Electronic Supplementary Material (ESI) for Chemical Science.


# Deciphering and Reprogramming the Cyclization Regioselectivity in Bifurcation of Indole Alkaloids Biosynthesis

Zhuo Wang,[1,2#] Yiren Xiao,[1,2#] Song Wu,[1,2] Jianghua Chen,[3] Ang Li,[4] EvangelosC. Tatsis[1, 5*]

[1]National Key Laboratory of Plant Molecular Genetics, CAS Center for Excellence in Molecular Plant Sciences, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, 200032, China

[2]University of Chinese Academy of Sciences, Beijing, China

[3]Key Laboratory of Tropical Plant Resources and Sustainable Use, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla County, China

[4]State Key Laboratory of Bioorganic and Natural Product Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai, China

[5]CEPAMS - CAS-JIC Centre of Excellence for Plant and Microbial Sciences

[#]Equal contribution

*Corresponding author: etatsis@cemps.ac.cn; evangelos.tatsis@jic.ac.uk

**Table of Contents**

## Materials and Methods

### Plant material, chemicals and molecular biology reagents.

During three different field trips in Xishuangbanna Tropical Botanical Garden (Yunnan, China), we collected different tissues (leaves, stems, twigs) from *A. scholaris* trees. The collected plant material from XTBG were sent for RNA sequencing (OE Biotech, Shanghai). Before the genome of *A. scholaris* became available, transcript sequence was obtained by using the splicing method of paired-end in Trinity software.[1] According to sequence similarity and length, the longest one was selected as Unigene. Then TGICL software[2] was used for clustering to eliminate redundant extension to obtain a final set of Unigene, which was used as a reference sequence for subsequent analysis. Features of RNA sequencing data are shown in Table S1. Raw data were processed using Trimmomatic. The function of the unigenes was annotated by alignment of the unigenes with the NCBI non-redundant (NR), SwissProt, and Clusters of orthologous groups for eukaryotic complete genomes (KOG) databases using Blastx with a threshold E-value of $10^{-5}$. The unigenes were mapped to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database to annotate their potential metabolic pathways. After the genome of *A. scholaris* became available, trimmed RNA-seq clean reads were mapped to the draft genome of *A. scholaris* by hisat2[3] and the overall alignment rate is 92.0%.

Rhazimol and akuammiline standard were synthesized by the group of Prof. Ang Li in Shanghai Institute of Organic Chemistry.[4] Strictosidine was enzymatically synthesized by incubating tryptamine (23) and secologanin (24) strictosidine synthase (STS) and purified according to previous work and literature.[5-7]

### Geissoschizine preparation.

Geissoschizine was isolated according to previous reported method by Dang et al Nat Chem Bio 2018.[8] In detail, we acquired dried plant material of *Uncaria rhynchophylla* through Taobao to isolate geissoschizine methyl ether. 30 g of fine grounded woody material were used for extraction with 1L of EtOH three times. The ethanolic extract was dried (~ 30 g) and resuspended in 300 mL of 0.01M HCl and washed with 300 mL of EtOAc. As in mentioned in literature, the EtOAc fraction was re-extracted with 250 mL of 0.01M HCl. The pH of combined aqueous phases was adjusted to 8.5 with solution 1M NaOH. Alkaloids were extracted from basified aqueous phase with 200 mL $CHCl_3$ for three times. The chloroform extracts were combined and concentrated under vacuum. The crude alkaloid extract was separated on a silica column SNAP KP-Sil 50 g with a Biotage Isolera Flash chromatography system using a gradient mobile phase consisted of dichloromethane and methanol with flow rate 60 mL/min. The gradient elution was achieved using 2% MeOH for 1 column volume (CV), from 2% to 20% methanol in 10 CV, 20% methanol in 2 CV, 20-95% methanol in 3CV, and 95% in 4 CV. The elution from flash chromatography was monitored by UV at 254 and 280 nm. All the fractions were analysed by HPLC and geissoschizine methyl ether purchased from www.biopurify.cn was used as standard. Analytical HPLC was performed on an Thermo Scientific

Dionex Ulitmate 3000 series (binary pump HPG-3x00RS, auto sampler WPS-3000(RS); diode array detector DAD-3000RS, 200–800 nm) using a Hypersil Gold C18 column (5 µm, 250 × 4 mm; injection volume 10 µL). For the analytical HPLC a linear gradient 0.05% $NH_4OH$- Acetonitrile from 25% to 98% Acetonitrile in 21 min (flow rate 1.25 mL $min^{-1}$; UV detection 280 nm) was applied. Fractions containing the geissoschizine methyl ether were collected, dried and applied for isolation of geissoschizine methyl ether with a semi-preparative HPLC separation in a Kinetex EVO C18 100Å (5 µm, 250 × 10 mm; injection volume 100 µL) with a linear gradient 0.05% $NH_4OH$- Acetonitrile from 40% to 80% Acetonitrile in 2 min, 80% to 90% Acetonitrile in 20 min and 90 to 98% acetonitrile in 2 minutes (flow rate 2 mL $min^{-1}$; UV detection 280 nm) applied. Geissoschizine methyl ether (7 mg) was eluted at 10.4 min and the structure was confirmed by $^1$H-NMR. Demethylation of methyl ether of geissoschizine was performed according to the literature. 2.4 mg of compound were dissolved in 80 µL of glacial acetic acid and 20 µL of concentrated HCl were added. The mixture was left to react at room temperature under magnetic stirring for 2 days. Reaction was monitored by LCMS to observe the accumulation of geissoschizine (loss of 14 amu) and quenched by adding slowly 720 µL 4 M $K_2CO_3$. Geissoschizine was extracted 5 times with 1 mL of ethyl acetate. Ethyl acetate fractions were concentrated in vacuum and dissolve in MeOH. Geissoschizine was eluted at 11.4 min by semi-preparative LC using the same method for isolation of methyl ether. 0.9 mg of geissoschizine were isolated and structure was confirmed by $^1$H-NMR (see figure S3).

**Cloning of cytochrome P450 genes.** The full-length coding regions of P450 candidates were amplified using cDNA of *A. scholaris*, as template which prepared using SuperScript™ IV First-Strand Synthesis System. Primers used for candidate genes amplification are listed in Table S5. A 50 µL system was used for candidate genes amplification contain 25 µL Thermo Scientific Phusion High-Fidelity PCR Master Mix, 2.5 µL forward primer, 2.5 µL reverse primer, 1 µL cDNA of *A. scholaris*, with 50 µL final volume. The pESC-HIS vector was digested with SpeI, P450 candidates were cloned into pESC-HIS vector using In-Fusion cloning (ClonExpress® II One Step Cloning Kit). All the In-Fusion products were transformed into DH5α competent cells and spread on LB plate which contain appropriate antibiotic (carbenicillin). Colony PCR was used for positive colonies identification. Plasmids were extracted using TIANprep Mini Plasmid Kit (DP103) and send for sequencing.

**Cytochrome P450 yeast expression, microsome isolation and immunoblot analysis.** We expressed the candidate p450 genes into *S. cerevisiae* WAT11 strain, which is engineered for heterologous expression of cytochrome p450s from plants.[9] The positive transgenic yeast colony or cryostock was inoculated in 5 mL of synthetic complete (SC) medium lacking histidine (SC-His) containing 2% (w/v) glucose. It was cultured at 30 ℃ and 200 rpm for 16-18h to make sure the $OD_{600}$ value was greater than 5.0. The culture was diluted to an $OD_{600}$ value of 0.05 in 100 mL SC-His supplemented with 2% (w/v) glucose. After 16h, yeast was subsequently collected and washed by $ddH_2O$ once. Yeast cells were resuspended and cultured in 100 mL SC-His supplement containing

2% galactose for a time period between 24-48h for recombination protein induction. Yeast cells were harvested by centrifugation at 4000 rpm for 5 minutes. It was resuspended 10 mL TEK (0.1M KCl in TE buffer) for 5 minutes at room temperature. The TEK buffer was removed by centrifugation at 4000 rpm for 5 minutes. Yeast cells were kept at low temperature (4 ℃). Yeast cells were resuspended in 9 mL TES-B (0.6M sorbitol in TE buffer) and lysed with an approximate volume of 2 mL glass beads (diam. 0.5 mm) using vortex. The mixture was kept in ice for 1 minute and afterwards was vortexed for 1 minutes, repeatedly for 8 times, the supernatant was collected by centrifugation at 4000 rpm for 3 minutes at 4 ℃. The pellet was resuspended by 8 mL TES-B and lysed twice again. The total supernatant was then centrifuged at 10000rpm for 10 minutes at 4 ℃. The supernatant was subsequently centrifuged at 110000x g for 1h at 4 ℃. The supernatant was removed and the pellet containing microsomes was resuspended by 500 µL TEG (20% (v/v) glycerol in TE). The recombination protein was aliquoted and stored in -80 ℃ freezer. Recombinant enzymes were detected by immunoblot analysis using α-Flag M2 (Abmart) detected with Meilunbio fg super sensitive ECL luminescence reagent.

**Enzyme assays for cytochrome p450s.** Standard enzyme assays were performed at 37 °C for 30 min in 100 µl total volume reaction, contain 50 mM phosphate buffer (pH 7.4), 4 µM substrate (geissoschizine), 20 µL microsomal protein, 0.5 mM NADPH. Reactions were quenched by two methods, one was quenched by adding 100 µL methanol directly, another was quenched by the addition of 50 µL 5mM $NaBH_4$ firstly and then with the addition of 50 µL methanol. The reaction mixture was centrifuged at 12000 rpm for 10 minutes at 4 ℃, the supernatant was collected and filtered, and the filtered aliquot was later analysed by LC-MS/MS (Q-Exactive).

**Site-directed mutagenesis of P450s.** The site-directed mutagenesis was performed with overlapping primers (Table S5). 100 ng wildtype template plasmids were used in 50 µL PCR system contain 2.5 forward primer, 2.5 µL reverse primer, 2.5 µL Thermo Scientific Phusion High-Fidelity PCR Master Mix, with 50 µL final volume. 44 µL PCR product was used in a 50 µL system contains 5 µL 10X CutSmart buffer (NEB) and 1 µL DpnI to remove template plasmids. After incubating at 37°C for 4 hours, 10 µL products were transformed into DH5α competent cells and spread on LB plate which contain appropriate antibiotic (carbenicillin). Plasmids were extracted using TIANprep Mini Plasmid Kit (DP103) and send for sequencing.

**Cloning of RHR and AKS candidate genes.** The full-length coding regions of RHR and AKS candidates were amplified using cDNA of *A. scholaris*, which prepared using SuperScript™ IV First-Strand Synthesis System. Primers used for candidate genes amplification are listed in supplementary Table S5. A 50 µL system was used for candidate genes amplification contain 25 µL Thermo Scientific Phusion High-Fidelity PCR Master Mix, 2.5 µL forward primer, 2.5 µL reverse primer, 1 µL cDNA of *A. scholaris*, with 50 µL final volume. The pOPIN vectors were digested with HindIII and KpnI, RHR candidates were cloned into pOPINF vector and AKS candidates were cloned

in pOPINJ vector using In-Fusion cloning (ClonExpress® II One Step Cloning Kit). All the In-Fusion products were transformed into DH5α competent cells and spread on LB plate which contain appropriate antibiotic (carbenicillin). Colony PCR was used for positive colonies identification. Plasmids were extracted using TIANprep Mini Plasmid Kit (DP103) and send for sequencing.

**Protein expression in *E. coli*.** Candidate genes for RHR and AKS activities were expressed either in C43(DE3) or Rosetta2(DE3) pLysS *E. coli*. A starter culture was grown overnight at 37°C in LB medium supplemented with carbenicillin (50 mg/mL) and chloramphenicol (34 mg/mL for Rosseta strain). A 1:100 dilution in fresh 2xYT (100 mL) media supplemented with antibiotics was prepared and grew at 37°C to an $OD_{600}$ of 0.8. Protein production was induced by addition of 0.5 mM IPTG and the cultures were shaken at 18°C for 16h. Cell were collected by centrifugation, lysed by sonication in buffer A1 (50 mM Tris-HCl pH 8.0, 50 mM glycine, 500 mM NaCl, 5 % v/v glycerol, 20 mM imidazole) supplemented with 0.2 mg/mL lysozyme and 100 mM PMSF. Cells were lysed using sonication for 8 min on ice using 2 s pulses and then centrifuged at 35000 *g* for 20 min at 4°C. The supernatant was collected and added 250 µl NiNTA. The mixture was mixed gently on rocking platform for 1.5h in cold room (4°C) and spin at 1000 g, 30 s to remove supernatant. 10 mL Buffer A1 was used to resuspend and rinse 2 times. 300 µl buffer B1 (50 mM Tris-HCl pH 8.0, 50 mM glycine, 500 mM NaCl, 5 % v/v glycerol, 500mM imidazole) was used to elute for 3 times. Protein was collected and filtered by 15 mL centrifugal filters with 10mM PBS buffer at least 6 times. Protein concentration was measured at 280 nm with Nanodrop and divided in 50-100 mL, fast-frozen in liquid nitrogen and stored at -20°C.

**Enzyme assays for GS.** Standard enzyme assays were performed at 37 °C in a 100 µl total volume reaction, contain 50 mM phosphate buffer (pH 7.4), 100 µM substrate (strictosidine), 1µl strictosidine glucosidase (SGD), 0.5 mM NADPH, incubate for 15 min. Then 5 µM of GS protein were added and incubate at 37 °C for 1.5 hour. Reactions were stopped by adding 100 µl methanol. The reaction mixture was centrifuged at 12000 rpm for 10 minutes at 4 ℃, the supernatant was collected and filtered, and the filtered aliquot was later analyzed by LC-MS/MS (Q-Exactive).

**Enzyme assays for RHR candidates.** Standard enzyme assays were performed at 37 °C for 30 min in 100 µl total volume reaction, contain 50 mM phosphate buffer (pH 7.4), 4 µM substrate (geissoschizine), 20 µl microsomal protein (AsRHS), 0.5 mM NADPH. Then 5 µM of ADH protein and 0.5 mM NADPH were added and incubate at 37 °C for 1.5 hour. Reactions were stopped by adding 100 µl methanol. The reaction mixture was centrifuged at 12000 rpm for 10 minutes at 4 ℃, the supernatant was collected and filtered, and the filtered aliquot was later analysed by LC-MS/MS (Q-Exactive and Q-Trap).

**Enzyme assays for AKS candidates.** Standard enzyme assays were performed at 37 °C for 2 hours in 100 µl total volume reaction, contain 50 mM phosphate buffer (pH 7.4), contain 100 µg substrate (rhazimol), 5 µM BAHD protein, 100 µM acetyl-CoA. Reactions were stopped by adding

100 µl methanol. The reaction mixture was centrifuged at 12000 rpm for 10 minutes at 4 ℃, the supernatant was collected and filtered, and the filtered aliquot was later analysed by LC-MS/MS (Q-Exactive and Q-Trap).

**Pathway reconstitution in yeast (*Saccharomyces cerevisiae*).** The AsRHS and AsRHR2 genes were cloned into pESC-His vector. AsRHS was cloned into Multiple Cloning Site (MCS) 1 which digested with SpeI and RHR2 into MCS2 which digested with SalI and BamHI by In-Fusion cloning (ClonExpress® II One Step Cloning Kit). AKS1 was cloned into pESC-Leu vector using In-Fusion cloning (ClonExpress® II One Step Cloning Kit) in MCS1 digested with SpeI. All the In-Fusion products were transformed into DH5α competent cells and spread on LB plate which contain appropriate antibiotic (carbenicillin). Colony PCR was used for positive colonies identification. Plasmids were extracted using TIANprep Mini Plasmid Kit (DP103) and send for sequencing. The two plasmids that contain three genes were transformed into yeast strain together. The positive transgenic yeast colony was inoculated in 5 mL of synthetic complete (SC) medium lacking histidine (SC-His-Leu) containing 2% (w/v) glucose. It was cultured at 30°C and 200 rpm for 36-48h to make sure the $OD_{600}$ value was greater than 5.0. Then the starter culture was inoculated to 100 mL SC-His-Leu medium containing 2% (w/v) glucose and cultured at 30°C and 200 rpm for 2 days. Yeast cell was subsequently collected and washed by dd$H_2O$, resuspended, and cultured in 100 mL SC-His-Leu supplement containing 2% galactose 48h for recombination protein induction. Yeast cells were harvested by centrifugation at 4000 rpm for 5 minutes. It was resuspended in 5 mL 50 mM sodium phosphate buffer (pH 7.4) and lysed with an approximate volume of 2 mL glass beads (diam. 0.5 mm) using vortex 20 times (1 min vortex, 1 min on ice). The supernatant was collected by centrifugation at 4000rpm for 5 minutes at 4 ℃ and followed by centrifugation at 10000 rpm for 20 minutes at 4 °C. The supernatant was collected for enzyme assay. A 200 µl reaction was performed at 37 °C for 20 ~ 24 hours which contain 20 µM geissoschizine, 1mM NADPH, 100µM acetyl-CoA. Reactions were stopped by adding 200 µl methanol. The reaction mixture was centrifuged at 12000 rpm for 10 minutes at 4 ℃, the supernatant was collected and filtered, and the filtered aliquot was later analysed by LC-MS/MS (Q-Exactive and Q-Trap).

**LC-MS/MS analysis**

**Method 1.** Enzyme assays were analysed by UPLC (ultrahigh performance liquid chromatograph) coupled to a Q Exactive hybrid quadrupole–orbitrap mass spectrometer (Thermo Fisher). Metabolites were separated with a Poroshell 120 SB-Aq 2.7 µm, 3.0 x 100 mm (Agilent) column. The solvents used were 2 mM ammonium acetate (pH~6.3) ($NH_4Ac$) in water as Solvent A and 100% methanol as solvent B, with a flow rate of 0.3 mL/min. Injection volume was 2 µl. The gradient profile was 0 min, 40% B; 0 – 1.0 min, 40% B, from 1.0 to 8.0 min, linear gradient to 95% B; from 8.0 min to 10.0 min, 95% B; from 10.0 – 10.1 min, back to 40% B; from 10.1 – 12.0 min, 40% B to re-equilibrate the column. The mass spectrometer with a heated electrospray ionization source was

operated in positive mode. The key parameters were as follows: ionization voltage, +3.5 kV/-2.8kV, sheath gas pressure, 40 arbitrary units; auxiliary gas, 10 arbitrary units; auxiliary gas heater temperature, 350 °C; capillary temperature, 275 °C. The mass spectrometer was run in full scan mode at *m/z* 200-700 scan range and 70,000 resolution. Data processing and ion annotation based on accurate mass performed in Thermo Xcalibur Qual Browser. MS/MS spectra were performed by higher-anergy collisional dissociation at 20, 40 and 60 eV collision energy. Fragment-ion spectra were recorded with 17,500 resolution.

**Method 2.** Mutant enzyme reactions were analysed by QTRAP® 6500$^{+}$ coupled to ExionLC™ system using the following parameters: 1 µL of sample was injected in column Poroshell 120 SB-Aq 2.7 µm, 3.0 x 100 mm (Agilent) with a flowrate of 0.4 mL/min. The column oven was set at 40 °C. The gradient [solvent A: 2 mM ammonium acetate ($NH_4Ac$) in water; B: 2 mM ammonium acetate ($NH_4Ac$) in methanol] was set as follows: 12 min method; 0 to 1.0 min 40% solvent B, followed by linear gradient from 40% to 95% solvent B in 7.0 min, then 95% to 99% solvent B in 2 min, from 10.0 to 10.1 min, back to 40% B, then equilibration from 10.1 to 12 min with 40% solvent B. Mass spectra were acquired in positive mode with these parameters: multiple reaction monitoring (MRM) scan; ion spray voltage: 4500 V; turbo heater temperature (TEM): 450 °C; curtain gas: 35 psi; collision gas: medium; ion source gas (ISG) 1 and 2 was 50 psi. Peak integration and the calculations were performed using Analyst® 1.6.3 software. All data were analysed by MultiQuant 3.0.2.

**Method 3.** The enzyme assay using ajmalicine as substrate were performed on an ACQUITY UPLC@BEH C18 column (2.1 x 50 mm, 1.7 µm) at flow rate of 0.3 mL/min and 40 °C column oven temperature. The column was equilibrated in solvent A (0.1% [v/v] formic acid in water and the following gradient was applied: 0-0.5 min 10% B (aceto-nitrile); 0.5-5.5 min 45% B; 5.5-7 min 90% B; and 7.1-9 min at 10% B for column re-equilibration.

**Phylogenetic analysis.** Rooted neighbour-joining phylogenetic tree for CYP candidates from this study and other characterized CYPs in monoterpene indole alkaloids biosynthesis pathway were performed using MEGA-X software (version 10.2.2). The accession numbers of amino acid sequences used to generate phylogenetic tree are as follows: *Solanum lycopersicum ent*-kaurene oxidase (XP_004238197.1), AhCYP71AY2 (*A. hubrichtii* CYP71AY2-like protein, AIR93639.1), Cr7DLH CYP72A224 (*C. roseus* 7-deoxyloganic acid hydroxylase Cytochrome P450 72A224, U5NE19.1), Cr16T3O (*C. roseus* 16T3O protein, CQR79400.1), CrAS CYP71AY1(*C. roseus* Alstonine synthase Cytochrome P450 71AY1, W8JDE2.1), CrC4H CYP73 (*C. roseus* Cinnamic acid 4-hydroxylase Cytochrome P450 73, P48522.1), CrCYP71BT1 (*C. roseus* Cytochrome P450 71BT1, W8JCE8.1), CrCYP71D1 (*C. roseus* cytochrome P450 71D1, ACD42776.1), CrCYP71D2 (*C. roseus* Cytochrome P450 71D2, W8JIT4.1), CrCYP76A26 (*C. roseus*CYP76A26, AHX24370.1), CrCYP76T24 (*C. roseus* Cytochrome P450 76T24, W8JMV1.1), CrCYP81C13 (*C. roseus* Cytochrome P450 81C13, W8JIS9.1), CrCYP81Q32 (*C. roseus* Cytochrome P450 81Q32, W8JMU7.1), CrF35H (*C. roseus* flavonoid 3',5'-hydroxylase, CAA09850.1), CrG8H CYP76B6 (*C.*

*roseus* Geraniol 8-hydroxylase Cytochrome P450 76B6, Q8VWZ7.1), CrGO CYP71D1V1 (*C. roseus* Geissoschizine oxidase Cytochrome P450 71D1V1, I1TEM0.1), CrIrOx (*C. roseus* Iridoid oxidase Cytochrome P450 76A26, W8JIS5.1), CrSLS CYP72A1(*C. roseus* Secologanin synthase Cytochrome P450 72A1, Q05047.1), CrT3O CYP71D1V2 (*C. roseus* Tabersonine 3-oxygenase Cytochrome P450 71D1V2, I1TEM1.1), CrT67EO1 (*C. roseus* Tabersonine 6,7-epoxidase isoform 1 Cytochrome P450 71D521, A0A343URW6.1), CrT67EO2 (*C. roseus* Tabersonine 6,7-epoxidase isoform 2 Cytochrome P450 71D347, A0A343URW7.1), CrT16H CYP71D12 (*C. roseus* Tabersonine 16-hydroxylase 1 Cytochrome P450 71D12, P98183.2), CrT16H CYP71D351 (*C. roseus* Tabersonine 16-hydroxylase 2 Cytochrome P450 71D351, U5HKE8.1), CrT19H (*C. roseus* Tabersonine/lochnericine 19-hydroxylase Cytochrome P450 71BJ1, F5BHA2.1), GsR/H11H1 (*G. sempervirens* rankinidine/humantenine-11-hydroxylase 1, AXK92566.1), GsR/H11H2 (*G. sempervirens* rankinidine/humantenine-11-hydroxylase 2, AXK92567.1), GsR/H11H3 (*G. sempervirens* rankinidine/humantenine-11-hydroxylase 3, AXK92568.1), GsSBE (*G. sempervirens* Sarpagan bridge enzyme Cytochrome P450 71AY5, P0DO14.1), Rs7DLH (*R. serpentina* 7-deoxyloganic acid 7-hydroxylase-like protein, AGX93059.1), RsCYP71AY2 (*R. serpentina* CYP71AY2-like protein, AIR93638.1), RsCYP76A26 (*R. serpentina* CYP76A26-like protein, AHX24367.1), RsG10H (*R. serpentina* geraniol 10-hydroxylase-like protein, AGX93053.1), RsSBECYP71AY4 (*R. serpentina* Sarpagan bridge enzyme Cytochrome P450 71AY4, P0DO13.1), RsVH (*R. serpentina* vinorine hydroxylase, ASG81458.1), Te7DLH (*T. elegans* 7-deoxyloganic acid 7-hydroxylase-like protein, AGX93060.1), TeCYP71AY2 (*T. elegans* CYP71AY2-like protein, AIR93636.1), TeG10H (*T. elegans* geraniol 10-hydroxylase-like protein, AGX93054.1), TiIA10H (*T. iboga* ibogamine 10-hydroxylase, AXF35974.1), VmCYP71AY (*V. minor* CYP71AY2-like protein, AIR93637.1).

**Alignment.** The amino acid sequences of proteins were used for alignment. A multiple sequence alignment was used by MUSCLE (Multiple Sequence Comparison by Log-Expectation) and exported as a clw file. The clw file was uploaded to ESPript 3.0 for further beatify.

**Homology modelling and AlphaFold.** The putative structures of AsRHS and AsGO were predicted by SWISS-MODEL software (https://swissmodel.expasy.org/interactive).[10] The software identified Ferruginol synthase CYP76AH1 from *Salvia miltiorrhiza* at 1.90 Å resolution (SMTL ID: 7CB9)[11] to be the best template for both RHS and AsGO1. The Seq Similarity and Coverage of RHS are 0.36 and 0.93. These Seq Similarity and Coverage of AsGO1 are 0.35 and 0.93. AlphaFold open source code is available in Github (https://github.com/kalininalab/alphafold_non_docker). AlphaFold was also used for AsRHS and AsGO protein structure prediction.

**Docking.** Molecular docking was carried out by using AutoDock software (http://autodock.scripps.edu/). The first step is to prepare the ligand and receptor coordinate files which includes polar hydrogen atoms, partial charges, atom types and information on the articulation of flexible molecules. The format of the coordinate files is PDBQT. Docking calculations is using rigid

receptor coordinates. The second step is to create grid parameter files in AutoDockTools. These parameters are created in “grid” menu. Parameters of center grid box: X center 51.269, Y center 0.175, Z center 6.008. The third step is to use the maps generated by AutoGrid to evaluate the ligand-protein interaction at each point in the docking simulation.The interactions between substrates and proteins are visualized by PyMOL(https://pymol.org/2/).

