

Elucidating the enzyme network driving Amaryllidaceae Alkaloids Biosynthesis

Isabel Desgagne-Penix

Isabel.Desgagne-Penix@uqtr.ca

Université du Québec à Trois-Rivières <https://orcid.org/0000-0002-4355-5503>

Basanta Lamichhane

Université du Québec à Trois-Rivières

Sarah-Eve Gélinas

Université du Québec à Trois-Rivières

Natacha Merindol

Université du Québec à Trois-Rivières

Manoj Koirala

Université du Québec à Trois-Rivières

Karen Cristine Gonçalves dos Santos

Université du Québec à Trois-Rivières

Hugo Germain

Université du Québec à Trois-Rivières

Article

Keywords: Amaryllidaceae alkaloids, cytochrome P450, metabolic engineering, phenol coupling, biosynthetic pathway, biotechnological applications, pharmaceuticals, enzyme characterization

Posted Date: September 9th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-4803623/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Version of Record: A version of this preprint was published at Plant Biotechnology Journal on March 4th, 2025. See the published version at <https://doi.org/10.1111/pbi.70026>.

Abstract

Amaryllidaceae alkaloids (AAs) are diverse bioactive metabolites with significant pharmaceutical potential, derived from 4'-O-methylnorbelladine (4'OM). The biosynthesis of these compounds involves the condensation of tyramine and 3,4-dihydroxybenzaldehyde by norbelladine synthase (NBS) and/or noroxomaritidine/norcrocusodine reductase (NR), followed by O-methylation. Cytochrome P450 enzymes, particularly the CYP96T family, introduce further structural diversity through C-C couplings, resulting in lycorine, galanthamine, and crinine cores. Despite their importance, the exact biosynthetic pathways remain poorly defined. In this study, we characterized key enzymes from *Leucojum aestivum* (La), providing crucial insight into AA biosynthesis. Transient expression in *Nicotiana benthamiana* demonstrated that LaNBS and LaNRII catalyze the conversion of tyramine and 3,4-dihydroxybenzaldehyde to norbelladine, which is subsequently O-methylated by a norbelladine-4'-O-methyltransferase (LaN4'OMT). Co-agroinfiltration of LaNBS, LaNRII, LaN4'OMT, and LaCYP96T1 resulted in the production of various phenol-coupled products, with lycorine as the predominant compound, alongside with haemanthamine, crinine/vittatine, and norgalanthamine. This study identifies LaCYP96T1 and LaCYP96T2 as the first monocot enzymes capable of catalyzing all three regioselective C-C phenol couplings, highlighting the substrate promiscuity of LaNRII. The findings not only elucidate critical steps in AA biosynthesis but also open new avenues for biotechnological application in producing valuable alkaloids, offering potential for novel drug development.

1. Introduction

Amaryllidoideae from the *Amaryllidaceae* family produce a class of diverse specialized metabolites (SMs) with potent biological and therapeutic effects, known as the *Amaryllidaceae* alkaloids (AAs). Although galanthamine is the only AA commercialized as a treatment of mild symptoms of Alzheimer's disease owing to its acetylcholinesterase inhibitory properties, many other AAs have potential pharmacological applications. For instance, lycorine is extensively studied for its strong antiviral activities, and haemanthamine is known for its anticancer properties ^{1, 2}.

However, AAs potential application is hindered by their limited commercial availability ³. On the one hand, the current supply relies on AAs extracted from wild plants, such as *Leucojum aestivum* (summer snowflake), with varying yields ⁴. On the other hand, total chemical synthesis is challenging due to the intricate carbon ring structures of these metabolites, compounded by environmentally unsustainable purification processes. *In vitro* cultures of galanthamine-producing species, primarily *Narcissus confusus* and *L. aestivum*, have gained attention as an alternative strategy ^{5, 6, 7, 8} but are not widely used to date due to the resulting limited accumulation of AAs.

Metabolic engineering presents a promising alternative avenue for future AA production, contingent upon a thorough understanding of the biosynthetic pathway. Yet, while our comprehension of the chemistry and pharmacological properties of AAs is rapidly advancing, there remains a scarcity of

knowledge regarding their synthesis ^{3, 9, 10, 11}. Only a few enzymes/genes have been discovered and characterized.

Roughly, the biosynthesis of AAs is governed by an extensive network of genes that code for enzymes, orchestrating a cascade of biochemical transformations starting from basic amino acid precursors to yield the structurally complex alkaloids characteristic of *Amaryllidoideae* (Fig. 1). Specifically, from the plastidic aromatic amino acids pathway, L-tyrosine undergoes decarboxylation to yield tyramine ¹², while L-phenylalanine, is converted to 3,4-dihydroxybenzaldehyde (3,4-DHBA) most likely through the phenylpropanoid pathway ^{8, 10}. Previous *in vitro* studies using norbelladine/norcaugsodine synthase (NBS) and noroxomaritidine/norcaugsodine reductase (NR) from *Narcissus pseudonarcissus*, *N. papyraceus* and *L. aestivum* suggested that the condensation of tyramine and 3,4-DHBA yields norcaugsodine, further reduced to form norbelladine (Fig. 1) ^{13, 14, 15, 16, 17}. A 4'-O-methyltransferase, identified in *Narcissus* sp. aff. *pseudonarcissus*, *Lycoris aurea* and *Lycoris radiata*, catalyzes the methylation of norbelladine *in vitro* into the metabolic intermediate of AAs, 4'-O-methylnorbelladine (4'OM) ^{18, 19, 20}. Radiolabeling studies in *L. aestivum* revealed that deuterium-labelled 4'OM was further incorporated into galanthamine and lycorine through a C-C coupling cyclization step, as observed in the biosynthesis pathway of berberine-type alkaloids ^{21, 22, 23}. Later studies uncovered that a new class of cytochrome P450, CYP96T1 catalyzes mainly *para-para'* (*p-p'*) C-C oxidative coupling of 4'OM in *N. sp. aff. pseudonarcissus* but also yields the *para-ortho'* (*p-o'*) product *N*-demethylnarwedine as a minor product ²⁴. Several homologs identified from *Narcissus* cv. Tête-à-Tête were suggested to exhibit specific catalytic activity in either *o-p'*, *p-o'* or *p-p'* coupling reactions in a preliminary manuscript ²⁵, but *Amaryllidoideae* CYP96T functionalities remain unclear. The reconstitution of the pathway has not been attempted yet and it is unclear whether assembling the four proposed enzymes (NBS, NR, OMT and CYP96T) would lead to the production of therapeutic AAs in a heterologous host.

Nicotiana benthamiana contains all necessary cofactors, intracellular compartments along with protein-folding mechanisms essential for efficient synthesis of plant proteins ²⁶. It has been extensively used to investigate plant metabolic pathways, from characterization of enzymes to elucidation of pathway, such as the monoterpene indole alkaloids ^{27, 28}, strictosidine ²⁹, or colchicine ³⁰, to name but a few.

In this study, we reconstituted the biosynthetic pathways of AAs through heterologous expression in *N. benthamiana* using *Agrobacterium*-mediated transient T-DNA expression. By leveraging the genetic tractability of *N. benthamiana*, we hypothesize that it is possible to produce therapeutically valuable AAs. This research provides insights into the complex biosynthetic pathways of these alkaloids and facilitates the development of a metabolic engineering platform.

2. Materials And Methods

2.1 Phylogenetic analysis

The candidate transcripts of *L. aestivum* were aligned against PR10/BetIV family proteins (in the case of *LaNBS*), short chain dehydrogenase/reductases (for *LaNR*), *O*-methyltransferases (for *LaN4'OMT*) and cytochrome P450 proteins (for *LaCYP96T*) obtained from UniProt, SwissProt, or the National Center for Biotechnology Information (NCBI) databases (accession numbers available in Supplementary Table S2) using ClustalW³¹ (default parameters). Catalytic sites and conserved motifs were identified on basis of published literature or NCBI conserved domain search tool or InterPro database following the alignment³². Phylogenetic trees were constructed using ModelRevelator software tool "Iqtree" with 1000 ultrafast bootstraps, 1000 branch test, as well as 1000 iterations with minimum correlation coefficient 0.99, followed by standard IQ tree search parameters^{33, 34, 35}. ModelRevelator is a model selection tool underpinned by two deep neural networks. The models resulting from neural network estimates are closer to the ground truth than maximum likelihood trees³⁶. Ete-view software was used for tree visualization³⁷.

2.2 Molecular modelling

Structural prediction of *L. aestivum* enzymes including norbelladine 4-*O*-methyltransferase (*LaN4'OMT*), norbelladine/norcaugsodine synthase (*LaNBS*), noroxomaritidine/norcaugsodine reductase (*LaNR*), and cytochrome P450 *LaCyp96T1* and cytochrome P450 *LaCyp96T2* was performed using CollabFold v1.5.3 (AlphaFold2 via MMseqs2)³⁸. Models were visualized and superimposed using Pymol 3.0.0 (Shrödinger). The MOE2022.09 software (Chemical Computing Group) was employed for analyzing model conformations and preparing receptors for docking, as outlined in¹⁴. Pocket residues of predicted enzymes were superimposed with reference crystal structures in MOE. Crystal structures of *Rattus norvegicus* and *Homo sapiens* CatOMTs (PDB:1H1D³⁹ and 3A7E, respectively) were used as references for *LaN4'OMT*, norcoclaurine synthase from *Thalictrum flavum* for *LaNBS* (PDB:2VQ5;⁴⁰), *N. pseudonarcissus* King Alfred noroxomaritidine/norcaugsodine reductase for *LaNR* (PDB:5FF9;⁴¹), and *Priestia megaterium* CYP102A1 (PDB:6JMH;⁴²) and *Mycobacterium tuberculosis* CYP121 (PDB:3G5H;⁴³) for *LaCYP96T1* and for *LaCYP96T2*. Methyl donor S-adenosylmethionine (SAM for *LaN4'OMT*), metal ion (Mg²⁺ for *LaN4'OMT*), cofactor nicotinamide adenine dinucleotide phosphate (NADPH for *LaNR*), as well as heme and two vicinal H₂O (*LaCyp96T1* and 2) were included in the active sites at coordinates of orthologous crystal structures. Structures underwent preparation including correcting issues, capping, charging termini, selecting appropriate alternate, protonation, calculating optimal hydrogen position and charges, and an energy minimization step with tethered active site residues and fixed NADPH, H₂O and heme, SAM and Mg²⁺, in corresponding enzymes, before docking. The docking site was predicted based on residues in interaction with ligands of orthologous crystal structures, as described in¹⁴, or using dummies atoms added to the receptor with MOE Site Finder tool. The docking site included Mg²⁺ and SAM, H₂O and heme, or NADPH, where applicable. Ligands (3,4-DHBA, tyramine (*LaNBS*), norcaugsodine (*LaNR*), norbelladine (*LaN4'OMT*), and 4-*O*-methylnorbelladine (*LaCYP96T1* and *LaCYP96T2*)) isomeric smiles codes were retrieved from PubChem and submitted to the ZINC20

database to download 3D data files ⁴⁴. Every protomer predicted at pH=7 was included as a possible ligand.

