond, hydrophobic, and P-stack) interactions of norbelladine with Trp82 from the anionic site, Tyr332 from the PAS, and Tyr440 from the binding site. **(c)** Non-covalent interactions of 3'-*O*-methylnorbelladine with W82 from the anionic site, Asp70 from the PAS, His438 from the catalytic triad, and Trp430 from the binding site. **(d)** Non-covalent interactions of 4'-*O*-methylnorbelladine with Trp82 from the anionic site. **(e)** Non-covalent interactions of 3'4'-*O*-dimethylnorbelladine with Trp82 from the anionic site and Thr120 and Tyr440 from the binding site.

4. Discussion

Norbelladine and its *O*-methylated forms are mandatory intermediates in the biosynthesis of AAs [309, 340], yet their biological properties remain poorly studied. Norbelladine itself was shown to possess anti-inflammatory and cyclooxygenase inhibitory effects [322], while 3'-*O*-methylnorbelladine and a few complex synthetic derivatives of norbelladine and belladine were shown to display anti-cholinesterase activity [341].

We obtained the precursor imines norcraugsodine, 3'-*O*-methylnorcraugsodine, 4'-*O*-methylnorcraugsodine, 3',4'-*O*-dimethylnorcraugsodine from which we made norbelladine [333], 3'-*O*-methylnorbelladine, 4'-*O*-methylnorbelladine, and 3',4'-*O*-dimethylnorbelladine by catalytic hydrogenation. Their chemical synthesis was a straightforward process, and the products were generally good to excellent.

Cytotoxic assays revealed that DHBA, norcraugsodine, and norbelladine were the most cytotoxic compounds both in monocytic leukemia THP-1 cells and hepatocarcinoma Huh7 cells (Figure A4-2; Table A4-1). Interestingly, the 3'-*O*-methylnorbelladine was selectively toxic to THP-1 cells, while 3',4'-*O*-dimethylnorbelladine was selectively toxic to Huh7, hinting towards different mechanisms of cytotoxicity between the two cell lines.

Norbelladine, norcraugsodine, and DHBA also strongly impeded HIV-1_{GFP} and $DENV_{GFP}$ infections (Table A4-1, Figure A4-3, Figure A4-4). The concentrations required to inhibit HIV-1 infection were cytotoxic to THP-1 cells (Table A4-1, Figure A4-3). Hence, the decrease in HIV-1 infection caused by these alkaloids is more consistent with a progressive depletion of cell viability rather than a specific antiviral effect. In the case of DENV, viral inhibition was more specific, occurring at non-cytotoxic doses. Some non-cytotoxic alkaloids, such as 3'- and 4'-*O*-methylnorbelladine, also efficiently inhibited $DENV_{GFP}$ replication at >100 μM. The 3'- *O*-methylnorcraugsodine, 4'-*O*-methylnorcraugsodine, and 3',4'-*O*-dimethylnorcraugsodine show little or no antiviral activity in contrast to norcraugsodine, revealing the importance of *O*-methylation to the toxic and antiviral nature of alkaloids (Table A4-1, Figure A4-4).

Several studies have demonstrated the antiviral effect of AAs against several types of viruses. Here, the virus used to perform $DENV_{GFP}$ infections corresponds to dengue serotype 2. Recently, we also uncovered that AAs haemanthamine, pancracine, and haemanthidine, isolated from *Pancratium maritimum*, inhibited DENV-2 infection [320]. This study adds to the growing evidence that AAs structure could be optimized to develop potent inhibitors against this potentially fatal disease. Future studies should address the potency of these compounds towards other serotypes (DENV-1, 3, and 4), and other flaviviruses [342, 343].

Investigation of the inhibitory properties of alkaloids on POP, AChE, and BuChE activity confirmed that AAs' O-methylation and reduction state weighs heavily on their potency. POP is a post-proline cleaving enzyme of the central nervous system, whose alteration is implicated in memory loss, Alzheimer's, Parkinson's, and Huntington's diseases, as well as other neurodegenerative diseases [344]. Norcraugsodine, 3'-*O*-methylnorcraugsodine, and 3'-*O*-methylnorbelladine were the only compounds to inhibit POP at high concentrations >500 μM (Figure A4-5). Mamun *et al.* also reported limited efficiency of synthetic norbelladine and belladine derivatives to inhibit this enzyme [341].

AChE activity is dominant in regulating acetylcholine levels in healthy brains, while in Alzheimer's disease (AD) patients' brains, the activity of BuChE is increased [345]. Thus, both enzymes are considered major therapeutic targets in the fight against AD. Except 3',4'-*O*-dimethylnorbelladine, none of the molecules inhibited AChE activity in our experiments. Using DMSO as a solvent could have masked their activity up to a certain level. However, norbelladine and *O*-methylated derivatives consistently inhibited BuChE-catalyzed hydrolysis of butyrylcholine and acetylcholine. Norbelladine was the most potent, while there was a 3-fold loss of inhibition in the case of 3',4'-*O*-dimethylnorbelladine. These results, in line with Mamun *et al* screening [341], emphasizes the interest in norbelladine as the backbone for developing butyrylcholinesterase inhibitors.

BuChE is a serine hydrolase enzyme that can hydrolyze several choline esters, including acetylcholine, succinylcholine, and butyrylcholine [324]. Although its role is not fully understood, its levels are increased in AD patients, and it could promote amyloid plaque formation [346]. The active site of this enzyme consists of 1) the catalytic site composed of Ser198, His438, and Glu325 [347], 2) an acyl pocket with Ala199, Leu286 and Val288 interacting with acyl group of the esters [348], 3) an anionic site with Trp82 that binds to quaternary nitrogen of choline, and 4) an enzyme gorge lip consisting of Asp70 and Tyr332 guiding substrate toward the catalytic site [349].

In a study conducted by Nachon *et al. (2013),* human BuChE was crystallized in complex with tacrine, a strong inhibitor of BuChE. According to this study, an aromatic π- π stacking between tacrine and Trp82 is essential for its inhibitory effect [350]. A similar interaction was predicted by docking norbelladine and its *O*-methylated forms with BuChE, consistent with the relatively strong inhibitory effect observed in our *in vitro* enzymatic assays (Figure A4-6). Furthermore, an additional interaction between Tyr332 and the hydroxyl group of the norbelladine ring was obtained in our model that potentially can reduce the accessibility of BuChE to the substrate since Tyr332 has been proposed to direct the substrate to the active site of the enzyme. It should be noted that a weak interaction between Tyr332 and tacrine was also observed [350].

5. Conclusions

In summary, precursors and norbelladine derivatives do not exhibit antiviral effects against HIV-1_{GFP} infections. However, they do possess appreciable inhibitory activity against $DENV_{GFP}$ infections and butyrylcholinesterase. The results obtained in this study increase our knowledge of

the structure-activity relationship of alkaloids with a norbelladine backbone. They provide further insight into the biological potency of norbelladine-type molecules, depending on the nature and location of the different *O*-methyl groups. This new knowledge could help to better guide the selection and optimization of AAs to develop new inhibitors to fight Alzheimer's disease and infections caused by flaviviruses.

Author Contributions: Conceptualization, N.M. and I.D-P.; methodology, M-P.G., V.K., M.H., F.C., G.B., L.B., N.M., and I.D-P.; software, M-P.G., and N.M.; validation, M-P.G., V.K., M.H., F.C., G.B., L.B., N.M., and I.D-P.; formal analysis, M-P.G., V.K., and N.M.; investigation, M-P.G., V.K., M.H., F.C., and N.M.; resources, L.B., G.B. and I.D-P.; writing—original draft preparation, M-P.G., V.K., M.H., N.M. and I.D-P.; writing—review and editing, L.B., G.B., N.M. and I.D-P.; supervision, L.B., G.B., N.M. and I.D-P.; project administration, N.M.; funding acquisition, I.D-P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Natural Sciences and Engineering Research Council of Canada – Discovery Grants Program Award No RGPIN-2021-03218 to I.D-P and by the Canada Research Chair on Plant Specialized Metabolism Award No 950-232164 to I.D-P. Many thanks are extended to the Canadian taxpayers and to the Canadian government for supporting the Discovery and the Canada Research Chairs Program.

Acknowledgements: The authors would like to thank all the lab members for their technical support and useful advice. They also warmly thank Professor Hugo Germain and his team for kindly providing advice, materials, and equipment.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix V; Supplementary Data of A IV

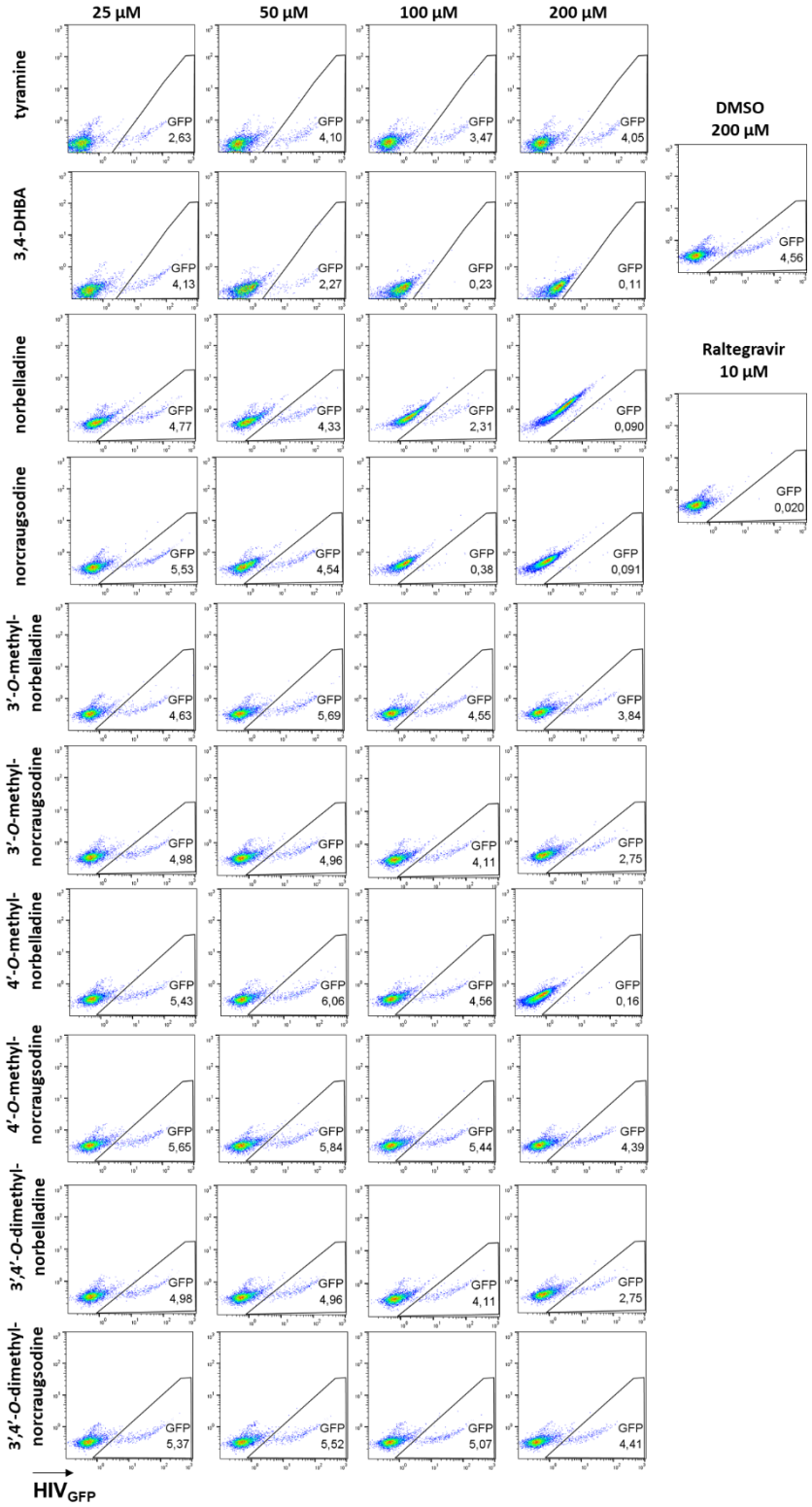


Figure A5-1. Antiretroviral effect of four concentrations (25, 50, 100, and 200 µM) of norbelladine precursors and derivatives on HIV-1$_{GFP}$ in THP1 cells. Representative dot plots are presented. DMSO (200 µM) and raltegravir (10 µM) are negative and positive controls, respectively.