**Supplementary Tables S1-S9**

**Table S1.** Features of RNA sequencing data of *Alstonia scholaris.*

| Assembly feature | Statistics |
|---|---|
| Unigene numbers(>=1000bp) | 32692 |
| N50 of Unigenes | 1,743bp |
| Total length | 103.45Mb |
| Max length of Unigenes | 15,861bp |
| Average length of Unigenes | 1,064.62bp |

**Table S2.** The amino acid sequence identity of alcohol dehydrogenases (ADH) candidates and characterized ADHs in MIAs pathway.

| | CrRedox2 | RHR1 | RHR2 | ADH11 | ADH12 | AsGS | CrGS1 | CrGS2 | ADH4 | CrDPAS | ADH2 | ADH3 | ADH8 | CrRedox1 | ADH9 | ADH5 | ADH7 | CrT3R | ADH1 | CrTHAS1 | CrHYS | CrTHAS | ADH6 | CrTHAS2 | ADH13 | CrTHAS3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CrRedox2 | | 80.2 | 75.0 | 62.3 | 66.3 | 9.8 | 9.8 | 9.3 | 11.1 | 11.3 | 10.3 | 10.6 | 11.1 | 10.9 | 9.8 | 8.2 | 10.6 | 9.8 | 11.1 | 9.8 | 10.6 | 10.8 | 10.4 | 10.1 | 11.4 | 11.6 |
| RHR1 | 80.2 | | 84.9 | 67.6 | 71.2 | 10.9 | 11.4 | 10.9 | 12.1 | 11.1 | 11.3 | 11.1 | 11.4 | 13.0 | 11.7 | 10.1 | 13.0 | 10.4 | 12.5 | 10.6 | 11.1 | 11.6 | 10.9 | 10.6 | 12.2 | 13.4 |
| RHR2 | 75.0 | 84.9 | | 67.9 | 69.7 | 10.1 | 10.3 | 9.8 | 10.5 | 11.1 | 10.0 | 10.3 | 11.1 | 12.2 | 10.1 | 9.3 | 11.4 | 9.8 | 11.1 | 9.6 | 10.1 | 10.8 | 10.1 | 9.6 | 11.4 | 12.1 |
| ADH11 | 62.3 | 67.6 | 67.9 | | 76.2 | 10.8 | 11.3 | 10.3 | 10.5 | 10.8 | 10.0 | 10.2 | 11.9 | 13.2 | 10.8 | 9.7 | 11.1 | 10.8 | 12.1 | 10.6 | 11.8 | 11.8 | 9.8 | 10.1 | 11.1 | 11.0 |
| ADH12 | 66.3 | 71.2 | 69.7 | 76.2 | | 11.0 | 11.7 | 11.2 | 11.1 | 11.2 | 9.9 | 10.1 | 11.0 | 12.6 | 11.0 | 9.6 | 11.7 | 10.2 | 12.3 | 9.9 | 11.5 | 11.7 | 9.7 | 10.0 | 11.5 | 11.9 |
| AsGS | 9.8 | 10.9 | 10.1 | 10.8 | 11.0 | | 91.5 | 89.0 | 54.1 | 49.0 | 52.3 | 51.8 | 54.5 | 53.1 | 56.9 | 57.4 | 59.8 | 49.3 | 58.5 | 57.7 | 55.2 | 55.5 | 52.3 | 51.3 | 51.2 | 48.4 |
| CrGS1 | 9.8 | 11.4 | 10.3 | 11.3 | 11.7 | 91.5 | | 91.8 | 51.1 | 47.4 | 50.4 | 49.9 | 53.1 | 52.5 | 56.6 | 54.9 | 57.9 | 47.9 | 56.3 | 55.5 | 53.0 | 53.6 | 51.7 | 50.5 | 49.6 | 47.9 |
| CrGS2 | 9.3 | 10.9 | 9.8 | 10.3 | 11.2 | 89.0 | 91.8 | | 50.5 | 47.7 | 49.3 | 49.0 | 53.8 | 52.8 | 55.8 | 55.5 | 57.9 | 47.4 | 55.5 | 55.5 | 53.0 | 53.6 | 51.2 | 50.3 | 49.6 | 47.6 |
| ADH4 | 11.1 | 12.1 | 10.5 | 10.5 | 11.1 | 54.1 | 51.1 | 50.5 | | 67.4 | 70.4 | 70.7 | 62.1 | 62.4 | 67.1 | 67.5 | 66.2 | 54.8 | 62.4 | 59.9 | 56.2 | 55.6 | 54.5 | 54.4 | 56.7 | 54.1 |
| CrDPAS | 11.3 | 11.1 | 11.1 | 10.8 | 11.2 | 49.0 | 47.4 | 47.7 | 67.4 | | 83.6 | 82.2 | 55.5 | 54.6 | 59.5 | 62.6 | 59.9 | 48.6 | 56.7 | 54.0 | 51.6 | 51.9 | 50.4 | 51.3 | 50.9 | 51.1 |
| ADH2 | 10.3 | 11.3 | 10.0 | 10.0 | 9.9 | 52.3 | 50.4 | 49.3 | 70.4 | 83.6 | | 85.5 | 58.1 | 57.7 | 63.4 | 63.5 | 63.5 | 51.1 | 59.8 | 58.4 | 54.1 | 54.7 | 52.0 | 52.4 | 52.3 | 50.8 |
| ADH3 | 10.6 | 11.1 | 10.3 | 10.2 | 10.1 | 51.8 | 49.9 | 49.0 | 70.7 | 82.2 | 85.5 | | 58.5 | 57.5 | 61.7 | 64.3 | 62.7 | 50.3 | 59.2 | 56.2 | 53.6 | 53.3 | 52.8 | 52.2 | 52.0 | 51.9 |
| ADH8 | 11.1 | 11.4 | 11.1 | 11.9 | 11.0 | 54.5 | 53.1 | 53.8 | 62.1 | 55.5 | 58.1 | 58.5 | | 84.3 | 64.8 | 68.1 | 65.1 | 53.3 | 60.8 | 59.0 | 55.5 | 55.5 | 53.4 | 52.7 | 54.0 | 51.6 |
| CrRedox1 | 10.9 | 13.0 | 12.2 | 13.2 | 12.6 | 53.1 | 52.5 | 52.8 | 62.4 | 54.6 | 57.7 | 57.5 | 84.3 | | 64.4 | 68.6 | 63.6 | 54.0 | 58.2 | 58.2 | 54.8 | 54.2 | 52.7 | 52.6 | 53.0 | 53.0 |
| ADH9 | 9.8 | 11.7 | 10.1 | 10.8 | 11.0 | 56.9 | 56.6 | 55.8 | 67.1 | 59.5 | 63.4 | 61.7 | 64.8 | 64.4 | | 72.9 | 70.4 | 52.9 | 61.5 | 61.0 | 58.0 | 57.7 | 55.6 | 55.0 | 55.9 | 52.0 |
| ADH5 | 8.2 | 10.1 | 9.3 | 9.7 | 9.6 | 57.4 | 54.9 | 55.5 | 67.5 | 62.6 | 63.5 | 64.3 | 68.1 | 68.6 | 72.9 | | 77.6 | 55.4 | 65.2 | 61.6 | 59.2 | 59.0 | 56.7 | 56.3 | 57.8 | 54.6 |
| ADH7 | 10.6 | 13.0 | 11.4 | 11.1 | 11.7 | 59.8 | 57.9 | 57.9 | 66.2 | 59.9 | 63.5 | 62.7 | 65.1 | 63.6 | 70.4 | 77.6 | | 60.3 | 69.8 | 62.9 | 62.0 | 61.5 | 63.0 | 59.3 | 63.0 | 57.4 |
| CrT3R | 9.8 | 10.4 | 9.8 | 10.8 | 10.2 | 49.3 | 47.9 | 47.4 | 54.8 | 48.6 | 51.1 | 50.3 | 53.3 | 54.0 | 52.9 | 55.4 | 60.3 | | 63.7 | 67.1 | 64.8 | 65.4 | 53.6 | 52.3 | 52.5 | 50.7 |
| ADH1 | 11.1 | 12.5 | 11.1 | 12.1 | 12.3 | 58.5 | 56.3 | 55.5 | 62.4 | 56.7 | 59.8 | 59.2 | 60.8 | 58.2 | 61.5 | 65.2 | 69.8 | 63.7 | | 70.4 | 72.7 | 70.5 | 62.0 | 59.8 | 62.3 | 58.1 |
| CrTHAS1 | 9.8 | 10.6 | 9.6 | 10.6 | 9.9 | 57.7 | 55.5 | 55.5 | 59.9 | 54.0 | 58.4 | 56.2 | 59.0 | 58.2 | 61.0 | 61.6 | 62.9 | 67.1 | 70.4 | | 76.0 | 75.4 | 56.4 | 54.6 | 55.1 | 52.6 |
| CrHYS | 10.6 | 11.1 | 10.1 | 11.8 | 11.5 | 55.2 | 53.0 | 53.0 | 56.2 | 51.6 | 54.1 | 53.6 | 55.5 | 54.8 | 58.0 | 59.2 | 62.0 | 64.8 | 72.7 | 76.0 | | 93.9 | 54.5 | 56.6 | 56.4 | 52.5 |
| CrTHAS4 | 10.8 | 11.6 | 10.8 | 11.8 | 11.7 | 55.5 | 53.6 | 53.6 | 55.6 | 51.9 | 54.7 | 53.3 | 55.5 | 54.2 | 57.7 | 59.0 | 61.5 | 65.4 | 70.5 | 75.4 | 93.9 | | 55.1 | 56.3 | 55.6 | 53.1 |
| ADH6 | 10.4 | 10.9 | 10.1 | 9.8 | 9.7 | 52.3 | 51.7 | 51.2 | 54.5 | 50.4 | 52.0 | 52.8 | 53.4 | 52.7 | 55.6 | 56.7 | 63.0 | 53.6 | 62.0 | 56.4 | 54.5 | 55.1 | | 61.2 | 66.0 | 64.5 |
| CrTHAS2 | 10.1 | 10.6 | 9.6 | 10.1 | 10.0 | 51.3 | 50.5 | 50.3 | 54.4 | 51.3 | 52.4 | 52.2 | 52.7 | 52.6 | 55.0 | 56.3 | 59.3 | 52.3 | 59.8 | 54.6 | 56.6 | 56.3 | 61.2 | | 69.5 | 63.9 |
| ADH13 | 11.4 | 12.2 | 11.4 | 11.1 | 11.5 | 51.2 | 49.6 | 49.6 | 56.7 | 50.9 | 52.3 | 52.0 | 54.0 | 53.0 | 55.9 | 57.8 | 63.0 | 52.5 | 62.3 | 55.1 | 56.4 | 55.6 | 66.0 | 69.5 | | 80.9 |
| CrTHAS3 | 11.6 | 13.4 | 12.1 | 11.0 | 11.9 | 48.4 | 47.9 | 47.6 | 54.1 | 51.1 | 50.8 | 51.9 | 51.6 | 53.0 | 52.0 | 54.6 | 57.4 | 50.7 | 58.1 | 52.6 | 52.5 | 53.1 | 64.5 | 63.9 | 80.9 | |

**Table S3.** The amino acid sequence identity of cytochrome P450 candidates and cytochrome P450 enzymes in MIAs pathway that act on strictosidine aglycon reduced products (geissoschizine and tetrahydroalstonine).

| | AsRHS | CrGO CYP71D1V1 | AsGO | AsAS | GsSBE CYP71AY5 | CrAS CYP71AY1 | RsSBE CYP71AY4 |
|---|---|---|---|---|---|---|---|
| AsRHS | | 61.8 | 62.4 | 50.2 | 53.6 | 54.3 | 55.4 |
| CrGO CYP71D1V1 | 61.8 | | 90.2 | 48.3 | 51.2 | 50.6 | 52.3 |
| AsGO | 62.4 | 90.2 | | 49.4 | 51.9 | 51.8 | 52.6 |
| AsAS | 50.2 | 48.3 | 49.4 | | 57.1 | 56.5 | 57.8 |
| GsSBE CYP71AY5 | 53.6 | 51.2 | 51.9 | 57.1 | | 58.4 | 60.4 |
| CrAS CYP71AY1 | 54.3 | 50.6 | 51.8 | 56.5 | 58.4 | | 74.4 |
| RsSBE CYP71AY4 | 55.4 | 52.3 | 52.6 | 57.8 | 60.4 | 74.4 | |

**Table S4.** The amino acid sequence identity of BAHD acetyl transferase candidates and reported BAHD acetyl transferases in MIAs pathway.

| | BAHD8 | AKS2 | CrDAT | AKS1 | BAHD4 | BAHD9 | BAHD6 | BAHD7 | RsVS | BAHD1 | BAHD2 | CrSAT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BAHD8 | | 25.8 | 25.2 | 30.0 | 27.9 | 30.8 | 29.1 | 28.6 | 29.1 | 31.3 | 29.9 | 31.5 |
| AKS2 | 25.8 | | 25.3 | 28.5 | 26.9 | 33.9 | 28.8 | 29.2 | 30.4 | 32.1 | 31.5 | 32.1 |
| CrDAT | 25.2 | 25.3 | | 50.8 | 53.4 | 28.2 | 28.3 | 27.4 | 27.9 | 32.2 | 29.2 | 27.9 |
| AKS1 | 30.0 | 28.5 | 50.8 | | 65.7 | 30.9 | 32.2 | 32.0 | 32.8 | 34.6 | 32.6 | 32.1 |
| BAHD4 | 27.9 | 26.9 | 53.4 | 65.7 | | 29.6 | 27.2 | 26.9 | 30.8 | 31.3 | 29.9 | 28.7 |
| BAHD9 | 30.8 | 33.9 | 28.2 | 30.9 | 29.6 | | 32.3 | 32.5 | 32.1 | 33.0 | 32.7 | 33.9 |
| BAHD6 | 29.1 | 28.8 | 28.3 | 32.2 | 27.2 | 32.3 | | 92.4 | 41.6 | 45.6 | 45.9 | 44.3 |
| BAHD7 | 28.6 | 29.2 | 27.4 | 32.0 | 26.9 | 32.5 | 92.4 | | 41.2 | 44.7 | 44.8 | 43.4 |
| RsVS | 29.1 | 30.4 | 27.9 | 32.8 | 30.8 | 32.1 | 41.6 | 41.2 | | 67.6 | 73.7 | 69.8 |
| BAHD1 | 31.3 | 32.1 | 32.2 | 34.6 | 31.3 | 33.0 | 45.6 | 44.7 | 67.6 | | 78.6 | 73.9 |
| BAHD2 | 29.9 | 31.5 | 29.2 | 32.6 | 29.9 | 32.7 | 45.9 | 44.8 | 73.7 | 78.6 | | 82.4 |
| CrSAT | 31.5 | 32.1 | 27.9 | 32.1 | 28.7 | 33.9 | 44.3 | 43.4 | 69.8 | 73.9 | 82.4 | |

**Table S5.** Akuammiline quantification from pathway reconstitution in *Saccharomyces cerevisiae*.

| | Substrate (geissoschizine) | Product (akuammiline) | product percentage |
|---|---|---|---|
| RHS+RHR1+AKS1 | 20 μM | 1.03 μM ± 0.29 | 5.2% ± 1.5% |
| RHS+RHR2+AKS1 | 20 μM | 1.59 μM ± 0.35 | 8.0% ± 1.8% |

**Table S6.** Quantification of rhazimol by LC-tQMS as enzymatic product of AsRHS wildtype and single mutants, AsGO wildtype and single mutants enzyme assays.

| Enzyme | | calculated rhazimol concentration (ng/mL) |
|---|---|---|
| AsRHS | *wild type* | 509.66±159.78 |
| | V112I | 264.98±80.74 |
| | C205V | 233.02±62.56 |
| | V208T | 25.77±7.00 |
| | S212A | 118.26±27.42 |
| | F372V | 44.96±2.29 |
| | F372Y | 2.41±0.06 |
| | F372L | 177.04±39.25 |
| | F372A | 8.73±1.72 |
| | F372W | 271.76±50.50 |
| AsGO | *wild type* | nd |
| | I109V | nd |
| | V202C | nd |
| | T205V | nd |
| | S209A | nd |
| | V372F | 2.47±0.96 |
| | V372Q | nd |
| | V372A | nd |
| | V372I | 1.02±0.24 |
| | V372E | nd |
| | V372L | 2.63±0.27 |

**Table S7.** Relative intergartion (peak area) of akuammicine by LC-tQMS as enzymatic product of AsRHS wildtype and single mutants, AsGO wildtype and single mutants enzyme assays.

| Enzyme | | relative peak area (akuammicine) |
|---|---|---|
| AsRHS | *wild type* | 3.34E+05 ± 2.74E+05 |
| | V112I | 8.71E+04 ± 2.43E+04 |
| | C205V | 1.00E+05 ± 3.75E+04 |
| | V208T | 9.51E+03 ± 5.06E+03 |
| | S212A | 3.86E+04 ± 1.58E+04 |
| | F372V | 8.09E+04 ± 1.02E+04 |
| | F372Y | 2.25E+03 ± 1.34E+03 |
| | F372L | 1.18E+05 ± 1.27E+04 |
| | F372A | 5.48E+03 ± 2.18E+03 |
| | F372W | 1.50E+05 ± 3.13E+04 |
| AsGO | *wild type* | 5.37E+05 ± 1.77E+05 |
| | I109V | 4.22E+05 ± 2.04E05 |
| | V202C | 2.78E+05 ± 9.21E+04 |
| | T205V | 6.85E+05 ± 2.69E+05 |
| | S209A | 4.34E+05 ± 2.02E+05 |
| | V372F | 8.62E+03 ± 2.73E+03 |
| | V372Q | 7.50E+03 ± 2.48E+04 |
| | V372A | 7.70E+05 ± 3.24E+05 |
| | V372I | 4.72E+05 ± 1.21E+05 |
| | V372E | 1.53E+03 ± 5.60E+02 |
| | V372L | 5.75E+05 ± 1.85E+05 |

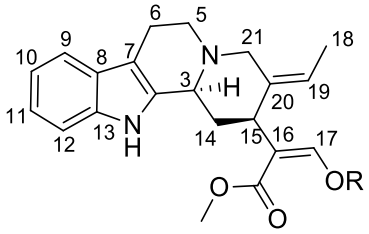
**Table S8.** List of primer sequences used for the assembly of expression vectors and creating mutants.

| Vector Name | Forward Primer | Reverse Primer |
|---|---|---|
| pOPINF-ADH1 | AAGTTCTGTTTCAGGGCCCGGCTGG TGAAATGGCCGCAAAGTC | ATGGTCTAGAAAGCTTTATGCAGATT TCGGCGAATTTCCAATA |
| pOPINF-ADH2 | AAGTTCTGTTTCAGGGCCCGGCTGG AAAATCACCAGAAGAGGAAC | ATGGTCTAGAAAGCTTTACGATTCTG GTGGAGTCAAAGTGTTT |
| pOPINF-ADH3 | AAGTTCTGTTTCAGGGCCCGGCTGG AAAATCACCTGAAGAAGAAC | ATGGTCTAGAAAGCTTTAGGGCTCTG GTGGAGTAAGGG |
| pOPINF-ADH4 | AAGTTCTGTTTCAGGGCCCGGCTGC AAAATCACCAGAAGAG | ATGGTCTAGAAAGCTTTAGGACTCGG CTGGAGGAG |
| pOPINF-ADH5 | AAGTTCTGTTTCAGGGCCCGGCTGC AAAAGCACCAGAAGAAGAGC | ATGGTCTAGAAAGCTTTAGGATGATG CTACTAGTGTGTTTCCG |
| pOPINF-ADH6 | AAGTTCTGTTTCAGGGCCCGGCAGA AAAATCAGAAGAAAGGGCAC | ATGGTCTAGAAAGCTTTAAGAAGATT TTAATGTATTGCCAATG |
| pOPINF-ADH7 | AAGTTCTGTTTCAGGGCCCGGCGAA GTCACCGGAAGTTGAGC | ATGGTCTAGAAAGCTTTAAGCAGATT TCAATGTGTTGCCAA |
| pOPINF-ADH8 | AAGTTCTGTTTCAGGGCCCGGGTTA CACTTCGTACCCTTGTGTTC | ATGGTCTAGAAAGCTTTACGCAGCGA CGGTAGCATTTC |
| pOPINF-ADH9 | AAGTTCTGTTTCAGGGCCCGCTGCT ATCTCCAGTTCTATACAC | ATGGTCTAGAAAGCTTTAAGGAGCTT TCAAGGTATTGC |
| pOPINF-ADH10 | AAGTTCTGTTTCAGGGCCCGGAGAA ACAAGTGCAAATGCCAG | ATGGTCTAGAAAGCTTTACAAGTCTC CGTCCCAAAGTTC |
| pOPINF-ADH11 | AAGTTCTGTTTCAGGGCCCGGAAAG TACAGTACAAATCCCAGAAA | ATGGTCTAGAAAGCTTTAAAAGTCTC CATCCCAAAATTCCTCA |
| pOPINF-ADH12 | AAGTTCTGTTTCAGGGCCCGCCAGA ACGAATGGGAAAAACAGTAG | ATGGTCTAGAAAGCTTTACACATCCC CATCCCAAAGTTCCTCC |
| pOPINF-ADH13 | AAGTTCTGTTTCAGGGCCCGGCTGC AGCATCAGCTGAAACAATAAAG | ATGGTCTAGAAAGCTTTAAGCAGATT TCAATGTATTCTCTATATAAATCACG |
| pOPINF-BAHD1 | AAGTTCTGTTTCAGGGCCCGGCACC CCAGGTGGAGATAGTATC | ATGGTCTAGAAAGCTTTATTTGCTAAA ATCACTGTTTACAAGTGAAAG |
| pOPINF-BAHD2 | AAGTTCTGTTTCAGGGCCCGGTAGT AAGCAAGATGGCGTAC | ATGGTCTAGAAAGCTTTACTTGCTAA AATCACTGTCTACAAG |
| pOPINJ-BAHD3 | AAGTTCTGTTTCAGGGCCCGGCACC CATCCTTCAGAGAATTTC | ATGGTCTAGAAAGCTTTATTTGGCAG AGTACAAAGATGCAAAG |
| pOPINJ-BAHD4 | AAGTTCTGTTTCAGGGCCCGTTCGT CAAACCCTCTTCACCC | ATGGTCTAGAAAGCTTTAAAGCTTCG ACTTCCACTTGGC |
| pOPINJ-BAHD5 | AAGTTCTGTTTCAGGGCCCGAATAT CACAATTCTTTCAAAGGAAAC | ATGGTCTAGAAAGCTTTAGTTTGTAC CATCTGAATCAAGCC |

| | | |
|---|---|---|
| pOPINJ-BAHD6 | AAGTTCTGTTTCAGGGCCCGGCGGC TCAGGTGGTGGTAC | ATGGTCTAGAAAGCTTTAGCGAAAAA GAGAAGCAAAAGG |
| pOPINJ-BAHD7 | AAGTTCTGTTTCAGGGCCCGGCGGC TCAGGTGGTGGTAC | ATGGTCTAGAAAGCTTTAACGAAAAA GAGAAGTAAAAGG |
| pOPINJ-BAHD8 | AAGTTCTGTTTCAGGGCCCGGATAG TTTTGAAGTTAGAGTCCTTTG | ATGGTCTAGAAAGCTTTATCCAATTTC TATGGAAGAAGTGAAGG |
| pOPINJ-BAHD9 | AAGTTCTGTTTCAGGGCCCGGAAGT GCAAATCATTTCTCAAGAGATC | ATGGTCTAGAAAGCTTTAAATCACAG CAACATTTGCAAGC |
| pESC-HIS-AS | ACCCTCACTAAAGGGCGGCCGCAAC CATGGCTAGCTCACTCCCACATC | GTCATCCTTGTAATCCATCGATACTTT GCTAGATGTGGAACCTGC |
| pESC-HIS-GO | ACCCTCACTAAAGGGCGGCCGCAAC CATGGAGTTTTCTTTCTCCTCACCC | GTCATCCTTGTAATCCATCGATACCT CATTTGCAAGATGTGGAAC |
| pESC-HIS-RHS | ACCCTCACTAAAGGGCGGCCGCAAC CATGGCTAGCTCACTCCCACATC | GTCATCCTTGTAATCCATCGATACTTT GCTAGATGTGGAACCTGC |
| pESC-HIS-GO_ I109V | CTGTTGCTAGCGTTATGACTTACAAC | GTTGTAAGTCATAACGCTAGCAACAG |
| pESC-HIS-GO_ V202C | GATGATCAAGATATGTAAAAGAACAT CCT | AGGATGTTCTTTTACATATCTTGATCA TC |
| pESC-HIS-GO_ T205V | ATAGTGAAAAGAGTTTCCTTGCTTTC C | GGAAAGCAAGGAAACTCTTTTCACTA T |
| pESC-HIS-GO_ S209A | CATCCTTGCTTGCTGGAACCCCTC | GAGGGGTTCCAGCAAGCAAGGATG |
| pESC-HIS-GO_ V372F | CTCCTCTCCTATTCCCCAGAGAATG | CATTCTCTGGGGAATAGGAGAGGAG |
| pESC-HIS-GO _ V202C/T205V | GATCAAGATATGTAAAAGAGTTTCCT TGCTT | AAGCAAGGAAACTCTTTTACATATCTT GATC |
| pESC-HIS-GO _ T205V/S209A | GTTTCCTTGCTTGCTGGAACCCCTC | GAGGGGTTCCAGCAAGCAAGGAAAC |
| pESC-HIS-RHS_ V112I | CAGTGTGGCGAAAATCATCTGTTAC AAC | GTTGTAACAGATGATTTTCGCCACAC TG |
| pESC-HIS-RHS_ C205V | GAGAAACTGATCAAGATAGTTAAAAA AGTATCATTT | AAATGATACTTTTTTAACTATCTTGAT CAGTTTCTC |
| pESC-HIS-RHS_ V208T | AAGATATGCAAAAAAACATCATTTTA TTCCGCCG | GGCGGAATAAAATGATGTTTTTTTGC ATATCTT |
| pESC-HIS-RHS_ S212A | AGTATCATTTTATGCTGCCGCCCCC AT | AATGGGGGCGGCAGCATAAAATGAT ACT |
| pESC-HIS-RHS_ F372V | GCTCCCCTCTTAGTTCCCAGAGAAG CC | GGCTTCTCTGGGAACTAAGAGGGGA GC |

| | | |
|---|---|---|
| pESC-HIS RHS_ C205V/V208T | GAGAAACTGATCAAGATAGTTAAAAA AACATCATTT | AAATGATGTTTTTTTAACTATCTTGAT CAGTTTCTC |
| pESC-HIS-RHS_ V208T/S212A | AACATCATTTTATGCTGCCGCCCCC AT | AATGGGGGCGGCAGCATAAAATGAT GTT |

**Table S9.** $^{1}$H NMR spectral data (δ) for geissoschizine and *O*-methyl geissoschizine ($CDCl_3$, 400 MHz).




Geissoschizine: R=H
*O*-methyl-geissoschizine R=Me

| | **Geissoschizine** | ***O*-Methyl-Geissoschizine** |
|---|---|---|
| **Position** | **$^{1}$H (mult., *J* (Hz)** | **$^{1}$H (mult., *J* (Hz)** |
| **1** | 7.72 (br s) | 7.86 (br s) |
| **3** | 3.86 (m overlapped) | 3.56 (br d, 12.4) |
| **5** | α 2.75 (td, 11.6, 4.0) | |
| | β 3.22 (m) | β 3.11 (m) |
| **6** | α 3.10 (m) | α 3.00 (m) |
| | β 2.85 (m) | B 2.82 (br d overlapped, 12.1) |
| **9** | 7.51 (d, 7.7) | 7.48 (d, 7.6) |
| **10** | 7.14 (td, 7.5, 1.1) | 7.14 (8.2, 1.2) |
| **11** | 7.20 (td, 7.5, 1.3) | 7.20 (td 8.0, 1.4) |
| **12** | 7.35 (d, 8.0) | 7.35 (d, 8.0) |
| **14** | α 2.64 (ddd, 13.5, 11.4, 5.7) | α 1.91 (m overlapped) |
| | β 2.14 (m) | β 1.97 (m overlapped) |
| **15** | 4.51 (d, 12.1, 1.4) | 3.7 (m overlapped) |
| **17** | 7.54 (s) | 7.37 (s) |
| **18** | 1.85 (dd, 6.6, 1.8) | 1.55 (br d, 7.0) |
| **19** | 5.42 (q, 7.3) | 5.49 (q, 7.1) |
| **21** | 3.98 (dt, 13.6, 2.0) | 3.32 (d, 12.3) |
| | 3.21 (d overlapped, 14.4) | 3.56 (d 12.3) |
| **COOMe** | 3.69 (s, 3H) | 3.73 (s, 3H) |
| **COMe** | | 3.86 (s, 3H) |

**Geissoschizine**

**$^{1}$H-NMR (400 MHz, $CDCl_3$):** *δ* (p.p.m) 7.72 (1H, br s, H-N), 7.54 (1H, s, H-17), 7.51 (1H, d, *J*= 7.7 Hz, H-9), 7.35 (1H, d, *J*= 8.0 Hz, H-12), 7.37 (1H, s, H-17) 7.20 (1H, td, *J*= 7.5 Hz, *J*= 1.3 Hz, H-11), 7.14 (1H, td, *J*= 7.5 Hz, *J*= 1.1 Hz, H-10), 5.42 (1H, q, *J*= 7.3 Hz, H-19), 4.51 (1H, d, *J*= 12.1 Hz, *J*= 1.4 Hz, H-15), 3.98 (dt, *J*= 13.6, *J*= 2.0, H-21β) 3.86 (1H, m overlapped, H-3), 3.69 (3H, s, COOMe), 3.22 (1H, m, H-5β), 3.21 (1H, d overlapped, *J*= 14.4 Hz, H-21α), 3.10 (1H, m, H-6α), 2.85 (1H, m, H-6β), 2.75 (1H, td, *J*= 11.6

Hz, *J*= 4.0 Hz, H-5α), 2.64 (1H, ddd, *J*= 13.5 Hz, *J*= 11.4 Hz, *J*= 5.7 Hz, H-14α), 2.14 (1H, m, H-14β), 1.85 (3H, dd, *J*= 6.6 Hz, *J*= 1.8 Hz, Me).