The AMBER10:EHT force field was employed. Triangle Matcher was used as the placement method for 200 poses and flexible docking was performed with tethered induced fit refinement, yielding 10 best poses. Docking poses were analyzed by comparison with crystalized protein-ligand complexes, and the first pose (with the best docking score) coherent with crystalized enzymes ligand positioning and with the catalyzed reaction was selected for each ligand. PLIP was utilized to analyze the interactions between ligands and receptor residues ⁴⁵, and the images were further processed using PyMOL along with prediction of hydrogen bonds between the metal ion or the water molecules and the docked ligand, where applicable.

2.3 RNA extraction and cDNA synthesis

RNA was extracted from 80mg of snap frozen *L. aestivum* plant leaves using the RNeasy® Plant Mini Kit (Qiagen) using manufacturers protocol escaping the optional steps. DNA synthesis was performed using the SensiFAST cDNA synthesis kit by Meridian Bioscience. Briefly, SensiFAST cDNA Synthesis Mix containing random hexamers and anchored oligo-dT primers was mixed with RNA followed by the addition of reverse transcriptase enzyme and reaction buffer, and incubated at 42°C. After the reaction, the enzyme was inactivated by heating at 85°C for 5 minutes. The resulting cDNA was then purified and stored at -20°C or -80°C.

2.4 PCR amplification of key candidate genes *LaNBS*, *LaNR11*, *LaN4'OMT*, *LaCYP96T1* and *LaCYP96T2*.

The coding sequences of norbelladine synthase (*LaNBS*), noroxomaritidine reductase (*LaNR11*), 4' *O*-methyltransferase (*LaN4'OMT*) and cytochrome P450 (*LaCYP96T1* and *LaCYP96T2*) of *L. aestivum* were amplified using primers designed with *attB1* and *attB2* overhangs for Gateway cloning in expression vector pK7WG2 ⁴⁶. Primers for *LaNBS*, *LaNR11*, *LaN4'OMT*, and *LaCYP96T1/2* were designed for expression with x3 AU1, x3 FLAG, x3 AU5, and x3 Myc tag at their C-terminal, respectively, tag sequences were included in the reverse primers (Supplementary Table S3). Primers for localization study of the candidate gene *LaN4'OMT* were designed for cloning into expression vectors pB7WGF2 (N-terminal EGFP fusion) and pB7FWG2 (C-terminal EGFP fusion) vectors, while *LaCYP96T2* primers were designed to clone into pB7FWG2 (C-terminal EGFP fusion) (Supplementary Table S3). PCR amplifications were conducted PrimeStar GXL premix (TaKaRa Bio) in T100 Thermal cycler of Bio-Rad. Amplification was checked by running 2µl of PCR product in 0.9% agarose gel and images were taken in Bio-Rad ChemiDoc Imaging system. PCR clean-up and gel purification was done according to manufacturer's protocol by GeneAid-PCR & Gel Kit.

2.5 Cloning and *Agrobacterium* transformation

Amplified PCR fragments were cloned with Gateway cloning, performed according to manufacturer's protocol (Thermo Fisher scientific). Briefly, the Gateway-adapted *attP*-flanked pDONR 221 vectors with kanamycin resistance gene was used for BP recombination reaction to generate an *attL*-flanked entry clone with *attB*-flanked-*LaNBS*, *-LaNRll*, *-LaN4'OMT*, *-LaCYP96T1*, *-LaCYP96T2*, *-LaN4'OMT* (C-terminal EGFP fusion), *-LaN4'OMT* (N-terminal EGFP fusion) and *-LaCYP96T2* (C-terminal EGFP fusion) DNA fragment. These entry vectors were transformed into *E. coli* DH5α chemically competent cells, subsequently grown on a kanamycin (50 µg/ml) LB agar plate. Following colony PCR, miniprep of positive clones were performed according to the manufacturer's protocol of Presto Mini Plasmid Kit from Geneaid. A positive entry clone of each construct was sent for Sanger sequencing (Genome Quebec). Once integrity was confirmed, the selected plasmids were incubated overnight with LR reagents to perform a LR recombination reaction between an *attL* containing entry clone and *attR*-containing destination vectors (pK7WG2, pB7FWG2 and pB7WGF2 vectors). The LR reaction mix was transformed into *E. coli* DH5α chemically competent cells and grown in spectinomycin (50 µg/ml) agar plates. The positive clones were confirmed by colony PCR, and plasmid extraction was done as mentioned above, growing positive colonies in 5 ml LB broth with spectinomycin (50 µg/ml). These expression clones were transformed in *Agrobacterium* GV3101. After 48 h of incubation in LB plates at 28 °C, the presence of insert in the construct was once again confirmed with colony PCR.

2.6 Transient expression of candidate genes in *N. benthamiana* (in vivo assay)

Individual *Agrobacterium* GV3101 colonies, carrying the desired gene constructs (*LaNBS*, *LaNR*, and *LaN4'OMT*, *LaCYP96T1* and *LaCYP96T2*), were streaked on LB plates with 50 µg/ml spectinomycin, 50 µg/ml rifampicin and 30 µg/ml gentamicin, and grown at 28 °C for 2 days. With the cell scraper, colonies were scraped from plate, resuspended in *Agrobacterium* induction buffer (10 mM MES, pH 5.6, 10 mM MgCl₂, and 150 µM acetosyringone) and incubated for 2 hours at room temperature. 4–5-week-old *N. benthamiana* grown in growth chamber with 22 °C temperature and 16 h of light/8 h of dark conditions were used for the experiments. *N. benthamiana* leaves were infiltrated or co-infiltrated with *Agrobacterium* suspensions of 0.6 of optical density at 600 nm (OD₆₀₀) on the leaf abaxial side using a needleless syringe. Unless otherwise mentioned, three leaves each of three independently infiltrated plants of each experimental condition were used to test the hypothesis. At first, *Agrobacterium* suspensions harboring the genes were infiltrated in the plant leaves and incubated for 48 hours. In the second step, leaves were infiltrated with 100 µM substrates. Following the incubation for 24 hours, the leaves were harvested, snap-frozen and preserved at -80 °C for further analysis. The snap-frozen samples were crushed into powder using 5-mm diameter stainless steel beads, shaking at 25 Hz for 1.5 min in TissueLyser (Qiagen). Expression of agroinfiltrated enzymes in *N. benthamiana* leaves was verified by western blot. Briefly, 100 mg of tissue was resuspended in 0.5 ml of lysis buffer (100 mM Tris-

HCl, pH 7.5, 1 mM phenylmethylsulfonyl fluoride, 10% glycerol, 10 mM β -mercaptoethanol and 2% polyvinylpolypyrrolidone) and incubated for 30 min on ice with periodic, gentle inversion. Extracts were centrifuged at 12 000 *g* at 4°C for 30 min to pellet debris, after which the supernatant were collected in 3 tubes and preserved at -80°C for future experiments. Specific primary antibodies were used to verify expression on PVDF membranes after semi-dry transfer and blocking by 5% milk overnight at 4 °C. For *LaNBS* and *LaN4' OMT*, goat anti-AU1 and anti-AU2 polyclonal antibodies from Thermo Fisher Scientific were employed, respectively. A rabbit anti-FLAG monoclonal antibody from Cell Signaling Technology was used for *LaNR11*, while an anti-Myc polyclonal antibody raised in rabbit from Millipore Sigma Aldrich was applied for *LaCYP96T1* and *T2*. For the *LaN4' OMT-EGFP* fusion protein and *LaCYP96T2-EGFP* fusion protein, a primary anti-GFP/CFP/YFP monoclonal antibody raised in mouse was used. Host specific secondary antibodies conjugated with Horseradish peroxidase (HRP) were utilized to detect the bands in the membrane i.e., Immun-Star goat anti-mouse (GAM)-HRP conjugate and Immun-Star goat anti-rabbit (GAR)-HRP conjugate from BioRad, and anti-goat-HRP from Jakson Immuno Research Laboratories.

2.7 Sample preparation for metabolite analysis

Metabolites were extracted from crushed snap-frozen samples using methanol-water extraction solvent (8:2, *v/v*) with a constant ratio of 20 μ L/mg of the samples. A concentration of 2 mg/L of papaverine was added as internal standard prior extraction. Samples were placed on a shaker at 200 rpm overnight and then heated at 65°C for 20 minutes with closed lid. Supernatants were collected after centrifuging at 10000 rpm for 10 minutes and dried using a rotary evaporator. Sample extracts were reconstituted in the mobile phase (Milli-Q water and methanol, both containing formic acid 0.1% (90:10)) at a constant ratio of 20 μ L/mg unless stated, vortexed and filtered through 0.22 μ m PTFE filter prior analysis conducted by HPLC-MS/MS.

2.8 Instrumentation and chromatographic conditions

High-performance liquid chromatography coupled with tandem mass spectrometry (HPLC-MS/MS) (Agilent Technologies, Santa Clara, California, USA) system equipped with an Agilent Jet Stream ionization source, a binary pump, an autosampler and a column compartment were used for the analysis. Compounds separation was achieved using a Kinetex EVO C18 column (150 $\times$ 4.6 mm, 5 mm, 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, 10% B; 10.0 min, 10% B; 20.0 min, 100% B; 25.0 min, 100% B; 26.0 min, 10% B. The total run time was 30 min per sample to allow the reconditioning of the column prior to 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) software was used to control the HPLC-MS/MS. MassHunter Qualitative Analysis (version 10.0) and MassHunter Quantitative QQQ Analysis (version 10.0) softwares were used for data processing. Compounds identification was performed using authentic reference standards or by employing collision-induced dissociation (CID) mass spectra published in literature^{24, 25, 41} in cases where reference standards were not available. MRM transitions and instrument parameters used for identifications in *N4'OMT* experiments and in *CYP96* experiments are included in Supplementary Table S4 and Supplementary Table S5 respectively. Instrument parameters used in Product Ion analysis mode to obtain the full CID mass spectra of signals detected in MRM mode are presented in Supplementary Table S6.