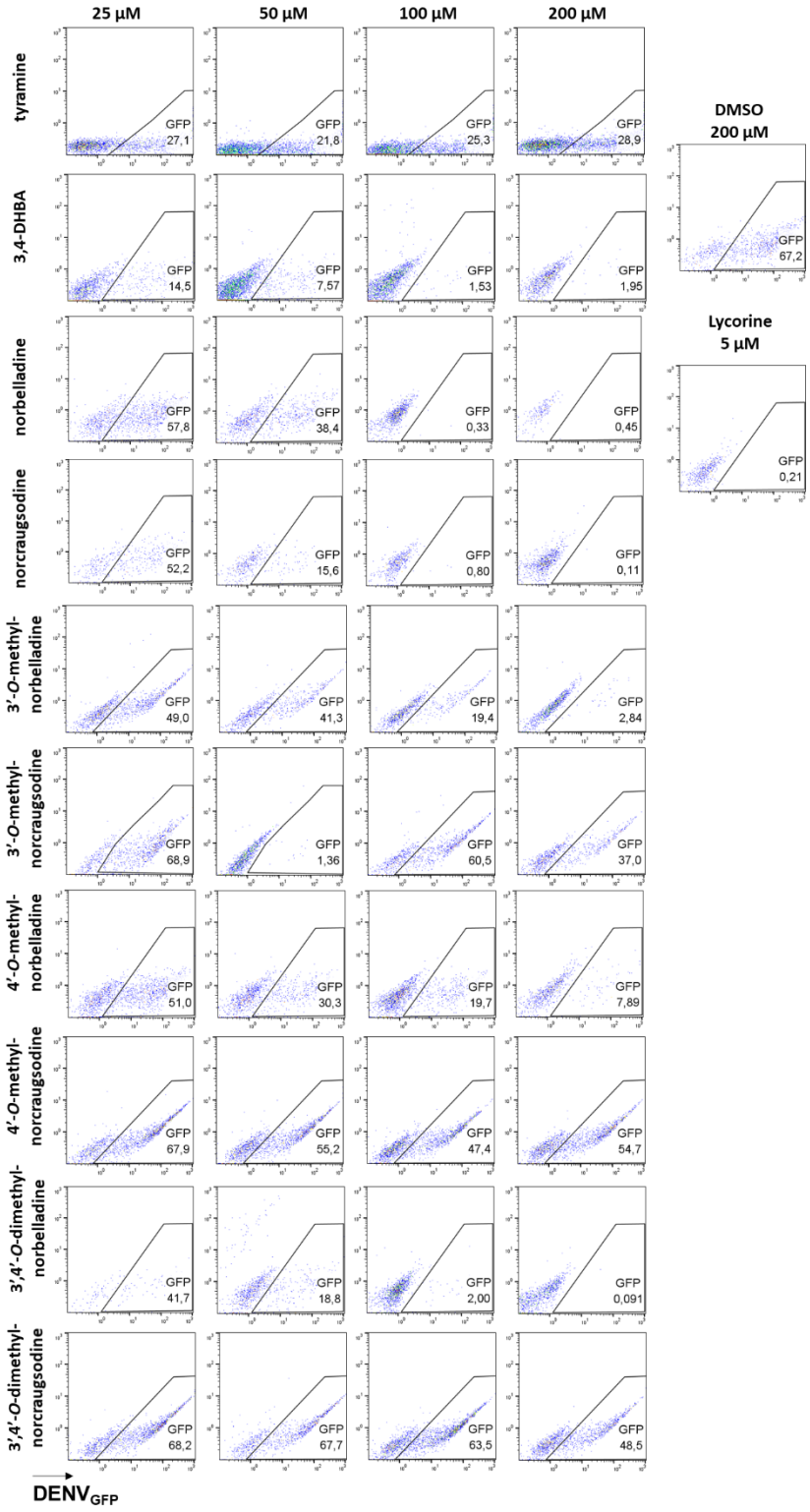


Figure A5-2. Antiflaviviral effect of norbelladine precursors and derivatives on $DENV_{GFP}$. The antiviral activity against $DENV_{GFP}$ was evaluated 72 hours post-infection using Huh7 cells by flow cytometry at four concentrations (25, 50, 100, and 200 µM) of norbelladine precursors and derivatives. Lycorine (5 µM) was included as a positive control, and DMSO (200 µM) as a negative control. Representative dot plots are presented.

Table A5-1. Inhibitory properties (IC_{50}) of norbelladine precursors and derivatives towards enzymes implicated in Alzheimer's disease.

	POP DPP	**AChE** ATCh	**BuChE** BuTCh	ATCh
DHBA	nid	nid	nid	nid
tyramine	nd	nid	nid	nid
norcraugsodine	567.30	nid	nid	nid
norbelladine	nid	nid	26.13	33.26
3′-*O*-methylnorcraugsodine	1146.01	nid	nid	nid
4′-O-methylnorcraugsodine	nid	nid	nid	nid
3′,4′-*O*-dimethylnorcraugsodine	nid	nid	nid	nid
3′-*O*-methylnorbelladine	1230.60	nid	58.95	40.22
4′-*O*-methylnorbelladine	nid	nid	91.61	42.92
3′,4′-*O*-dimethylnorbelladine	nid	319.6	87.86	87.80

IC_{50} is expressed in µM, nid stands for no inhibition detected at 1 mM. POP: prolyl oligopeptidase, DPP: Ala-Pro-AMC dipeptide, AChE: acetylcholinesterase, ATCh: acetylthiocholine, BuTCh: butyrylthiocholine, BuChE: butyrylcholinesterase.

References

[1] S.C. Bhatla, M.A. Lal, Secondary metabolites, in: Plant physiology, development and metabolism, Springer, 2023, pp. 765-808.
[2] A. Ortiz, E. Sansinenea, Phenylpropanoid derivatives and their role in plants' health and as antimicrobials, Current Microbiology, 80 (2023) 380.
[3] B.A. Pandian, R. Sathishraj, M. Djanaguiraman, P.V. Prasad, M. Jugulam, Role of cytochrome P450 enzymes in plant stress response, Antioxidants, 9 (2020) 454.
[4] A.V. Alber, H. Renault, A. Basilio-Lopes, J.E. Bassard, Z. Liu, P. Ullmann, A. Lesot, F. Bihel, M. Schmitt, D. Werck-Reichhart, Evolution of coumaroyl conjugate 3-hydroxylases in land plants: lignin biosynthesis and defense, The Plant Journal, 99 (2019) 924-936.
[5] S. Baek, J.C. Utomo, J.Y. Lee, K. Dalal, Y.J. Yoon, D.-K. Ro, The yeast platform engineered for synthetic gRNA-landing pads enables multiple gene integrations by a single gRNA/Cas9 system, Metabolic Engineering, 64 (2021) 111-121.
[6] J.C. Utomo, H.B. Barrell, R. Kumar, J. Smith, M.S. Brant, H. De la Hoz Siegler, D.-K. Ro, Reconstructing curcumin biosynthesis in yeast reveals the implication of caffeoyl-shikimate esterase in phenylpropanoid metabolic flux, Metabolic Engineering, (2024).
[7] X. Huang, R. Li, J. Fu, N. Dudareva, A peroxisomal heterodimeric enzyme is involved in benzaldehyde synthesis in plants. Nat Commun. 2022: 13 (1): 1352, in.
[8] I. Desgagné-Penix, Biosynthesis of alkaloids in Amaryllidaceae plants: A review, Phytochemistry Reviews, 20 (2021) 409-431.
[9] L.-M. Huang, H. Huang, Y.-C. Chuang, W.-H. Chen, C.-N. Wang, H.-H. Chen, Evolution of terpene synthases in Orchidaceae, International journal of molecular sciences, 22 (2021) 6947.
[10] J. Barros, L. Escamilla-Trevino, L. Song, X. Rao, J.C. Serrani-Yarce, M.D. Palacios, N. Engle, F.K. Choudhury, T.J. Tschaplinski, B.J. Venables, 4-Coumarate 3-hydroxylase in the lignin biosynthesis pathway is a cytosolic ascorbate peroxidase, Nature Communications, 10 (2019) 1994.
[11] Y. Zhao, G. Liu, F. Yang, Y. Liang, Q. Gao, C. Xiang, X. Li, R. Yang, G. Zhang, H. Jiang, Multilayered regulation of secondary metabolism in medicinal plants, Molecular Horticulture, 3 (2023) 11.
[12] A. Samanta, G. Das, S.K. Das, Roles of flavonoids in plants, Carbon, 100 (2011) 12-35.
[13] M.L. Falcone Ferreyra, S.P. Rius, P. Casati, Flavonoids: biosynthesis, biological functions, and biotechnological applications, Frontiers in plant science, 3 (2012) 222.
[14] H. Haraguchi, N. Yoshida, H. Ishikawa, Y. Tamura, K. Mizutani, T. Kinoshita, Protection of mitochondrial functions against oxidative stresses by isoflavans from Glycyrrhiza glabra, Journal of pharmacy and pharmacology, 52 (2000) 219-223.
[15] G. Agati, S. Biricolti, L. Guidi, F. Ferrini, A. Fini, M. Tattini, The biosynthesis of flavonoids is enhanced similarly by UV radiation and root zone salinity in L. vulgare leaves, Journal of plant physiology, 168 (2011) 204-212.
[16] L.M. Trujillo-Chacón, E.R. Pastene-Navarrete, L. Bustamante, M. Baeza, J.E. Alarcón-Enos, C.L. Cespedes-Acuña, In vitro micropropagation and alkaloids analysis by GC–MS of Chilean Amaryllidaceae plants: Rhodophiala pratensis, Phytochemical analysis, 31 (2020) 46-56.
[17] C. Koutouan, V.L. Clerc, R. Baltenweck, P. Claudel, D. Halter, P. Hugueney, L. Hamama, A. Suel, S. Huet, M.-H.B. Merlet, Link between carrot leaf secondary metabolites and resistance to Alternaria dauci, Scientific reports, 8 (2018) 13746.
[18] A. Ziouti, C. El Modafar, A. Fleuriet, S. El Boustani, J. Macheix, Phenolic compounds in date palm cultivars sensitive and resistant to Fusarium oxysporum, Biologia Plantarum, 38 (1996) 451-457.

[19] Z. Song, W. Chen, X. Du, H. Zhang, L. Lin, H. Xu, Chemical constituents of Picea neoveitchii, Phytochemistry, 72 (2011) 490-494.
[20] K.A. Leiss, F. Maltese, Y.H. Choi, R. Verpoorte, P.G. Klinkhamer, Identification of chlorogenic acid as a resistance factor for thrips in chrysanthemum, Plant Physiology, 150 (2009) 1567-1575.
[21] C. Garagounis, N. Delkis, K.K. Papadopoulou, Unraveling the roles of plant specialized metabolites: using synthetic biology to design molecular biosensors, New Phytologist, 231 (2021) 1338-1352.
[22] D.J. Newman, G.M. Cragg, Natural products as sources of new drugs from 1981 to 2014, Journal of natural products, 79 (2016) 629-661.
[23] H.N. Matsuura, S. Malik, F. de Costa, M. Yousefzadi, M.H. Mirjalili, R. Arroo, A.S. Bhambra, M. Strnad, M. Bonfill, A.G. Fett-Neto, Specialized plant metabolism characteristics and impact on target molecule biotechnological production, Molecular biotechnology, 60 (2018) 169-183.
[24] E. Ono, J. Murata, Exploring the evolvability of plant specialized metabolism: uniqueness out of uniformity and uniqueness behind uniformity, Plant and Cell Physiology, 64 (2023) 1449-1465.
[25] S. de Vries, J.M. Fürst-Jansen, I. Irisarri, A. Dhabalia Ashok, T. Ischebeck, K. Feussner, I.N. Abreu, M. Petersen, I. Feussner, J. de Vries, The evolution of the phenylpropanoid pathway entailed pronounced radiations and divergences of enzyme families, The Plant Journal, 107 (2021) 975-1002.
[26] Y. Kawai, E. Ono, M. Mizutani, Evolution and diversity of the 2–oxoglutarate-dependent dioxygenase superfamily in plants, The Plant Journal, 78 (2014) 328-343.
[27] A.E. Wilson, L. Tian, Phylogenomic analysis of UDP-dependent glycosyltransferases provides insights into the evolutionary landscape of glycosylation in plant metabolism, The Plant Journal, 100 (2019) 1273-1288.
[28] C.C. Hansen, D.R. Nelson, B.L. Møller, D. Werck-Reichhart, Plant cytochrome P450 plasticity and evolution, Molecular Plant, 14 (2021) 1244-1265.
[29] R. Akiyama, B. Watanabe, M. Nakayasu, H.J. Lee, J. Kato, N. Umemoto, T. Muranaka, K. Saito, Y. Sugimoto, M. Mizutani, The biosynthetic pathway of potato solanidanes diverged from that of spirosolanes due to evolution of a dioxygenase, Nature communications, 12 (2021) 1300.
[30] F.M. Afendi, T. Okada, M. Yamazaki, A. Hirai-Morita, Y. Nakamura, K. Nakamura, S. Ikeda, H. Takahashi, M. Altaf-Ul-Amin, L.K. Darusman, KNApSAcK family databases: integrated metabolite–plant species databases for multifaceted plant research, Plant and Cell Physiology, 53 (2012) e1-e1.
[31] E. Ono, M. Nakai, Y. Fukui, N. Tomimori, M. Fukuchi-Mizutani, M. Saito, H. Satake, T. Tanaka, M. Katsuta, T. Umezawa, Formation of two methylenedioxy bridges by a Sesamum CYP81Q protein yielding a furofuran lignan,(+)-sesamin, Proceedings of the National Academy of Sciences, 103 (2006) 10116-10121.
[32] D. Nelson, D. Werck-Reichhart, A P450-centric view of plant evolution, The Plant Journal, 66 (2011) 194-211.
[33] Y. Lin, Industrial applications of plant secondary metabolites, in, The Ohio State University, 2017.
[34] J. Dai, R.J. Mumper, Plant phenolics: extraction, analysis and their antioxidant and anticancer properties, Molecules, 15 (2010) 7313-7352.
[35] D. Tholl, Terpene synthases and the regulation, diversity and biological roles of terpene metabolism, Current opinion in plant biology, 9 (2006) 297-304.
[36] E. Pichersky, J.P. Noel, N. Dudareva, Biosynthesis of plant volatiles: nature's diversity and ingenuity, Science, 311 (2006) 808-811.
[37] M. Ashour, M. Wink, J. Gershenzon, Biochemistry of terpenoids: monoterpenes, sesquiterpenes and diterpenes, Annual plant reviews volume 40: biochemistry of plant secondary metabolism, (2010) 258-303.
[38] R. Kaur, S. Arora, Alkaloids-important therapeutic secondary metabolites of plant origin, Journal of Critical Reviews, 2 (2015) 1-8.
[39] H.-G. Boit, Ergebnisse der Alkaloid-Chemie bis 1960: unter besonderer Berücksichtigung der Fortschritte seit 1950, (No Title), (1961).