***O*-Methyl-Geissoschizine**

**$^1$H-NMR (400 MHz, $CDCl_3$):** *δ* (p.p.m) 7.86 (1H, br s, H-N), 7.48 (1H, d, *J*= 7.6 Hz, H-9), 7.37 (1H, s, H-17), 7.31 (1H, d, *J*= 8.0 Hz, H-12) 7.14 (1H, td, *J*= 8.0 Hz, *J*= 1.2 Hz, H-11), 7.09 (1H, td, *J*= 8.2 Hz, *J*= 1.2 Hz, H-10), 5.49 (1H, q, *J*= 7.1 Hz, H-19), 3.86 (1H, m overlapped, H-3), 3.86 (3H, s, OMe), 3.73 (3H, s, COOMe), 3.70 (1H, m overlapped, H-15), 3.56 (1H, d, *J*= 12.3, H-21β), 3.32 (1H, d, *J*= 12.3 Hz, H-21α), 3.11 (1H, m, H-5β), 3.00 (1H, m, H-6α), 2.82 (1H, m, H-6β), 1.97-1.91 (3H, m, H-14α, H-14β), 1.69 (3H, d, *J*= 6.9 Hz, Me).

## Supplementary Figures S1-S43

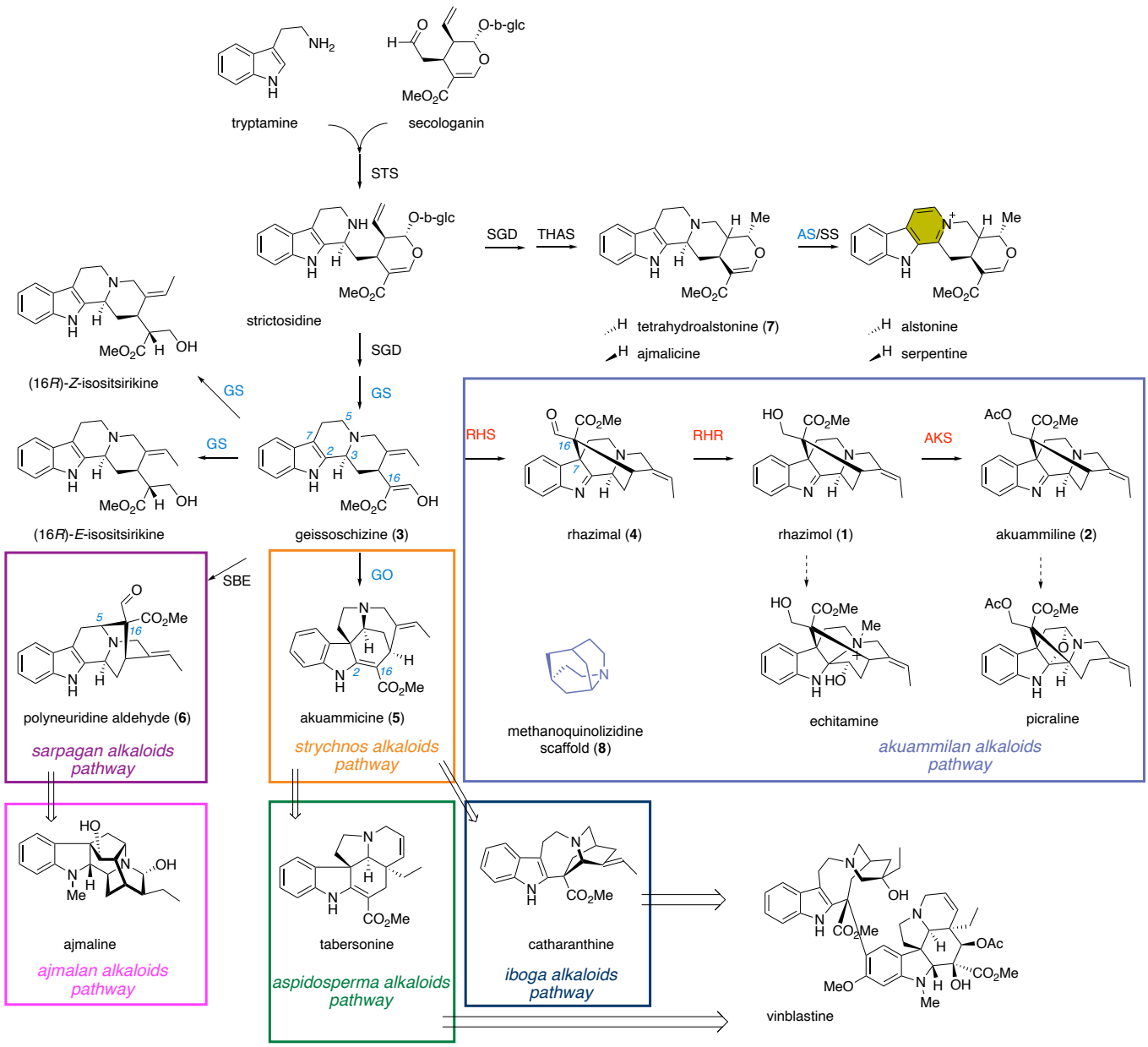



**Figure S1. Schematic overview of monoterpene indole alkaloids (MIAs) metabolic pathways in *Alstonia scholaris*.** Strictosidine is the common precursor to MIAs, formed by condensation of tryptamine and secologanin by strictosidine synthase (STS). Strictosidine aglycon generated by strictosidine glucosidase (SGD) is subject to structural rearrangements by the concerted activity of alcohol dehydrogenases (GS, THAS) and P450s (GO, RHS, AS) to generate diverse chemical scaffolds.

The cytochrome p450 enzymes acting on reduced strictosidine aglycon compounds geissoschizine (**3**) and tetrahydrolastonine (**7**), catalysing either the cyclization between C2-C16 (highlighted in orange) of **3** to form the *strychnos* alkaloid **5,**[7] the cyclization between C5-C16 of **6** (highlighted in magenta) to form the *sarpagan* alkaloid **6**,[8] the cyclization between C7-C16 of **3** (highlighted in purple) to form the *akuammilan* alkaloid **1**, the aromatization (highlighted in green) of tetrahydroalstonine to alstonine.[8] Downstream steps RHR and AKS are described in detail in the manuscript. The enzymes highlighted in red colour (RHS, RHR and AKS) were first characterized in this paper. The enzymes highlighted in blue colour (GS, GO

and AS) were reported enzymes that identified in *A. scholaris*. Abbreviations: STS, strictosidine synthase; SGD, strictosidine glucosidase; THAS, tetrahydroalstonine synthase; SS, serpentine synthase; GS, geissoschizine synthase; GO, geissoschizine oxidase; AS, alstonine synthase; RHS, rhazimal synthase; RHR, rhazimal reductase; AKS, akuammiline synthase.

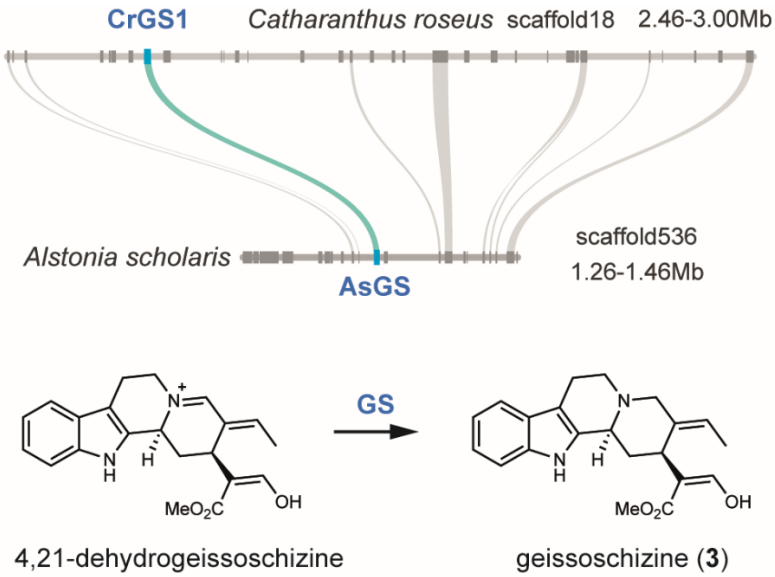



**Figure S2. Synteny of geissoschizine synthase (GS) from *C. roseus* and *A. scholaris*.** Geissoschizine synthase (GS) can catalyse the reduction of iminium 4,21-dehydrogeissoschizine to geissoschizine (**3**).

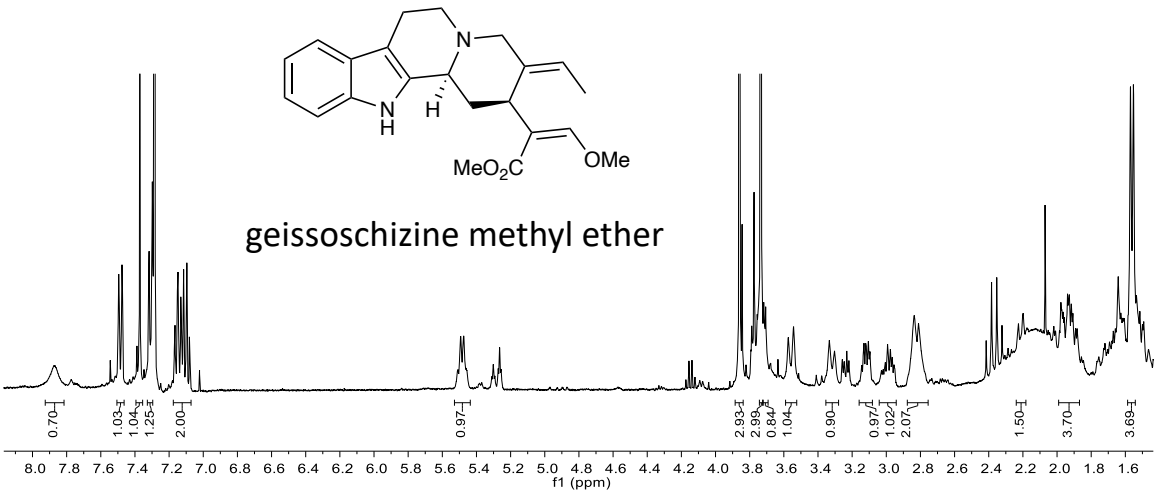



**Figure S3. $^{1}$H-NMR spectrum of *O*-methyl-geissoschizine ($CDCl_3$, 400 MHz) (67% purity).** The isolation and synthesis have been performed as it is described by Dang et. al Nat Chem Bio 2018.[8]

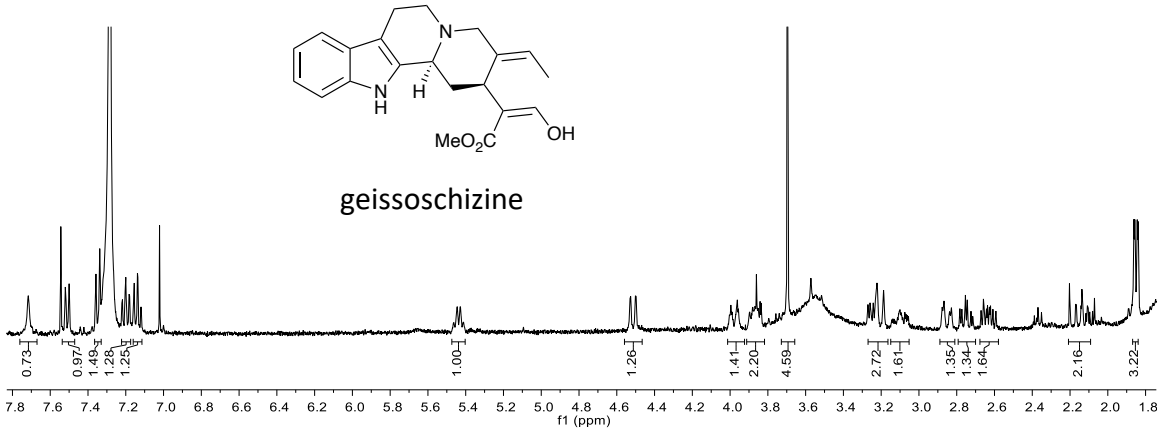



**Figure S4. $^{1}$H-NMR spectrum of geissoschizine (CDCl3, 400 MHz).** The isolation and synthesis have been performed as it is described by Dang et. al Nat Chem Bio 2018.[8]

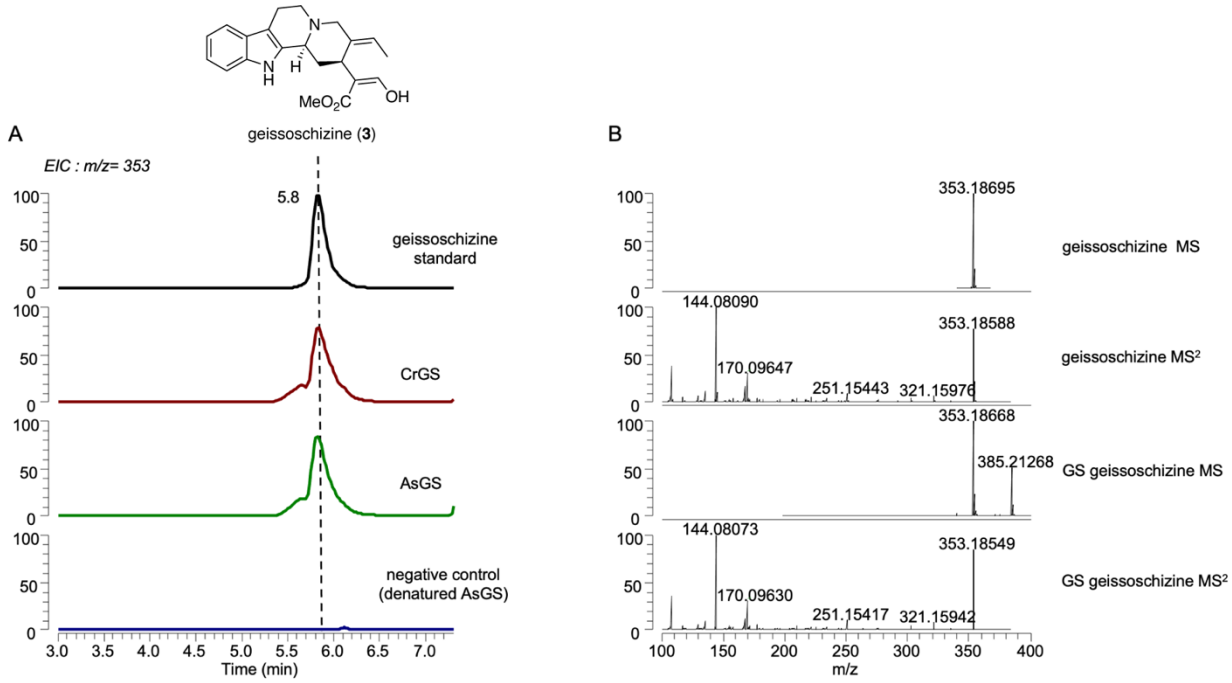



**Figure S5.** ***In vitro*** **catalytic activity of CrGS and AsGS against strictosidine aglycone. (A)** The extracted ion chromatograms EIC at *m/z* 353 for geissoschizine (**3**) from geissoschizine semisynthetic standard, CrGS and AsGS enzyme assays. Strictosidine was incubated with heterologously expressed and purified *C. roseus* strictosidine glucosidase (SGD) for 30 min. To generate the strictosidine aglycon followed by the addition of AsGS (42928_t) enzyme and NADPH.[5, 7] The analysis by LCMS of the enzyme assays demonstrates the geissoschizine synthase activity of 42928_t in comparison to semisynthetic **3.** In the enzyme assay of 42928_t where a peak with *m/z* 353.18 is detected at 5.8 min with an $MS^2$ spectra with major fragments at *m/z* 170.1, *m/z* 144.1 and *m/z* 108.1, identical to geissoschizine. **(B)** The MS and $MS^2$ spectra of geissoschizine (**3**) (rt 5.8 min) from geissoschizine semisynthetic standard and enzyme assay of AsGS.

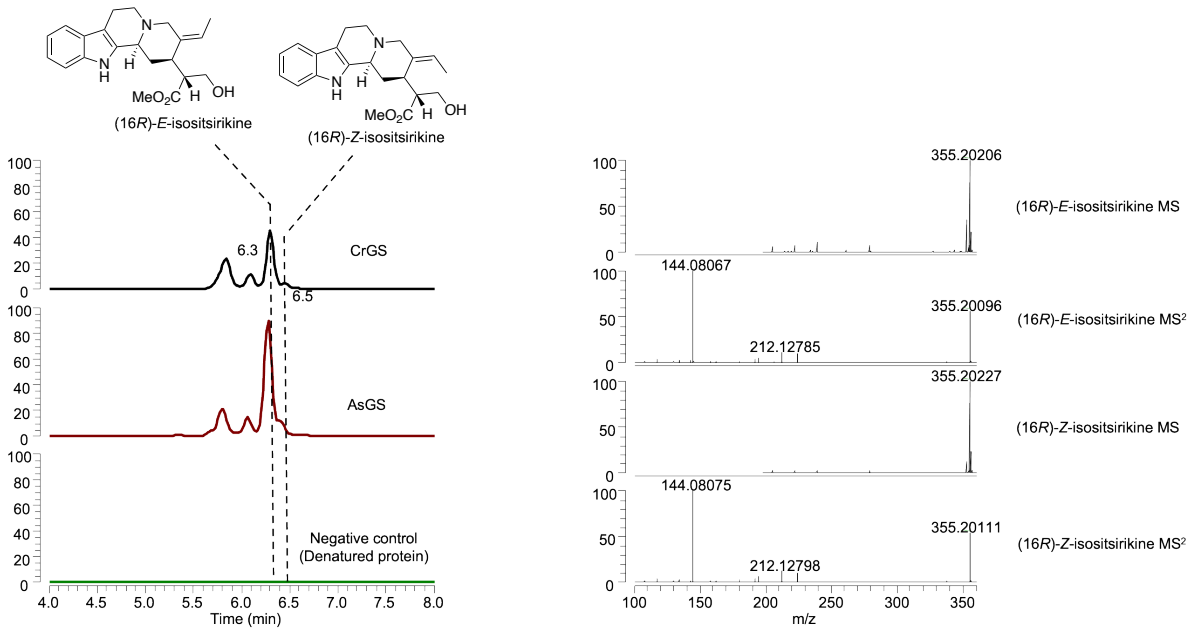



**Figure S6.** ***In vitro*** **catalytic activity of CrGS and AsGS against strictosidine aglycone. (A)** The extracted ion chromatograms EIC at *m/z* 355 for (16*R*)-*Z*-isositsirikine and 16*R*-*E*-isositsirikine from CrGS and AsGS enzyme assays. In the enzyme assay of 42928_t, one peaks with *m/z* 355.20 was detected at 6.3 min, with two identical $MS^2$ spectra with major fragments at *m/z* 144.08 and *m/z* 212.13. Based on the comparison to enzyme assay of 42928_t with CrGS, the peak was identified as (16*R*)-*E*-isositsirikine and (16*R*)-*Z*-isositsirikine due to identical retention times, MS and $MS^2$ spectra. **(B)** The MS and $MS^2$ spectra of (16*R*)-*E*-isositsirikine (rt 6.3 min) and (16*R*)-*Z*-isositsirikine (rt 6.5 min) from enzyme assay of CrGS and AsGS.

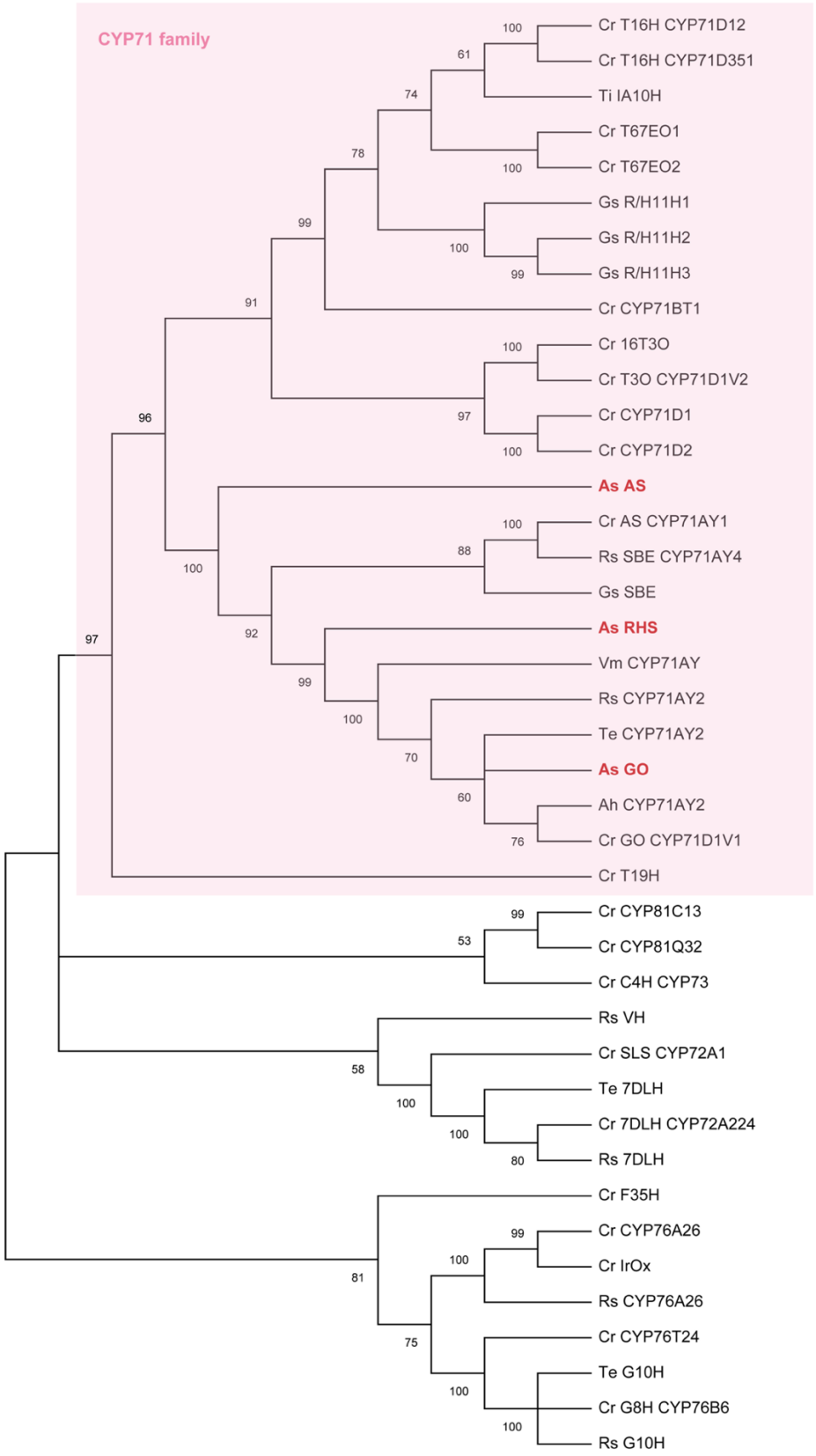



**Figure S7. The neighbour-joining phylogenetic tree for cytochrome p450 enzymes from *A. scholaris* and previously reported cytochrome p450 enzymes involved in MIAs metabolic pathways from plants of Apocynaceae family.** The enzymes from *A. scholaris* AS, AsGO and AsRHS that are characterised in present study are highlighted in red. The accession numbers of enzymes amino acid sequences used to generate phylogenetic tree are listed in methods. Bootstrap frequencies for each clade were based on 1000 iterations. Phylogenetic tree was built using the MEGA software.