2.9 Subcellular localization study

Subcellular localization of *LaN4'OMT* and *LaCYP96T2* was determined using confocal microscopy by infiltrating culture of *A. tumefaciens* (GV3101) harboring pB7FWG2-*LaN4'OMT*, pB7WGF2-*LaN4'OMT*, or pB7FWG2-*LaCYP96T2* into 3-4 weeks old *N. benthamiana* plants. *Agrobacterium* suspensions were prepared and infiltrated as described above. For co-localization assay, co-infiltration of the suspensions aforementioned along with *Agrobacterium* harbouring RFP or ER-mCherry plasmids⁴⁷ were mixed to a final OD₆₀₀ of 0.6, and infiltrated. After 48 hours of infiltration, the abaxial epidermis of leaf discs was placed on a microscopic slide in a water drop, shielded by a cover slip, and photographed immediately. A confocal laser scanning microscope (Leica TCS SP8; Leica Microsystems) with a 40X/1.30 oil immersion objective was used to observe protein localization. Las X software was used to create the composite photos (Leica Microsystems). Colocalization graph was created using florescence intensity.

3. Results

3.1 In silico characteristics of pathway candidates

We initially examined the sequences of the proposed enzymes involved in the pathway of AAs (Fig. 1), focusing on the candidates identified in the *L. aestivum* transcriptome¹⁶. Prospects for norbelladine synthesis included *LaNBS* and *LaNRII*. *LaN4'OMT* was proposed to catalyze the formation of 4'-*O*-methylnorbelladine, and two CYPs candidates, *CYP96T1* and *T2* were identified.

LaNBS belongs to the PR10/Bet v1-like protein superfamily similar to previously characterized norcoclaurine synthase (NCS) (Supplementary Fig. 1a). The *LaNBS* protein has 159 amino acids and shares 31%, 38%, 41%, and 72% similarity with NCS from *Papaver somniferum*, *T. flavum* subsp. *glaucum*, *Coptis japonica*, and *Nelumbo nucifera*, respectively. Additionally, *LaNBS* sequence and structure included a conserved glycine-rich P-loop, proximal to the entrance of the ligand-binding pocket, a characteristic feature of the Bet v1 protein (Supplementary Fig. 1b and 1c). Furthermore, eight ligand-binding residues were conserved in all Bet v1 transcripts from various species

(Supplementary Fig. 1b), including hydrophobic residues Tyr68, Tyr100 for wider opening of substrate binding tunnel and Lys83 for intercepting the carbonyl group of the substrate. AlphaFold modeling of *LaNBS* revealed a seven-stranded antiparallel β -sheets with two long N- and C-terminal helices surrounding a cleft with a core of hydrophobic residues, confirming its Bet v1-like structure and its resemblance to *TfNCS* (Supplementary Fig. 1c). The similarity in sequence and structures with NCSs that catalyze the condensation of dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA), supports the hypothesis that *LaNBS* performs the condensation of 3,4-DHBA and tyramine (Fig. 1).

The candidate for norcraugsodine and other AAs reduction, *LaNRII* clustered with the tropinone reductase (TR) sub-family of the short-chain dehydrogenase/reductase (SDR) family (Supplementary Fig. 2a). SDRs are NAD- or NADP-dependent oxidoreductases^{48, 49}. *LaNRII* encodes 269 amino acids sharing 92%, 73% and 41% sequence similarity with *N. papyraceus*, *N. pseudonarcissus* and *Datura stramonium* TRs, respectively. *LaNRII* sequence corresponds to a classical SDR having TGXXXGXXG cofactor binding motif in position 29-33 and YXXXK active site motif in 173-177 position. The other conserved domain NNAG is required for stabilization of central β -sheets, and PGxxxT is suggested to play a structural role (Supplementary Fig. 2b and c). The AlphaFold predicted *LaNRII* structure was arranged as a seven-stranded parallel β -sheet embedded between a pair of three α -helices (Supplementary Fig. 2c).

Putative *LaN4' OMT* was identified to cluster with class I OMT encompassing many other previously characterized OMTs from different species like *C. japonica*, *P. somniferum*, *L. radiata* and *Narcissus* sp. (Supplementary Fig. 3a). Class I OMTs are metal-dependent enzymes characterized by a distinctive feature: a Rossmann fold-binding site for S-adenosyl-L-methionine (SAM) as the methyl donor (Supplementary Fig. 3b and c). *LaN4' OMT* comprises 239 amino acids, and displays conserved residues involved in interactions with SAM such as Val55, Glu77, Gly79, Val80, Tyr84, Ala132A, Asp157²⁰. Asp155, Asp181, and Asn182 are responsible for metal ion binding and Lys158 and Asp230 are implicated in the catalysis of the methyl transfer with Asn182 involved in the catalysis of *O*-methylation²⁰ (Supplementary Fig. 3b). The predicted *LaN4' OMT* candidate folded similarly to a catechol OMT, with 7 β -sheets encompassed between α -helices and loops (Supplementary Fig. 3c).

Transcript candidates encoding enzyme involved in the phenol coupling reactions, *LaCYP96T1* and *LaCYP96T2* formed a distinct clade with other CYP96 transcripts, separate from the CYP80, CYP71, and CYP76 subfamilies (Supplementary Fig. 4a and b). CYPs bind their diverse ligands within a buried, hydrophobic active site, accessed through a substrate channel formed by two flexible helices and their connecting loop (Supplementary Fig. 4c). *LaCYP96T1* and *LaCYP96T2* proteins shared 98% identity and exhibited 74% and 73% sequence homology with *N. sp. aff. pseudonarcissus* CYP96T1, respectively, which has been identified to cyclise 4'-*O*-methylnorbelladine²⁴. *LaCYP96T1* and *LaCYP96T2* encodes 512 amino acids long proteins. InterPro domain search suggested that residues 7 to 29 encodes a signal peptide to embed the protein in the membrane, usually anchoring the CYPs at the cytoplasmic surface of the ER. Residues Ala219, Leu220, Thr303, Leu306, Val307, Thr311, Pro382-Ile385 are predicted to be involved in chemical substrate binding pocket along with other 25 residues conserved throughout

various species (Supplementary Fig. 4b and c). The oxygen binding and activation motif (A/G)GX(D/E)TT is substantially different between *LaCYP96T1* and its variants when compared to CYP450 enzymes that catalyze hydroxylation reactions. *LaCYP96T1* and *LaCYP96T2* also contained the heme binding domain (FXXGXRXCXG) that secures the heme in the active site by a strongly preserved thiolate ligand (Cys457) which interacts with two molecular oxygens, with one oxygen molecule oxidizing the substrate and the other being reduced to form water⁵⁰ (Supplementary Fig. 4b). The AlphaFold predicted α -helix-rich shape of *LaCYP96T1* and *LaCYP96T2* was typical of a CYP450 (Supplementary Fig. 4c and d). They contained the N-terminal membrane-anchoring α -helix domain that enables binding to the cytoplasmic surface of the endoplasmic reticulum membrane (at the left in Supplementary Fig. 4c and d), and the heme pocket buried inside the enzyme large cavity (visible in the center of each prediction). Phylogenetic analysis, multiple sequence alignment, and folding prediction of the five candidate transcripts provide strong evidence to support their capability in executing the proposed biosynthetic pathway (Fig. 1).

3.2 Localization assay

We previously showed that *LaNBS* and *LaNR* both localize in the cytoplasm^{14, 16}. To further characterize the AA pathway, we studied the localization of *LaN4' OMT* and *LaCYP96T2* following agroinfiltration in *N. benthamiana*. The red fluorescent protein (RFP), which localizes to the cytoplasm and nucleus, and ER_M-cherry that localizes in endoplasmic reticulum (ER), were used as controls. Both N- and C-terminal *LaN4' OMT*-EGFP accumulated in the cytoplasm (Fig. 2a, Supplementary Fig. 5), while *LaCYP96T2* localized in the ER (Fig. 2b, Supplementary Fig. 5). Expression of the fusion proteins was verified by western blot. Both amino and carboxy terminal tagged EGFP fusion *LaN4' OMT* proteins were observed at around 54 kDa in blot, and C-terminal-EGFP fusion *LaCYP96T2* at around 87kDa consistent with a GFP fusion (Supplementary Fig. 6a and b). These results suggest that *LaN4' OMT* accumulates in the cytosol, the same compartment as *LaNBS* and *LaNR*, while *LaCYP96* localizes to the ER.

3.3 Reconstitution of the foundation steps of the AA pathway to produce 4'-O-methylnorbelladine in vivo.

The proposed pathway entails the condensation of 3,4-DHBA and tyramine by *LaNBS* to yield norcaugsodine, which is subsequently reduced by *LaNR* and *O*-methylated by *LaN4' OMT*, as crucial initial steps (Fig. 1). Having established the candidates' phylogeny and localization, we proceeded to characterize the interaction of enzymes with their suggested ligands, both *in silico* and *in planta*.

Tyramine and 3,4-DHBA were docked into the active site of *LaNBS* with binding scores of -5.1 and -5.5 kcal/mol, respectively. In the docked conformation, tyramine and 3,4-DHBA are positioned facing each other. Tyramine forms hydrogen bonds with Tyr59, Glu71, and Ile39, and 3,4-DHBA with Glu71 and Lys83. Additionally, both ligands engaged in hydrophobic interactions with surrounding residues: tyramine interacts with Ile39, Pro139, Ile143, and Val85, while 3,4-DHBA interacts with Leu104 and Leu140 (Fig.

3a, right panel, Table I). The near-parallel orientation of the aromatic rings of the two ligands, and the carbonyl group of 3,4-DHBA facing tyramine's amine group are reminiscent of dopamine and hydroxybenzaldehyde positioning in *TfNCS*⁵¹. Overall, their pose is consistent with Lys83- and Glu71-assisted condensation to form norcraugsodine. The latter is proposed to be further reduced by the noroxomaritidine/norcraugsodine reductase *LaNRII*.