[40] A.A. Mamun, F. Pidaný, D. Hulcová, J. Maříková, T. Kučera, M. Schmidt, M.C. Catapano, M. Hrabinová, D. Jun, L. Múčková, Amaryllidaceae alkaloids of norbelladine-type as inspiration for development of highly selective butyrylcholinesterase inhibitors: Synthesis, biological activity evaluation, and docking studies, International Journal of Molecular Sciences, 22 (2021) 8308.
[41] S. Bhambhani, K.R. Kondhare, A.P. Giri, Diversity in chemical structures and biological properties of plant alkaloids, Molecules, 26 (2021) 3374.
[42] J. Ziegler, P.J. Facchini, Alkaloid biosynthesis: metabolism and trafficking, Annual Review of Plant Biology, 59 (2008) 735-769.
[43] B. Casciaro, L. Mangiardi, F. Cappiello, I. Romeo, M.R. Loffredo, A. Iazzetti, A. Calcaterra, A. Goggiamani, F. Ghirga, M.L. Mangoni, Naturally-occurring alkaloids of plant origin as potential antimicrobials against antibiotic-resistant infections, Molecules, 25 (2020) 3619.
[44] S.E. O'Connor, J.J. Maresh, Chemistry and biology of monoterpene indole alkaloid biosynthesis, Natural Product Reports, 23 (2006) 532-547.
[45] N.S. Liyanage, F. Awwad, K.C. Gonçalves dos Santos, T.U. Jayawardena, N. Mérindol, I. Desgagné-Penix, Navigating Amaryllidaceae Alkaloids: Bridging Gaps and Charting Biosynthetic Territories–A Comprehensive Review, Journal of Experimental Botany, (2024) erae187.
[46] P. Wang, L.F. Li, Q.Y. Wang, L.Q. Shang, P.Y. Shi, Z. Yin, Anti-dengue-virus activity and structure–activity relationship studies of lycorine derivatives, ChemMedChem, 9 (2014) 1522-1533.
[47] S. Ka, N. Merindol, A.A. Sow, A. Singh, K. Landelouci, M.B. Plourde, G. Pépin, M. Masi, R. Di Lecce, A. Evidente, Amaryllidaceae alkaloid cherylline inhibits the replication of dengue and Zika viruses, Antimicrobial Agents and Chemotherapy, 65 (2021) 10.1128/aac. 00398-00321.
[48] J. Olin, L. Schneider, Galantamine for Alzheimer's disease, The Cochrane database of systematic reviews, (2002) CD001747-CD001747.
[49] M. Heinrich, H.L. Teoh, Galanthamine from snowdrop—the development of a modern drug against Alzheimer's disease from local Caucasian knowledge, Journal of ethnopharmacology, 92 (2004) 147-162.
[50] T.U. Jayawardena, N. Merindol, N.S. Liyanage, I. Desgagné-Penix, Unveiling Amaryllidaceae alkaloids: from biosynthesis to antiviral potential–a review, Natural Product Reports, (2024).
[51] B.B. Majhi, S.-E. Gélinas, N. Mérindol, S. Ricard, I. Desgagné-Penix, Characterization of norbelladine synthase and noroxomaritidine/norcraugsodine reductase reveals a novel catalytic route for the biosynthesis of Amaryllidaceae alkaloids including the Alzheimer's drug galanthamine, Frontiers in Plant Science, 14 (2023) 1231809.
[52] L. Tousignant, A.M. Diaz-Garza, B.B. Majhi, S.E. Gelinas, A. Singh, I. Desgagne-Penix, Transcriptome analysis of *Leucojum aestivum* and identification of genes involved in norbelladine biosynthesis, Planta, 255 (2022) 30.
[53] M.B. Kilgore, M.M. Augustin, C.M. Starks, M. O'Neil-Johnson, G.D. May, J.A. Crow, T.M. Kutchan, Cloning and characterization of a norbelladine 4'-O-methyltransferase involved in the biosynthesis of the Alzheimer's drug galanthamine in Narcissus sp. aff. pseudonarcissus, PloS one, 9 (2014) e103223.
[54] N. Mehta, Y. Meng, R. Zare, R. Kamenetsky-Goldstein, E. Sattely, A developmental gradient reveals biosynthetic pathways to eukaryotic toxins in monocot geophytes, bioRxiv, (2023).
[55] M.B. Kilgore, M.M. Augustin, G.D. May, J.A. Crow, T.M. Kutchan, CYP96T1 of Narcissus sp. aff. Pseudonarcissus catalyzes formation of the Para-Para'CC phenol couple in the Amaryllidaceae alkaloids, Frontiers in plant science, 7 (2016) 225.
[56] P.J. Facchini, V. De Luca, Differential and tissue-specific expression of a gene family for tyrosine/dopa decarboxylase in opium poppy, Journal of Biological Chemistry, 269 (1994) 26684-26690.
[57] P.S. Grewal, J.A. Samson, J.J. Baker, B. Choi, J.E. Dueber, Peroxisome compartmentalization of a toxic enzyme improves alkaloid production, Nature Chemical Biology, 17 (2021) 96-103.
[58] T. Lehmann, S. Pollmann, Gene expression and characterization of a stress-induced tyrosine decarboxylase from Arabidopsis thaliana, FEBS letters, 583 (2009) 1895-1900.

[59] R. Wang, X. Han, S. Xu, B. Xia, Y. Jiang, Y. Xue, R. Wang, Cloning and characterization of a tyrosine decarboxylase involved in the biosynthesis of galanthamine in Lycoris aurea. PeerJ 7: e6729, in, 2019.
[60] J. Hu, W. Li, Z. Liu, G. Zhang, Y. Luo, Molecular cloning and functional characterization of tyrosine decarboxylases from galanthamine-producing Lycoris radiata, Acta Physiologiae Plantarum, 43 (2021) 84.
[61] X. Zhang, C.-J. Liu, Multifaceted regulations of gateway enzyme phenylalanine ammonia-lyase in the biosynthesis of phenylpropanoids, Molecular plant, 8 (2015) 17-27.
[62] H. Zandavar, M.A. Babazad, Secondary metabolites: Alkaloids and flavonoids in medicinal plants, in: Herbs and Spices-New Advances, IntechOpen, 2023.
[63] R.S. Riseh, F. Fathi, A. Lagzian, M. Vatankhah, J.F. Kennedy, Modifying lignin: A promising strategy for plant disease control, International Journal of Biological Macromolecules, 271 (2024) 132696.
[64] T. Al-Warhi, A. Sabt, E.B. Elkaeed, W.M. Eldehna, Recent advancements of coumarin-based anticancer agents: An up-to-date review, Bioorganic Chemistry, 103 (2020) 104163.
[65] Y. Deng, S. Lu, Biosynthesis and regulation of phenylpropanoids in plants, Critical reviews in plant sciences, 36 (2017) 257-290.
[66] P.J. Rudall, Colourful cones: how did flower colour first evolve?, Journal of Experimental Botany, 71 (2020) 759-767.
[67] M.-Z. Shi, D.-Y. Xie, Biosynthesis and metabolic engineering of anthocyanins in Arabidopsis thaliana, Recent patents on biotechnology, 8 (2014) 47-60.
[68] S. Grunewald, S. Marillonnet, G. Hause, I. Haferkamp, H.E. Neuhaus, A. Veß, T. Hollemann, T. Vogt, The tapetal major facilitator NPF2. 8 is required for accumulation of flavonol glycosides on the pollen surface in Arabidopsis thaliana, The Plant Cell, 32 (2020) 1727-1748.
[69] N.Q. Dong, H.X. Lin, Contribution of phenylpropanoid metabolism to plant development and plant–environment interactions, Journal of integrative plant biology, 63 (2021) 180-209.
[70] M. Dastmalchi, S. Dhaubhadel, Proteomic insights into synthesis of isoflavonoids in soybean seeds, Proteomics, 15 (2015) 1646-1657.
[71] A. Hapeshi, J.M. Benarroch, D.J. Clarke, N.R. Waterfield, Iso-propyl stilbene: a life cycle signal?, Microbiology, 165 (2019) 516-526.
[72] D.-D. Huang, G. Shi, Y. Jiang, C. Yao, C. Zhu, A review on the potential of Resveratrol in prevention and therapy of diabetes and diabetic complications, Biomedicine & Pharmacotherapy, 125 (2020) 109767.
[73] M.H. Pan, J.C. Wu, C.T. Ho, C.S. Lai, Antiobesity molecular mechanisms of action: Resveratrol and pterostilbene, BioFactors, 44 (2018) 50-60.
[74] C.W. Jin, G.Y. You, Y.F. He, C. Tang, P. Wu, S.J. Zheng, Iron deficiency-induced secretion of phenolics facilitates the reutilization of root apoplastic iron in red clover, Plant Physiology, 144 (2007) 278-285.
[75] M.J. Stassen, S.-H. Hsu, C.M. Pieterse, I.A. Stringlis, Coumarin communication along the microbiome–root–shoot axis, Trends in Plant Science, 26 (2021) 169-183.
[76] L. Yang, W. Ding, Y. Xu, D. Wu, S. Li, J. Chen, B. Guo, New insights into the antibacterial activity of hydroxycoumarins against Ralstonia solanacearum, Molecules, 21 (2016) 468.
[77] L. Labeeuw, P.T. Martone, Y. Boucher, R.J. Case, Ancient origin of the biosynthesis of lignin precursors, Biology Direct, 10 (2015) 1-21.
[78] A. Añón, J.F. López, D. Hernando, I. Orriols, E. Revilla, M.M. Losada, Effect of five enological practices and of the general phenolic composition on fermentation-related aroma compounds in Mencia young red wines, Food chemistry, 148 (2014) 268-275.
[79] J.R. Widhalm, N. Dudareva, A familiar ring to it: biosynthesis of plant benzoic acids, Molecular plant, 8 (2015) 83-97.
[80] N. Kumar, N. Goel, Phenolic acids: Natural versatile molecules with promising therapeutic applications, Biotechnology reports, 24 (2019) e00370.

[81] J.-S. Xue, B. Zhang, H. Zhan, Y.-L. Lv, X.-L. Jia, T. Wang, N.-Y. Yang, Y.-X. Lou, Z.-B. Zhang, W.-J. Hu, Phenylpropanoid derivatives are essential components of sporopollenin in vascular plants, Molecular Plant, 13 (2020) 1644-1653.
[82] T. Vogt, Phenylpropanoid biosynthesis, Molecular plant, 3 (2010) 2-20.
[83] M.W. Hyun, Y.H. Yun, J.Y. Kim, S.H. Kim, Fungal and plant phenylalanine ammonia-lyase, Mycobiology, 39 (2011) 257-265.
[84] L. Poppe, J. Rétey, Friedel–crafts-type mechanism for the enzymatic elimination of ammonia from histidine and phenylalanine, Angewandte Chemie International Edition, 44 (2005) 3668-3688.
[85] L. Xiang, B.S. Moore, Biochemical characterization of a prokaryotic phenylalanine ammonia lyase, Journal of Bacteriology, 187 (2005) 4286-4289.
[86] R.R. Fritz, D. Hodgins, C. Abell, Phenylalanine ammonia-lyase. Induction and purification from yeast and clearance in mammals, Journal of Biological Chemistry, 251 (1976) 4646-4650.
[87] Y.-Y. Kao, S.A. Harding, C.-J. Tsai, Differential expression of two distinct phenylalanine ammonia-lyase genes in condensed tannin-accumulating and lignifying cells of quaking aspen, Plant physiology, 130 (2002) 796-807.
[88] A. Rohde, K. Morreel, J. Ralph, G. Goeminne, V. Hostyn, R. De Rycke, S. Kushnir, J. Van Doorsselaere, J.-P. Joseleau, M. Vuylsteke, Molecular phenotyping of the pal1 and pal2 mutants of Arabidopsis thaliana reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism, The Plant Cell, 16 (2004) 2749-2771.
[89] J.W. Blount, K.L. Korth, S.A. Masoud, S. Rasmussen, C. Lamb, R.A. Dixon, Altering expression of cinnamic acid 4-hydroxylase in transgenic plants provides evidence for a feedback loop at the entry point into the phenylpropanoid pathway, Plant Physiology, 122 (2000) 107-116.
[90] G.A. Bubna, R.B. Lima, D.Y.L. Zanardo, W.D. Dos Santos, M.d.L.L. Ferrarese, O. Ferrarese-Filho, Exogenous caffeic acid inhibits the growth and enhances the lignification of the roots of soybean (Glycine max), Journal of Plant Physiology, 168 (2011) 1627-1633.
[91] R. Yin, B. Messner, T. Faus-Kessler, T. Hoffmann, W. Schwab, M.-R. Hajirezaei, V. von Saint Paul, W. Heller, A.R. Schäffner, Feedback inhibition of the general phenylpropanoid and flavonol biosynthetic pathways upon a compromised flavonol-3-O-glycosylation, Journal of experimental botany, 63 (2012) 2465-2478.
[92] C. Appert, E. Logemann, K. Hahlbrock, J. Schmid, N. Amrhein, Structural and catalytic properties of the four phenylalanine ammonia-lyase isoenzymes from parsley (Petroselinum crispum Nym.), European Journal of Biochemistry, 225 (1994) 491-499.
[93] M. Mizutani, D. Ohta, Diversification of P450 genes during land plant evolution, Annual review of plant biology, 61 (2010) 291-315.
[94] K. Goiris, K. Muylaert, S. Voorspoels, B. Noten, D. De Paepe, G.J. E Baart, L. De Cooman, Detection of flavonoids in microalgae from different evolutionary lineages, Journal of phycology, 50 (2014) 483-492.
[95] J. Barros, R.A. Dixon, Plant phenylalanine/tyrosine ammonia-lyases, Trends in plant science, 25 (2020) 66-79.
[96] D. Werck-Reichhart, R. Feyereisen, Cytochromes P450: a success story, Genome biology, 1 (2000) 1-9.
[97] L. Achnine, E.B. Blancaflor, S. Rasmussen, R.A. Dixon, Colocalization of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis, The Plant Cell, 16 (2004) 3098-3109.
[98] M. Dastmalchi, M.A. Bernards, S. Dhaubhadel, Twin anchors of the soybean isoflavonoid metabolon: evidence for tethering of the complex to the endoplasmic reticulum by IFS and C4H, The Plant Journal, 85 (2016) 689-706.
[99] S. Lu, Y. Zhou, L. Li, V.L. Chiang, Distinct roles of cinnamate 4-hydroxylase genes in Populus, Plant and Cell Physiology, 47 (2006) 905-914.