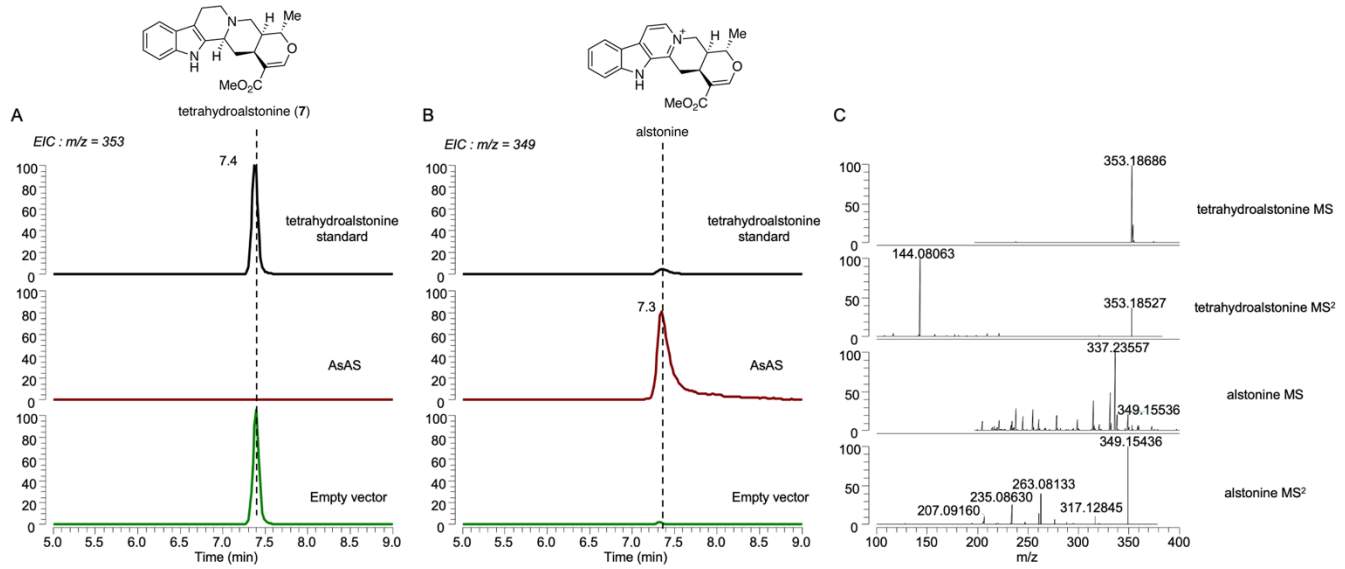



**Figure S8. *In vitro* catalytic activity of AsAS against tetrahydroalstonine (7). (A)** The extracted ion chromatograms EIC at *m/z* 353 for tetrahydroalstonine (**7**) from tetrahydroalstonine standard, AsAS enzyme assay and empty vector as negative control. When AsAS (45522_t) was assayed with tetrahydroalstonine (**7**) substrate, substrate consumption was observed and a new peak with *m/z* 349.19 was observed at 7.3 min. The $MS^2$ spectrum of the new peak exhibits major fragments at *m/z* 317.13, 263.08, 235.09 and 207.09 similar $MS^2$ spectrum to alstonine. **(B)** The extracted ion chromatograms EIC at *m/z* 349 for alstonine from tetrahydroalstonine standard, AsAS enzyme assays and empty vector. **(C)** The MS and $MS^2$ spectra of tetrahydroalstonine (**7**) (rt 7.4 min) and alstonine (rt 7.3 min).

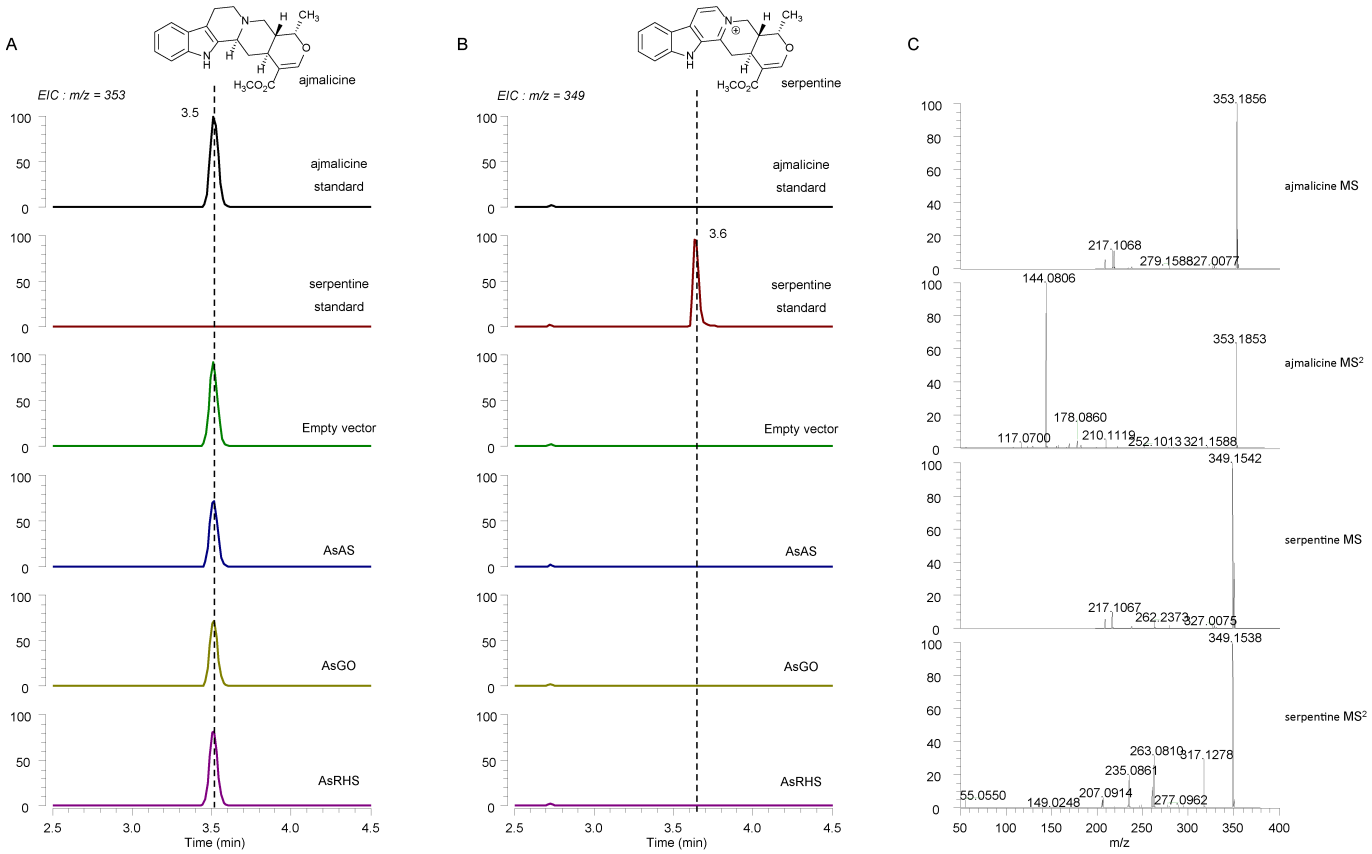



**Figure S9.** ***In vitro*** **catalytic activity of AsAS, AsGO, AsRHS against ajmalicine. (A)** The extracted ion chromatograms EIC at *m/z* 353 for ajmalicine standard, AsAS enzyme assay, AsGO enzyme assay, and AsGO enzyme assay and empty vector. **(B)** The extracted ion chromatograms EIC at *m/z* 349 for serpentine standard, AsAS enzyme assay, AsGO enzyme assay, and AsGO enzyme assay and empty vector. **(C)** The MS and $MS^2$ spectra of ajmalicine (rt 3.5 min) and serpentine (rt 3.6 min).

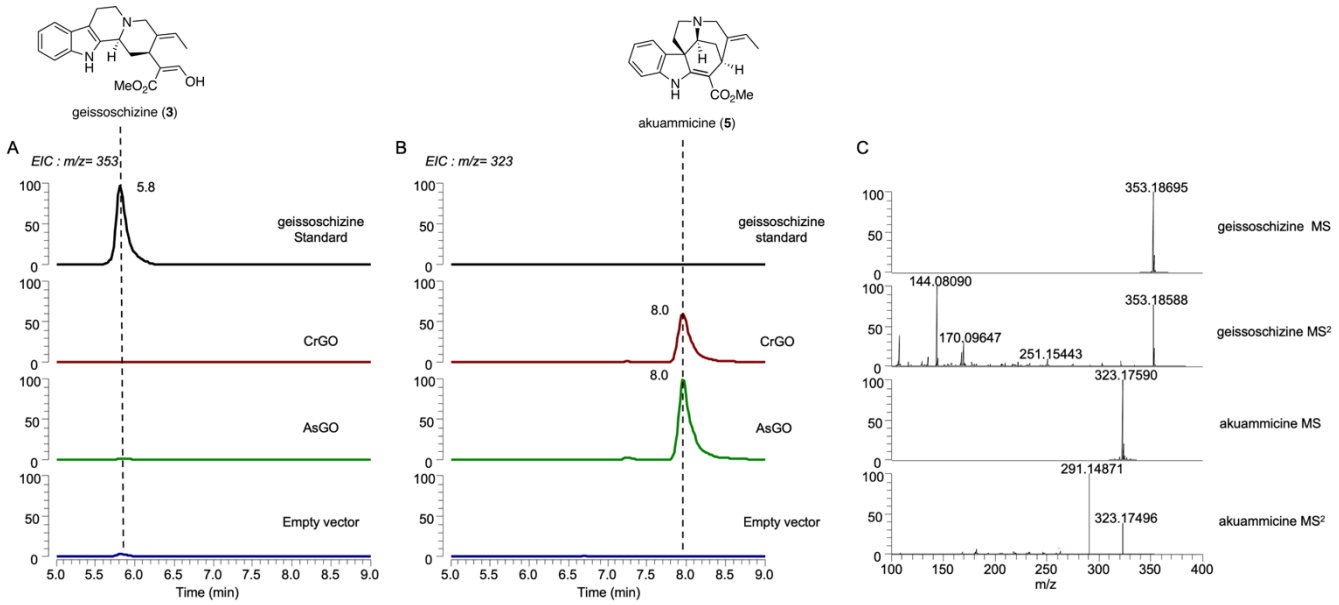



**Figure S10.** ***In vitro*** **catalytic activity of AsGO against semisynthetic geissoschizine (3). (A)** The extracted ion chromatograms EIC at *m/z* 353 for geissoschizine (**3**) from semisynthetic geissoschizine standard, CrGO and AsGO enzyme assays. Microsomal protein bearing the AsGO (43978_t) against geissoschizine (**3**), the analysis by LCMS of enzymatic reactions for AsGO shows that similar product profile to the enzyme assay of CrGO,[7] with the consumption of substrate (geissoschizine) and the appearance of a peak at 8.1 min with *m/z* 323.17 and a major fragment *m/z* 291.15 at $MS^2$ spectrum, identified as akuammicine (**5**). **(B)** The extracted ion chromatograms EIC at *m/z* 323 for akuammicine (**5**) from geissoschizine standard, CrGO and AsGO enzyme assays. **(C)** The MS and $MS^2$ spectra of geissochizine (**3**) (rt 5.8 min) and akuammicine (**5**) (rt 8.1 min).

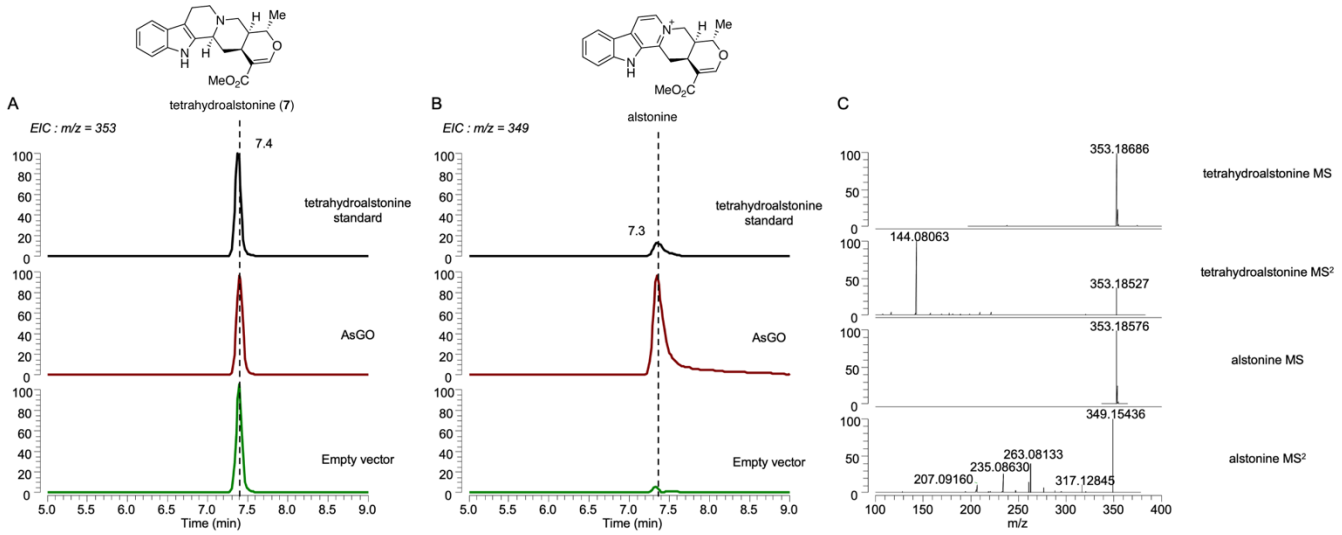



**Figure S11.** ***In vitro*** **catalytic activity of AsGO against tetrahydroalstonine (7). (A)** The extracted ion chromatograms EIC at *m/z* 353 for tetrahydroalstonine (**7**) from tetrahydroalstonine standard, denatured CrGO and AsGO enzyme assays. **(B)** The extracted ion chromatograms EIC at *m/z* 349 for alstonine from tetrahydroalstonine standard, denatured AsAS and AsAs enzyme assays. **(C)** The MS and $MS^2$ spectra of tetrahydroalstonine (**7**) (rt 7.4 min) and alstonine (rt 7.3 min).

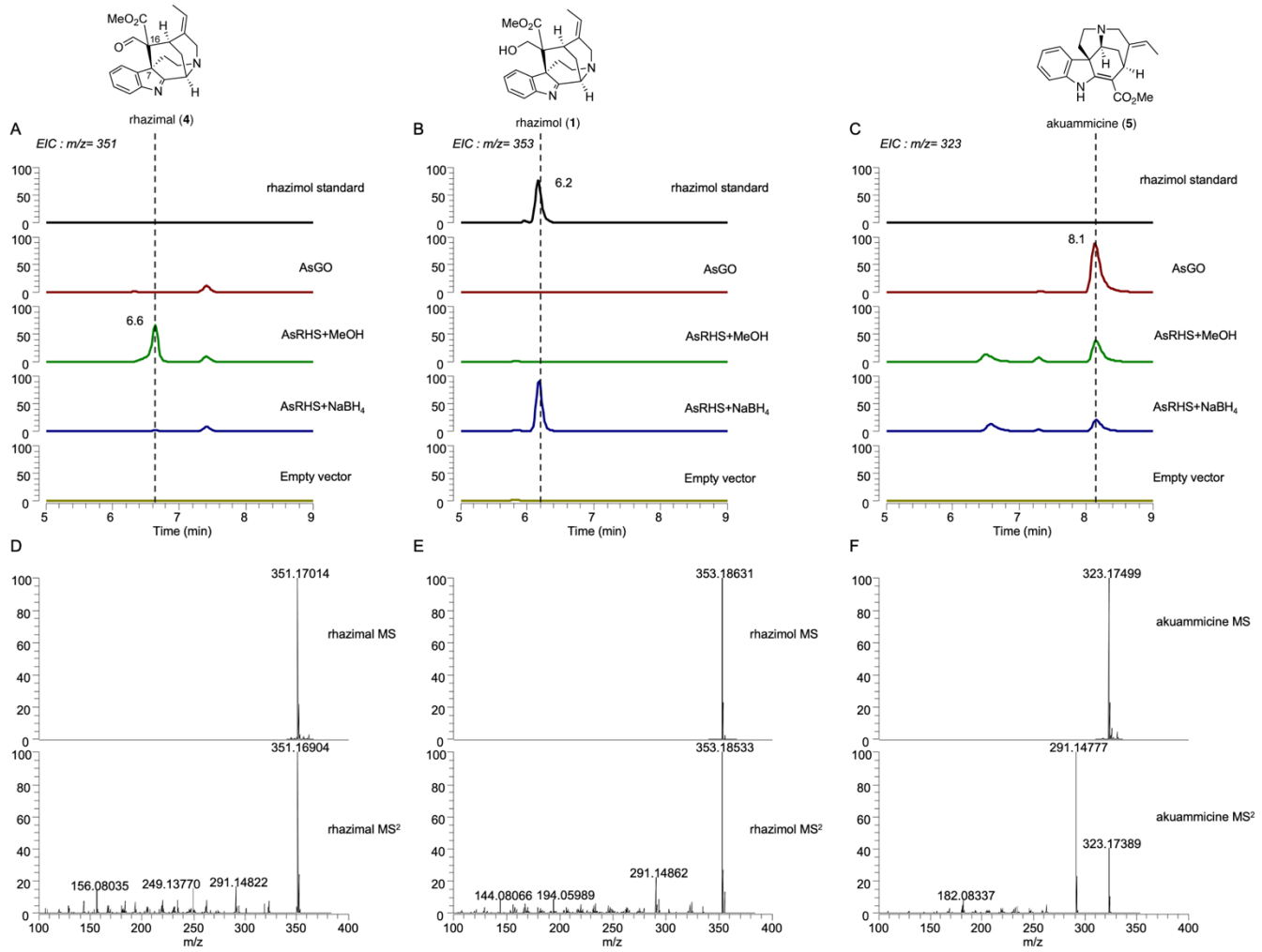



**Figure S12.** ***In vitro*** **catalytic activity of AsGO and AsRHS against semisynthetic geissoschizine (3).** **(A)** The extracted ion chromatograms EIC at *m/z* 351 for rhazimal (**4**) from rhazimol standard, AsGO enzyme assay, RHS enzyme assay stopped by MeOH and by $NaBH_4$. **(B)** The extracted ion chromatograms EIC at *m/z* 353 for rhazimol (**1**) from rhazimol standard, AsGO enzyme assay, RHS enzyme assay stopped by MeOH and by $NaBH_4$. The microsomal protein fraction harbouring the RHS (46693_t) showed activity towards **3** resulting to formation of **5** (*m/z* 323.17 at 8.1 min) and a second a peak at 6.6 min with a *m/z* 351 and major fragments *m/z* 291.15 in $MS^2$ spectra, when the enzymatic reaction mixture was quenched with methanol. The synthetic standard of **1** elutes at 6.2 min with a *m/z* 353 corresponding to the molecular ion $[M+H]^+$ and with a major fragment in $MS^2$ at *m/z* 291.15. Since the aldehyde function of **4** could be easily reduced by mild reducing agents to alcohol and therefore converted to **1**, we decided to stop the reaction using $NaBH_4$ a widely used mild reducing agent. When the enzyme assay of microsomal protein with RHS was quenched by $NaBH_4$, a new peak at 6.2 min appeared with *m/z* 353 and $MS^2$ spectra identical to **1**. **(C)** The extracted ion chromatograms EIC at *m/z* 323 for **5** from rhazimol standard, AsGO enzyme assay, RHS enzyme assay stopped by MeOH and by $NaBH_4$. **(D)** The MS and $MS^2$ spectra of **4** (rt 6.6 min) from RHS enzyme assay stopped by MeOH. **(E)** The MS and $MS^2$ spectra of **1** (rt 6.2 min) from AsRHS enzyme assay stopped by $NaBH_4$. **(F)** The MS and $MS^2$ spectra of **5** (rt 8.1 min) from GO enzyme assay.

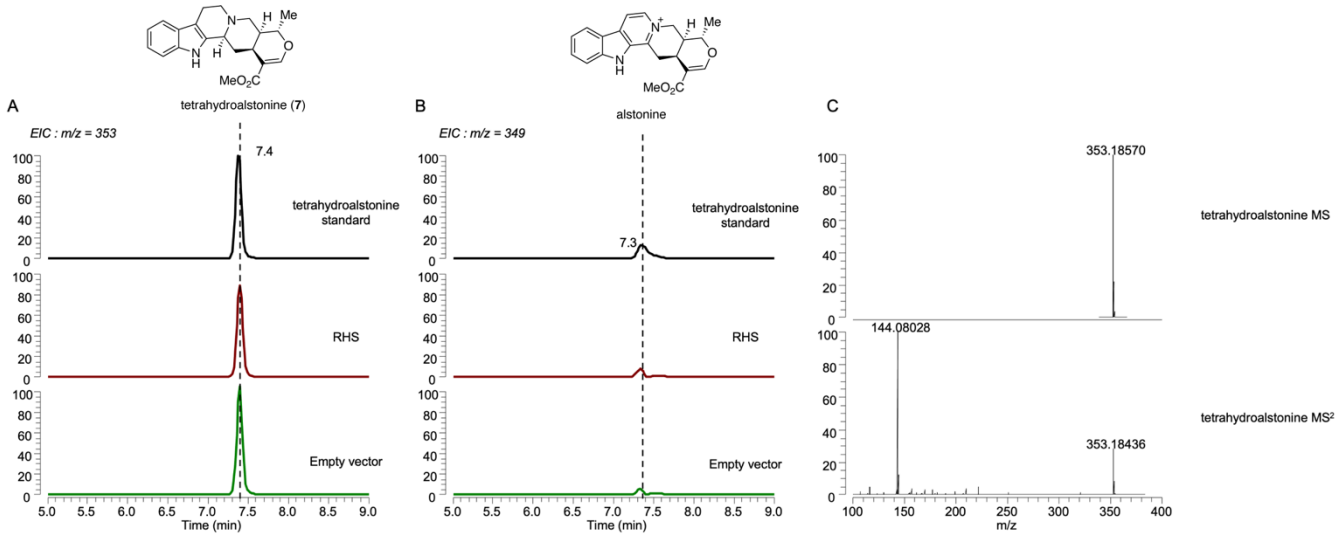



**Figure S13.** ***In vitro*** **catalytic activity of AsRHS against tetrahydroalstonine (7). (A)** The extracted ion chromatograms EIC at *m/z* 353 for tetrahydroalstonine (**7**) from tetrahydroalstonine standard, RHS enzyme assay and empty vector. **(B)** The extracted ion chromatograms EIC at *m/z* 349 for alstonine from tetrahydroalstonine standard, RHS enzyme assay and empty vector. **(C)** The MS and $MS^2$ spectra of tetrahydroalstonine (**7**) (rt 7.4 min).

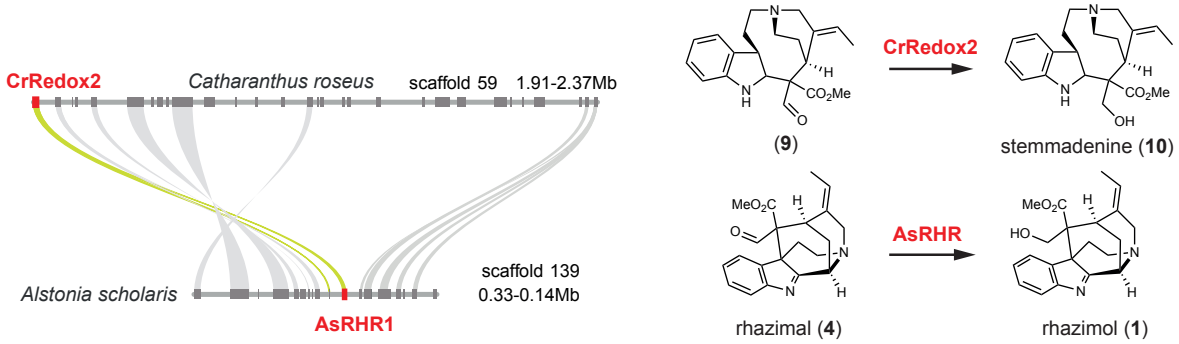



**Figure S14. Synteny between alcohol dehydrogenase Redox2 from *C. roseus* and rhazimal reductase AsRHR1 from *A. scholaris*.** CrRedox2 reduces the aldehyde (**9**) to alcohol stemmadenine (**10**). According to plausible hypothesis for the first steps in metabolic pathway of akuammiline alkaloids (Figure S1), following the formation of rhazimal (**4**), a reduction is expected to convert the aldehyde group of **4** to alcohol **1**. A similar transformation is catalysed by the NADPH dependent oxidoreductase enzyme Redox 2[12] in metabolic pathway of vinblastine in *C. roseus* catalysing the reduction of aldehyde (**9**) to alcohol resulting to formation of stemadenine (**10**). In the synteny between the genomes of *C. roseus* and *A. scholaris*, it was found an NADPH dependent oxidoreductase gene *15881_t* in syntenic relationship with *C. roseus* Redox2 with 80.2% sequence identity. Therefore, we decided to test the 15881_t for rhazimal reduction. In this paper, we report for first time rhazimal reductase activity of AsRHR1 to reduce the aldehyde rhazimal (**4**) to alcohol rhazimol (**1**).