Norcraugsodine docked with a score of -6.3 kCal/mol in the active site of *LaNRII*, its N-C double bond laid closest to the nicotinamide ring of the electron donor NADPH and to Tyr173, interacting with Leu114, Val161, Val211, Phe214 and Tyr173 of the enzyme active site (Supplementary Table S1, Fig. 3b). This positioning aligns with the proposed NADPH-dependent reduction of norcraugsodine to produce norbelladine. In the next step yielding to 4'OM, norbelladine docked in *LaN4'OMT* active site with a score of -8.8 kCal/mol, forming interactions with Mg²⁺, Lys158, as well as Ile11, Ser52, Asp155, Trp185, Glu200 and Glu201. Mg²⁺ was positioned between the ligand 3' and 4'OH groups, the 4'OH located proximal to the SAM methyl group (Supplementary Table S1, Fig. 3c). The docking results are consistent with the catalysis of a 4'-O-methylation of norbelladine.

To our knowledge, the AA biosynthetic pathway has never been reconstituted *in vivo*. To achieve this, we agroinfiltrated the candidate genes encoding the pathway enzymes into *N. benthamiana*. Forty-eight hours post agroinfiltration, we infiltrated the substrates and harvested the leaves after 24 hours for metabolite analysis. First, we confirmed protein accumulation by western blotting (WB) (Supplementary Fig. 6c-f), showing bands corresponding to the expected sizes (~20 kDa for *LaNBS*, ~30 kDa for *LaNR* and the EGFP control, ~35 kDa for *LaN4'OMT*, and ~60 kDa for *LaCYP96T1* as well as *LaCYP96T2*). Previous *in vitro* studies suggested that norcraugsodine is unstable and the yield of norbelladine synthesis increases when *LaNBS* and *LaNR* interact¹⁴. Therefore, we co-agroinfiltrated *LaNBS* and *LaNRII* with tyramine and 3,4-DHBA (Fig. 3d, Supplementary Fig. 7a). Under these conditions, norbelladine was detected in the metabolite extracts. In contrast, feeding EGFP-infiltrated leaves, with the substrate, feeding substrates to leaves alone, or non-fed *LaNBS* and *LaNRII*-infiltrated leaves did not produce detectable levels of norbelladine. Subsequently, agroinfiltration of *LaN4'OMT* with norbelladine confirmed the production of 4'-O-methylnorbelladine *in planta* (Fig. 3e, Supplementary Fig. 7b). Finally, co-infiltration of the genes encoding the three enzymes *LaNBS*, *LaNRII* and *LaN4'OMT* with tyramine and 3,4-DHBA as substrates led to the production of both norbelladine and 4'-O-methylnorbelladine, which were undetected in the negative controls (Fig. 3f). Overall, these results corroborate the functionality of the proposed biosynthetic pathway.

3.4 Multiple C-C phenol couplings by a single CYP96T

In the subsequent steps of the proposed biosynthetic pathway (Fig. 1), 4'OM undergoes significant transformations, leading to the generation of a diverse array of alkaloid structures catalyzed by CYP96T isoforms (Fig. 4a). The active site of these enzymes includes a heme group and two water molecules, essential for correct ligand positioning and catalysis. 4'OM docked with its 4-OH in the vicinity of the

iron-bonded water molecule (Fig. 4b). The ligand interacts with Leu112, Phe116, Phe124, Leu224, Val307, Val381, Leu495, and Lys109 in CYP96T1, and with Phe116, Phe124, Leu220, Val307, Leu495, and Lys109 in CYP96T2, with the *O*-methylated ring stabilized by aromatic residues (Supplementary Table S1, Fig. 4b, Supplementary Fig. 8a and b). This positioning supports the catalysis of a C-C-phenol coupling reaction, potentially leading to the synthesis of noroxopluviine, *N*-demethylnarwedine, or noroxomaritidine, in both active sites (Fig. 4b).

To experimentally validate *in vivo* enzyme activity, we expressed *LaCYP96T1* or *LaCYP96T2*, and fed 4'-OM as the substrate. Although we did not detect the *o*-*p'* phenol coupled product noroxopluviine, lycorine was intriguingly detected with *LaCYP96T2* and 4'-OM (Fig. 4f). In leaves agroinfiltrated with *LaCYP96T1* or *LaCYP96T2* and 4'-OM as substrate, we also observed the *p*-*o'* coupled product *N*-demethylnarwedine (Fig. 4g, Supplementary Fig. 9a and b), which was absent in all five negative controls. The accumulation of *N*-demethylnarwedine was nearly twice as high with *LaCYP96T1* compared to *LaCYP96T2* ($p < 0.0001$, Tukey's multiple comparison post hoc test, one way ANOVA). Additionally, we detected the *p*-*p'* phenol coupled product noroxomaritidine and its reduced form oxomaritinamine in leaves agroinfiltrated with *LaCYP96T1* or *LaCYP96T2* gene and 4'-OM (Fig. 4h, Supplementary Fig. 10a and b, Supplementary Fig. 11a and b), whereas these products were undetectable in leaves agroinfiltrated with *EGFP* and 4'-OM. Therefore, *LaCYP96T1* and *LaCYP96T2* catalyzed the synthesis of the three types of phenol coupled products, namely, *p*-*o'* (galanthamine-type), *p*-*p'* (crinine-type) and *o*-*p'* (lycorine-type) involved in AAs biosynthesis pathway *in planta*.

3.5 Multiple reduction steps by LaNR1I.

The reduction step is key for stabilizing metabolites and producing end-of-pathway AAs. We hypothesized that *LaNR1I* could accept *N*-demethylnarwedine, noroxopluviine and noroxomaritidine as substrates. Docking results supported this hypothesis, showing noroxopluviine and *N*-demethylnarwedine positioned to reduce their C-O double bonds through interaction with Tyr173 near the nicotinamide ring of NADPH at the enzyme's active site (Supplementary Table S1, Fig. 4c and d, Supplementary Fig. 8d and c). Noroxomaritidine docked in two distinct orientations, favoring the reduction of the C-O or the C-C double bond, in interaction with Tyr73 (Supplementary Table S1, Fig. 4e, Supplementary Fig. 8e and f).

To further investigate the pathway reconstitution, we co-expressed *LaCYP96T1* or *LaCYP96T2* with *LaNR1I*, and fed 4'-*O*-methylnorbelladine as substrate. This resulted in the detection of the reduced *o*-*p'* phenol coupled product norpluviine (Fig. 4f). The norpluviine signal was significantly higher in the leaves infiltrated with *LaCYP96T1* and *LaNR1I* (~1.4 times, $p < 0.0001$) compared to *LaCYP96T2* and *LaNR1I* (Fig. 4f, right panel, Supplementary Fig. 12a and b). Norpluviine was not detected in any negative controls or samples lacking the reductase enzyme. Lycorine was also detected in samples infiltrated with either of the *LaCYP96T* and the reductase (Fig. 4f, Supplementary Fig. 13). Interestingly, co-infiltrating *LaCYP96T1* and *LaNR1I* resulted in a decreased accumulation of *N*-demethylnarwedine to the

level of *LaCYP96T2* ($p < 0.0001$, Tukey's multiple comparison post-hoc test, one way ANOVA, Fig. 4g). On the other hand, the yield of *N*-demethylnarwedine rose slightly (by 1.3-fold), but significantly ($p = 0.0105$) when *LaNRII* was co-infiltrated with *LaCYP96T2*.

Co-infiltrating *LaCYP96T1* or *LaCYP96T2* and *LaNRII* also led to an increase in oxomaritinamine by nearly 4- and 6-fold respectively, compared to when infiltrated without *LaNRII* ($p < 0.0001$, Tukey's multiple comparison test by ordinary two-way ANOVA, Fig. 4h). Strikingly, agroinfiltration of *LaCYP96T1* or *LaCYP96T2* with *LaNRII* resulted in a strong signal of normaritidine, which had not been detected in samples with *LaCYP96T1* or *LaCYP96T2* alone (Fig. 4h, Supplementary Fig. 14). Normaritidine accumulated significantly more following the infiltration of *LaCYP96T1* compared to *LaCYP96T2* (1.4-fold, $p < 0.0001$). Noroxomaritidine, oxomaritinamine, and normaritidine were not detected in any negative controls. Thus, in addition to norbelladine biosynthesis, *LaNRII* catalyzes the reduction of all phenol coupled products *in planta*.

3.6 Full pathway reconstitution with four enzymes

To confirm if the four candidates could reconstitute the complete pathway from the aromatic amino acid and the phenylpropanoid precursors, we conducted co- agroinfiltration experiments. We infiltrated *LaNBS*, *LaNRII*, *LaN4'OMT*, *LaCYP96T1*, or *LaCYP96T2* and supplemented with tyramine and 3,4-DHBA. This resulted in the detection of norbelladine, 4' OM, the *o-p'* coupled products norpluviine and lycorine, the *p-p'* products noroxomaritidine, normaritidine, haemanthamine, and vittatine/crinine, and the *p-o'* products *N*-demethylnarwedine and norgalanthamine (Supplementary Table S7). Lycorine was the major product formed by both *LaCYP96T* enzymes, with a yield nearly three times higher for *LaCYP96T1* compared to *LaCYP96T2* (1.16 ± 0.077 ng/mg and 0.41 ± 0.016 ng/mg of biomass, respectively).

While noroxomaritidine and normaritidine were consistently detected in all samples, vittatine and haemanthamine were only confirmed when *LaCYP96T1* was included, with a yield of 4.85 ng/ml. Norgalanthamine was detected, but only confirmed in 1 out of 3 samples for *LaCYP96T1*, and 2 out of 3 samples for *LaCYP96T2* (Supplementary Table S7), indicating that *LaCYP96T2* more efficiently yields norgalanthamine compared to *LaCYP96T1*. In summary, the results support the previous stepwise findings and validate the reconstructed pathway with the four enzymes, as illustrated in Supplementary Fig. 15.