[100] D. Cukovica, J. Ehlting, J.A.V. Ziffle, C.J. Douglas, Structure and evolution of 4-coumarate: coenzyme A ligase (4CL) gene families, (2001).
[101] B. Hamberger, K. Hahlbrock, The 4-coumarate: CoA ligase gene family in Arabidopsis thaliana comprises one rare, sinapate-activating and three commonly occurring isoenzymes, Proceedings of the National Academy of Sciences, 101 (2004) 2209-2214.
[102] P. Khatri, L. Chen, I. Rajcan, S. Dhaubhadel, Functional characterization of Cinnamate 4-hydroxylase gene family in soybean (*Glycine max*), PLoS One, 18 (2023) e0285698.
[103] S.G. Lavhale, R.M. Kalunke, A.P. Giri, Structural, functional and evolutionary diversity of 4-coumarate-CoA ligase in plants, Planta, 248 (2018) 1063-1078.
[104] L. Hoffmann, S. Maury, F. Martz, P. Geoffroy, M. Legrand, Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism, Journal of biological chemistry, 278 (2003) 95-103.
[105] A.M. Anterola, N.G. Lewis, Trends in lignin modification: a comprehensive analysis of the effects of genetic manipulations/mutations on lignification and vascular integrity, Phytochemistry, 61 (2002) 221-294.
[106] M. Matsuno, V. Compagnon, G.A. Schoch, M. Schmitt, D. Debayle, J.-E. Bassard, B. Pollet, A. Hehn, D. Heintz, P. Ullmann, Evolution of a novel phenolic pathway for pollen development, Science, 325 (2009) 1688-1692.
[107] M. Petersen, Y. Abdullah, J. Benner, D. Eberle, K. Gehlen, S. Hücherig, V. Janiak, K.H. Kim, M. Sander, C. Weitzel, Evolution of rosmarinic acid biosynthesis, Phytochemistry, 70 (2009) 1663-1679.
[108] V. Mahesh, R. Million-Rousseau, P. Ullmann, N. Chabrillange, J. Bustamante, L. Mondolot, M. Morant, M. Noirot, S. Hamon, A. de Kochko, Functional characterization of two p-coumaroyl ester 3′-hydroxylase genes from coffee tree: evidence of a candidate for chlorogenic acid biosynthesis, Plant Molecular Biology, 64 (2007) 145-159.
[109] A.V. Alber, H. Renault, A. Basilio-Lopes, J.E. Bassard, Z. Liu, P. Ullmann, A. Lesot, F. Bihel, M. Schmitt, D. Werck-Reichhart, J. Ehlting, Evolution of coumaroyl conjugate 3-hydroxylases in land plants: lignin biosynthesis and defense, Plant J, 99 (2019) 924-936.
[110] R. Vanholme, I. Cesarino, K. Rataj, Y. Xiao, L. Sundin, G. Goeminne, H. Kim, J. Cross, K. Morreel, P. Araujo, Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in Arabidopsis, Science, 341 (2013) 1103-1106.
[111] N. Kitaoka, T. Nomura, S. Ogita, Y. Kato, Bioproduction of glucose conjugates of 4-hydroxybenzoic and vanillic acids using bamboo cells transformed to express bacterial 4-hydroxycinnamoyl-CoA hydratase/lyase, Journal of bioscience and bioengineering, 130 (2020) 89-97.
[112] A. Mitra, Y. Kitamura, M.J. Gasson, A. Narbad, A.J. Parr, J. Payne, M.J. Rhodes, C. Sewter, N.J. Walton, 4-Hydroxycinnamoyl-CoA Hydratase/lyase (HCHL)—An Enzyme of Phenylpropanoid Chain Cleavage fromPseudomonas, Archives of Biochemistry and Biophysics, 365 (1999) 10-16.
[113] A. Podstolski, D. Havkin-Frenkel, J. Malinowski, J.W. Blount, G. Kourteva, R.A. Dixon, Unusual 4-hydroxybenzaldehyde synthase activity from tissue cultures of the vanilla orchid Vanilla planifolia, Phytochemistry, 61 (2002) 611-620.
[114] H. Shen, M. Mazarei, H. Hisano, L. Escamilla-Trevino, C. Fu, Y. Pu, M.R. Rudis, Y. Tang, X. Xiao, L. Jackson, A genomics approach to deciphering lignin biosynthesis in switchgrass, The Plant Cell, 25 (2013) 4342-4361.
[115] B. Zhang, J.A. Lewis, W. Vermerris, S.E. Sattler, C. Kang, A sorghum ascorbate peroxidase with four binding sites has activity against ascorbate and phenylpropanoids, Plant Physiology, 192 (2023) 102-118.
[116] S.C. Demir, A.B. Yildirim, A.U. Turker, I. Eker, Seasonal variation in alkaloid content, phenolic constituent and biological activities of some Leucojum aestivum L. populations in Turkey, South African Journal of Botany, 147 (2022) 713-723.

[117] L. Szlávik, Á. Gyuris, J. Minárovits, P. Forgo, J. Molnár, J. Hohmann, Alkaloids from Leucojum vernum and antiretroviral activity of Amaryllidaceae alkaloids, Planta medica, 70 (2004) 871-873.
[118] M. Kukhanova, L. Victorova, A. Krayevsky, Peptidyltransferase center of ribosomes: on the mechanism of action of alkaloid lycorine, FEBS letters, 160 (1983) 129-133.
[119] A. Ptak, A. El Tahchy, F. Dupire, M. Boisbrun, M. Henry, Y. Chapleur, M. Mos, D. Laurain-Mattar, LCMS and GCMS for the screening of alkaloids in natural and in vitro extracts of Leucojum aestivum, Journal of natural products, 72 (2009) 142-147.
[120] D.-K. Ro, E.M. Paradise, M. Ouellet, K.J. Fisher, K.L. Newman, J.M. Ndungu, K.A. Ho, R.A. Eachus, T.S. Ham, J. Kirby, Production of the antimalarial drug precursor artemisinic acid in engineered yeast, Nature, 440 (2006) 940-943.
[121] S. Galanie, K. Thodey, I.J. Trenchard, M. Filsinger Interrante, C.D. Smolke, Complete biosynthesis of opioids in yeast, Science, 349 (2015) 1095-1100.
[122] J. Zhang, L.G. Hansen, O. Gudich, K. Viehrig, L.M. Lassen, L. Schrübbers, K.B. Adhikari, P. Rubaszka, E. Carrasquer-Alvarez, L. Chen, A microbial supply chain for production of the anti-cancer drug vinblastine, Nature, 609 (2022) 341-347.
[123] J.E. DiCarlo, J.E. Norville, P. Mali, X. Rios, J. Aach, G.M. Church, Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems, Nucleic acids research, 41 (2013) 4336-4343.
[124] J.C. Utomo, C.L. Hodgins, D.-K. Ro, Multiplex genome editing in yeast by CRISPR/Cas9–a potent and agile tool to reconstruct complex metabolic pathways, Frontiers in Plant Science, 12 (2021) 719148.
[125] S. Shi, Y. Liang, M.M. Zhang, E.L. Ang, H. Zhao, A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in Saccharomyces cerevisiae, Metabolic engineering, 33 (2016) 19-27.
[126] L. Wang, A. Deng, Y. Zhang, S. Liu, Y. Liang, H. Bai, D. Cui, Q. Qiu, X. Shang, Z. Yang, Efficient CRISPR–Cas9 mediated multiplex genome editing in yeasts, Biotechnology for biofuels, 11 (2018) 1-16.
[127] M.M. Jessop-Fabre, T. Jakočiūnas, V. Stovicek, Z. Dai, M.K. Jensen, J.D. Keasling, I. Borodina, EasyClone-MarkerFree: A vector toolkit for marker-less integration of genes into Saccharomyces cerevisiae via CRISPR-Cas9, Biotechnology journal, 11 (2016) 1110-1117.
[128] A. Reider Apel, L. d'Espaux, M. Wehrs, D. Sachs, R.A. Li, G.J. Tong, M. Garber, O. Nnadi, W. Zhuang, N.J. Hillson, A Cas9-based toolkit to program gene expression in Saccharomyces cerevisiae, Nucleic acids research, 45 (2017) 496-508.
[129] S. Hou, Q. Qin, J. Dai, Wicket: a versatile tool for the integration and optimization of exogenous pathways in Saccharomyces cerevisiae, ACS synthetic biology, 7 (2018) 782-788.
[130] L. Bourgeois, M.E. Pyne, V.J. Martin, A highly characterized synthetic landing pad system for precise multicopy gene integration in yeast, ACS Synthetic Biology, 7 (2018) 2675-2685.
[131] M. Shahsavarani, J.C. Utomo, R. Kumar, M. Paz-Galeano, J.J.O. Garza-García, Z. Mai, D.-K. Ro, Y. Qu, Improved protein glycosylation enabled heterologous biosynthesis of monoterpenoid indole alkaloids and their unnatural derivatives in yeast, Metabolic engineering communications, 16 (2023) e00215.
[132] J.C. Utomo, H.B. Barrell, R. Kumar, J. Smith, M.S. Brant, H. De la Hoz Siegler, D.-K. Ro, Reconstructing curcumin biosynthesis in yeast reveals the implication of caffeoyl-shikimate esterase in phenylpropanoid metabolic flux, Metabolic Engineering, 82 (2024) 286-296.
[133] N. Mehta, Y. Meng, R.N. Zare, R. Kamenetsky, E. Sattely, A developmental gradient reveals biosynthetic pathways to eukaryotic toxins in monocot geophytes, bioRxiv, (2023) 2023.2005. 2012.540595.
[134] W. Li, Y. Yang, C. Qiao, G. Zhang, Y. Luo, Functional characterization of phenylalanine ammonia-lyase- and cinnamate 4-hydroxylase-encoding genes from Lycoris radiata, a galanthamine-producing plant, International journal of biological macromolecules, 117 (2018) 1264-1279.

[135] S. d'Oelsnitz, D.J. Diaz, W. Kim, D.J. Acosta, T.L. Dangerfield, M.W. Schechter, M.B. Minus, J.R. Howard, H. Do, J.M. Loy, Biosensor and machine learning-aided engineering of an amaryllidaceae enzyme, Nature Communications, 15 (2024) 2084.
[136] T. Hotchandani, I. Desgagne-Penix, Heterocyclic Amaryllidaceae Alkaloids: Biosynthesis and Pharmacological Applications, Curr Top Med Chem, 17 (2017) 418-427.
[137] L.R. Tallini, E.H. Osorio, V.D.D. Santos, W.S. Borges, M. Kaiser, F. Viladomat, J.A.S. Zuanazzi, J. Bastida, *Hippeastrum reticulatum* (Amaryllidaceae): Alkaloid Profiling, Biological Activities and Molecular Docking, Molecules, 22 (2017) 4901.
[138] T.U. Jayawardena, N. Merindol, N.S. Liyanage, I. Desgagné-Penix, Unveiling Amaryllidaceae alkaloids: from biosynthesis to antiviral potential – a review, Natural Product Reports, (2024).
[139] S. Ka, N. Merindol, M. Koirala, I. Desgagne-Penix, Biosynthesis and biological activities of newly discovered Amaryllidaceae alkaloids., Molecules, 25 (2020) 4901.
[140] S. Berkov, L. Georgieva, V. Kondakova, A. Atanassov, F. Viladomat, J. Bastida, C. Codina, Plant sources of galanthamine: phytochemical and biotechnological aspects, Biotechnology & Biotechnological Equipment, 23 (2009) 1170-1176.
[141] J. Liu, W.X. Hu, L.F. He, M. Ye, Y. Li, Effects of lycorine on HL-60 cells via arresting cell cycle and inducing apoptosis, FEBS Lett, 578 (2004) 245-250.
[142] R. Havelek, M. Seifrtova, K. Kralovec, L. Bruckova, L. Cahlikova, M. Dalecka, J. Vavrova, M. Rezacova, L. Opletal, Z. Bilkova, The effect of Amaryllidaceae alkaloids haemanthamine and haemanthidine on cell cycle progression and apoptosis in p53-negative human leukemic Jurkat cells, Phytomedicine, 21 (2014) 479-490.
[143] M.P. Girard, N. Merindol, L. Berthoux, I. Desgagne-Penix, [Antiviral properties of plant alkaloids against RNA viruses], Virologie (Montrouge), 26 (2022) 431-450.
[144] M. Koirala, V. Karimzadegan, N.S. Liyanage, N. Mérindol, I. Desgagné-Penix, Biotechnological approaches to optimize the production of Amaryllidaceae alkaloids, Biomolecules, 12 (2022) 893.
[145] A. Diamond, I. Desgagne-Penix, Metabolic engineering for the production of plant isoquinoline alkaloids, Plant biotechnology journal, 14 (2016) 1319-1328.
[146] J. Zhang, L.G. Hansen, O. Gudich, K. Viehrig, L.M.M. Lassen, L. Schrubbers, K.B. Adhikari, P. Rubaszka, E. Carrasquer-Alvarez, L. Chen, V. D'Ambrosio, B. Lehka, A.K. Haidar, S. Nallapareddy, K. Giannakou, M. Laloux, D. Arsovska, M.A.K. Jorgensen, L.J.G. Chan, M. Kristensen, H.B. Christensen, S. Sudarsan, E.A. Stander, E. Baidoo, C.J. Petzold, T. Wulff, S.E. O'Connor, V. Courdavault, M.K. Jensen, J.D. Keasling, A microbial supply chain for production of the anti-cancer drug vinblastine, Nature, 609 (2022) 341-347.
[147] B. Zhang, K.M. Lewis, A. Abril, D.R. Davydov, W. Vermerris, S.E. Sattler, C. Kang, Structure and Function of the Cytochrome P450 Monooxygenase Cinnamate 4-hydroxylase from Sorghum bicolor, Plant Physiol, 183 (2020) 957-973.
[148] X.Y. Liu, H.N. Yu, S. Gao, Y.F. Wu, A.X. Cheng, H.X. Lou, The isolation and functional characterization of three liverwort genes encoding cinnamate 4-hydroxylase, Plant Physiol Biochem, 117 (2017) 42-50.
[149] M. Hotze, G. Schröder, J. Schröder, Cinnamate 4-hydroxylase from Catharanthus roseus and a strategy for the functional expression of plant cytochrome P450 proteins as translational fusions with P450 reductase in Escherichia coli, FEBS letters, 374 (1995) 345-350.
[150] D.K. Ro, N. Mah, B.E. Ellis, C.J. Douglas, Functional characterization and subcellular localization of poplar (Populus trichocarpa× Populus deltoides) cinnamate 4-hydroxylase, Plant Physiology, 126 (2001) 317-329.
[151] A.L. Schilmiller, J. Stout, J.K. Weng, J. Humphreys, M.O. Ruegger, C. Chapple, Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in Arabidopsis, The Plant Journal, 60 (2009) 771-782.