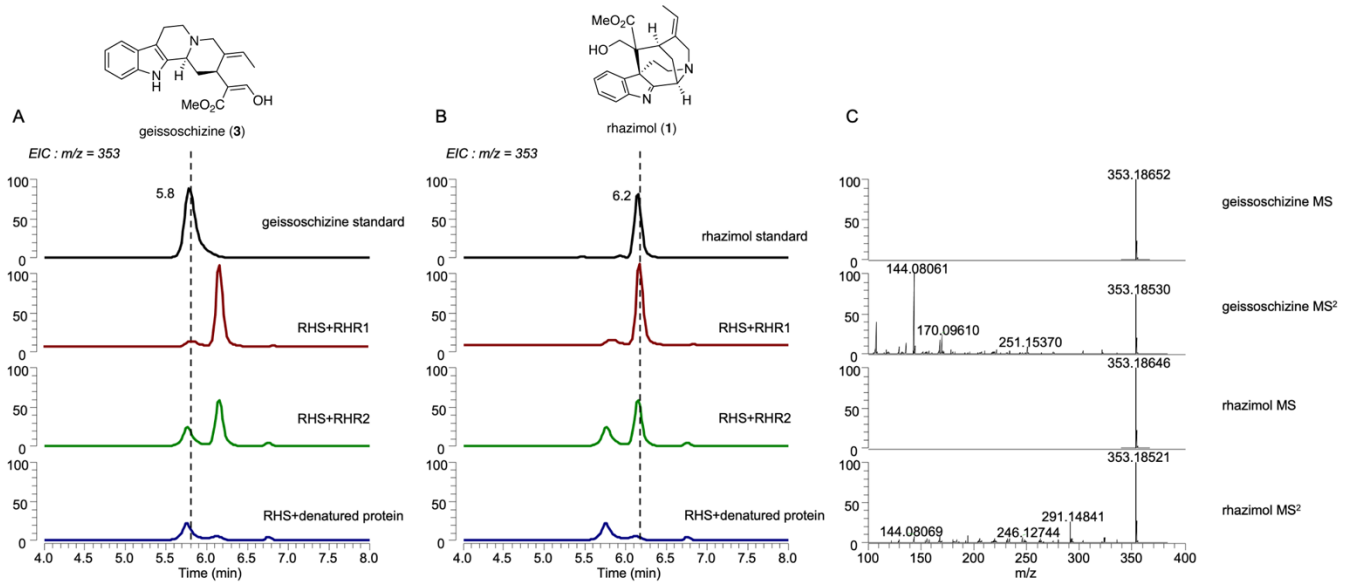



**Figure S15**. ***In vitro*** **catalytic activity of RHR against rhazimal (4). (A)** The extracted ion chromatograms EIC at *m/z* 353 for geissoschizine (**3**) from geissoschizine standard, RHR enzyme assays and negative control (denatured protein). To assay the enzymatic activity for reduction of rhazimal, geissoschizine (**3**) was incubate with microsomal protein RHS to generate the substrate rhazimal (**4**) (*m/z* 351) and the purified heterologous oxidoreductase protein was added after 30 minutes. The enzyme assays were analysed by LC-MS and we compare the chromatographic and MS data with those of the synthetic standard rhazimol (**1**), we identified a peak at *m/z* 353 at the same retention time and the same MS and $MS^2$ spectra to synthetic **1**. **(B)** The extracted ion chromatograms EIC at *m/z* 353 for rhazimol (**1**) from rhazimol standard, RHR1 enzyme assay, RHR2 enzyme assay and negative control (denatured protein). **(C)** The MS and $MS^2$ spectra of geissoschizine (**3**) (rt 5.8 min) and rhazimol (**1**) (rt 6.2 min) from AsRHR1 enzyme assay.

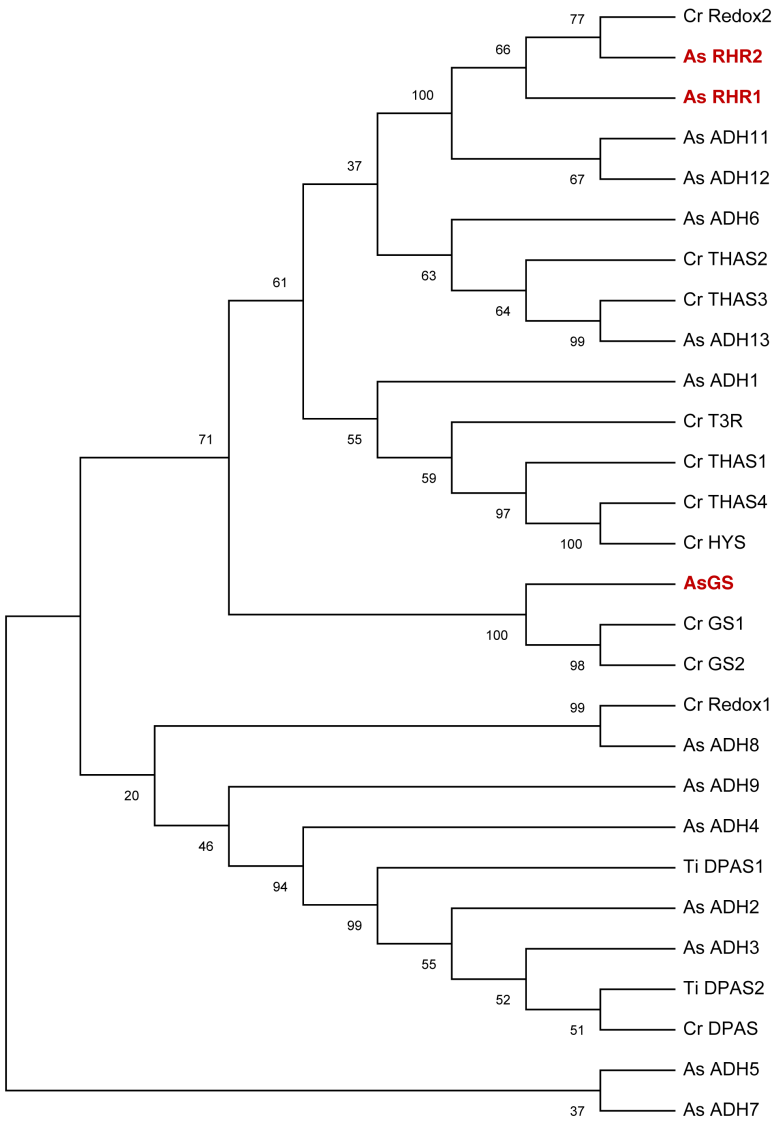



**Figure S16. The neighbour-joining phylogenetic tree for alcohol dehydrogenase/reductase (ADH) candidate enzymes from *A. scholaris* and characterized ADHs involved in MIAs metabolic pathways from plants of Apocynaceae family.** The enzymes from *A. scholaris* GS and RHR1/2 that are characterised in present study are highlighted in red. Bootstrap frequencies for each clade were based on 1000 iterations. Phylogenetic tree was built using the MEGA software. The accession numbers of amino acid sequences used to generate phylogenetic tree are as follows: TiDPAS1 (*T. iboga* dihydroprecondylocarpine acetate synthase 1, QED20594.1), TiDPAS2 (*T. iboga* dihydroprecondylocarpine acetate synthase 2, QED20595.1), CrGS1 (*C. roseus* geissoschizine synthase 1 CrADH14, AHK60846.1), CrGS2 (*C. roseus* geissoschizine synthase 1 CrADH13, AHK60845.1), CrTHAS1 (*C. roseus* tetrahydroalstonine synthase, AKF02528.1), CrTHAS2 (*C. roseus* tetrahydroalstonine synthase 2, ANQ45223.1), CrTHAS3 (*C. roseus* tetrahydroalstonine synthase 3, ANQ45222.1), CrTHAS4 (*C. roseus* tetrahydroalstonine synthase 4,ANQ45224.1), CrDPAS (*C. roseus* dihydroprecondylocarpine acetate synthase, ANQ45231.1), CrHYS (*C. roseus* heteroyohimbine synthase, ANQ45225.1), CrRedox1 (*C. roseus* Redox 1, AVM85917.1), CrRedox2 (*C. roseus* Redox 2, AVM85918.1), CrT3R (*C. roseus* Tabersonine 3-reductase, A0A161CAI1.1).

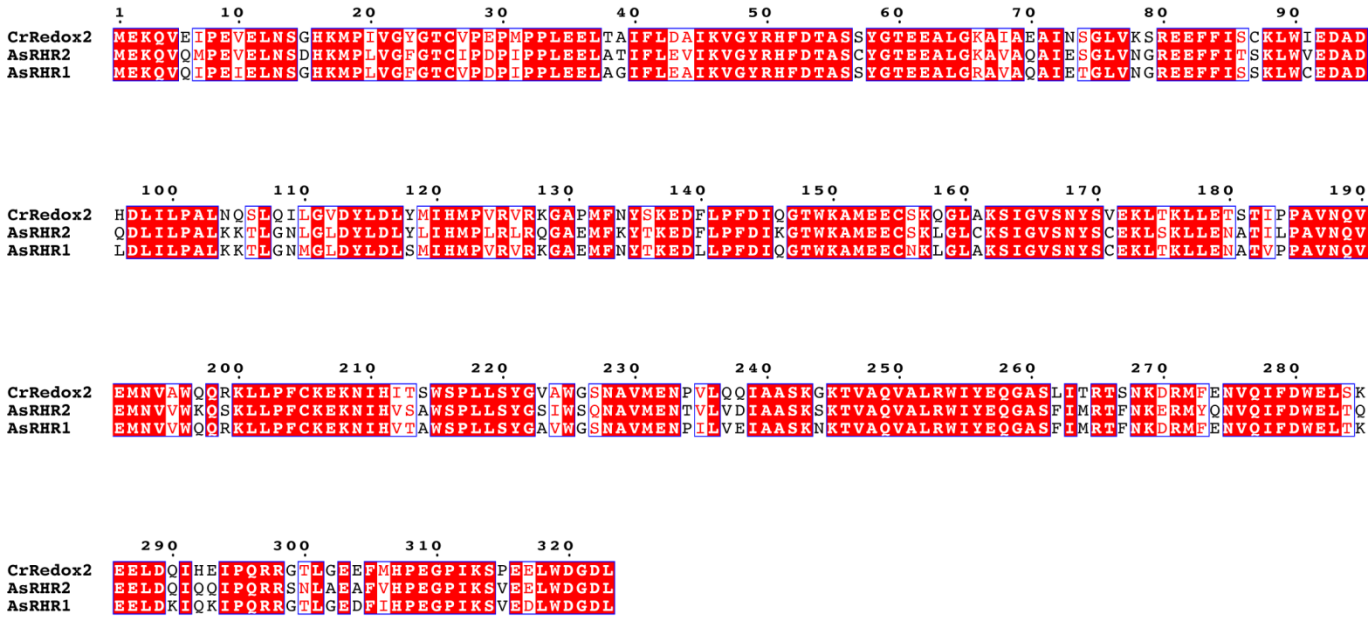



**Figure S17. Sequence alignment of alcohol dehydrogenase Redox2 from *C. roseus* and RHRs from *A. scholaris*.** Numbering corresponds to CrRedox2. Identical and similar amino acids are highlighted in red, respectively.

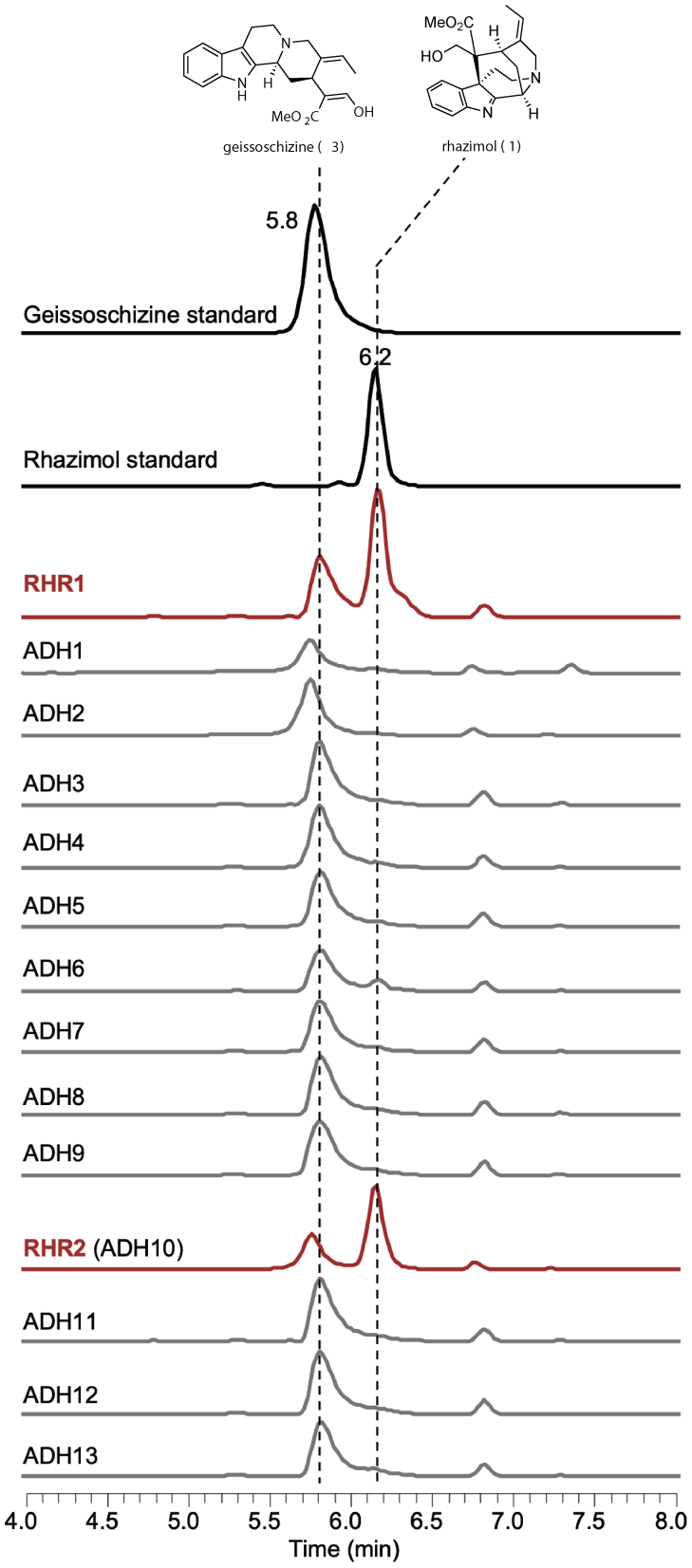



**Figure S18. Selected LC-MS chromatograms at *m/z* 353 of geissoschizine (3) and rhazimol (1) standards and the screened inactive alcohol dehydrogenase (ADH) candidates against rhazimal (4).** All ADH homologues that were assayed in this study against geissoschizine (**3**) incubated with AsRHS failed to show significant turnover as shown in this figure. The function of these enzymes remains unidentified, except for RHR2 (ADH10, *A. scholaris* rhazimal reductase AsRHR2), which was characterized in this paper.

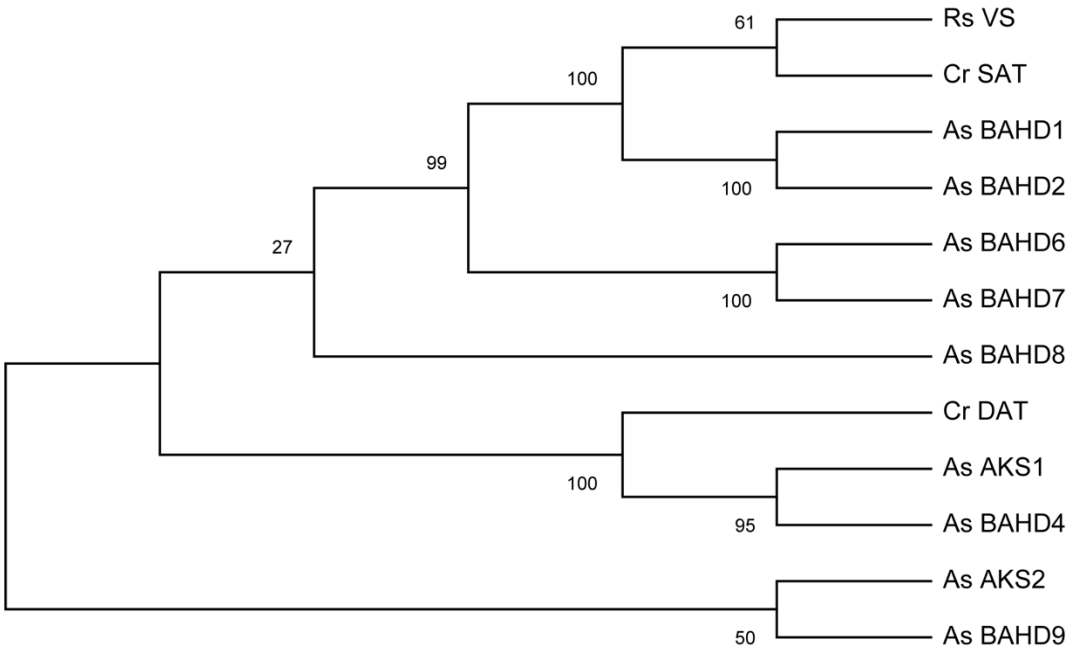



**Figure S19. The neighbour-joining phylogenetic tree for BAHD candidates from *A. scholaris* and characterized BAHD acetyl transferases from monoterpene indole alkaloids biosynthesis pathway in Apocynaceae family.** The enzymes from *A. scholaris* AKS1/2 that are characterised in present study are highlighted in red. Bootstrap frequencies for each clade were based on 1000 iterations. Phylogenetic tree was built using the MEGA software. The accession numbers of amino acid sequences used to generate phylogenetic tree are as follows: RsVS (*R. serpentina* vinorine synthase, CAD89104.2), CrSAT (*C. roseus* stemmadenine O-acetyltransferase, AVM85919.1), CrDAT (*C. roseus* deacetylvindoline 4-O-acetyltransferase, AAC99311.1).

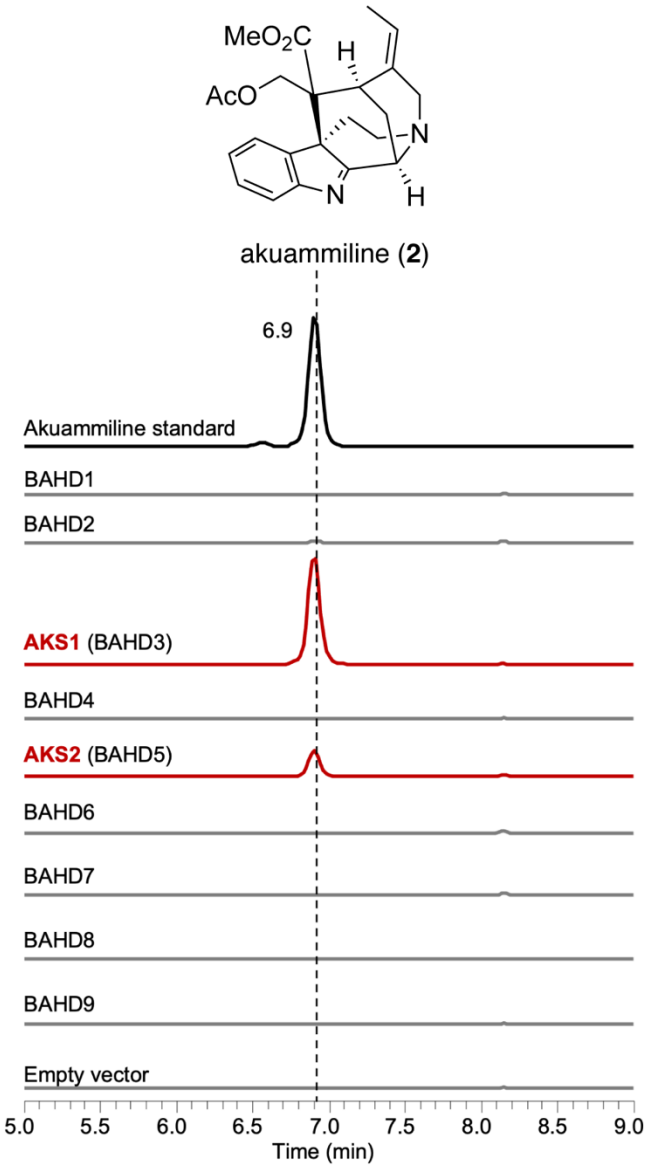



**Figure S20. Selected LC-MS chromatograms at *m/z* 395 of akuammiline (2) standard and the screened BAHD acetyltransferase (BAHD) candidates against rhazimol (1).** All BAHD homologues that were assayed in this study but failed to show significant turnover with rhazimol (**1**) are shown in this figure. The function of these enzymes remains unidentified, except for *A.scholaris* akuammiline synthase 1 AKS1 (BAHD3) and akuammiline synthase 2 AKS2 (BAHD5), which were characterized in this paper.

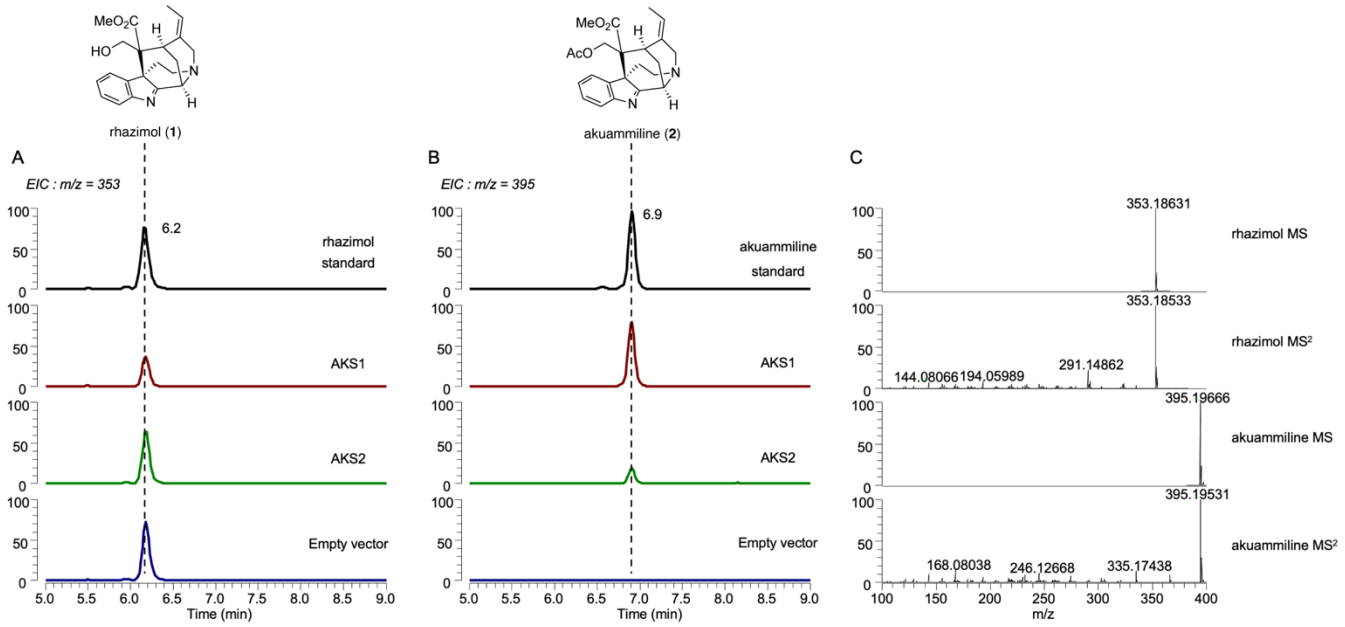



**Figure S21. *In vitro* catalytic activity of AKS1 and AKS2 against rhazimol (1). (A)** The extracted ion chromatograms EIC at *m/z* 353 for rhazimol (**1**) from rhazimol standard, AKS1 enzyme assay, AKS2 enzyme assay and empty vector. The purified heterologous BAHD proteins were incubated with rhazimol (*m/z* 353) and acetyl-CoA for 1 hour. In the enzyme reactions of acetyltransferases BAHD3 and BAHD5 a new peak at *m/z* 395 detected with the same retention time and the same $MS^2$ fragmentation pattern to synthetic akuammiline (**2**). **(B)** The extracted ion chromatograms EIC at *m/z* 395 for akuammiline (**2**) from akuammiline standard, AKS1 enzyme assay, AKS2 enzyme assay and empty vector. **(C)** The MS and $MS^2$ spectra of rhazimol (**1**) (rt 6.2 min) and akuammiline (**2**) (rt 6.9 min).

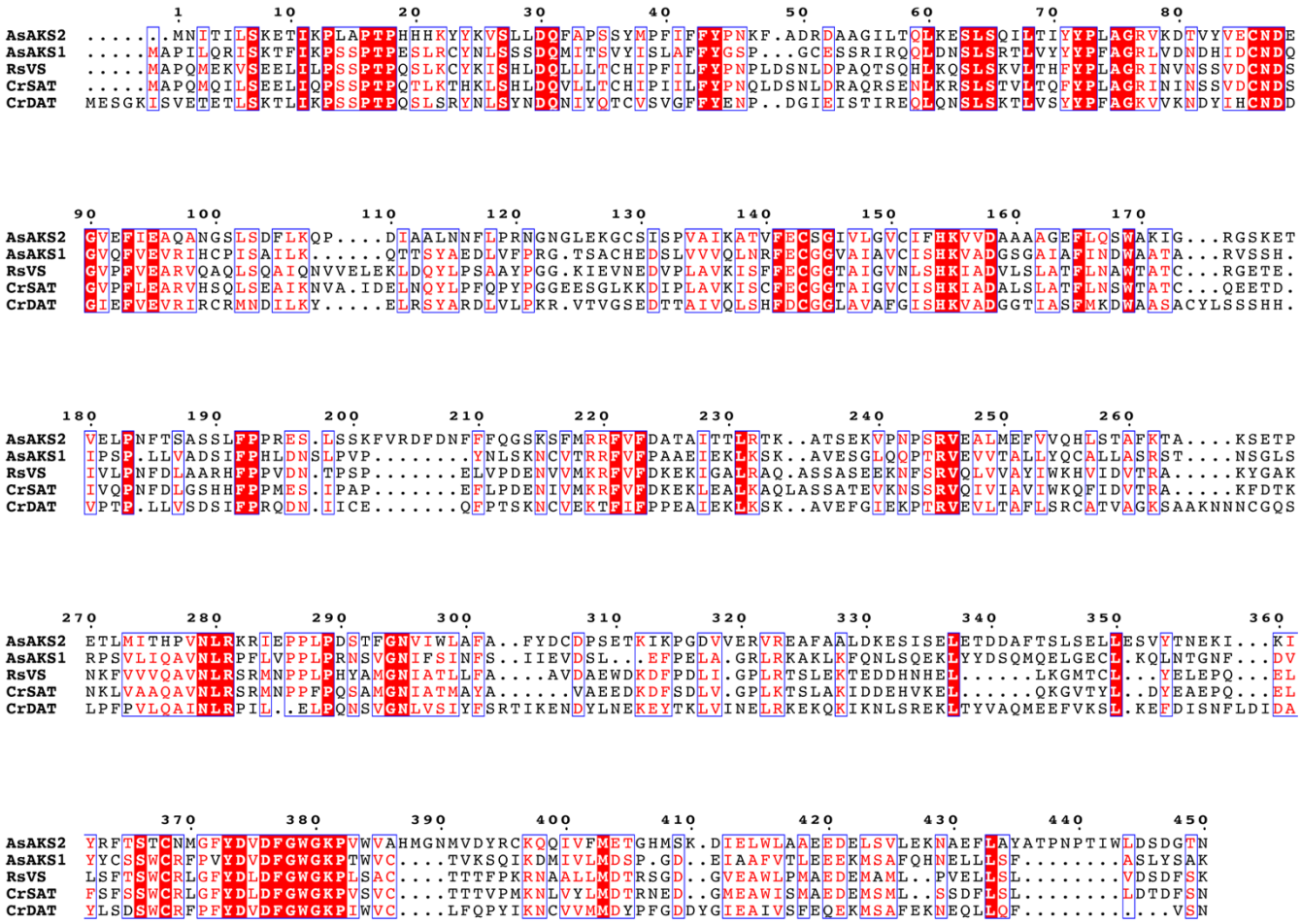



**Figure S22. Sequence alignment of BAHD acetyl transferase from *C. roseus* and AKSs from *A. scholaris*.** Numbering corresponds to AsAKS2. Identical and similar amino acids are highlighted in red, respectively.