4. Discussion

In this study, we identified and characterized the key enzymes involved in the biosynthesis of the pharmaceutically potent AA using homology-based sequence comparison, phylogenetic reconstruction, subcellular localization studies, docking and *in planta* assays. The synthesis of norbelladine, the first AA of the pathway, is a critical step. The identified *LaNBS* candidate belongs to the pathogenesis-related 10 (PR10)/Bet v1 protein superfamily, which has gained significant attention for its role in condensation

reactions in the initial steps of benzyloisoquinoline alkaloids (BIA) and AA biosynthesis pathways. A well-known member of this superfamily, (*S*)-norcoclaurine synthase (NCS) catalyzes condensation of dopamine and 4-HPAA to produce (*S*)-norcoclaurine, the first committed step in the biosynthesis of BIA such as morphine and codeine^{51, 52}. A truncated NCS gene from *C. japonica* condensed 4-hydroxydihydrocinnamaldehyde (4-HDCA) and dopamine, derived from L -phenylalanine and L -tyrosine respectively, to yield 1-phenethylisoquinoline, a precursor in the colchicine alkaloid biosynthetic pathway^{30, 53}. Similarly, *LaNBS* encodes an enzyme catalyzing the condensation of tyramine and 3,4-DHBA into norbelladine^{15, 16}. Although a norcraugsodine/noroxomaritidine reductase has also been shown to catalyze this step⁴¹, the condensation of tyramine and 3,4-DHBA to yield norbelladine in yeast requires both NBS and NR enzymes¹⁴, suggesting that NBS catalyzes the condensation to form norcraugsodine, which is then reduced to norbelladine by NR. Docking analysis indicates that the condensation of the amine and the aldehyde is facilitated by catalytic residues Tyr59, Glu71 and Lys83, and stabilized by hydrophobic interactions, while norcraugsodine is positioned in the active site of NR to expose its C-N double bond to Tyr173 and to NADPH. Consistent with previous *in vitro* results and with the current docking analysis, we demonstrate here the formation of norbelladine by *LaNBS* and *LaNRII* in *planta*.

Norbelladine must then be 4'-*O*-methylated for further diversification in various AAs (Fig. 1). Some plant OMTs, especially those involved in the synthesis of phenylpropanoids and flavonoids, can methylate a range of substrates, showing versatility in their function⁵⁴. *In vitro* studies in *N. sp. aff. pseudonarcissus* have shown that OMT regiospecifically methylates norbelladine at the 4'-*O* position¹⁸. However, the OMT from *N. papyraceus*, *Galanthus elwesii* and *Lycoris* species are promiscuous, methylating at the 3' or 4' position depending on the ligand^{19, 20, 55, 56}. Docking results suggest that the *LaN4'OMT* candidate identified here likely yields 4'-*O*-methylnorbelladine, with norbelladine 4'-*O* to face the SAM methyl group in interaction with Mg^{2+} , Asp155 and Lys158. *In planta* agroinfiltration confirmed the regiospecificity, yielding 4'-OM as sole product. Co-infiltration of *LaNBS*, *LaNRII*, *LaN4'OMT*, along with supplementation of tyramine and 3,4-DHBA, led to the detection of norbelladine and 4'-OM, reconstituting the foundational steps of the pathway *in vivo*, for the first time.

At this branching point of the pathway, 4'-OM undergoes three possible phenol couplings to generate over 650 AAs. The CYP superfamily plays a key role in metabolic diversification, catalyzing a variety of reactions, including oxidative rearrangement of carbon skeletons, oxidative C-C bond cleavage, dehydration, dealkylation, desaturation, decarboxylation, phenol coupling, ring expansion, ring contraction, formation of allene oxides, *N*-oxygenations, isomerization, and aryl migration reactions⁵⁷. Phenol couplings catalyzed by CYPs in the BIA pathway of several plants like *Gloriosa superba*, *Berberis stolonifera*, *C. japonica*, *P. somniferum*, and *Argemone mexicana* lead to the production of pharmaceutically potent compounds. The discovery of a C-C phenol coupling enzyme (CYP96T1) in *N. sp. aff. pseudonarcissus* in 2016 was a breakthrough in AA biosynthesis²⁴. This enzyme, identified for the first time in a monocot, was responsible for catalyzing the ring closure of 4'-OM,

dramatically favoring the synthesis of the *p-p'* coupled product noroxomaritidine over the *p-o'* product *N*-demethylnarwedine. There is growing evidence that CYPs exhibit significant substrate and product promiscuity⁵⁷. A recent preprint study reported that three distinct isoforms *NtCYP96T1*, *NtCYP96T6* and *NtCYP96T5* from *N.cv. Tête-à-Tête*, yield specific *p-p'*, *p-o'* and *o-p'* products, respectively²⁵. All these isoforms share at least 78% similarity with CYP96T1. Although this study is not yet reviewed, it raises questions about the multifunctionality of CYP96Ts in performing all three steps.

Here, we used *LaCYP96T1* and *LaCYP96T2*, which share 72% identity with CYP96T1 of *N. sp. aff. pseudonarcissus*. They accumulated in the ER membrane, as suggested by the N-terminal membrane-anchoring α -helix domain which hints that they could recruit additional cytosolic metabolic enzymes, such as *LaNBS*, *LaNR* and *LaN4'OMT*, to form a metabolon. Upon feeding 4'OM to *N. benthamiana* leaves, *LaCYP96T1* and *LaCYP96T2* catalyzed the formation of noroxomaritidine and *N*-demethylnarwedine, and CYP96T2 yielded detectable amounts lycorine. These results also demonstrate that both *LaCYP96T1* and *LaCYP96T2* accomplished *o-p'*, *p-o'* and *p-p'* phenol couplings, to yield various metabolites from 4'OM, suggesting that they are promiscuous with regards to the product they can form. While Mehta *et al*/unreviewed manuscript suggested that CYP96T1 Ala308 was key to *p-p'* coupling, while Leu308 was required for the *p-o'* coupling of *NtCYP96T6*, our candidates instead, contain a valine at this position, which is predicted to interact with 4'OM in the docking analysis. This suggests that the residue at position 308 could be key to the regio-promiscuity of the identified *LaCYP96Ts*.

Docking results suggested that *LaNRII*, which previously reduced norcraugsodine to norbelladine, is a promiscuous enzyme, accepting noroxopluiine, *N*-demethylnarwedine, and noroxomaritidine (in two positions) with similar scores, and conserving the key interaction with Tyr173. Co-infiltration of *LaCYP96T1* or *LaCYP96T2* with *LaNRII* supplemented with 4'OM yielded six detectable products from the three coupling branches, namely *p'-o* *N*-demethylnarwedine, the *p-p'* coupled product, noroxomaritidine, oxomaritinamine and normaritidine, together with the *o-p'* phenol coupling norpluviine and lycorine. These results unravel that *NRII* is involved in the synthesis of all three types of phenol coupled products formed in the AA pathway. The ability of *LaNRII* to participate in norbelladine synthesis and in the reduction of all three phenol coupled products at different positions, as reported here, is novel but consistent with the promiscuity reported for other SDRs⁵⁸. This finding also aligns with previous report on *NR* from *N. sp. aff. pseudonarcissus*, which reduced norcraugsodine to norbelladine and noroxomaritidine to oxomaritinamine⁴¹

Finally, whole pathway co-infiltration (*LaNBS*, *LaNRII*, *LaN4'OMT*, and *LaCYP96T1* or *LaCYP96T2*), along with precursors tyramine and 3,4-DHBA, yielded lycorine as the major product, with *LaCYP96T1* being almost three times more efficient than *LaCYP96T2*. While norpluviine likely resulted from noroxopluiine reduction by *LaNRII*, the synthesis of lycorine requires, in theory, more enzymatic steps, such as oxygen bridging and hydroxylation^{10, 59}, possibly carried out by endogenous enzymes in *N. benthamiana* as observed in other studies^{60, 61, 62, 63}.

While only a decrease in *N*-demethylnarwedine levels was detected when co-infiltrating *LaCYP96T1* or *LaCYP96T2* with *LaNRII* and adding 4'OM, norgalanthamine was observed in leaves co-infiltrated with *LaNBS*, *LaNRII*, *LaN4'OMT*, *LaCYP96T1* or *LaCYP96T2*, supplemented with tyramine and 3,4-DHBA, confirming the possible reduction of *N*-demethylnarwedine.

Infiltration of the whole pathway also yielded traces of crinine and haemanthamine when *LaCYP96T1* was included, though these products were not detected with *LaCYP96T2*. These results indicate that *LaCYP96T1* is more efficient at producing lycorine as well as crinine/haemanthamine pathway, whereas *LaCYP96T2* yielded lycorine and norgalanthamine. The proposed biosynthetic pathway for haemanthamine suggests that a CYP71 family enzyme facilitates the methyl bridging, leading to the formation of vittatine/crinine. Subsequently, vittatine/crinine undergoes hydroxylation to produce 11-hydroxyvittatine, followed by *O*-methylation. These transformation steps were likely carried out by endogenous enzymes present in *N. benthamiana*, as observed for the monoterpene indole alkaloid (MIA) pathway reconstitution that lead to strictosidine biosynthesis²⁷.

5. Conclusion

In conclusion, our study elucidates the enzymatic steps involved in the biosynthesis of *Amaryllidaceae* alkaloids, demonstrating that four enzymes, NBS, NR, N4'OMT, and CYP96T, are sufficient to yield the three types of phenol coupling reaction in *N. benthamiana*, ultimately producing the potent antiviral and anticancer alkaloid, lycorine. Furthermore, our results highlight the substrate as well as product promiscuity of *LaNRII*, *LaCYP96T1* and *LaCYP96T2*. The investigation of the residues responsible for this promiscuity may pave the way towards understanding enzyme flexibility, developing novel biocatalysts and enhance our knowledge of evolutionary processes in protein function.. Overall, we provide robust evidence for the synthesis of *Amaryllidaceae* alkaloids *in planta* and deliver new knowledge for metabolic engineering approaches to achieve sustainable production of SMs through synthetic biology. The elucidation of the biosynthetic pathway of *Amaryllidaceae* alkaloids in *L. aestivum* represents a significant advancement in our understanding of plant specialized metabolism.

Declarations

Funding

This research was financially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) award number RGPIN-2021-03218 to I.D-P. This work was also supported by the Canada Research Chair on plant specialized metabolism Award No CRC-2018-00137 to I.D-P. Thanks are extended to the Canadian taxpayers and to the Canadian government for supporting the Canada Research Chairs Program.

Acknowledgement:

We thank Patrick Lagüe from Université Laval for sharing his expertise on molecular docking. We thank Bharat Bhusan Majhi for providing *LaNBS*, and *LaNR11* gene, Fatma Meddeb for timely help in obtaining the lab materials and helpful discussions. Melodie B. Plourde for helping with confocal imaging.

Authors Contribution:

Basanta Lamichhane: Conceptualization, Methodology, Investigation, Formal analysis, Validation, Visualization, Writing- Original draft; Sarah-Eve Gélinas: Methodology, Resources, and Editing, Natacha Merindol: Conceptualization, Methodology, Formal analysis, Project administration, Resources, Writing- Reviewing and Editing; Manoj Koirala: Methodology, Resources; Karen Cristine Goncalves dos Santos: Methodology, Resources; Isabel Desgagné-Penix: Conceptualization, Methodology, Resources, Funding acquisition, Supervision, Writing- Reviewing and Editing.