[152] H.C. Chen, Q. Li, C.M. Shuford, J. Liu, D.C. Muddiman, R.R. Sederoff, V.L. Chiang, Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis, Proc Natl Acad Sci U S A, 108 (2011) 21253-21258.
[153] Y. Li, J. Li, B. Qian, L. Cheng, S. Xu, R. Wang, De Novo Biosynthesis of p-Coumaric Acid in E. coli with a trans-Cinnamic Acid 4-Hydroxylase from the Amaryllidaceae Plant Lycoris aurea, Molecules, 23 (2018).
[154] Y. Xue, Y. Zhang, S. Grace, Q. He, Functional expression of an Arabidopsis p450 enzyme, p-coumarate-3-hydroxylase, in the cyanobacterium Synechocystis PCC 6803 for the biosynthesis of caffeic acid, Journal of applied phycology, 26 (2014) 219-226.
[155] G. Schoch, S. Goepfert, M. Morant, A. Hehn, D. Meyer, P. Ullmann, D. Werck-Reichhart, CYP98A3 from Arabidopsis thaliana is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway, Journal of Biological Chemistry, 276 (2001) 36566-36574.
[156] B. Zhang, J.A. Lewis, W. Vermerris, S.E. Sattler, C. Kang, A sorghum ascorbate peroxidase with four binding sites has activity against ascorbate and phenylpropanoids, Plant Physiol, 192 (2023) 102-118.
[157] S. Duvaud, C. Gabella, F. Lisacek, H. Stockinger, V. Ioannidis, C. Durinx, Expasy, the Swiss bioinformatics resource portal, as designed by its users, Nucleic Acids Res, 49 (2021) W216-W227.
[158] K. Tamura, G. Stecher, S. Kumar, MEGA11: molecular evolutionary genetics analysis version 11, Molecular biology and evolution, 38 (2021) 3022-3027.
[159] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis, Nature protocols, 10 (2015) 845-858.
[160] M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G.R. Lee, J. Wang, Q. Cong, L.N. Kinch, R.D. Schaeffer, C. Millan, H. Park, C. Adams, C.R. Glassman, A. DeGiovanni, J.H. Pereira, A.V. Rodrigues, A.A. van Dijk, A.C. Ebrecht, D.J. Opperman, T. Sagmeister, C. Buhlheller, T. Pavkov-Keller, M.K. Rathinaswamy, U. Dalwadi, C.K. Yip, J.E. Burke, K.C. Garcia, N.V. Grishin, P.D. Adams, R.J. Read, D. Baker, Accurate prediction of protein structures and interactions using a three-track neural network, Science, 373 (2021) 871-876.
[161] B.B. Majhi, S.E. Gelinas, N. Merindol, S. Ricard, I. Desgagne-Penix, Characterization of norbelladine synthase and noroxomaritidine/norcraugsodine reductase reveals a novel catalytic route for the biosynthesis of Amaryllidaceae alkaloids including the Alzheimer's drug galanthamine, Front Plant Sci, 14 (2023) 1231809.
[162] T. Sterling, J.J. Irwin, ZINC 15–ligand discovery for everyone, Journal of chemical information and modeling, 55 (2015) 2324-2337.
[163] M.F. Adasme, K.L. Linnemann, S.N. Bolz, F. Kaiser, S. Salentin, V.J. Haupt, M. Schroeder, PLIP 2021: Expanding the scope of the protein–ligand interaction profiler to DNA and RNA, Nucleic acids research, 49 (2021) W530-W534.
[164] M. Holsters, B. Silva, F. Van Vliet, C. Genetello, M. De Block, P. Dhaese, A. Depicker, D. Inze, G. Engler, R. Villarroel, et al., The functional organization of the nopaline *A. tumefaciens* plasmid pTiC58, Plasmid, 3 (1980) 212-230.
[165] D. Padmakshan, V.I. Timokhin, F. Lu, P.F. Schatz, R. Vanholme, W. Boerjan, J. Ralph, Synthesis of hydroxycinnamoyl shikimates and their role in monolignol biosynthesis, Holzforschung, 76 (2022) 133-144.
[166] M.W. Pfaffl, A new mathematical model for relative quantification in real-time RT–PCR, Nucleic acids research, 29 (2001) e45-e45.
[167] D.K. Ro, N. Mah, B.E. Ellis, C.J. Douglas, Functional characterization and subcellular localization of poplar (*Populus trichocarpa* x *Populus deltoides*) cinnamate 4-hydroxylase, Plant Physiol, 126 (2001) 317-329.
[168] R.D. Frederick, R.L. Thilmony, G. Sessa, G.B. Martin, Recognition specificity for the bacterial avirulence protein AvrPto is determined by Thr-204 in the activation loop of the tomato Pto kinase, Molecular cell, 2 (1998) 241-245.

[169] D.R. Nelson, H.W. Strobel, On the membrane topology of vertebrate cytochrome P-450 proteins, J Biol Chem, 263 (1988) 6038-6050.
[170] W. Li, L. Yang, L. Jiang, G. Zhang, Y. Luo, Molecular cloning and functional characterization of a cinnamate 4-hydroxylase-encoding gene from Camptotheca acuminata, Acta Physiologiae Plantarum, 38 (2016) 1-9.
[171] E. Szczesnaskorupa, P. Straub, B. Kemper, Deletion of a conserved tetrapeptide, PPGP, in P450 2C2 results in loss of enzymatic activity without a change in its cellular location, Archives of biochemistry and biophysics, 304 (1993) 170-175.
[172] H.G. Teutsch, M.P. Hasenfratz, A. Lesot, C. Stoltz, J.-M. Garnier, J.-M. Jeltsch, F. Durst, D. Werck-Reichhart, Isolation and sequence of a cDNA encoding the Jerusalem artichoke cinnamate 4-hydroxylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway, Proceedings of the National Academy of Sciences, 90 (1993) 4102-4106.
[173] K. Sen, W. Thiel, Role of two alternate water networks in Compound I formation in P450eryF, J Phys Chem B, 118 (2014) 2810-2820.
[174] L. Kriegshauser, S. Knosp, E. Grienenberger, K. Tatsumi, D.D. Gütle, I. Sørensen, L. Herrgott, J. Zumsteg, J.K. Rose, R. Reski, Function of the HYDROXYCINNAMOYL-CoA: SHIKIMATE HYDROXYCINNAMOYL TRANSFERASE is evolutionarily conserved in embryophytes, The Plant Cell, 33 (2021) 1472-1491.
[175] D. Plewczynski, M. Łaźniewski, R. Augustyniak, K. Ginalski, Can we trust docking results? Evaluation of seven commonly used programs on PDBbind database, Journal of computational chemistry, 32 (2011) 742-755.
[176] M. Morant, G.A. Schoch, P. Ullmann, T. Ertunç, D. Little, C.E. Olsen, M. Petersen, J. Negrel, D. Werck-Reichhart, Catalytic activity, duplication and evolution of the CYP98 cytochrome P450 family in wheat, Plant Molecular Biology, 63 (2007) 1-19.
[177] W. Li, Y. Yang, C. Qiao, G. Zhang, Y. Luo, Functional characterization of phenylalanine ammonia-lyase- and cinnamate 4-hydroxylase-encoding genes from Lycoris radiata, a galanthamine-producing plant, Int J Biol Macromol, 117 (2018) 1264-1279.
[178] H.-C. Chen, Q. Li, C.M. Shuford, J. Liu, D.C. Muddiman, R.R. Sederoff, V.L. Chiang, Membrane protein complexes catalyze both 4-and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis, Proceedings of the National Academy of Sciences, 108 (2011) 21253-21258.
[179] G. Li, X. Liu, Y. Zhang, A. Muhammad, W. Han, D. Li, X. Cheng, Y. Cai, Cloning and functional characterization of two cinnamate 4-hydroxylase genes from Pyrus bretschneideri, Plant Physiology and Biochemistry, 156 (2020) 135-145.
[180] D. Werck-Reichhart, R. Feyereisen, Cytochromes P450: a success story, Genome Biol, 1 (2000) REVIEWS3003.
[181] L. Achnine, E.B. Blancaflor, S. Rasmussen, R.A. Dixon, Colocalization of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis, Plant Cell, 16 (2004) 3098-3109.
[182] M. Matsuno, V. Compagnon, G.A. Schoch, M. Schmitt, D. Debayle, J.E. Bassard, B. Pollet, A. Hehn, D. Heintz, P. Ullmann, C. Lapierre, F. Bernier, J. Ehlting, D. Werck-Reichhart, Evolution of a novel phenolic pathway for pollen development, Science, 325 (2009) 1688-1692.
[183] Z. Liu, R. Tavares, E.S. Forsythe, F. André, R. Lugan, G. Jonasson, S. Boutet-Mercey, T. Tohge, M.A. Beilstein, D. Werck-Reichhart, Evolutionary interplay between sister cytochrome P450 genes shapes plasticity in plant metabolism, Nature Communications, 7 (2016) 13026.
[184] M. Morant, G.A. Schoch, P. Ullmann, T. Ertunc, D. Little, C.E. Olsen, M. Petersen, J. Negrel, D. Werck-Reichhart, Catalytic activity, duplication and evolution of the CYP98 cytochrome P450 family in wheat, Plant Mol Biol, 63 (2007) 1-19.

[185] A. Uhlmann, J. Ebel, Molecular cloning and expression of 4-coumarate: coenzyme A ligase, an enzyme involved in the resistance response of soybean (Glycine max L.) against pathogen attack, Plant Physiology, 102 (1993) 1147-1156.
[186] Y. Hu, Y. Gai, L. Yin, X. Wang, C. Feng, L. Feng, D. Li, X.-N. Jiang, D.-C. Wang, Crystal structures of a Populus tomentosa 4-coumarate: CoA ligase shed light on its enzymatic mechanisms, The Plant Cell, 22 (2010) 3093-3104.
[187] J.E. Bassard, L. Richert, J. Geerinck, H. Renault, F. Duval, P. Ullmann, M. Schmitt, E. Meyer, J. Mutterer, W. Boerjan, G. De Jaeger, Y. Mely, A. Goossens, D. Werck-Reichhart, Protein-protein and protein-membrane associations in the lignin pathway, Plant Cell, 24 (2012) 4465-4482.
[188] W. Biała, M. Jasiński, The phenylpropanoid case–it is transport that matters, Frontiers in Plant Science, 9 (2018) 1610.
[189] R. Havelek, M. Seifrtova, K. Kralovec, L. Bruckova, L. Cahlikova, M. Dalecka, J. Vavrova, M. Rezacova, L. Opletal, Z. Bilkova, The effect of Amaryllidaceae alkaloids haemanthamine and haemanthidine on cell cycle progression and apoptosis in p53-negative human leukemic Jurkat cells, Phytomedicine, 21 (2014) 479-490.
[190] S. Ka, N. Merindol, A.A. Sow, A. Singh, K. Landelouci, M.B. Plourde, G. Pépin, M. Masi, R. Di Lecce, A. Evidente, M. Seck, L. Berthoux, L. Chatel-Chaix, I. Desgagné-Penix, Amaryllidaceae Alkaloid Cherylline Inhibits the Replication of Dengue and Zika Viruses, Antimicrob. Agents Chemother., 65 (2021) e0039821.
[191] D.G. Gibson, L. Young, R.-Y. Chuang, J.C. Venter, C.A. Hutchison, H.O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases, Nature methods, 6 (2009) 343-345.
[192] B.B. Majhi, S.-E. Gélinas, N. Mérindol, S. Ricard, I. Desgagné-Penix, Characterization of norbelladine synthase and noroxomaritidine/norcraugsodine reductase reveals a novel catalytic route for the biosynthesis of Amaryllidaceae alkaloids including the Alzheimer's drug galanthamine, Frontiers in Plant Science, 14 (2023).
[193] S.Y.Y. Hnin, Y. Nakashima, H. Morita, Substrate Flexibilities of Norbelladine Synthase and Noroxomaritidine/Norcraugsodine Reductase for Hydroxylated and/or Methoxylated Aldehydes, Chemical and Pharmaceutical Bulletin, 72 (2024) 507-511.
[194] N. Samanani, P.J. Facchini, Isolation and partial characterization of norcoclaurine synthase, the first committed step in benzylisoquinoline alkaloid biosynthesis, from opium poppy, Planta, 213 (2001) 898-906.
[195] E.-J. Lee, P. Facchini, Norcoclaurine synthase is a member of the pathogenesis-related 10/Bet v1 protein family, The Plant Cell, 22 (2010) 3489-3503.
[196] J.J. Sramek, E.J. Frackiewicz, N.R. Cutler, Review of the acetylcholinesterase inhibitor galanthamine, Expert Opinion on Investigational Drugs, 9 (2000) 2393-2402.
[197] A.M. Takos, F. Rook, Towards a molecular understanding of the biosynthesis of Amaryllidaceae alkaloids in support of their expanding medical use, International journal of molecular sciences, 14 (2013) 11713-11741.
[198] S. Kawai, A. Mori, T. Shiokawa, S. Kajita, Y. Katayama, N. Morohoshi, Isolation and analysis of cinnamic acid 4-hydroxylase homologous genes from a hybrid aspen, Populus kitakamiensis, Bioscience, biotechnology, and biochemistry, 60 (1996) 1586-1597.
[199] P. Urban, D. WERCK-REICHHART, H.G. Teutsch, F. Durst, S. Regnier, M. Kazmaier, D. Pompon, Characterization of recombinant plant cinnamate 4-hydroxylase produced in yeast: Kinetic and spectral properties of the major plant P450 of the phenylpropanoid pathway, European Journal of Biochemistry, 222 (1994) 843-850.
[200] K. Jørgensen, A.V. Rasmussen, M. Morant, A.H. Nielsen, N. Bjarnholt, M. Zagrobelny, S. Bak, B.L. Møller, Metabolon formation and metabolic channeling in the biosynthesis of plant natural products, Current opinion in plant biology, 8 (2005) 280-291.