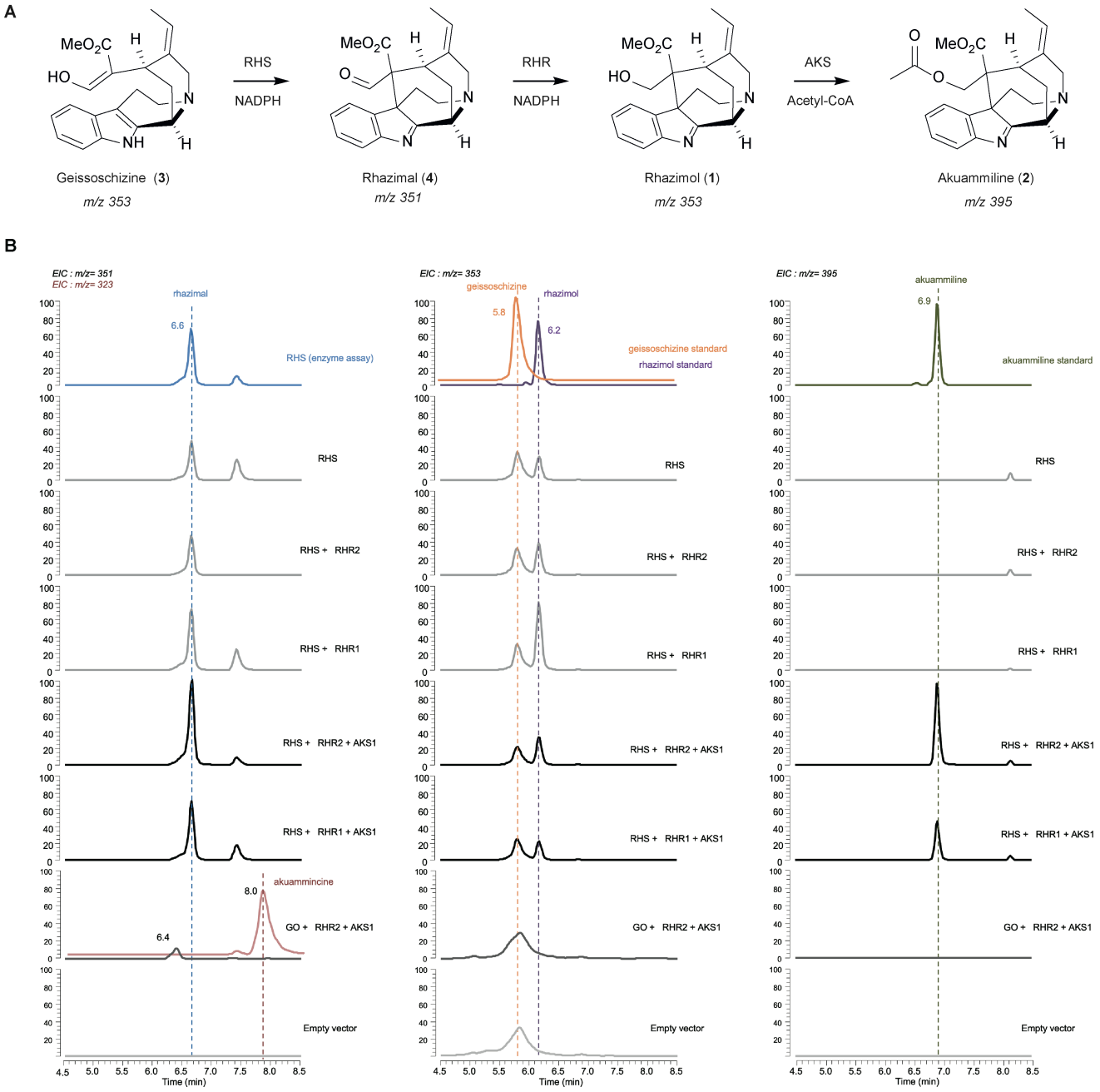



**Figure S23. Reconstitution of Akuammiline biosynthesis in *Saccharomyces cerevisiae*. (A)** Schematic representation of metabolic pathway from geissoschzine (**3**) to akuammiline (**2**). **(B)** The extracted ion chromatograms EIC at *m/z* 353 for rhazimal (**4**) (highlighted in blue) from enzyme assay and cell free engineered yeast. **(C)** The extracted ion chromatograms EIC at *m/z* 353 for geissoschizine (**3**) standard (highlighted in orange), rhazimol (**1**) standard (highlighted in purple) and cell free engineered yeast. **(D)** The extracted ion chromatograms EIC at *m/z* 395 for akuammiline (**2**) standard (highlighted in dark green) and cell free engineered yeast.

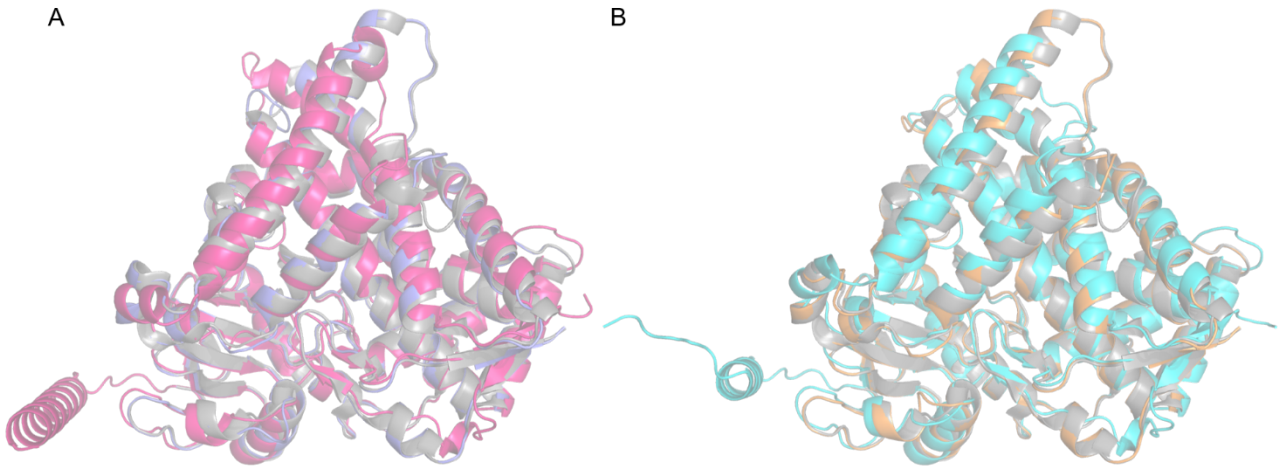



**Figure S24. Homology modelling structures of AsRHS and AsGO based on X-ray structure 7CB9.** **(A)** Structural overlay of the SWISS-MODEL derived homology model AsRHS (purple), SmFS crystal structure (template - grey) and predicted structure of AsRHS based on AlphaFold2 (magenta). **(B)** Structural overlay of the SWISS-MODEL derived homology model AsGO (orange), SmFS crystal structure (template - grey) and predicted structure of AsGO based on AlphaFold2 (cyan).

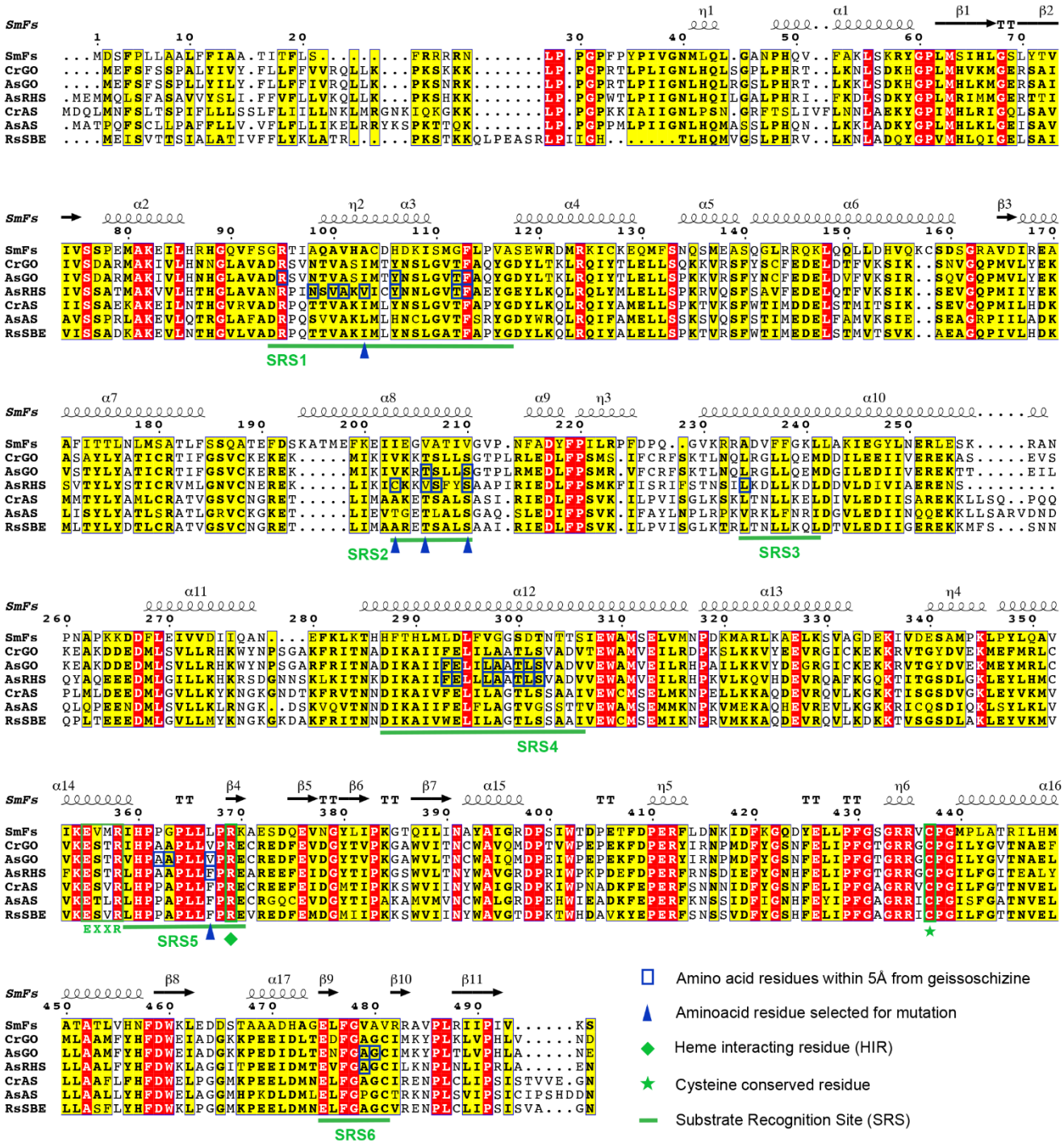



**Figure S25**. **Sequence alignment of cytochrome P450 enzymes in MIAs pathway that act on strictosidine aglycon reduced products**. We identified (a) the heme-coordinating cysteine residue that acts as fifth ligand in the P450 signature motif CXG[13] (green star); (b) the substrate recognition site (SRS) domains (highlighted in green ribbons) that are part of the cytochrome p450s active site surrounding the heme[14-16] and contribute to the catalytic activity, substrate recognition or substrate binding specificity. (c) the highly conserved P450 signature motif EXXR;[17] and (d) the arginine residue R374 of SRS5 at position 11 after the EXXR motif serves as the heme interacting residue (HIR) with carboxylate group of heme. Numbering corresponds to SmFS (*Salvia miltiorrhiza* Ferruginol synthase). Identical and similar amino acids are highlighted in red and yellow, respectively. Secondary structure elements of the SmFs crystal structure are displayed.

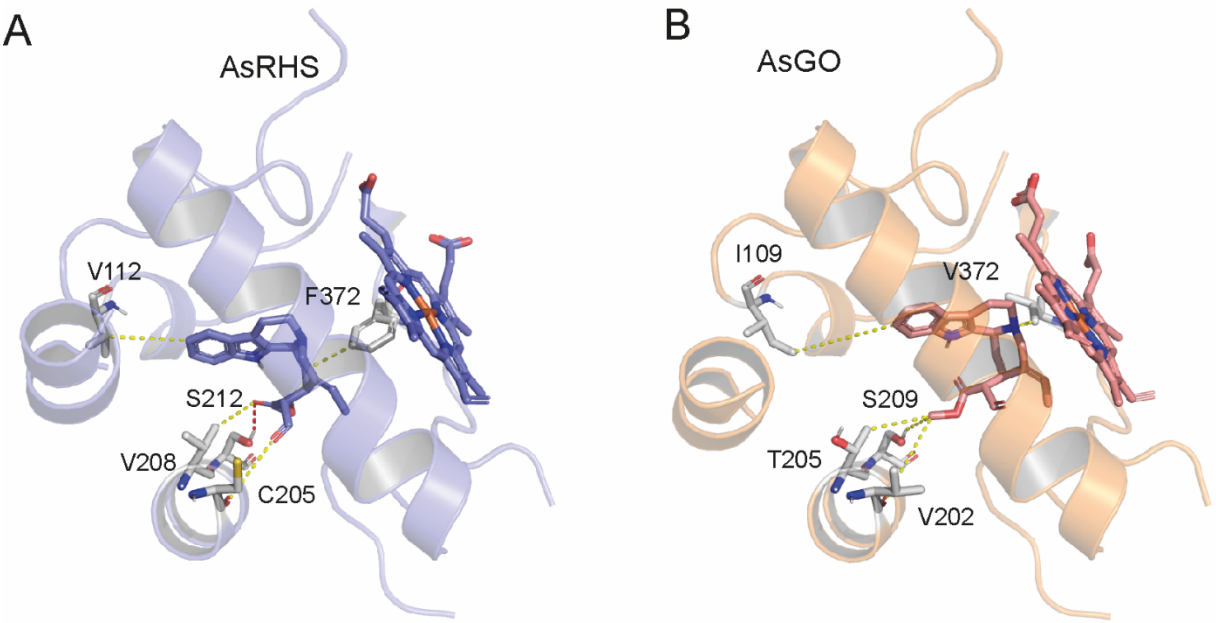



**Figure S26. Homology modelling of AsRHS, AsGO and docking of geissoschizine.** AsRHS (in purple) **(A)** and AsGO (in orange) **(B)** protein structure homology models derived by SWISS-MODEL with heme and geissoschizine (substrate) after applying docking using AutoDock Vina (highlighted in blue in AsRHS, and in pink for AsGO); target amino acids for mutation and their positions from the substrate are depicted in sticks.

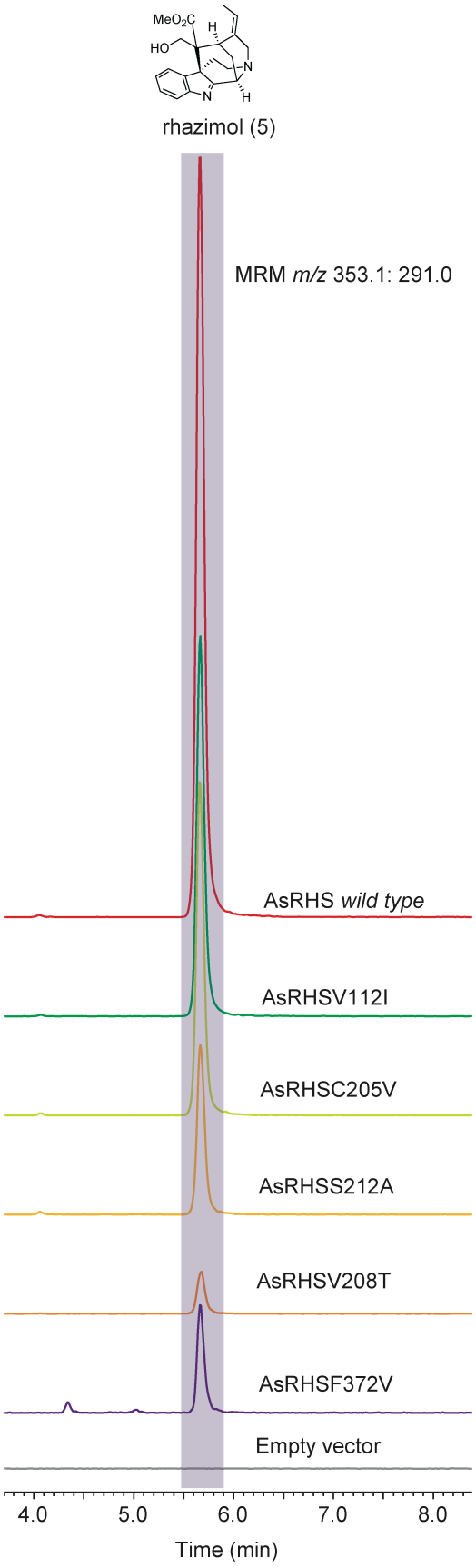



**Figure S27. Characterization of rhazimal synthase activity of AsRHS *wildtype* and single mutants using geissoschizine as substrate.** The enzyme reactions were analysed by QTRAP® 6500+ coupled to ExionLC™ system.

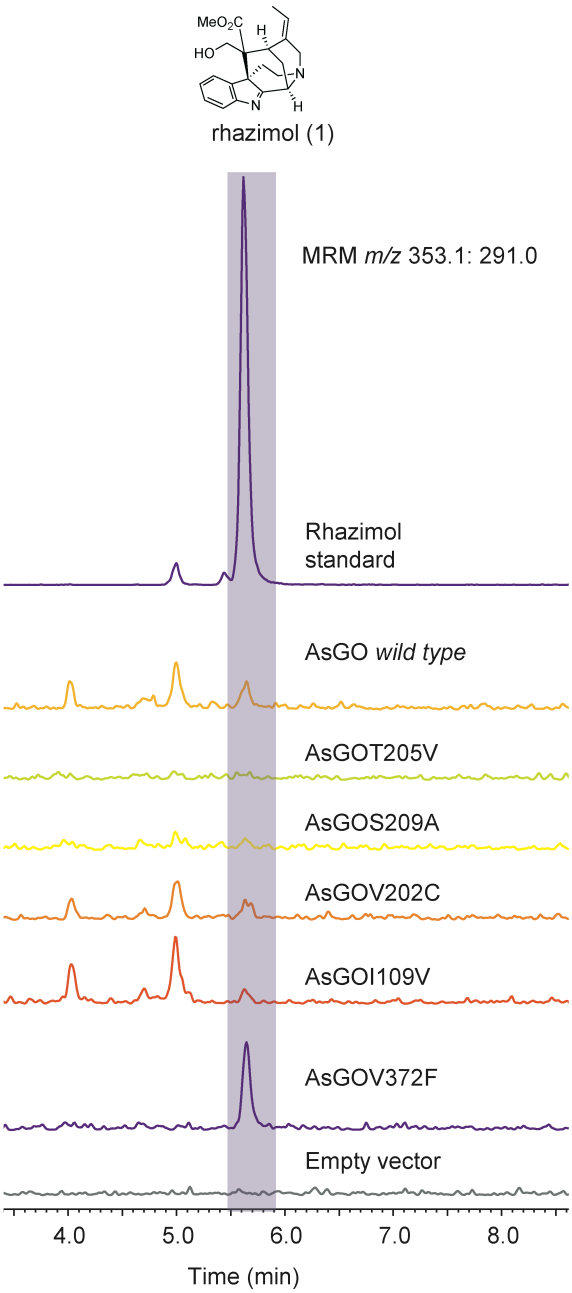



**Figure S28. Characterization of rhazimal synthase activity of AsGO *wildtype* and single mutants using geissoschizine as substrate.** The enzyme reactions were analysed by QTRAP® 6500+ coupled to ExionLC™ system.

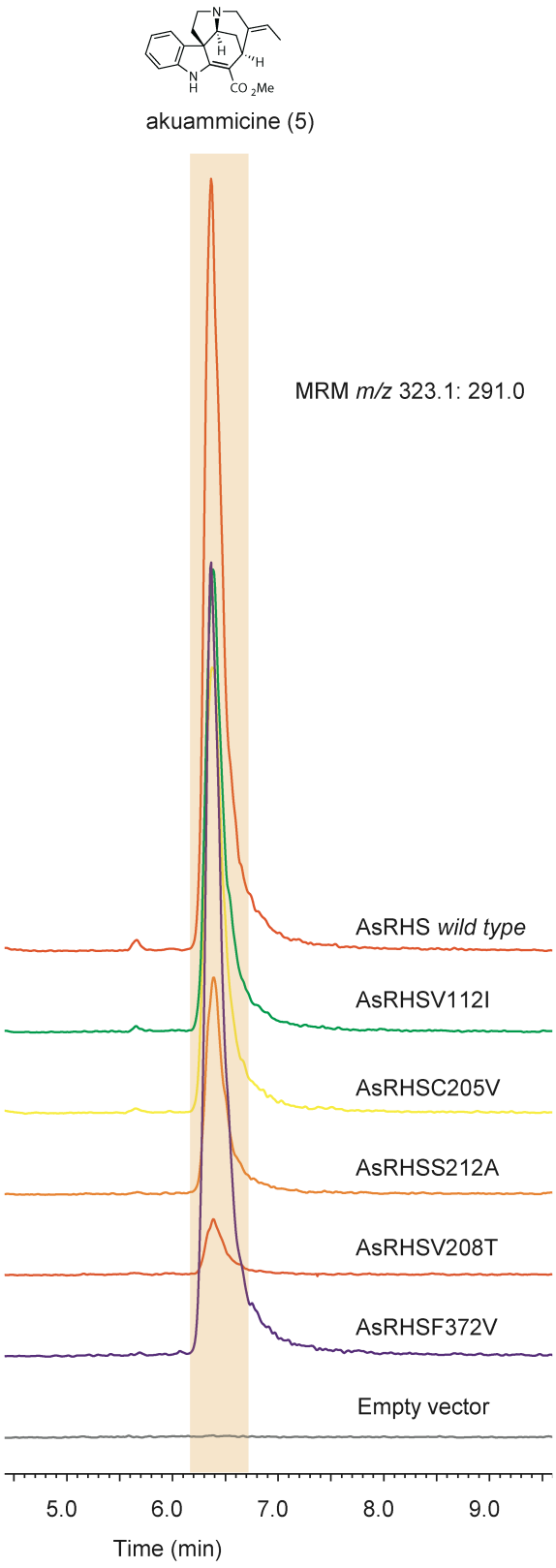



**Figure S29. Characterization of geissoschizine oxidase activity of AsRHS *wildtype* and single mutants.** The enzyme reactions were analysed by QTRAP® 6500⁺ coupled to ExionLC™ system.

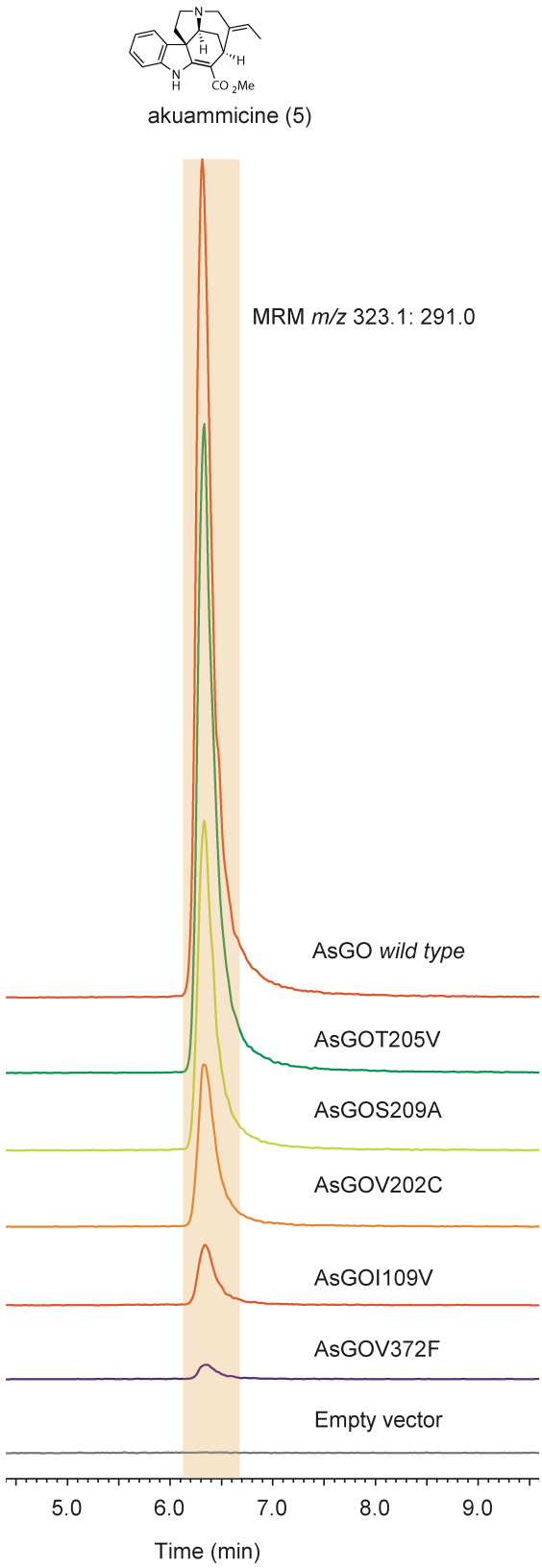



**Figure S30. Characterization of geissoschizine oxidase activity of AsGO *wildtype* and single mutants.** The enzyme reactions were analysed by QTRAP® 6500$^{+}$ coupled to ExionLC™ system.