Declaration of Interest:

The authors declare that there is no situation real, potential, or apparent conflict of interest.

References

1. Cahlíková L, Breiterová K, Opletal L. Chemistry and biological activity of alkaloids from the genus *Lycoris* (Amaryllidaceae). *Molecules* **25**, 4797 (2020).
2. Girard M-P, *et al.* Chemical Synthesis and Biological Activities of Amaryllidaceae Alkaloid Norbelladine Derivatives and Precursors. *Molecules* **27**, 5621 (2022).
3. Takos AM, Rook F. Towards a molecular understanding of the biosynthesis of Amaryllidaceae alkaloids in support of their expanding medical use. *International journal of molecular sciences* **14**, 11713-11741 (2013).
4. Georgievaa L, *et al.* Alkaloid variability in *Leucojum aestivum* from wild populations. *Zeitschrift für Naturforschung C* **62**, 627-635 (2007).
5. Pavlov A, *et al.* Galanthamine production by *Leucojum aestivum* in vitro systems. *Process Biochemistry* **42**, 734-739 (2007).
6. Kohut E, Ördögh M, Jámbor-Benczúr E, Máthé Á. Results with the establishment of in vitro culture of *Leucojum aestivum*. *International Journal of Horticultural Science* **13**, 67-71. (2007).
7. Berkov S, *et al.* Alkaloid synthesis and accumulation in *Leucojum aestivum* in vitro cultures. *Natural Product Communications* **4**, 1934578X0900400328 (2009).
8. Koirala M, Karimzadegan V, Liyanage NS, Mérindol N, Desgagné-Penix I. Biotechnological Approaches to Optimize the Production of Amaryllidaceae Alkaloids. *Biomolecules* **12**, 893 (2022).
9. Kilgore MB, Kutchan TM. The Amaryllidaceae alkaloids: biosynthesis and methods for enzyme discovery. *Phytochemistry reviews* **15**, 317-337 (2016b).
10. Desgagné-Penix I. Biosynthesis of alkaloids in Amaryllidaceae plants: A review. *Phytochemistry Reviews* **20**, 409-431 (2021).

11. Liyanage NS, Awwad F, Gonçalves dos Santos KC, Jayawardena TU, Mérindol N, Desgagné-Penix I. Navigating Amaryllidaceae Alkaloids: Bridging Gaps and Charting Biosynthetic Territories – A Comprehensive Review. *Journal of experimental botany*, (2024).
12. Xu J-J, Fang X, Li C-Y, Yang L, Chen X-Y. General and specialized tyrosine metabolism pathways in plants. *Abiotech* **1**, 97-105 (2020).
13. Liu J, Han L, Li G, Zhang A, Liu X, Zhao M. Transcriptome and metabolome profiling of the medicinal plant *Veratrum mengtzeanum* reveal key components of the alkaloid biosynthesis. *Frontiers in Genetics* **14**, 1023433 (2023).
14. Majhi BB, Gélinas S-E, Mérindol N, Ricard S, Desgagné-Penix I. 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**, 1231809 (2023).
15. Singh A, *et al.* Cloning and characterization of norbelladine synthase catalyzing the first committed reaction in Amaryllidaceae alkaloid biosynthesis. *BMC plant biology* **18**, 1-12 (2018).
16. Tousignant L, Diaz-Garza AM, Majhi BB, Gélinas S-E, Singh A, Desgagne-Penix I. Transcriptome analysis of *Leucojum aestivum* and identification of genes involved in norbelladine biosynthesis. *Planta* **255**, 30 (2022).
17. Hnin SYY, Nakashima Y, Morita H. Substrate Flexibilities of Norbelladine Synthase and Noroxomaritidine/Norcraugsodine Reductase for Hydroxylated and/or Methoxylated Aldehydes. *Chemical and Pharmaceutical Bulletin* **72**, 507-511 (2024).
18. Kilgore MB, *et al.* 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**, e103223 (2014).
19. Sun B, Wang P, Wang R, Li Y, Xu S. Molecular Cloning and Characterization of a meta/para-O-Methyltransferase from *Lycoris aurea*. *International journal of molecular sciences* **19**, 1911 (2018).
20. Li W, Qiao C, Pang J, Zhang G, Luo Y. The versatile O-methyltransferase LrOMT catalyzes multiple O-methylation reactions in amaryllidaceae alkaloids biosynthesis. *International journal of biological macromolecules* **141**, 680-692 (2019).
21. El Tahchy A, *et al.* New method for the study of Amaryllidaceae alkaloid biosynthesis using biotransformation of deuterium-labeled precursor in tissue cultures. *Acta Biochimica Polonica* **57**, (2010).
22. Iwasa K, Kim CW. Biotransformations of protoberberines in cell cultures of *Dicentra spectabilis*. *Phytochemistry* **46**, 1359-1363 (1997).
23. Cui W, *et al.* Biotransformation of phenolic 1-benzyl-N-methyltetrahydroisoquinolines in plant cell cultures followed by LC/NMR, LC/MS, and LC/CD. *Journal of natural products* **70**, 1771-1778 (2007).
24. Kilgore MB, Augustin MM, May GD, Crow JA, Kutchan TM. CYP96T1 of *Narcissus* sp. aff. *Pseudonarcissus* catalyzes formation of the Para-Para'CC phenol couple in the Amaryllidaceae alkaloids. *Frontiers in plant science* **7**, 225 (2016a).

25. Mehta N, Meng Y, Zare R, Kamenetsky-Goldstein R, Sattely E. A developmental gradient reveals biosynthetic pathways to eukaryotic toxins in monocot geophytes. *bioRxiv*, (2023, Preprint).
26. Reed J, Osbourn A. Engineering terpenoid production through transient expression in *Nicotiana benthamiana*. *Plant cell reports* **37**, 1431-1441 (2018).
27. Miettinen K, *et al.* The seco-iridoid pathway from *Catharanthus roseus*. *Nature communications* **5**, 3606 (2014).
28. Boccia M, Grzech D, Lopes AA, O'Connor SE, Caputi L. Directed biosynthesis of new to nature alkaloids in a heterologous *Nicotiana benthamiana* expression host. *Frontiers in Plant Science* **13**, 919443 (2022).
29. Dudley QM, *et al.* Reconstitution of monoterpene indole alkaloid biosynthesis in genome engineered *Nicotiana benthamiana*. *Communications biology* **5**, 949 (2022).
30. Nett RS, Lau W, Sattely ES. Discovery and engineering of colchicine alkaloid biosynthesis. *Nature* **584**, 148-153 (2020).
31. Sievers F, *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* **7**, 539 (2011).
32. Paysan-Lafosse T, *et al.* InterPro in 2022. *Nucleic acids research* **51**, D418-D427 (2023).
33. von Haeseler A, Schmidt HA, Bui MQ, Nguyen LT. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. (2014).
34. Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods* **14**, 587-589 (2017).
35. Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. *Molecular biology and evolution* **35**, 518-522 (2018).
36. Burgstaller-Muehlbacher S, Crotty SM, Schmidt HA, Reden F, Drucks T, von Haeseler A. ModelRevelator: Fast phylogenetic model estimation via deep learning. *Molecular Phylogenetics and Evolution* **188**, 107905 (2023).
37. Huerta-Cepas J, Serra F, Bork P. ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Molecular biology and evolution* **33**, 1635-1638 (2016).
38. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nature methods* **19**, 679-682 (2022).
39. Bonifácio MJ, *et al.* Kinetics and crystal structure of catechol-o-methyltransferase complex with co-substrate and a novel inhibitor with potential therapeutic application. *Molecular pharmacology* **62**, 795-805 (2002).
40. Ilari A, *et al.* Structural basis of enzymatic (S)-norcoclaurine biosynthesis. *Journal of biological chemistry* **284**, 897-904 (2009).
41. Kilgore MB, Holland CK, Jez JM, Kutchan TM. Identification of a noroxomaritidine reductase with Amaryllidaceae alkaloid biosynthesis related activities. *Journal of Biological Chemistry* **291**, 16740-16752 (2016c).

42. Stanfield JK, *et al.* Crystals in Minutes: Instant On-Site Microcrystallisation of Various Flavours of the CYP102A1 (P450BM3) Haem Domain. *Angewandte Chemie International Edition* **59**, 7611-7618 (2020).
43. Belin P, *et al.* Identification and structural basis of the reaction catalyzed by CYP121, an essential cytochrome P450 in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences* **106**, 7426-7431 (2009).
44. Irwin JJ, *et al.* ZINC20—a free ultralarge-scale chemical database for ligand discovery. *Journal of chemical information and modeling* **60**, 6065-6073 (2020).
45. Adasme MF, *et al.* PLIP 2021: Expanding the scope of the protein–ligand interaction profiler to DNA and RNA. *Nucleic acids research* **49**, W530-W534 (2021).
46. Karimi M, Depicker A, Hilson P. Recombinational cloning with plant gateway vectors. *Plant physiology* **145**, 1144-1154 (2007).
47. Nelson BK, Cai X, Nebenführ A. A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. *The Plant Journal* **51**, 1126-1136 (2007).
48. Kallberg Y, Oppermann U, Persson B. Classification of the short-chain dehydrogenase/reductase superfamily using hidden Markov models. *The FEBS journal* **277**, 2375-2386 (2010).
49. Franco G, Marco A, Tofanelli S. Gene Structure Evolution of the Short-Chain Dehydrogenase/Reductase (SDR) Family. *GENES* **14**, (2022).
50. Porter T, Coon M. Cytochrome P-450. Multiplicity of isoforms, substrates, and catalytic and regulatory mechanisms. *Journal of Biological Chemistry* **266**, 13469-13472 (1991).
51. Samanani N, Facchini PJ. Purification and characterization of norcoclaurine synthase: The first committed enzyme in benzyloquinoline alkaloid biosynthesis in plants. *Journal of Biological Chemistry* **277**, 33878-33883 (2002).
52. Samanani N, Liscombe DK, Facchini PJ. Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzyloquinoline alkaloid biosynthesis. *The Plant Journal* **40**, 302-313 (2004).
53. Qiao F, *et al.* Elucidation of the 1-phenethylisoquinoline pathway from an endemic conifer *Cephalotaxus hainanensis*. *Proceedings of the National Academy of Sciences* **120**, e2209339120 (2023).
54. Liscombe DK, Louie GV, Noel JP. Architectures, mechanisms and molecular evolution of natural product methyltransferases. *Natural product reports* **29**, 1238-1250 (2012).
55. Su Y, *et al.* Regioselectivity Inversion of an O-Methyltransferase via Semi-rational Mutagenesis Combined with Metal Ion Substitution. *ChemCatChem* **14**, e202200844 (2022).
56. Koirala M, *et al.* Kinetic and *in silico* structural characterization of norbelladine O-methyltransferase of Amaryllidaceae alkaloids biosynthesis. *Journal of Biological Chemistry* **Accepted - JBC-D-24-00879** (2024).