[201] O. Gotoh, Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences, J Biol Chem, 267 (1992) 83-90.
[202] G.A. Schoch, J.K. Yano, M.R. Wester, K.J. Griffin, C.D. Stout, E.F. Johnson, Structure of human microsomal cytochrome P450 2C8. Evidence for a peripheral fatty acid binding site, J Biol Chem, 279 (2004) 9497-9503.
[203] S.F. Martin, The amaryllidaceae alkaloids, in: The Alkaloids: Chemistry and Pharmacology, Elsevier, 1987, pp. 251-376.
[204] S. Berkov, E. Osorio, F. Viladomat, J. Bastida, Chemodiversity, chemotaxonomy and chemoecology of Amaryllidaceae alkaloids, Alkaloids Chem Biol, 83 (2020) 113-185.
[205] M. He, C. Qu, O. Gao, X. Hu, X. Hong, Biological and pharmacological activities of amaryllidaceae alkaloids, RSC Advances, 5 (2015) 16562-16574.
[206] I. Desgagné-Penix, Biosynthesis of alkaloids in Amaryllidaceae plants: a review, Phytochemistry Reviews, 20 (2020) 409-431.
[207] S. Ka, M. Koirala, N. Merindol, I. Desgagne-Penix, Biosynthesis and Biological Activities of Newly Discovered Amaryllidaceae Alkaloids, Molecules, 25 (2020).
[208] P.J. Houghton, Y. Ren, M.-J. Howes, Acetylcholinesterase inhibitors from plants and fungi, Natural product reports, 23 (2006) 181-199.
[209] U.K. Kalola, H. Nguyen, Galantamine, in: StatPearls, Treasure Island (FL), 2022.
[210] G. Razay, G.K. Wilcock, Galantamine in Alzheimer's disease, Expert Rev Neurother, 8 (2008) 9-17.
[211] G.S. Santos, S.B.P. Sinoti, F.T.C. de Almeida, D. Silveira, L.A. Simeoni, K.K.P. Gomes-Copeland, Use of galantamine in the treatment of Alzheimer's disease and strategies to optimize its biosynthesis using the in vitro culture technique, Plant Cell, Tissue and Organ Culture (PCTOC), 143 (2020) 13-29.
[212] B.M. Trost, W. Tang, An Efficient Enantioselective Synthesis of (−)-Galanthamine, Angewandte Chemie International Edition, 41 (2002) 2795-2797.
[213] J. Marco-Contelles, C. Rodríguez, A.G. García, Chemical synthesis of galantamine, an acetylcholinesterase inhibitor for treatment of Alzheimer’s disease, Expert Opinion on Therapeutic Patents, 15 (2005) 575-587.
[214] A.M. Takos, F. Rook, Towards a molecular understanding of the biosynthesis of amaryllidaceae alkaloids in support of their expanding medical use, Int J Mol Sci, 14 (2013) 11713-11741.
[215] M. Roy, L. Liang, X. Xiao, P. Feng, M. Ye, J. Liu, Lycorine: A prospective natural lead for anticancer drug discovery, Biomed Pharmacother, 107 (2018) 615-624.
[216] Y. Ding, D. Qu, K.M. Zhang, X.X. Cang, Z.N. Kou, W. Xiao, J.B. Zhu, Phytochemical and biological investigations of Amaryllidaceae alkaloids: a review, J. Asian Nat. Prod. Res., 19 (2017) 53-100.
[217] M. Nonn, A. Binder, B. Volk, L. Kiss, Stereo- and regiocontrolled synthesis of highly functionalized cyclopentanes with multiple chiral centers, Synthetic Communications, 50 (2020) 1199-1209.
[218] K.-i. Yamada, M. Yamashita, T. Sumiyoshi, K. Nishimura, K. Tomioka, Total synthesis of (−)-lycorine and (−)-2-epi-lycorine by asymmetric conjugate addition cascade, Organic Letters, 11 (2009) 1631-1633.
[219] H.-S. Shin, Y.-G. Jung, H.-K. Cho, Y.-G. Park, C.-G. Cho, Total Synthesis of (±)-Lycorine from the Endo-Cycloadduct of 3, 5-Dibromo-2-pyrone and (E)-β-Borylstyrene, Organic letters, 16 (2014) 5718-5720.
[220] S. Manolov, S. Nikolova, M. Ghate, I. Ivanov, A brief review of Cherylline synthesis, Indian Journal of Chemistry Section B, 54B (2015) 1301-1320.
[221] S. Lebrun, A. Couture, E. Deniau, P. Grandclaudon, A new synthesis of (+)-and (−)-cherylline, Organic & biomolecular chemistry, 1 (2003) 1701-1706.
[222] I.R. Miller, N.J. McLean, G.A. Moustafa, V. Ajavakom, S.C. Kemp, R.K. Bellingham, N.P. Camp, R.C. Brown, Transition-metal-mediated chemo-and stereoselective total synthesis of (−)-galanthamine, The Journal of Organic Chemistry, 87 (2022) 1325-1334.
[223] V. Satcharoen, N.J. McLean, S.C. Kemp, N.P. Camp, R.C. Brown, Stereocontrolled synthesis of (−)-galanthamine, Organic letters, 9 (2007) 1867-1869.

[224] A. Brossi, G. Grethe, S. Teitel, W.C. Wildman, D.T. Bailey, Cherylline, a 4-phenyl-1,2,3,4-tetrahydroisoquinoline alkaloid, The Journal of Organic Chemistry, 35 (2002) 1100-1104.
[225] L. Torras-Claveria, S. Berkov, C. Codina, F. Viladomat, J. Bastida, Metabolomic analysis of bioactive Amaryllidaceae alkaloids of ornamental varieties of Narcissus by GC–MS combined with k-means cluster analysis, Industrial Crops and Products, 56 (2014) 211-222.
[226] B. Ncube, J.J. Nair, L. Rárová, M. Strnad, J.F. Finnie, J. Van Staden, Seasonal pharmacological properties and alkaloid content in Cyrtanthus contractus N.E. Br, South African Journal of Botany, 97 (2015) 69-76.
[227] A. Santos, F. Fidalgo, I. Santos, R. Salema, In vitrobulb formation ofNarcissus asturiensis, a threatened species of theAmaryllidaceae, The Journal of Horticultural Science and Biotechnology, 77 (2015) 149-152.
[228] A. Pavlov, S. Berkov, E. Courot, T. Gocheva, D. Tuneva, B. Pandova, M. Georgiev, V. Georgiev, S. Yanev, M. Burrus, M. Ilieva, Galanthamine production by Leucojum aestivum in vitro systems, Process Biochemistry, 42 (2007) 734-739.
[229] M. Diop, A. Hehn, A. Ptak, F. Chrétien, S. Doerper, E. Gontier, F. Bourgaud, M. Henry, Y. Chapleur, D. Laurain-Mattar, Hairy root and tissue cultures of Leucojum aestivum L.—relationships to galanthamine content, Phytochemistry reviews, 6 (2007) 137-141.
[230] S. Berkov, A. Pavlov, V. Georgiev, J. Bastida, M. Burrus, M. Ilieva, C. Codina, Alkaloid synthesis and accumulation in Leucojum aestivum in vitro cultures, Nat Prod Commun, 4 (2009) 359-364.
[231] V. Georgiev, S. Berkov, M. Georgiev, M. Burrus, C. Codina, J. Bastida, M. Ilieva, A. Pavlov, Optimized nutrient medium for galanthamine production in Leucojum aestivum L. in vitro shoot system, Z Naturforsch C J Biosci, 64 (2009) 219-224.
[232] D. Laurain-Mattar, A. Ptak, Amaryllidaceae Alkaloid Accumulation by Plant In Vitro Systems, in: A. Pavlov, T. Bley (Eds.) Bioprocessing of Plant In Vitro Systems, Springer International Publishing, Cham, 2018, pp. 203-223.
[233] V. Georgiev, I. Ivanov, A. Pavlov, Recent Progress in Amaryllidaceae Biotechnology, Molecules, 25 (2020).
[234] H. Kaur, S. Chahal, P. Jha, M.M. Lekhak, M.S. Shekhawat, D. Naidoo, A.D. Arencibia, S.J. Ochatt, V. Kumar, Harnessing plant biotechnology-based strategies for in vitro galanthamine (GAL) biosynthesis: a potent drug against Alzheimer's disease, Plant Cell, Tissue and Organ Culture (PCTOC), 149 (2022) 81-103.
[235] F. Carbone, A. Preuss, R.C. De Vos, E. D'Amico, G. Perrotta, A.G. Bovy, S. Martens, C. Rosati, Developmental, genetic and environmental factors affect the expression of flavonoid genes, enzymes and metabolites in strawberry fruits, Plant Cell Environ, 32 (2009) 1117-1131.
[236] F. Bourgaud, A. Gravot, S. Milesi, E. Gontier, Production of plant secondary metabolites: a historical perspective, Plant Science, 161 (2001) 839-851.
[237] N. Singh, B. Sharma, Toxicological effects of berberine and sanguinarine, Frontiers in molecular biosciences, 5 (2018) 21.
[238] A.M. Gallego, L.F. Rojas, O. Parra, H.A. Rodriguez, J.C. Mazo Rivas, A.I. Urrea, L. Atehortúa, A.S. Fister, M.J. Guiltinan, S.N. Maximova, Transcriptomic analyses of cacao cell suspensions in light and dark provide target genes for controlled flavonoid production, Scientific reports, 8 (2018) 13575.
[239] L.M. Trujillo-Chacon, E.R. Pastene-Navarrete, L. Bustamante, M. Baeza, J.E. Alarcon-Enos, C.L. Cespedes-Acuna, In vitro micropropagation and alkaloids analysis by GC-MS of Chilean Amaryllidaceae plants: Rhodophiala pratensis, Phytochem Anal, 31 (2020) 46-56.
[240] J. Juan-Vicedo, A. Pavlov, S. Ríos, J.L. Casas, In vitro culture and micropropagation of the Baetic-Moroccan endemic plant Lapiedra martinezii Lag. (Amaryllidaceae), In Vitro Cellular & Developmental Biology - Plant, 55 (2019) 725-732.
[241] R. Colque, F. Viladomat, J. Bastida, C. Codina, Micropropagation of the rareEucrosia stricklandii(Amaryllidaceae) by twin-scaling and shake liquid culture, The Journal of Horticultural Science and Biotechnology, 77 (2015) 739-743.

[242] Z. Ren, Y. Lin, X. Lv, J. Zhang, D. Zhang, C. Gao, Y. Wu, Y. Xia, Clonal bulblet regeneration and endophytic communities profiling of Lycoris sprengeri, an economically valuable bulbous plant of pharmaceutical and ornamental value, Scientia Horticulturae, 279 (2021) 109856.
[243] T. Hotchandani, J. de Villers, I. Desgagne-Penix, Developmental Regulation of the Expression of Amaryllidaceae Alkaloid Biosynthetic Genes in Narcissus papyraceus, Genes (Basel), 10 (2019) 594.
[244] S. Ka, M. Masi, N. Merindol, R. Di Lecce, M.B. Plourde, M. Seck, M. Gorecki, G. Pescitelli, I. Desgagne-Penix, A. Evidente, Gigantelline, gigantellinine and gigancrinine, cherylline- and crinine-type alkaloids isolated from Crinum jagus with anti-acetylcholinesterase activity, Phytochemistry, 175 (2020) 112390.
[245] C.W. Fennell, E.E. Elgorashi, J. van Staden, Alkaloid production in Crinum moorei cultures, J Nat Prod, 66 (2003) 1524-1526.
[246] S. Berkov, L. Georgieva, B. Sidjimova, M. Nikolova, M. Stanilova, J. Bastida, In vitro propagation and biosynthesis of Sceletium-type alkaloids in Narcissus pallidulus andNarcissus cv. Hawera, South African Journal of Botany, 136 (2021) 190-194.
[247] M. Sellés, S. Bergoñón, F. Viladomat, J. Bastida, C. Codina, Effect of sucrose on growth and galanthamine production in shoot-clump cultures of Narcissus confusus in liquid-shake medium, Plant cell, tissue and organ culture, 49 (1997) 129-136.
[248] A. Ferdausi, X. Chang, A. Hall, M. Jones, Galanthamine production in tissue culture and metabolomic study on Amaryllidaceae alkaloids in Narcissus pseudonarcissus cv. Carlton, Industrial crops and products, 144 (2020) 112058.
[249] A. Tarakemeh, M. Azizi, V. Rowshan, H. Salehi, R. Spina, F. Dupire, H. Arouie, D. Laurain-Mattar, Screening of Amaryllidaceae alkaloids in bulbs and tissue cultures of Narcissus papyraceus and four varieties of N. tazetta, Journal of Pharmaceutical and Biomedical Analysis, 172 (2019) 230-237.
[250] A. Ferdausi, X. Chang, A. Hall, M. Jones, Galanthamine production in tissue culture and metabolomic study on Amaryllidaceae alkaloids in Narcissus pseudonarcissus cv. Carlton, Industrial Crops and Products, 144 (2020).
[251] R. Colque, F. Viladomat, J. Bastida, C. Codina, Improved production of galanthamine and related alkaloids by methyl jasmonate in Narcissus confusus shoot-clumps, Planta medica, 70 (2004) 1180-1188.
[252] A. Ferdausi, X. Chang, M. Jones, Enhancement of galanthamine production through elicitation and NMR-based metabolite profiling in Narcissus pseudonarcissus cv. Carlton in vitro callus cultures, In Vitro Cellular & Developmental Biology-Plant, 57 (2021) 435-446.
[253] Q. Li, J. Xu, Y. Zheng, Y. Zhang, Y. Cai, Transcriptomic and Metabolomic Analyses Reveals That Exogenous Methyl Jasmonate Regulates Galanthamine Biosynthesis in Lycoris longituba Seedlings, Frontiers in plant science, (2021) 1933.
[254] Q. Li, J. Xu, L. Yang, Y. Sun, X. Zhou, Y. Zheng, Y. Zhang, Y. Cai, LED Light Quality Affect Growth, Alkaloids Contents, and Expressions of Amaryllidaceae Alkaloids Biosynthetic Pathway Genes in Lycoris longituba, Journal of Plant Growth Regulation, 41 (2021) 257-270.
[255] H.-m. Mu, R. Wang, X.-d. Li, Y.-m. Jiang, C.-y. Wang, J.-p. Quan, F. Peng, B. Xia, Effect of abiotic and biotic elicitors on growth and alkaloid accumulation of Lycoris chinensis seedlings, Zeitschrift für Naturforschung C, 64 (2009) 541-550.
[256] A. Ptak, M. Simlat, E. Morańska, E. Skrzypek, M. Warchoł, A. Tarakemeh, D. Laurain-Mattar, Exogenous melatonin stimulated Amaryllidaceae alkaloid biosynthesis in in vitro cultures of Leucojum aestivum L, Industrial Crops and Products, 138 (2019) 111458.
[257] I. Ivanov, V. Georgiev, A. Pavlov, Elicitation of galanthamine biosynthesis by Leucojum aestivum liquid shoot cultures, Journal of plant physiology, 170 (2013) 1122-1129.
[258] A. Schumann, L. Torras-Claveria, S. Berkov, D. Claus, A. Gerth, J. Bastida, C. Codina, Elicitation of galanthamine production by Leucojum aestivum shoots grown in temporary immersion system, Biotechnology progress, 29 (2013) 311-318.