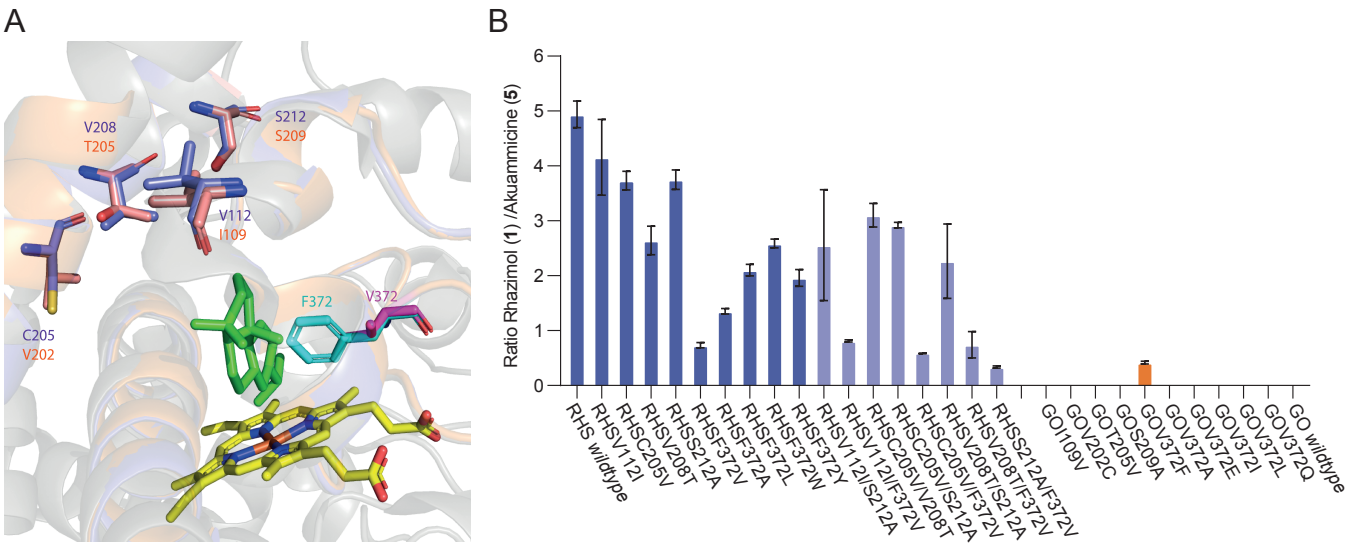



**Figure S31. Molecular modelling and mutagenesis on active site of AsRHS and AsGO. (A)** Overlay of the homology model structures of AsRHS, AsGO and the template X-ray structure 7CB9 of SmCYP76AH1 showing the P450s active site, with the heme (in yellow) and substrate militiradiene (in green) from template structure. In purple and orange colour are depicted the structural elements (SRSs and amino acids) from AsRHS and AsGO, respectively. The divergent amino acids selected for reciprocal mutations are shown in sticks. The key amino acid residues with side chains occupying space between the heme and substrate that direct the enzyme assay outcome are highlighted in cyan (AsRHS F372) and magenta (AsGO V372). (**B**) Relative ratio of rhazimol (**1**)/ akuammicine (**5**) as products released by enzyme assays (peak integrals) of wildtype and mutants of RHS (in purple) and GO (in orange). The enzyme assays were quenched by addition of $NaBH_4$. The bars represent the standard deviation between three individual enzyme replicates. The bars in purple denote single amino acid RHS mutant and in light purple double amino acid mutants.

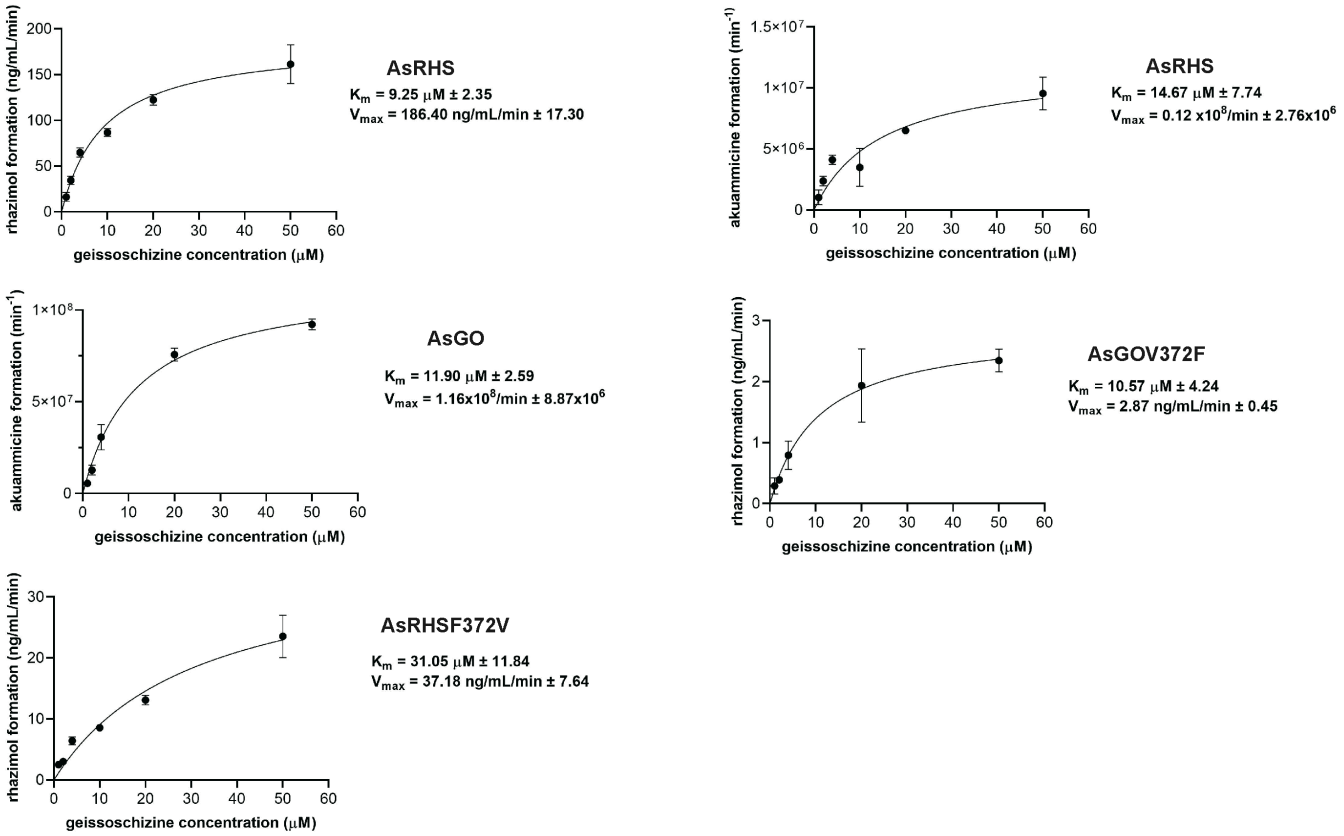



**Figure S32. End-point enzyme kinetics of AsRHS (*wildtype* and F372V) and AsGO (*wildtype* and V372F).** The enzyme reactions were analysed by QTRAP® 6500+ coupled to ExionLC™ system. The bars represent the standard deviation between four individual enzyme replicates.

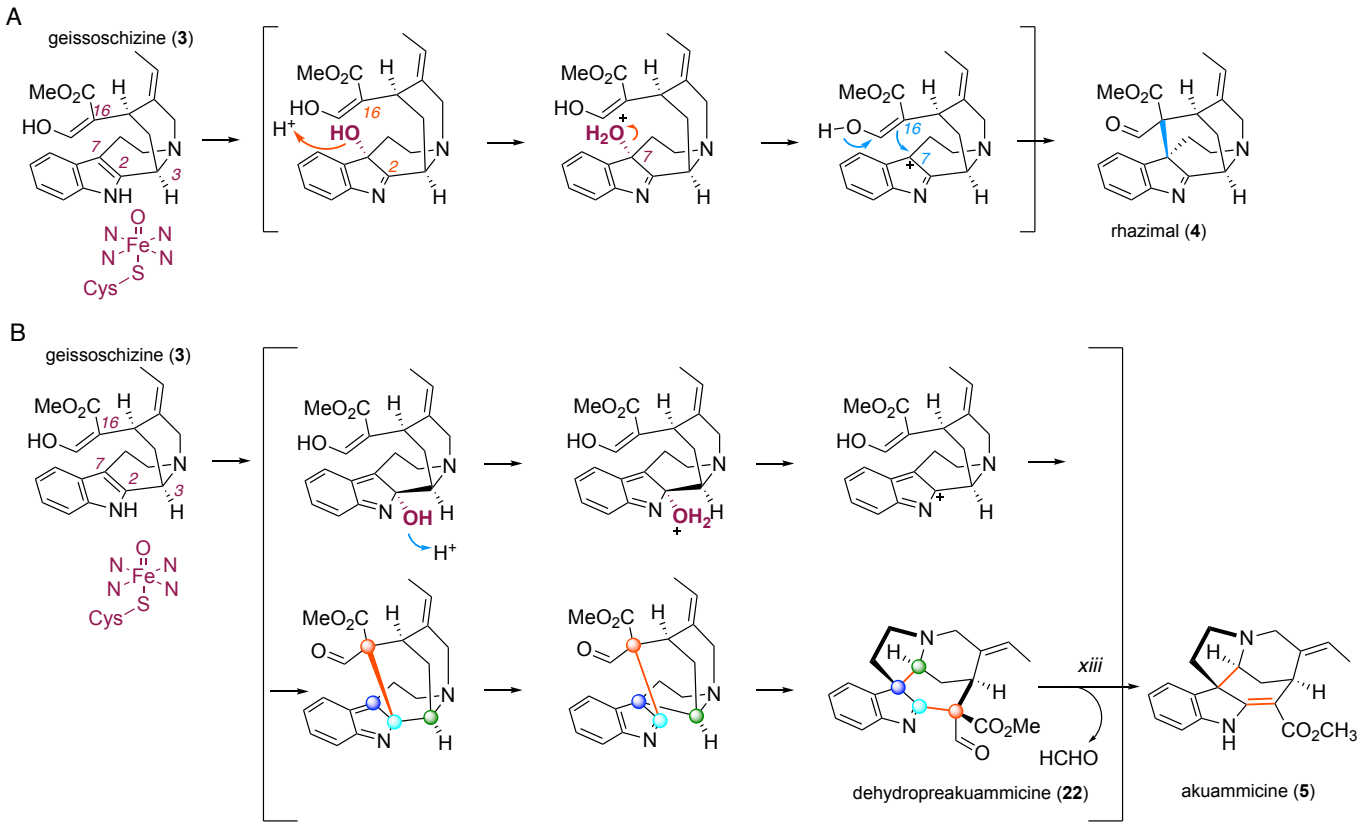



**Figure S33. Proposed unifying two-step mechanism for RHS and GO through hydroxylation and a later nucleophilic attack from the keto-enol function of geissoschizine. (A**) Plausible mechanism for formation of rhazimal (*akuammilan* scaffold) through the oxidation of geissoschizine on C7 followed by nucleophilic attack by C16. The new formed bond C7-C16 is highlighted in blue. (**B**) Plausible mechanism for formation of akuammicine (*strychnos* scaffold) through the oxidation of geissoschizine on C2 followed by nucleophilic attack by C16, 1,2 (sigmatropic) rearrangement by the breaking of C2-C3 bond and the formation of C3-C7 bond and finally the deformylation of dehydropreakuammicine. The new formed bonds C2-C16 and C3-C7 are highlighted in orange.

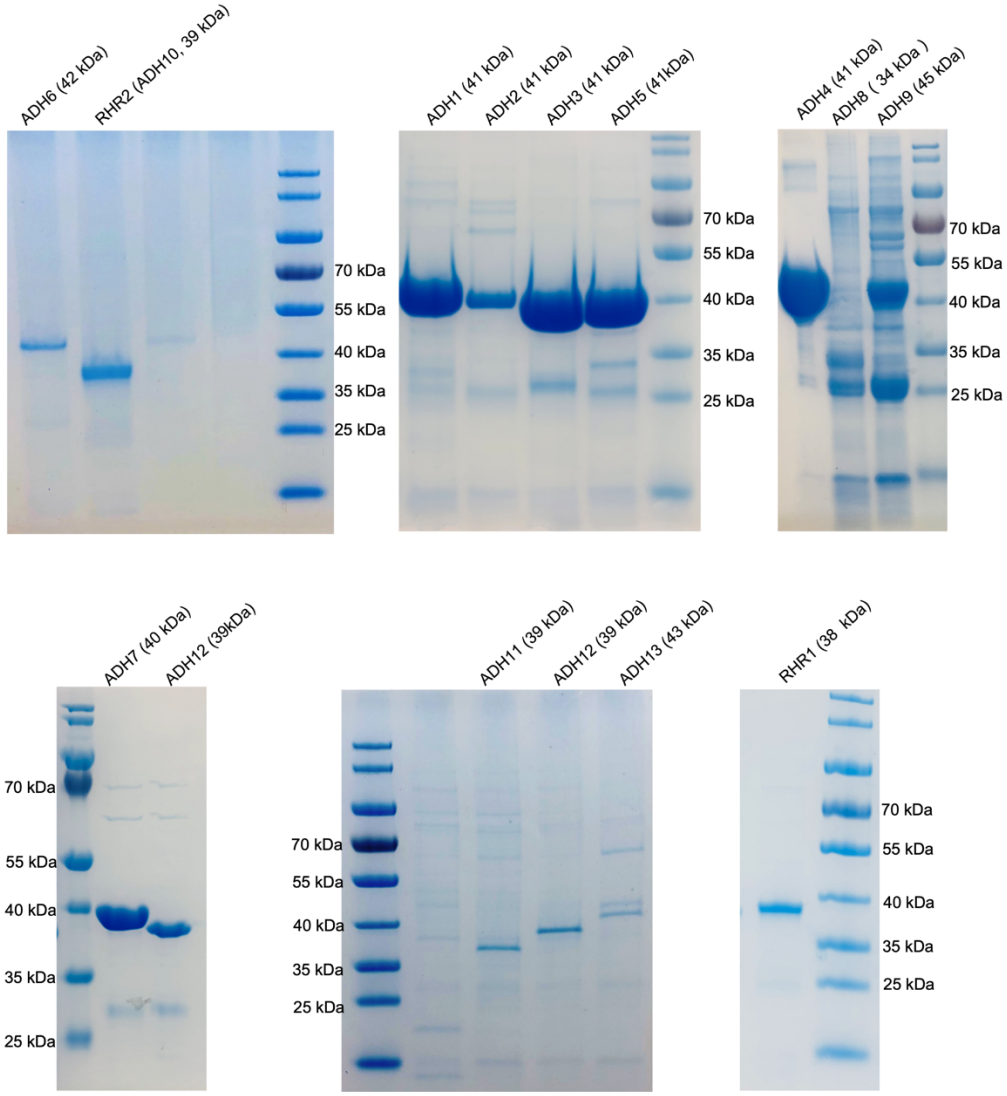



**Figure S34. SDS-PAGE analysis of purified alcohol dehydrogenase/reductase (ADH) candidate proteins from *A. scholaris*.**

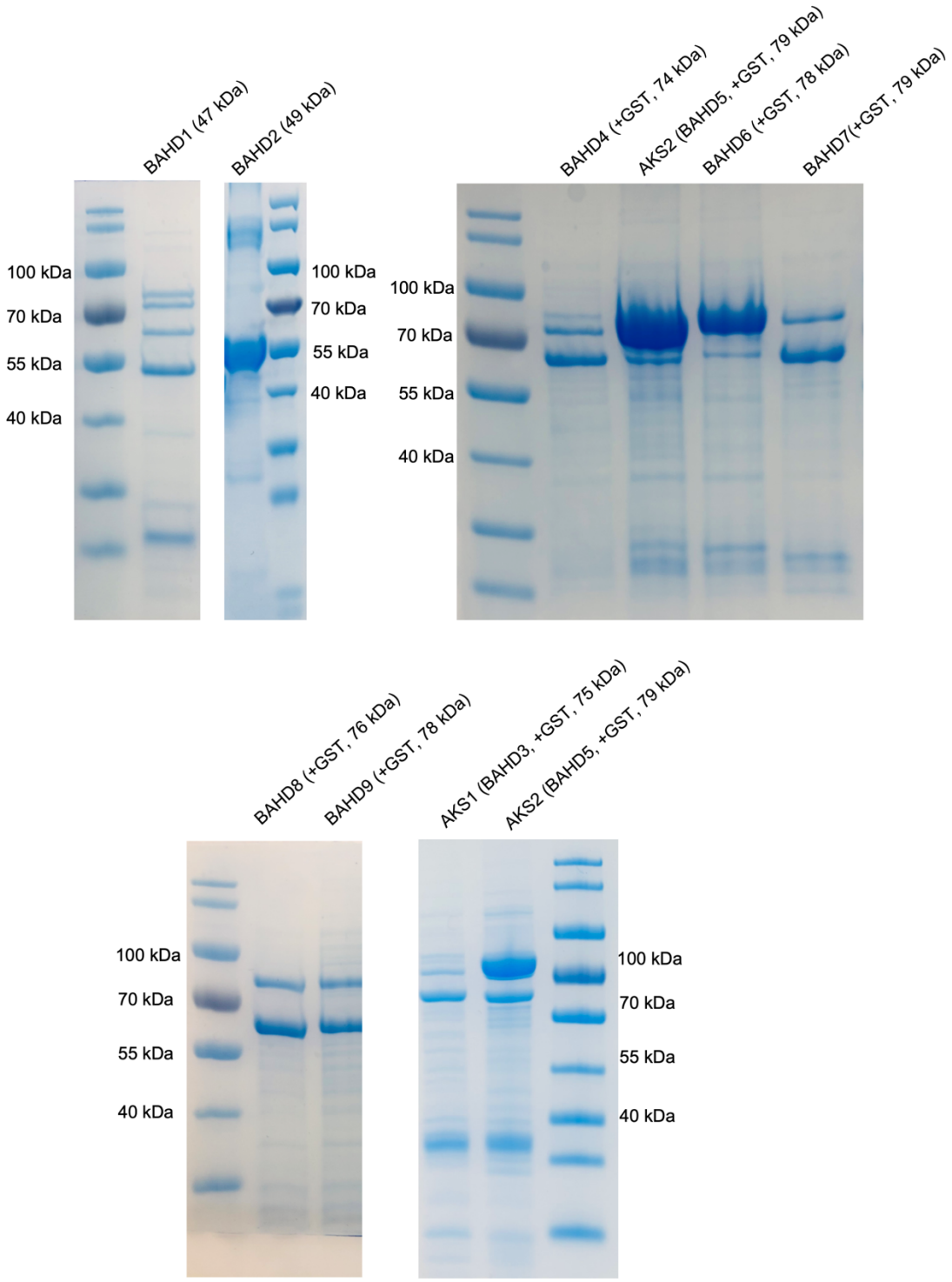



**Figure S35. SDS-PAGE analysis of purified BAHD acetyltransferase (BAHD) candidate proteins from *A. scholaris*.**

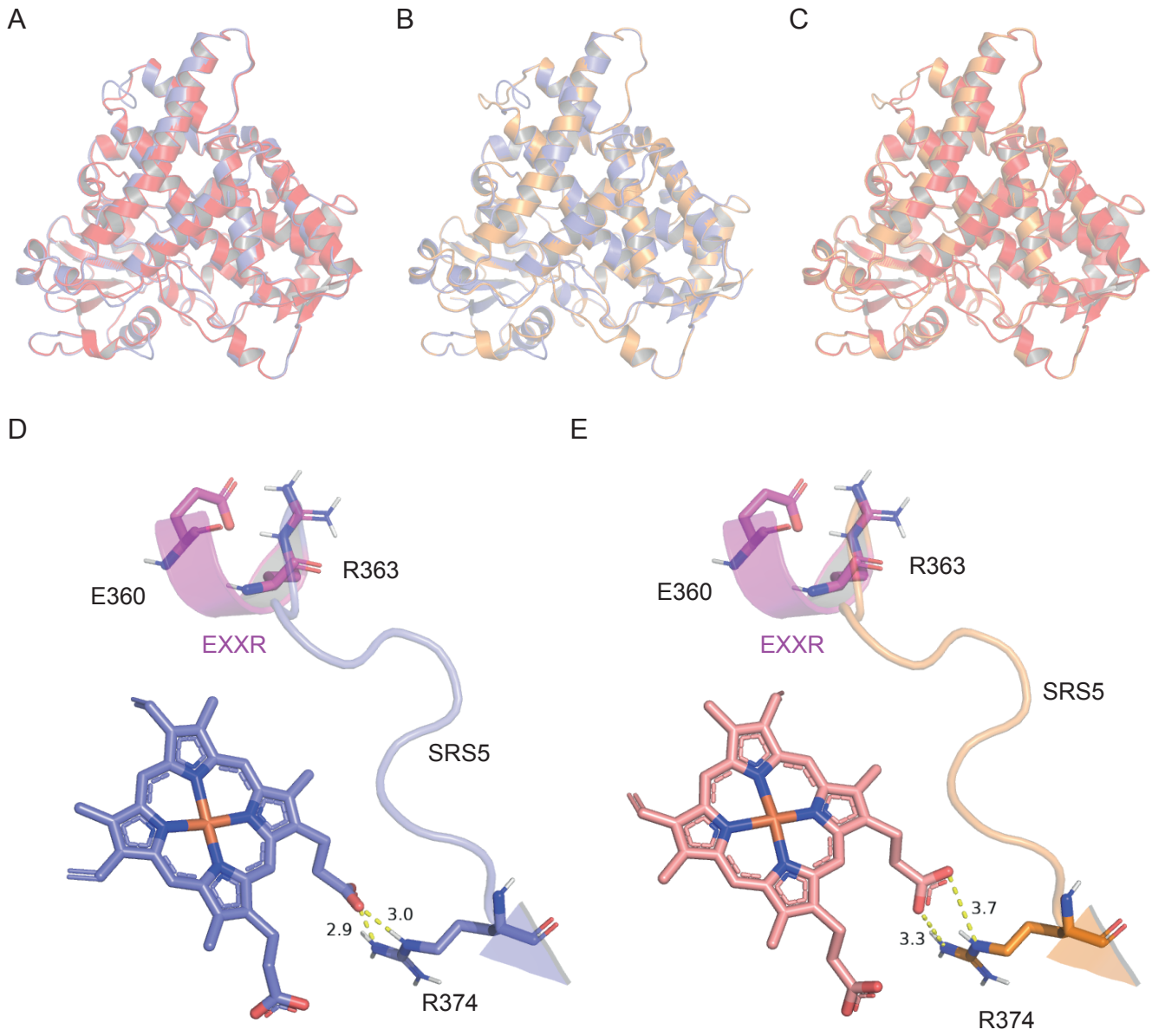



**Figure S36. Homology modelling structures of AsRHS and AsGO based on X-ray structure 7CB9.** **(A)** Structural overlay of the homology model AsRHS (purple) and SmFS crystal structure (red). **(B)** Structural overlay of the homology model AsRHS (purple) and AsGO (orange). **(C)** Structural overlay of the homology model AsGO (orange) and SmFS crystal structure (red). **(D&E)** Heme and SRS5 region in the two cytochrome AsRHS (in purple, D) and AsGO (in orange, E). The characteristic of cytochrome p450s conserved motif EXXR (highlighted in magenta) that is important for the stabilization of the meander loop and probably for the safeguarding of the CYP tertiary structure and activity.[13, 17] The SRS5 region is usually extended from the conserved ExxR motif to the strand β1-4. In both cases the R374 amino acid residue is identified as the heme interacting residue (HIR) and the distances between the heme 7'-propionate and the side chain amino groups of HIR are indicated.

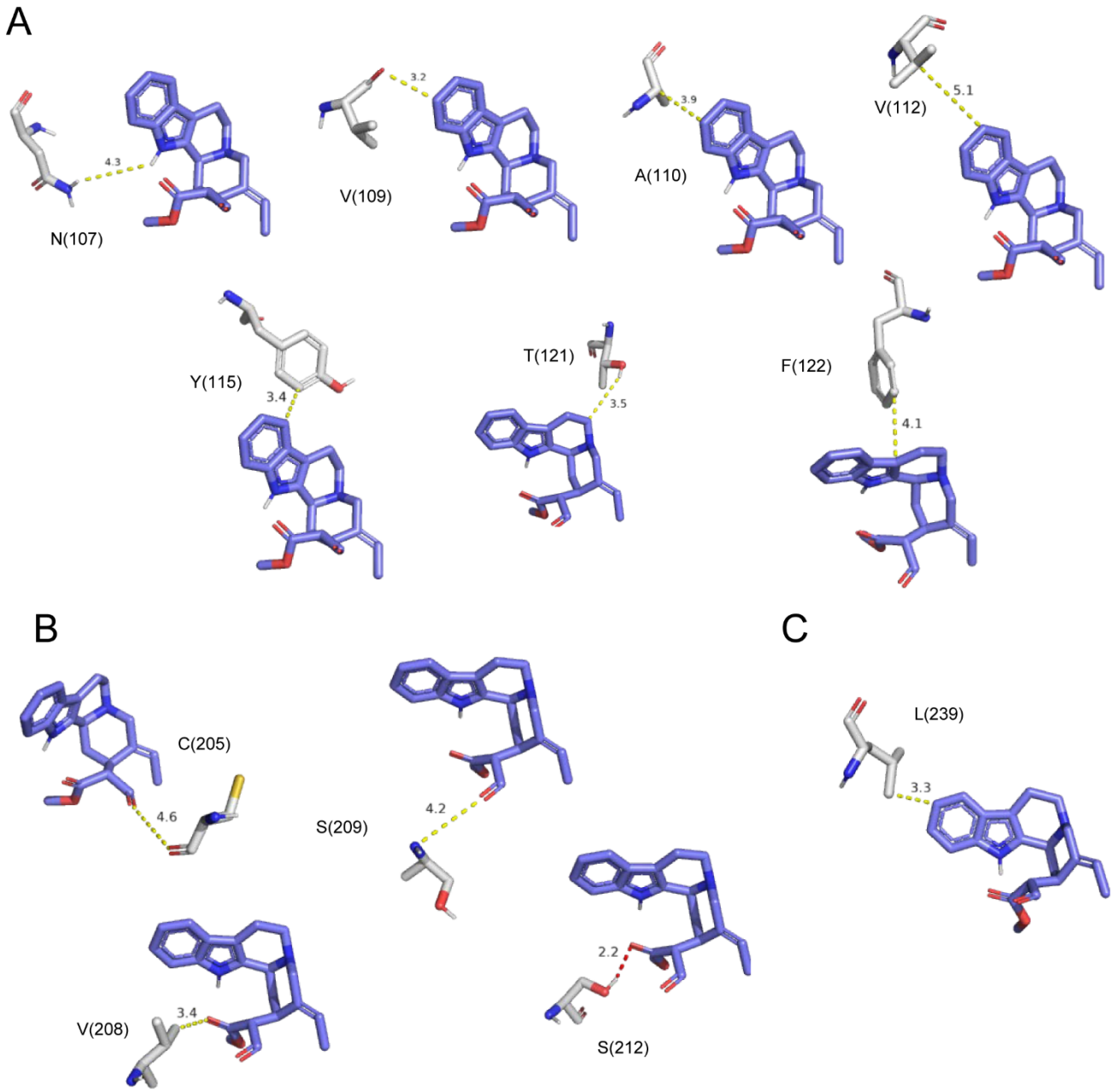



**Figure S37. Amino acid residues of AsRHS identified in a radius of 5 Å from geissoschizine according to homology modelling of AsRHS and docking of geissoschizine into AsRHS. (A)** Calculated distances of AsRHS SRS1 amino acids residues from geissoschizine. **(B)** Calculated distances of AsRHS SRS2 amino acids residues from geissoschizine. **(C)** Calculated distances of AsRHS SRS3 amino acids residues from geissoschizine.