57. Hansen CC, Nelson DR, Møller BL, Werck-Reichhart D. Plant cytochrome P450 plasticity and evolution. *Molecular Plant* **14**, 1244-1265 (2021).
58. Roth S, Kilgore MB, Kutchan TM, Müller M. Exploiting the Catalytic Diversity of Short-Chain Dehydrogenases/Reductases: Versatile Enzymes from Plants with Extended Imine Substrate Scope. *ChemBioChem* **19**, 1849-1852 (2018).
59. Jayawardena TU, Merindol N, Liyanage NS, Desgagné-Penix I. Unveiling Amaryllidaceae alkaloids: from biosynthesis to antiviral potential—a review. *Natural Product Reports*, (2024).
60. Dong L, *et al.* Characterization of two geraniol synthases from *Valeriana officinalis* and *Lippia dulcis*: similar activity but difference in subcellular localization. *Metabolic engineering* **20**, 198-211 (2013).
61. van Herpen TW, Cankar K, Nogueira M, Bosch D, Bouwmeester HJ, Beekwilder J. *Nicotiana benthamiana* as a production platform for artemisinin precursors. *PLoS One* **5**, e14222 (2010).
62. Caputi L, *et al.* Missing enzymes in the biosynthesis of the anticancer drug vinblastine in Madagascar periwinkle. *Science* **360**, 1235-1239 (2018).
63. Boccia M, Grzech D, Lopes AA, O'Connor SE, Caputi L. Directed Biosynthesis of New to Nature Alkaloids in a Heterologous *Nicotiana benthamiana* Expression Host. *Front Plant Sci* **13**, 919443 (2022).

Figures

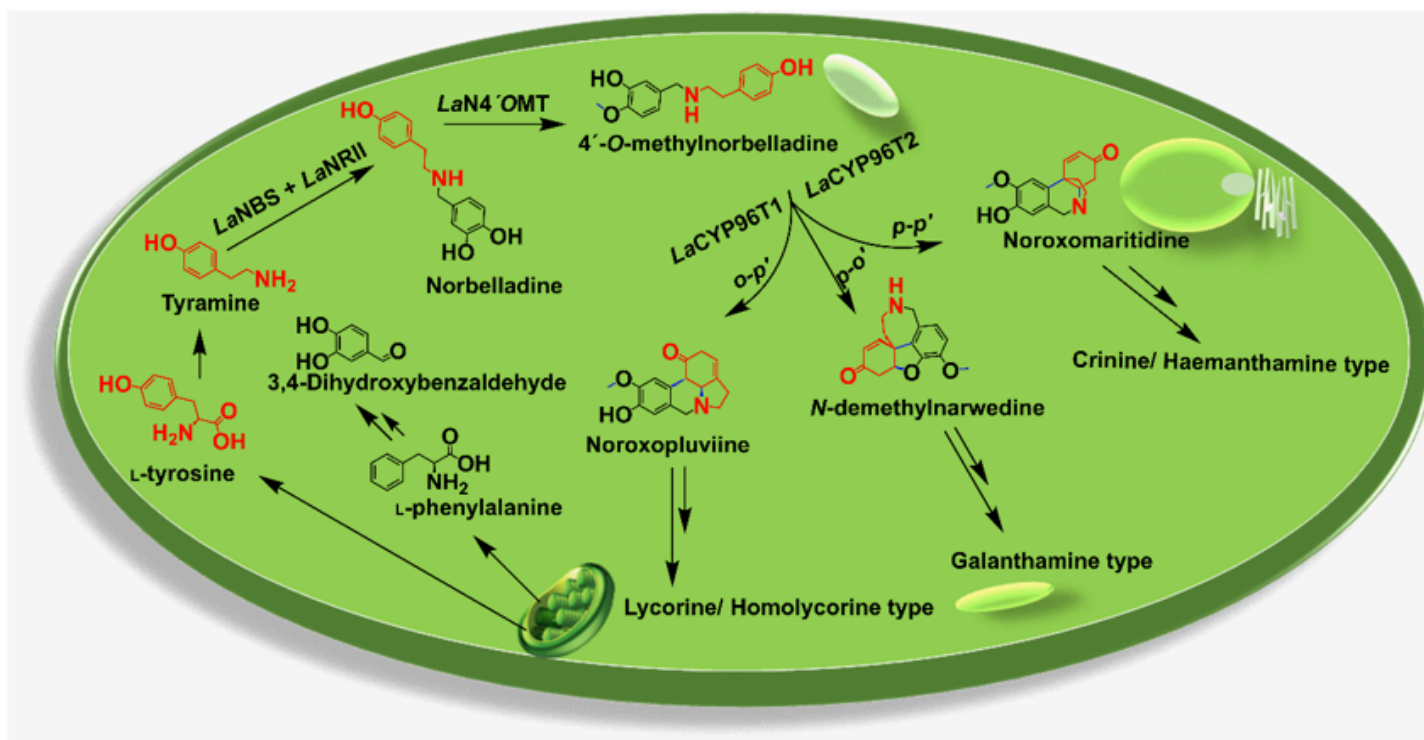


Figure 1

Proposed alkaloid biosynthesis pathway in *Leucojum aestivum*.

The diagram illustrates the hypothesized sequence of biochemical reactions leading to the synthesis of galanthamine-, lycorine/homolycorine- and crinine/haemanthamine-types of alkaloids in *Leucojum aestivum*, highlighting key enzymatic steps and intermediate compounds. *p-p'*: *para-para'* coupling, *p-o'*: *para-ortho'* coupling, *o-p'*: *ortho-para'* coupling.

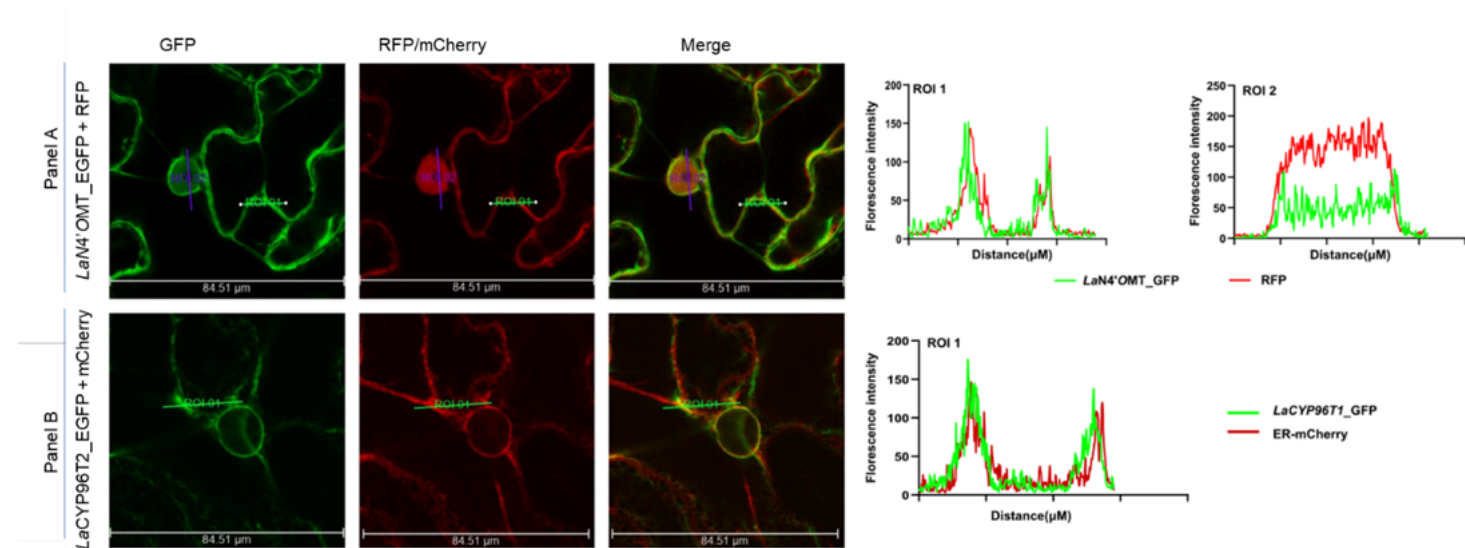


Figure 2

Subcellular localization of *LaN4'OMT* and *LaCYP96T1* proteins in *Nicotiana benthamiana* leaves.

Agroinfiltrated C-terminal-EGFP-fusion-*LaN4'OMT* and C-terminal-EGFP-fusion-*LaCYP96T1*. The histogram of the right panel shows fluorescence intensity profile of respective panel. Scale bars resemble 84.51 μm.