[259] A. Ptak, E. Morańska, S. Saliba, A. Zieliński, M. Simlat, D. Laurain-Mattar, Elicitation of galanthamine and lycorine biosynthesis by Leucojum aestivum L. and L. aestivum 'Gravety Giant'plants cultured in bioreactor RITA®, Plant Cell, Tissue and Organ Culture (PCTOC), 128 (2017) 335-345.
[260] A. Ptak, E. Morańska, S. Saliba, A. Zieliński, M. Simlat, D. Laurain-Mattar, Elicitation of galanthamine and lycorine biosynthesis by Leucojum aestivum L. and L. aestivum 'Gravety Giant'plants cultured in bioreactor RITA®, Plant Cell, Tissue and Organ Culture (PCTOC), 128 (2017) 335-345.
[261] J. Zhou, Z. Liu, S. Wang, J. Li, Y. Li, W.K. Chen, R. Wang, Fungal endophytes promote the accumulation of Amaryllidaceae alkaloids in Lycoris radiata, Environmental microbiology, 22 (2020) 1421-1434.
[262] S. Bergoñón, C. Codina, J. Bastida, F. Viladomat, E. Melé, Galanthamine production in "shoot-clump" cultures of Narcissus confusus in liquid-shake medium, Plant cell, tissue and organ culture, 45 (1996) 191-199.
[263] R. Colque, F. Viladomat, J. Bastida, C. Codina, Micropropagation of the rare Eucrosia stricklandii (Amaryllidaceae) by twin-scaling and shake liquid culture, The Journal of Horticultural Science and Biotechnology, 77 (2002) 739-743.
[264] S. Bergoñón, C. Codina, J. Bastida, F. Viladomat, E. Melé, Galanthamine production in "shoot-clump" cultures of Narcissus confusus in liquid-shake medium, Plant cell, tissue and organ culture, 45 (1996) 191-199.
[265] I. Ivanov, V. Georgiev, M. Georgiev, M. Ilieva, A. Pavlov, Galanthamine and related alkaloids production by Leucojum aestivum L. shoot culture using a temporary immersion technology, Applied Biochemistry and Biotechnology, 163 (2011) 268-277.
[266] V. Georgiev, I. Ivanov, S. Berkov, M. Ilieva, M. Georgiev, T. Gocheva, A. Pavlov, Galanthamine production byLeucojum aestivumL. shoot culture in a modified bubble column bioreactor with internal sections, Engineering in Life Sciences, 12 (2012) 534-543.
[267] I. Ivanov, V. Georgiev, S. Berkov, A. Pavlov, Alkaloid patterns in Leucojum aestivum shoot culture cultivated at temporary immersion conditions, Journal of Plant Physiology, 169 (2012) 206-211.
[268] I. Ivanov, V. Georgiev, S. Berkov, A. Pavlov, Alkaloid patterns in Leucojum aestivum shoot culture cultivated at temporary immersion conditions, J Plant Physiol, 169 (2012) 206-211.
[269] S. Berkov, A. Pavlov, V. Georgiev, J. Bastida, M. Burrus, M. Ilieva, C. Codina, Alkaloid synthesis and accumulation in Leucojum aestivum in vitro cultures, Natural Product Communications, 4 (2009) 1934578X0900400328.
[270] M. Rahimi Khonakdari, H. Rezadoost, R. Heydari, M.H. Mirjalili, Effect of photoperiod and plant growth regulators on in vitro mass bulblet proliferation of Narcissus tazzeta L.(Amaryllidaceae), a potential source of galantamine, Plant Cell, Tissue and Organ Culture (PCTOC), 142 (2020) 187-199.
[271] G. Fridborg, M.K. Pedersen, L. Landström, T.R. Eriksson, The Effect of Activated Charcoal on Tissue Cultures: Adsorption of Metabolites Inhibiting Morphogenesis, Physiologia Plantarum, 43 (1978) 104-106.
[272] C. Fennell, E. Elgorashi, J. Van Staden, Alkaloid production in Crinum moorei cultures, Journal of natural products, 66 (2003) 1524-1526.
[273] F.-s. Dong, M.-Y. Lv, J.-p. Wang, X.-p. Shi, X.-x. Liang, Y.-w. Liu, F. Yang, H. Zhao, J.-F. Chai, S. Zhou, Transcriptome analysis of activated charcoal-induced growth promotion of wheat seedlings in tissue culture, BMC genetics, 21 (2020) 1-11.
[274] A. Ptak, E. Moranska, E. Skrzypek, M. Warchol, R. Spina, D. Laurain-Mattar, M. Simlat, Carbohydrates stimulated Amaryllidaceae alkaloids biosynthesis in Leucojum aestivum L. plants cultured in RITA((R)) bioreactor, PeerJ, 8 (2020) e8688.
[275] T. Gaspar, C. Kevers, C. Penel, H. Greppin, D.M. Reid, T.A. Thorpe, Plant hormones and plant growth regulators in plant tissue culture, In Vitro Cellular & Developmental Biology-Plant, 32 (1996) 272-289.
[276] A. El Tahchy, S. Bordage, A. Ptak, F. Dupire, E. Barre, C. Guillou, M. Henry, Y. Chapleur, D. Laurain-Mattar, Effects of sucrose and plant growth regulators on acetylcholinesterase inhibitory activity of

alkaloids accumulated in shoot cultures of Amaryllidaceae, Plant Cell, Tissue and Organ Culture (PCTOC), 106 (2011) 381-390.
[277] A. Ptak, M. Simlat, E. Morańska, R. Spina, F. Dupire, S. Kozaczka, D. Laurain-Mattar, Potential of the Leucojum aestivum L. endophytic bacteria to promote Amaryllidaceae alkaloids biosynthesis, in: ACTA BIOLOGICA CRACOVIENSIA Series Botanica, 2019.
[278] R. Wang, S. Xu, N. Wang, B. Xia, Y. Jiang, R. Wang, Transcriptome analysis of secondary metabolism pathway, transcription factors, and transporters in response to methyl jasmonate in Lycoris aurea, Frontiers in plant science, 7 (2017) 1971.
[279] M.B. Kilgore, T.M. Kutchan, The Amaryllidaceae alkaloids: biosynthesis and methods for enzyme discovery, Phytochemistry reviews, 15 (2016) 317-337.
[280] L. Tousignant, A.M. Diaz-Garza, B.B. Majhi, S.-E. Gélinas, A. Singh, I. Desgagne-Penix, Transcriptome analysis of Leucojum aestivum and identification of genes involved in norbelladine biosynthesis, Planta, 255 (2022) 1-23.
[281] A. Singh, I. Desgagné-Penix, Transcriptome and metabolome profiling of Narcissus pseudonarcissus 'King Alfred'reveal components of Amaryllidaceae alkaloid metabolism, Scientific reports, 7 (2017) 17356.
[282] C.H. Park, H.J. Yeo, Y.E. Park, S.A. Baek, J.K. Kim, S.U. Park, Transcriptome Analysis and Metabolic Profiling of Lycoris Radiata, Biology (Basel), 8 (2019).
[283] J. Hu, W. Li, Z. Liu, G. Zhang, Y. Luo, Molecular cloning and functional characterization of tyrosine decarboxylases from galanthamine-producing Lycoris radiata, Acta Physiologiae Plantarum, 43 (2021) 1-12.
[284] M.B. Kilgore, C.K. Holland, J.M. Jez, T.M. Kutchan, Identification of a Noroxomaritidine Reductase with Amaryllidaceae Alkaloid Biosynthesis Related Activities, J. Biol. Chem., 291 (2016) 16740-16752.
[285] A. Singh, M.A. Massicotte, A. Garand, L. Tousignant, V. Ouellette, G. Berube, I. Desgagne-Penix, Cloning and characterization of norbelladine synthase catalyzing the first committed reaction in Amaryllidaceae alkaloid biosynthesis, BMC Plant Biol, 18 (2018) 338.
[286] M.B. Kilgore, M.M. Augustin, C.M. Starks, M. O'Neil-Johnson, G.D. May, J.A. Crow, T.M. Kutchan, Cloning and characterization of a norbelladine 4'-O-methyltransferase involved in the biosynthesis of the Alzheimer's drug galanthamine in Narcissus sp. aff. pseudonarcissus, PLoS One, 9 (2014) e103223.
[287] A. Katoh, H. Ohki, K. Inai, T. Hashimoto, Molecular regulation of nicotine biosynthesis, Plant Biotechnology, 22 (2005) 389-392.
[288] A. Onoyovwe, J.M. Hagel, X. Chen, M.F. Khan, D.C. Schriemer, P.J. Facchini, Morphine biosynthesis in opium poppy involves two cell types: sieve elements and laticifers, Plant Cell, 25 (2013) 4110-4122.
[289] A. El Tahchy, A. Ptak, M. Boisbrun, E. Barre, C. Guillou, F.o. Dupire, F.o. Chrétien, M. Henry, Y. Chapleur, D. Laurain-Mattar, Kinetic study of the rearrangement of deuterium-labeled 4'-O-methylnorbelladine in Leucojum aestivum shoot cultures by mass spectrometry. Influence of precursor feeding on Amaryllidaceae alkaloid accumulation, Journal of natural products, 74 (2011) 2356-2361.
[290] F. Aleya, C. Xianmin, H. Anthony, J. Meriel, Relative expression of putative genes involved in galanthamine and other Amaryllidaceae alkaloids biosynthesis in Narcissus field and in vitro tissues, Gene, 774 (2021) 145424.
[291] B. Sun, P. Wang, R. Wang, Y. Li, S. Xu, Molecular Cloning and Characterization of a meta/para-O-Methyltransferase from Lycoris aurea, International journal of molecular sciences, 19 (2018) 1911.
[292] Y. Li, S. Li, K. Thodey, I. Trenchard, A. Cravens, C.D. Smolke, Complete biosynthesis of noscapine and halogenated alkaloids in yeast, Proceedings of the National Academy of Sciences, 115 (2018) E3922-E3931.
[293] A. Diamond, I. Desgagné-Penix, Metabolic engineering for the production of plant isoquinoline alkaloids, Plant biotechnology journal, 14 (2016) 1319-1328.
[294] P.J. Facchini, J. Bohlmann, P.S. Covello, V. De Luca, R. Mahadevan, J.E. Page, D.-K. Ro, C.W. Sensen, R. Storms, V.J. Martin, Synthetic biosystems for the production of high-value plant metabolites, Trends in biotechnology, 30 (2012) 127-131.