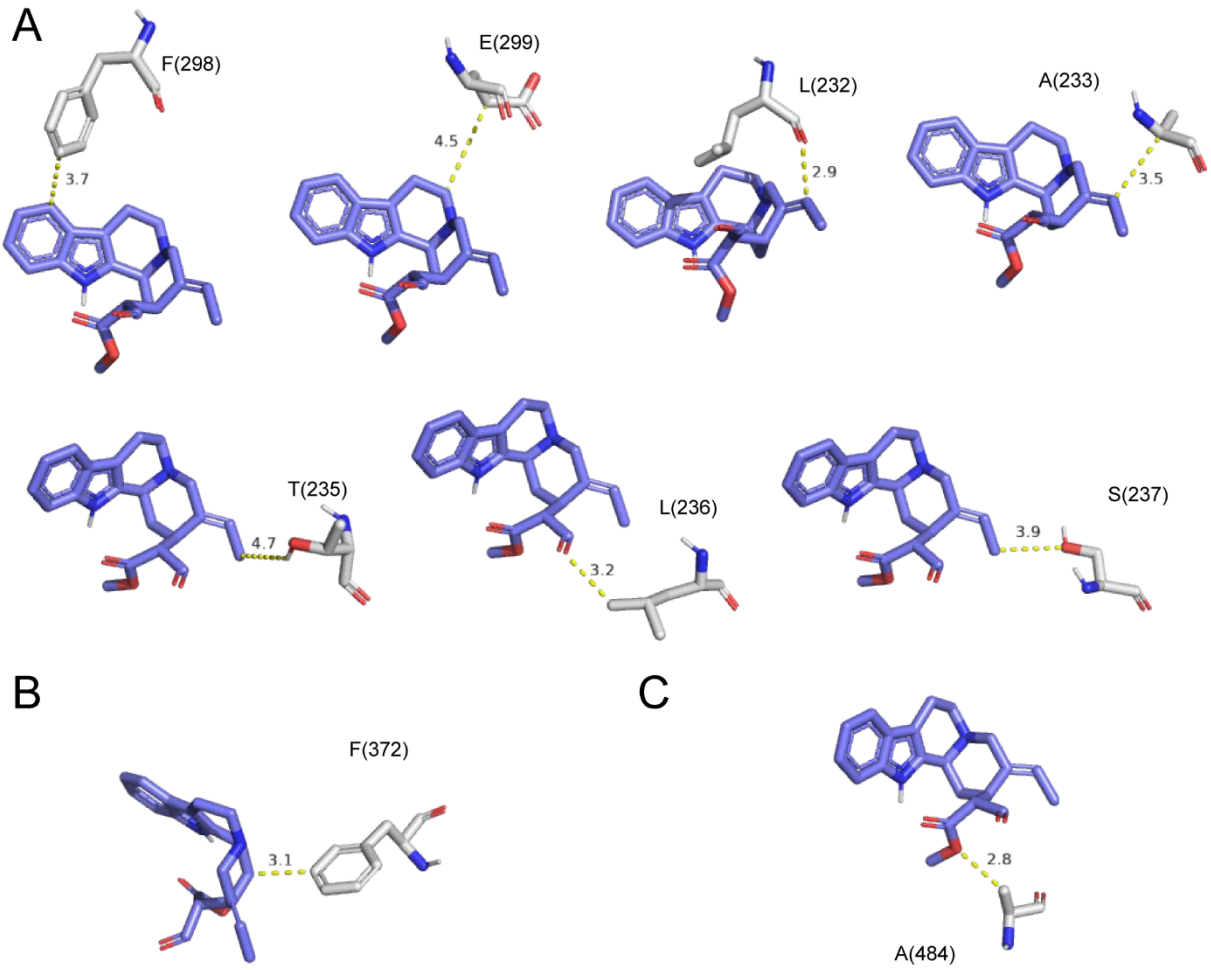



**Figure S38. Amino acid residues of AsRHS identified in a radius of 5 Å from geissoschizine according to homology modelling of AsRHS and docking of geissoschizine into AsRHS. (A)** Calculated distances of AsRHS SRS4 amino acids residues from geissoschizine. **(B)** Calculated distances of AsRHS SRS5 amino acids residues from geissoschizine. **(C)** Calculated distances of AsRHS SRS6 amino acids residues from geissoschizine.

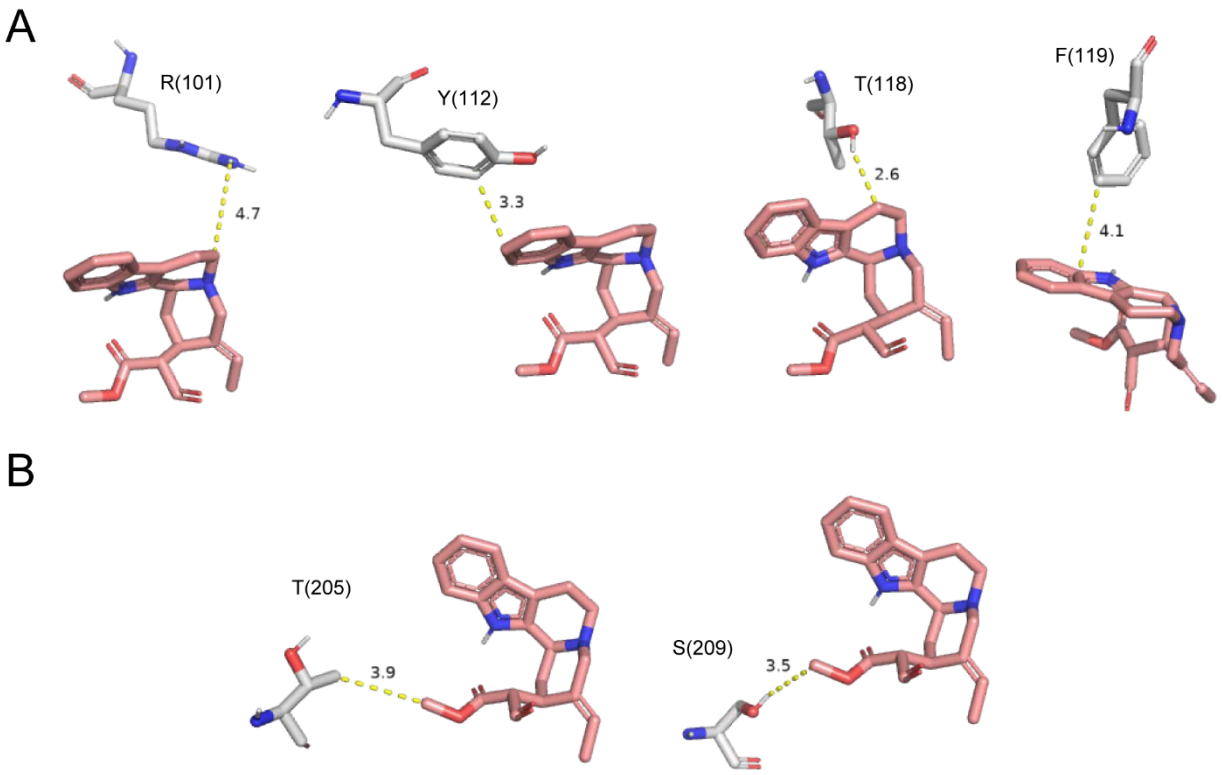



**Figure S39. Amino acid residues of AsGO identified in a radius of 5 Å from geissoschizine according to homology modelling of AsGO and docking of geissoschizine into AsGO. (A)** Calculated distances of AsGO SRS1 amino acids residues from geissoschizine. **(B)** Calculated distances of AsGO SRS2 amino acids residues from geissoschizine.

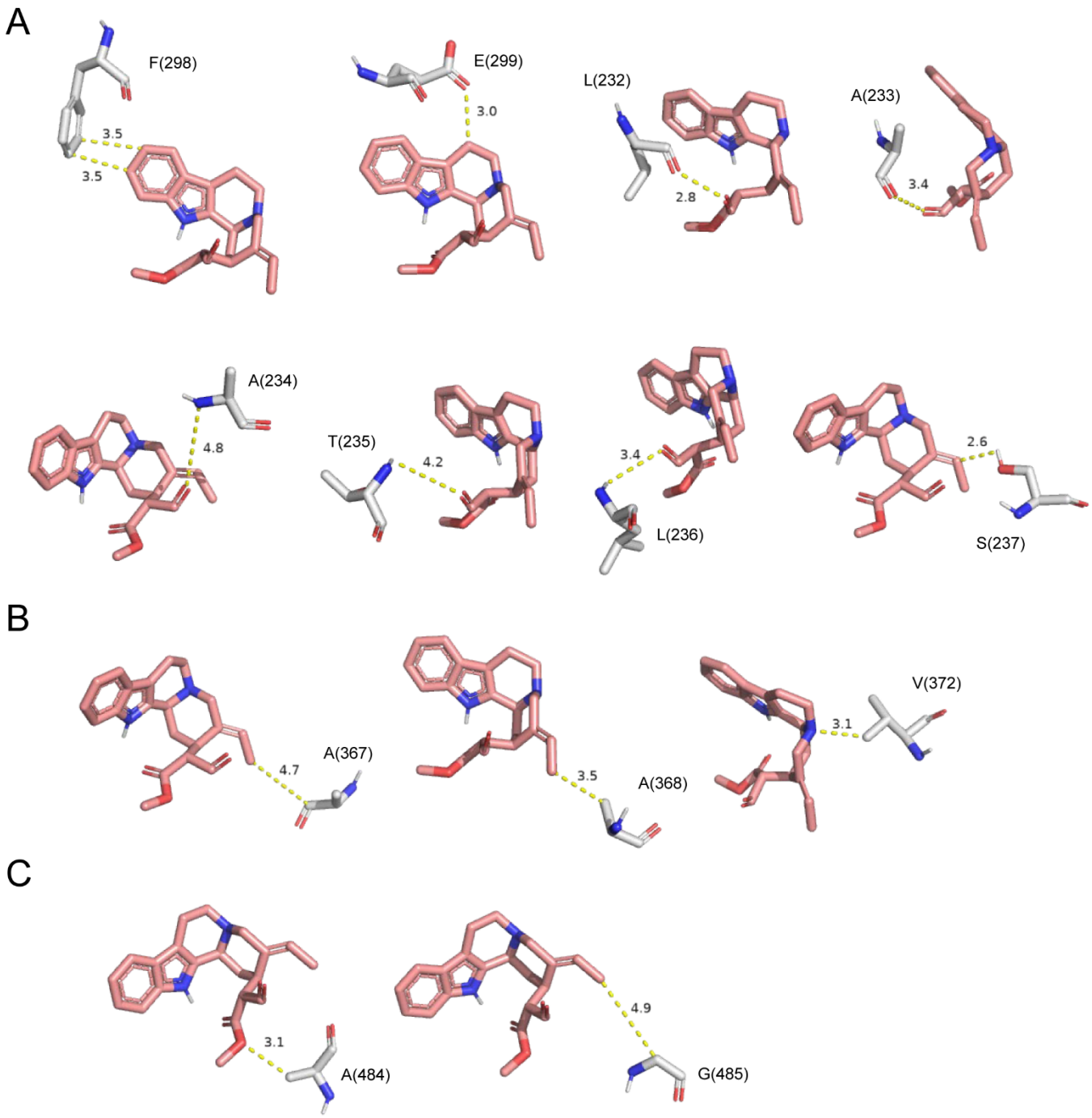



**Figure S40. Amino acid residues of AsGO identified in a radius of 5 Å from geissoschizine according to homology modelling of AsGO and docking of geissoschizine into AsGO. (A)** Calculated distances of AsGO SRS4 amino acids residues from geissoschizine. **(B)** Calculated distances of AsGO SRS5 amino acids residues from geissoschizine. **(C)** Calculated distances of AsGO SRS6 amino acids residues from geissoschizine.

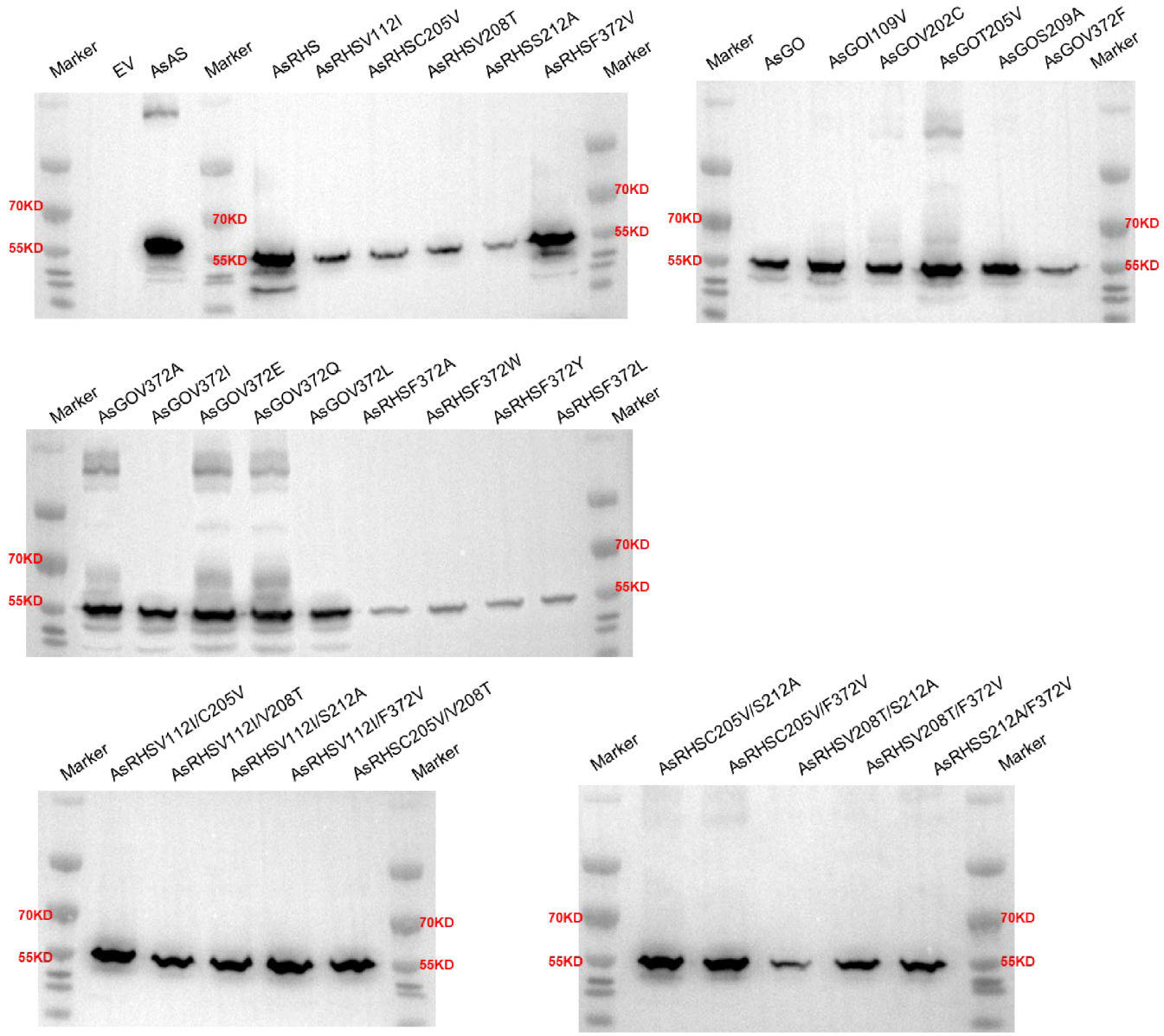



**Figure S41. Western blot showing the expression of AsAS, AsRHS, AsGO, AsRHS mutations and AsGO mutations.**

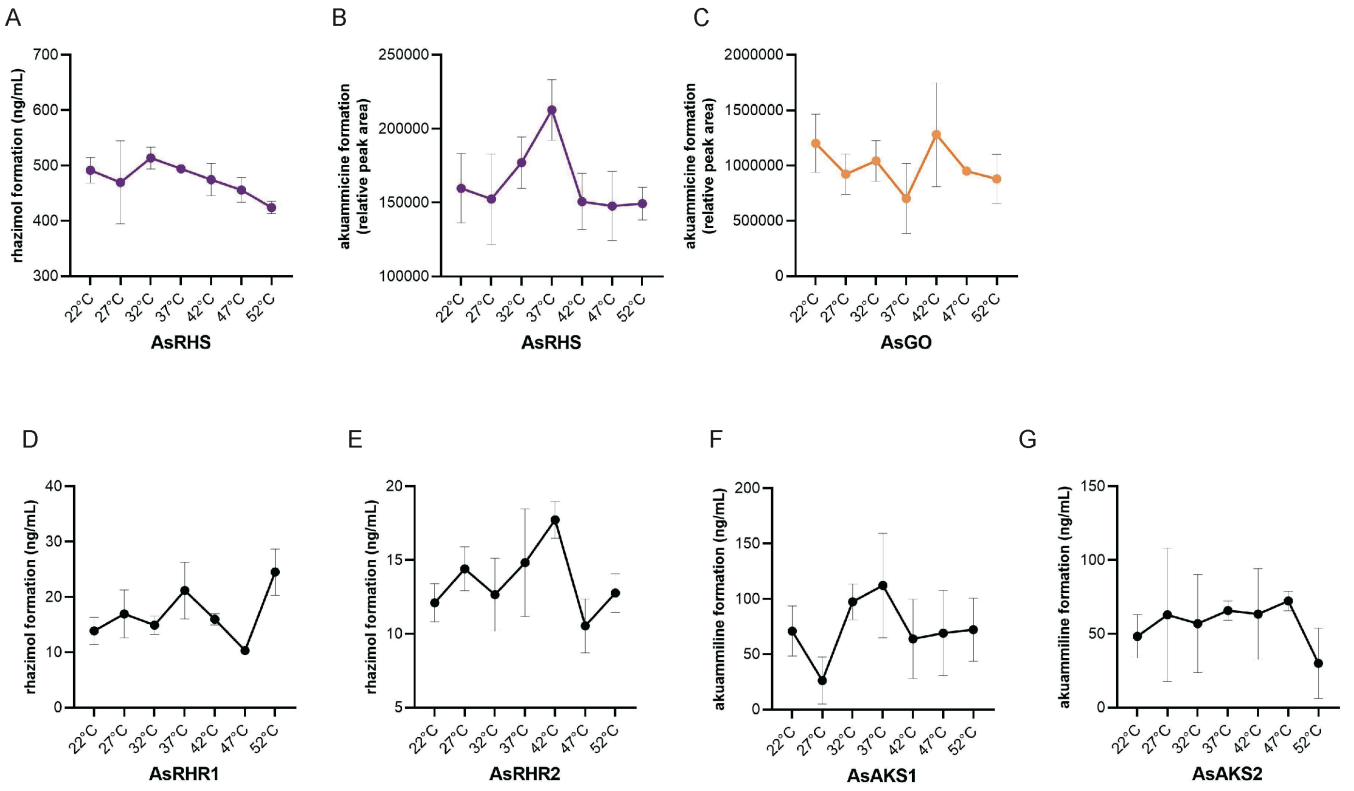



**Figure S42. Product accumulation of AsRHS, AsGO, AsRHR1, AsRHR2, AsAKS1 and AsAKS2 with different temperature.** The bars represent the standard deviation between three individual enzyme replicates.

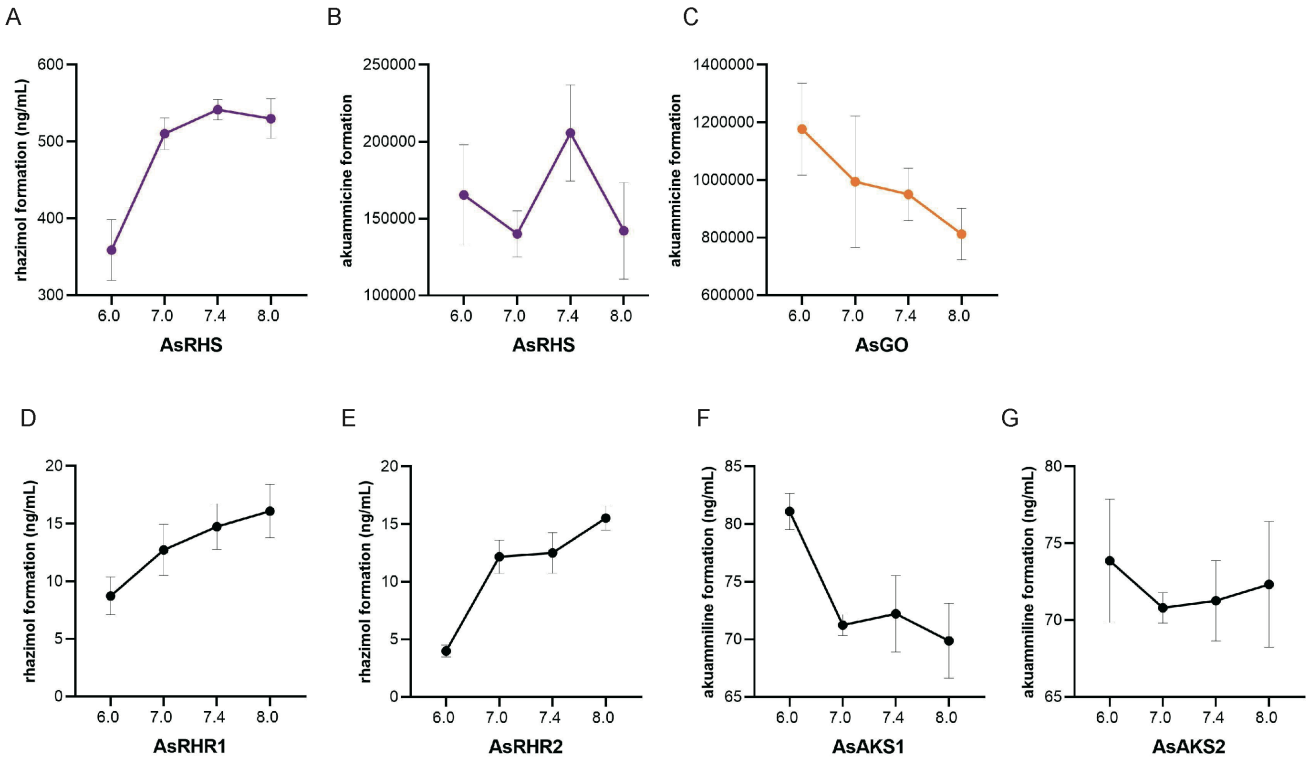



**Figure S43. Product accumulation of AsRHS, AsGO, AsRHR1, AsRHR2, AsAKS1 and AsAKS2 with different pH.** The enzyme reactions were analysed by QTRAP® 6500+ coupled to ExionLC™ system. The bars represent the standard deviation between three individual enzyme replicates.

## References


1. M. G. Grabherr, B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman and A. Regev, *Nat Biotechnol*, 2011, **29**, 644-652.
2. G. Pertea, X. Huang, F. Liang, V. Antonescu, R. Sultana, S. Karamycheva, Y. Lee, J. White, F. Cheung, B. Parvizi, J. Tsai and J. Quackenbush, *Bioinformatics*, 2003, **19**, 651-652.
3. Y. Zhang, C. Park, C. Bennett, M. Thornton and D. Kim, *Genome Res*, 2021, **31**, 1290-1295.
4. X. Zhang, B. N. Kakde, R. Guo, S. Yadav, Y. Gu and A. Li, *Angew Chem Int Ed Engl*, 2019, **58**, 6053-6058.
5. A. Stavrinides, E. C. Tatsis, L. Caputi, E. Foureau, C. E. Stevenson, D. M. Lawson, V. Courdavault and S. E. O'Connor, *Nat Commun*, 2016, **7**, 12116.
6. R. M. Payne, D. Xu, E. Foureau, M. I. Teto Carqueijeiro, A. Oudin, T. D. Bernonville, V. Novak, M. Burow, C. E. Olsen, D. M. Jones, E. C. Tatsis, A. Pendle, B. Ann Halkier, F. Geu-Flores, V. Courdavault, H. H. Nour-Eldin and S. E. O'Connor, *Nat Plants*, 2017, **3**, 16208.
7. E. C. Tatsis, I. Carqueijeiro, T. Dugé de Bernonville, J. Franke, T. T. Dang, A. Oudin, A. Lanoue, F. Lafontaine, A. K. Stavrinides, M. Clastre, V. Courdavault and S. E. O'Connor, *Nat Commun*, 2017, **8**, 316.
8. T. T. Dang, J. Franke, I. S. T. Carqueijeiro, C. Langley, V. Courdavault and S. E. O'Connor, *Nat Chem Biol*, 2018, **14**, 760-763.
9. D. Pompon, B. Louerat, A. Bronine and P. Urban, *Methods Enzymol*, 1996, **272**, 51-64.
10. A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore and T. Schwede, *Nucleic Acids Res*, 2018, **46**, W296-w303.
11. C. Shi, M. Gu, Z. Chen, X. Huang, J. Guo, L. Huang, J. Deng, K. He, L. Zhang, L. Huang and Z. Chang, *Biochemical and Biophysical Research Communications*, 2021, **582**, 125-130.
12. Y. Qu, M. Easson, R. Simionescu, J. Hajicek, A. M. K. Thamm, V. Salim and V. De Luca, *Proc Natl Acad Sci U S A*, 2018, **115**, 3180-3185.
13. K. Syed and S. S. Mashele, *PLOS ONE*, 2014, **9**, e95616.
14. D. Sirim, M. Widmann, F. Wagner and J. Pleiss, *BMC Structural Biology*, 2010, **10**, 34.
15. O. Gotoh, *J Biol Chem*, 1992, **267**, 83-90.
16. Ł. Gricman, C. Vogel and J. Pleiss, *Proteins: Structure, Function, and Bioinformatics*, 2015, **83**, 1593-1603.
17. A. Seifert and J. Pleiss, *Proteins: Structure, Function, and Bioinformatics*, 2009, **74**, 1028-1035.