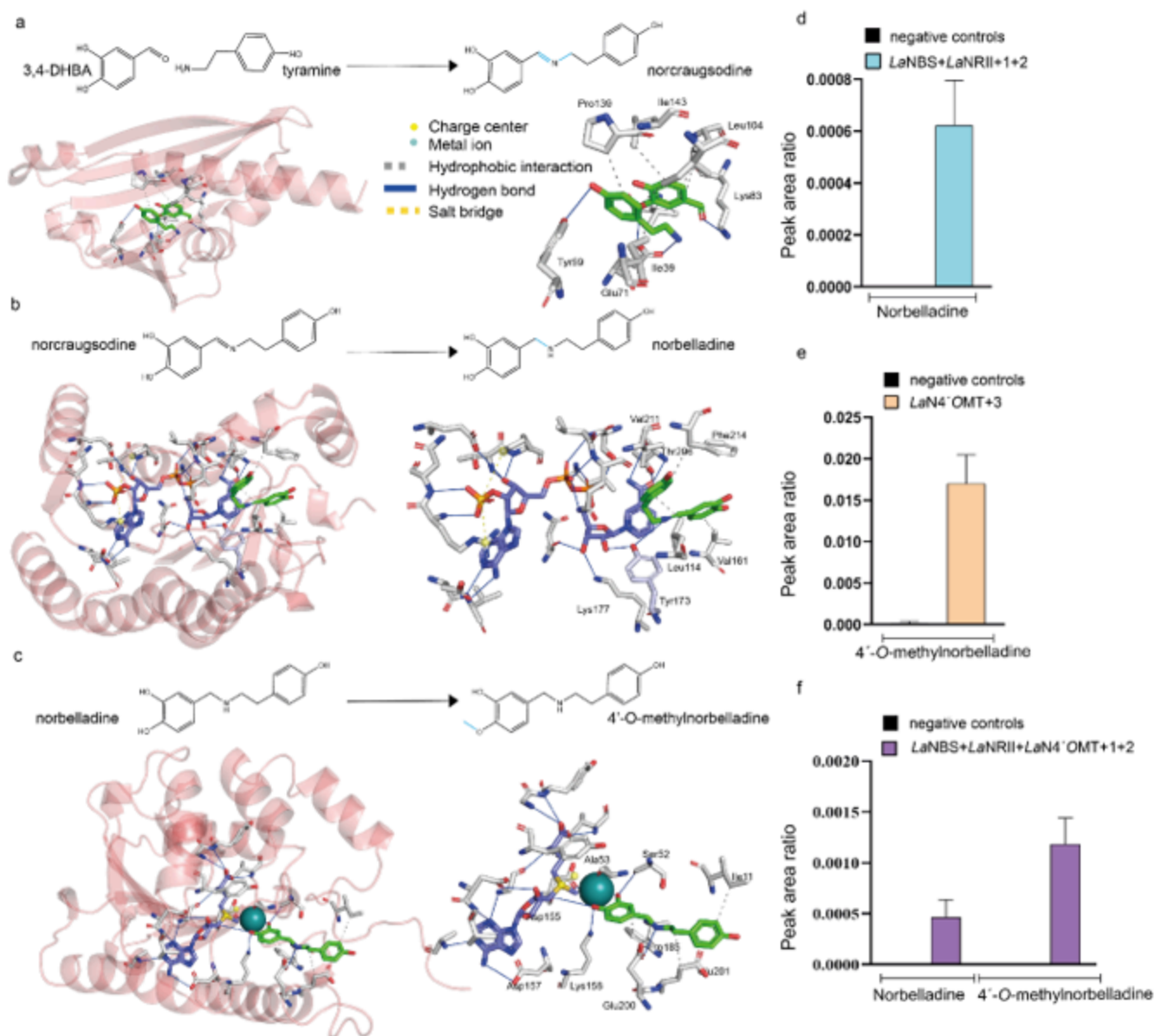


Figure 3

Reconstitution of the foundation steps of the Amaryllidaceae alkaloid pathway to produce 4'-O-methylnorbelladine *in vivo*.

a LaNBS-catalysed condensation of 3,4-DHBA and tyramine. **b** LaNRII-catalysed reduction of norcraugsodine. **c** LaN4'OMT-catalysed 4'-O-methylation of norbelladine. Ligands were docked in the candidate enzyme active site and are shown as green sticks. The AlphaFold2 predicted structure are shown as a transparent deep salmon cartoon (left) and active site residues are shown as light grey sticks. NADPH and SAM are shown as violet sticks. Mg^{2+} is displayed as a blue sphere. A specific view on the active site residues and ligands positioning with interaction is shown (right). The blue bonds in the schemes display the catalysed reactions. **d** Relative quantification of norbelladine production *in planta* following agroinfiltration with *LaNBS* and *LaNRII* and feeding with 3,4-DHBA and tyramine. **e** Relative quantification of 4'-O-methylnorbelladine production upon agroinfiltration of *LaN4'OMT* and

feeding with norbelladine. **f** Relative quantification of norbelladine and 4'-O-methylnorbelladine production by agroinfiltrated *La*NBS, *La*NRII and *La*N4'OMT upon feeding with 3,4-DHBA and tyramine. The product profiles obtained in different assays were analyzed by HPLC-MS/MS. The area ratio of each detected analyte was obtained by dividing the peak area of the targeted compound by the one of the internal standard papaverine. Relative quantification was performed using the area ratio of each analyte produced in the assay. Data are the means $\pm$ SE of three biological repeats with background subtracted. 1 is denoted for tyramine, 2 for 3,4-dihydroxybenzaldehyde and 3 is for norbelladine. *N. benthamiana* infiltrated with *Agrobacterium* harboring EGFP and respective substrate/s are the negative controls in bar plots. Other two negative controls for each experiments included *N. benthamiana* leaf infiltrated with gene/genes only (no substrate), leaf infiltrated with substrate only (no gene/genes) are not shown. Representative detection of at least three independent experiments with three leaves infiltrated in each replicate.

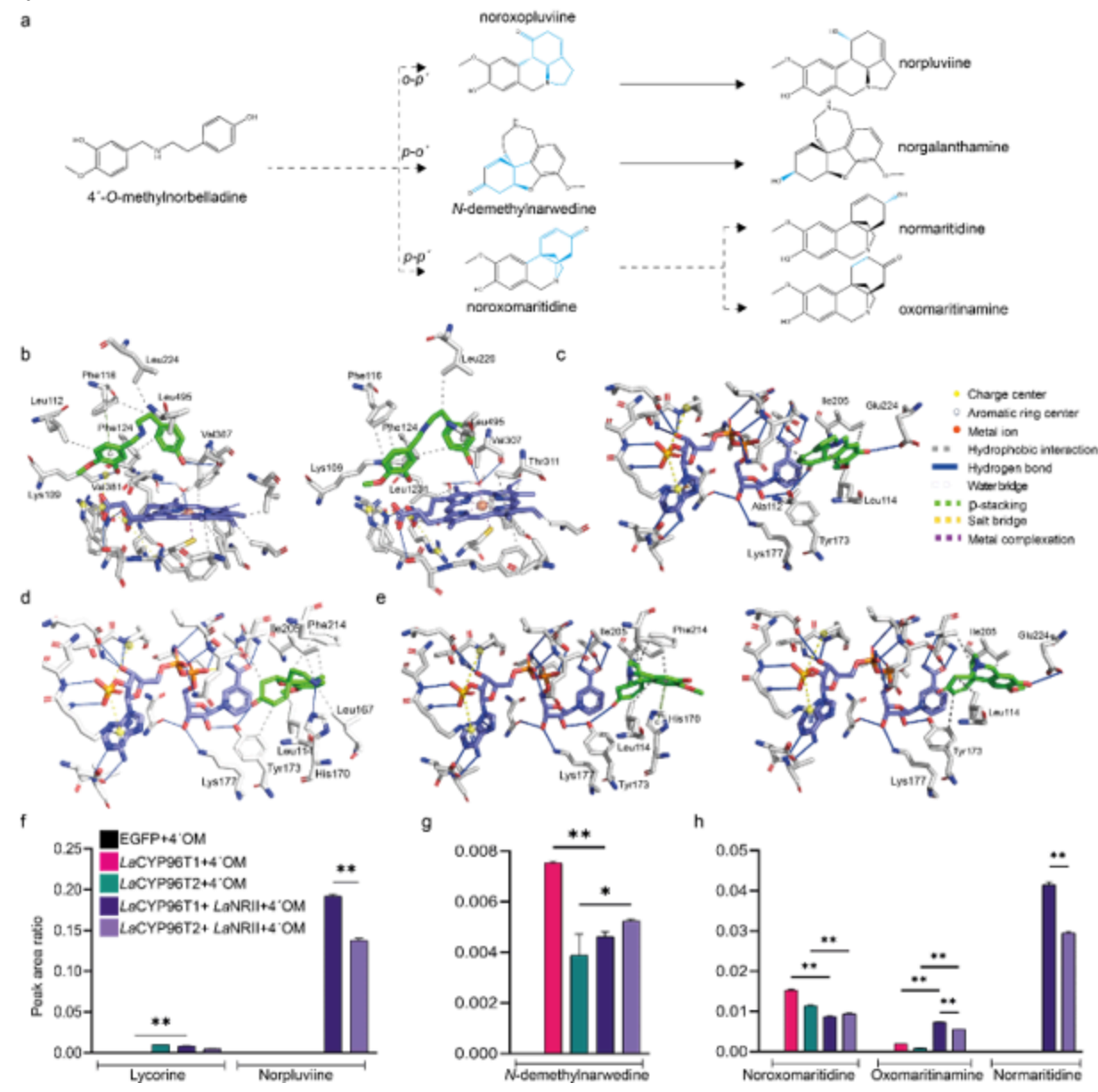


Figure 4

Reconstitution of the branching pathway steps to produce Amaryllidaceae alkaloids *in vivo*.

a Scheme of the proposed *LaCyp96T*-catalysed phenol couplings of 4'-*O*-methylnorbelladine yielding noroxopluviine, *N*-demethylnarwedine, noroxomaritidine, further reduced by *LaNRII* into norpluviine, norgalanthamine, normaritidine and oxomaritinamine. Blue bonds are the results of the proposed catalyzed reactions. **b** Docked 4'-*O*-methylnorbelladine in the active site of CYP96T1 (left) and CYP96T2 (right). **c** Docked noroxopluviine in the active site of *LaNRII* to yield norpluviine. **d** Docked *N*-demethylnarwedine in the active site of *LaNRII* to yield norgalanthamine. **e** Docked noroxomaritidine in the active site of *LaNRII* to yield normaritidine (left) or oxomaritinamine (right). Candidate active site residues are shown as grey sticks and ligands as green sticks. Heme and NADPH are shown as purple sticks. Iron as a red sphere. Interactions are shown as lines as displayed in the legend. **f** Relative quantification of *ortho-para*'products norpluviine and lycorine. **g** Relative quantification of *para-ortho*'product *N*-demethylnarwedine. **h** Relative quantification of *para-para*'products noroxomaritidine, oxomaritinamine and normaritidine. The area ratio of each detected analyte was obtained by dividing the peak area of the targeted compound by the internal standard papaverine. Relative quantification was performed using the area ratio of each analyte produced in the assay. Data are the means $\pm$ SE of three biological repeats with background subtracted. *N. benthamiana* infiltrated with *Agrobacterium* harboring EGFP and respective substrate/s is the negative control in each bar plot. Other three negative controls for each experiments included *N. benthamiana* leaf alone, leaf infiltrated with gene/genes only (no substrate), leaf infiltrated with substrate only (no gene/genes) are not shown. ** and * represent the statistically significant differences among means with $p < 0.0001$ and $p < 0.05$, respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Lamichhaneetal2024SupplementarytablesreadytosubmitNature202407151.xlsx](#)
- [SupplementaryFigures.docx](#)
- [GA.png](#)