[295] E. Fossati, L. Narcross, A. Ekins, J.-P. Falgueyret, V.J. Martin, Synthesis of morphinan alkaloids in Saccharomyces cerevisiae, PLoS One, 10 (2015) e0124459.
[296] M. Cao, M. Gao, M. Suastegui, Y. Mei, Z. Shao, Building microbial factories for the production of aromatic amino acid pathway derivatives: from commodity chemicals to plant-sourced natural products, Metabolic engineering, 58 (2020) 94-132.
[297] Y. Li, J. Li, B. Qian, L. Cheng, S. Xu, R. Wang, De novo biosynthesis of p-coumaric acid in E. coli with a trans-cinnamic acid 4-hydroxylase from the Amaryllidaceae plant Lycoris aurea, Molecules, 23 (2018) 3185.
[298] R. Patnaik, R.R. Zolandz, D.A. Green, D.F. Kraynie, L-tyrosine production by recombinant Escherichia coli: fermentation optimization and recovery, Biotechnology and bioengineering, 99 (2008) 741-752.
[299] M. Farhi, E. Marhevka, J. Ben-Ari, A. Algamas-Dimantov, Z. Liang, V. Zeevi, O. Edelbaum, B. Spitzer-Rimon, H. Abeliovich, B. Schwartz, Generation of the potent anti-malarial drug artemisinin in tobacco, Nature Biotechnology, 29 (2011) 1072-1074.
[300] R. Kitney, P. Freemont, Synthetic biology–the state of play, FEBS letters, 586 (2012) 2029-2036.
[301] N. Matasci, L.H. Hung, Z. Yan, E.J. Carpenter, N.J. Wickett, S. Mirarab, N. Nguyen, T. Warnow, S. Ayyampalayam, M. Barker, J.G. Burleigh, M.A. Gitzendanner, E. Wafula, J.P. Der, C.W. dePamphilis, B. Roure, H. Philippe, B.R. Ruhfel, N.W. Miles, S.W. Graham, S. Mathews, B. Surek, M. Melkonian, D.E. Soltis, P.S. Soltis, C. Rothfels, L. Pokorny, J.A. Shaw, L. DeGironimo, D.W. Stevenson, J.C. Villarreal, T. Chen, T.M. Kutchan, M. Rolf, R.S. Baucom, M.K. Deyholos, R. Samudrala, Z. Tian, X. Wu, X. Sun, Y. Zhang, J. Wang, J. Leebens-Mack, G.K. Wong, Data access for the 1,000 Plants (1KP) project, Gigascience, 3 (2014) 17.
[302] M. Xiao, Y. Zhang, X. Chen, E.-J. Lee, C.J. Barber, R. Chakrabarty, I. Desgagné-Penix, T.M. Haslam, Y.-B. Kim, E. Liu, Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest, Journal of biotechnology, 166 (2013) 122-134.
[303] S. Ka, M. Koirala, N. Mérindol, I. Desgagné-Penix, Biosynthesis and biological activities of newly discovered Amaryllidaceae alkaloids, Molecules, 25 (2020) 4901.
[304] Y. Ding, D. Qu, K.-M. Zhang, X.-X. Cang, Z.-N. Kou, W. Xiao, J.-B. Zhu, Phytochemical and biological investigations of Amaryllidaceae alkaloids: a review, Journal of Asian natural products research, 19 (2017) 53-100.
[305] L. Cahlíková, N. Benesová, K. Macáková, K. Urbanova, L. Opletal, GC/MS Analysis of Three Amaryllidaceae Species and Their Cholinesterase Activity, Natural Product Communications, 6 (2011).
[306] L. Cheesman, J.J. Nair, J. van Staden, Antibacterial activity of crinane alkaloids from Boophone disticha (Amaryllidaceae), Journal of ethnopharmacology, 140 2 (2012) 405-408.
[307] K. Habartová, L. Cahlíková, M. Řezáčová, R. Havelek, The Biological Activity of Alkaloids from the Amaryllidaceae: From Cholinesterases Inhibition to Anticancer Activity, Natural product communications, 11 (2016) 1587-1594.
[308] Z. Jin, G. Yao, Amaryllidaceae and Sceletium alkaloids, Natural Product Reports, 36 (2019) 1462-1488.
[309] A. El Tahchy, M. Boisbrun, A. Ptak, F. Dupire, F. Chretien, M. Henry, Y. Chapleur, D. Laurain-Mattar, New method for the study of Amaryllidaceae alkaloid biosynthesis using biotransformation of deuterium-labeled precursor in tissue cultures, Acta Biochim Pol, 57 (2010) 75-82.
[310] T. Hotchandani, I. Desgagne-Penix, Heterocyclic Amaryllidaceae alkaloids: biosynthesis and pharmacological applications, Current topics in medicinal chemistry, 17 (2017) 418-427.
[311] S.-Y. Han, S.C. Mayer, E.J. Schweiger, B.M. Davis, M.M. Joulleé, Synthesis and biological activity of galanthamine derivatives as acetylcholinesterase (AChE) inhibitors, Bioorganic & Medicinal Chemistry Letters, 1 (1991) 579-580.
[312] Y. Sun, P. Wu, Y. Sun, F.S. Sharopov, Q. Yang, F. Chen, P. Wang, Z. Liang, Lycorine possesses notable anticancer potentials in on-small cell lung carcinoma cells via blocking Wnt/β-catenin signaling and epithelial-mesenchymal transition (EMT), Biochemical and biophysical research communications, 495 (2018) 911-921.

[313] C. Wang, Q. Wang, X. Li, Z. Jin, P. Xu, N. Xu, A. Xu, Y. Xu, S. Zheng, J. Zheng, C. Liu, P. Huang, Lycorine induces apoptosis of bladder cancer T24 cells by inhibiting phospho-Akt and activating the intrinsic apoptotic cascade, Biochemical and biophysical research communications, 483 (2017) 197-202.
[314] Z. Cao, D. Yu, S. Fu, G. Zhang, Y. Pan, M. Bao, J. Tu, B. Shang, P. Guo, P. Yang, Q. Zhou, Lycorine hydrochloride selectively inhibits human ovarian cancer cell proliferation and tumor neovascularization with very low toxicity, Toxicology Letters, 218 (2013) 174-185.
[315] J. Yubin, Y. Miao, Q. Zheng, C. Di, X. Guosong, W. Bing, J. Weiling, C. Lin, Study on apoptosis effect of human breast cancer cell MCF-7 induced by lycorine hydrochloride via death receptor pathway, in: Saudi Pharmaceutical Journal, pp. 633-637.
[316] M. Hu, S. Peng, Y. He, M. Qin, X. Cong, Y. Xing, M. Liu, Z. Yi, Lycorine is a novel inhibitor of the growth and metastasis of hormone-refractory prostate cancer, Oncotarget, 6 (2015) 15348-15361.
[317] B. Gabrielsen, T.P. Monath, J.W. Huggins, D.F. Kefauver, G.R. Pettit, G. Groszek, M. Hollingshead, J.J. Kirsi, W.M. Shannon, E.M. Schubert, Antiviral (RNA) activity of selected Amaryllidaceae isoquinoline constituents and synthesis of related substances, Journal of natural products, 55 (1992) 1569-1581.
[318] S.-y. Li, C. Chen, H.-q. Zhang, H.-y. Guo, H. Wang, L. Wang, X. Zhang, S.-n. Hua, J. Yu, P.-g. Xiao, Identification of natural compounds with antiviral activities against SARS-associated coronavirus, Antiviral research, 67 (2005) 18-23.
[319] G. Zou, F. Puig-Basagoiti, B. Zhang, M. Qing, L. Chen, K.W. Pankiewicz, K. Felczak, Z. Yuan, P.Y. Shi, A single-amino acid substitution in West Nile virus 2K peptide between NS4A and NS4B confers resistance to lycorine, a flavivirus inhibitor, Virology, 384 (2009) 242-252.
[320] M. Masi, R. Di Lecce, N. Mérindol, M.-P. Girard, L. Berthoux, I. Desgagné-Penix, V. Calabrò, A. Evidente, Cytotoxicity and Antiviral Properties of Alkaloids Isolated from Pancratium maritimum, Toxins, 14 (2022) 262.
[321] S. Tan, M.G. Banwell, W.C. Ye, P. Lan, L.V. White, The Inhibition of RNA Viruses by Amaryllidaceae Alkaloids: Opportunities for the Development of Broad-Spectrum Anti-Coronavirus Drugs, Chemistry–An Asian Journal, 17 (2022) e202101215.
[322] J.B. Park, Synthesis and characterization of norbelladine, a precursor of Amaryllidaceae alkaloid, as an anti-inflammatory/anti-COX compound, Bioorganic & medicinal chemistry letters, 24 (2014) 5381-5384.
[323] K. Babkova, J. Korabecny, O. Soukup, E. Nepovimova, D. Jun, K. Kuca, Prolyl oligopeptidase and its role in the organism: Attention to the most promising and clinically relevant inhibitors, Future medicinal chemistry, 9 (2017) 1015-1038.
[324] S. Darvesh, D.A. Hopkins, C. Geula, Neurobiology of butyrylcholinesterase, Nature Reviews Neuroscience, 4 (2003) 131-138.
[325] W.L.F. Armarego, Purification of Laboratory Chemicals (Eight Edition), Oxford, 2017.
[326] D. Lamoral-Theys, A. Andolfi, G. Van Goietsenoven, A. Cimmino, B. Le Calvé, N. Wauthoz, V. Mégalizzi, T. Gras, C. Bruyère, J. Dubois, Lycorine, the main phenanthridine Amaryllidaceae alkaloid, exhibits significant antitumor activity in cancer cells that display resistance to proapoptotic stimuli: an investigation of structure− activity relationship and mechanistic insight, Journal of medicinal chemistry, 52 (2009) 6244-6256.
[327] W. Fischl, R. Bartenschlager, High-throughput screening using dengue virus reporter genomes, Antiviral methods and protocols, (2013) 205-219.
[328] J. He, Y. Chen, M. Farzan, H. Choe, A. Ohagen, S. Gartner, J. Busciglio, X. Yang, W. Hofmann, W. Newman, C.R. Mackay, J. Sodroski, D. Gabuzda, CCR3 and CCR5 are co-receptors for HIV-1 infection of microglia, Nature, 385 (1997) 645-649.
[329] L. Chatel-Chaix, R. Bartenschlager, Dengue virus-and hepatitis C virus-induced replication and assembly compartments: the enemy inside—caught in the web, Journal of virology, 88 (2014) 5907-5911.

[330] N. Merindol, M. El-Far, M. Sylla, N. Masroori, C. Dufour, J.-x. Li, P. Cherry, M.B. Plourde, C. Tremblay, L. Berthoux, HIV-1 capsids from B27/B57+ elite controllers escape Mx2 but are targeted by TRIM5α, leading to the induction of an antiviral state, PLoS pathogens, 14 (2018) e1007398.
[331] V. Summa, A. Petrocchi, F. Bonelli, B. Crescenzi, M. Donghi, M. Ferrara, F. Fiore, C. Gardelli, O. Gonzalez Paz, D.J. Hazuda, Discovery of raltegravir, a potent, selective orally bioavailable HIV-integrase inhibitor for the treatment of HIV-AIDS infection, Journal of medicinal chemistry, 51 (2008) 5843-5855.
[332] S. Salentin, S. Schreiber, V.J. Haupt, M.F. Adasme, M. Schroeder, PLIP: fully automated protein–ligand interaction profiler, Nucleic acids research, 43 (2015) W443-W447.
[333] A. Singh, M.-A. Massicotte, A. Garand, L. Tousignant, V. Ouellette, G. Bérubé, I. Desgagné-Penix, Cloning and characterization of norbelladine synthase catalyzing the first committed reaction in Amaryllidaceae alkaloid biosynthesis, BMC plant biology, 18 (2018) 1-12.
[334] J.J. Nair, J. Bastida, J. van Staden, In vivo cytotoxicity studies of Amaryllidaceae alkaloids, Natural Product Communications, 11 (2016) 1934578X1601100134.
[335] Z. Cao, P. Yang, Q. Zhou, Multiple biological functions and pharmacological effects of lycorine, Science China Chemistry, 56 (2013) 1382-1391.
[336] X. Chen, I.G. Tikhonova, M. Decker, Probing the mid-gorge of cholinesterases with spacer-modified bivalent quinazolinimines leads to highly potent and selective butyrylcholinesterase inhibitors, Bioorganic & medicinal chemistry, 19 (2011) 1222-1235.
[337] I.R. Macdonald, E. Martin, T.L. Rosenberry, S. Darvesh, Probing the peripheral site of human butyrylcholinesterase, Biochemistry, 51 (2012) 7046-7053.
[338] M. Bajda, A. Więckowska, M. Hebda, N. Guzior, C.A. Sotriffer, B. Malawska, Structure-based search for new inhibitors of cholinesterases, International journal of molecular sciences, 14 (2013) 5608-5632.
[339] Z. Zhang, F. Fan, W. Luo, Y. Zhao, C. Wang, Molecular dynamics revealing a detour-forward release mechanism of tacrine: implication for the specific binding characteristics in butyrylcholinesterase, Frontiers in Chemistry, 8 (2020) 730.
[340] J.D. Mann, H.M. Fales, S.H. Mudd, Alkaloids and Plant Metabolism: VI. O-methylation in vitro of norbelladine, a precursor of amaryllidaceae alkaloids, Journal of Biological Chemistry, 238 (1963) 3820-3823.
[341] A.A. Mamun, F. Pidany, D. Hulcova, J. Marikova, T. Kucera, M. Schmidt, M.C. Catapano, M. Hrabinova, D. Jun, L. Muckova, J. Kunes, J. Janousek, R. Andrys, L. Novakova, R. Perinova, N. Maafi, O. Soukup, J. Korabecny, L. Cahlikova, Amaryllidaceae Alkaloids of Norbelladine-Type as Inspiration for Development of Highly Selective Butyrylcholinesterase Inhibitors: Synthesis, Biological Activity Evaluation, and Docking Studies, Int J Mol Sci, 22 (2021).
[342] R. Zhao, M. Wang, J. Cao, J. Shen, X. Zhou, D. Wang, J. Cao, Flavivirus: From Structure to Therapeutics Development, Life, 11 (2021) 615.
[343] M.S. Mustafa, V. Rasotgi, S. Jain, V. Gupta, Discovery of fifth serotype of dengue virus (DENV-5): A new public health dilemma in dengue control, Medical Journal Armed Forces India, 71 (2015) 67-70.
[344] T. Siatka, M. Adamcova, L. Opletal, L. Cahlikova, D. Jun, M. Hrabinova, J. Kunes, J. Chlebek, Cholinesterase and Prolyl Oligopeptidase Inhibitory Activities of Alkaloids from Argemone platyceras (Papaveraceae), Molecules, 22 (2017).
[345] N.H. Greig, D.K. Lahiri, K. Sambamurti, Butyrylcholinesterase: an important new target in Alzheimer's disease therapy, International psychogeriatrics, 14 (2002) 77-91.
[346] S. Darvesh, M.K. Cash, G.A. Reid, E. Martin, A. Mitnitski, C. Geula, Butyrylcholinesterase is associated with β-amyloid plaques in the transgenic APPSWE/PSEN1dE9 mouse model of Alzheimer disease, Journal of Neuropathology & Experimental Neurology, 71 (2012) 2-14.
[347] O. Lockridge, C.F. Bartels, T.A. Vaughan, C.K. Wong, S.E. Norton, L.L. Johnson, Complete amino acid sequence of human serum cholinesterase, Journal of Biological Chemistry, 262 (1987) 549-557.

[348] D.C. Vellom, Z. Radic, Y. Li, N.A. Pickering, S. Camp, P. Taylor, Amino acid residues controlling acetylcholinesterase and butyrylcholinesterase specificity, Biochemistry, 32 (1993) 12-17.
[349] P. Masson, W. Xie, M.-T. Froment, V. Levitsky, P.-L. Fortier, C. Albaret, O. Lockridge, Interaction between the peripheral site residues of human butyrylcholinesterase, D70 and Y332, in binding and hydrolysis of substrates, Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology, 1433 (1999) 281-293.
[350] F. Nachon, E. Carletti, C. Ronco, M. Trovaslet, Y. Nicolet, L. Jean, P.-Y. Renard, Crystal structures of human cholinesterases in complex with huprine W and tacrine: elements of specificity for anti-Alzheimer's drugs targeting acetyl-and butyryl-cholinesterase, Biochemical Journal, 453 (2013) 393-399.