Supporting Information

Generation of Alternate Indole Diterpene Architectures in Two Species of Aspergilli

Rudranuj Bundela, Rosannah C. Cameron, A. Jonathan Singh, Rose M. McLellan, Alistair T. Richardson, Daniel Berry, Matthew J. Nicholson, Emily J. Parker*

[†]Ferrier Research Institute, Victoria University of Wellington, Wellington 6012, New Zealand

[‡]Maurice Wilkins Centre for Molecular Biodiscovery, Victoria University of Wellington, New Zealand

Table of Contents

List of Supplementary Tables

List of Supplementary Figures

Experimental procedures

Genome sequencing

*Aspergillus striatus (*ATCC 64988) genomic DNA was prepared by phenol-$CHCl_3$ extraction and treated with RNase A. One-half of a run with 250-bp paired-end fragment reads was done on an Illumina MiSeq instrument by New Zealand Genomics Limited (NZGL). Quality trimming of the sequencing adapter cleaned reads was performed using the SolexaQA++ package[4] with an error threshold of p≤0.05 and removal of all reads shorter than 100bp following trimming. ABySS[5] was used with the default parameters except for the k-mer length (k=), which was changed by an increment of 8 during a k-mer sweep between 47 and 119. Assembly statistics for each of the assemblies (using different k-mer sizes) were generated using QUAST (http://bioinf.spbau.ru/quast). The final assembly (k-mer = 111) consisted of 624 contigs over 500 bp with an average length of 50,918 nucleotides. The total number of nucleotide residues was 31,773,021 with a GC content of 50.75%. The largest contig was 1,042,198 bp, the N_{50} was 280,492 bp, and the L_{50} was 37.

Identification and annotation of *A. striatus* and *A. desertorum* gene clusters (GenBank accession numbers: OP493542, OP493543, OP493544 and BK062751)

The genome of *A. striatus* was searched using known indole diterpene biosynthetic genes[6] as the search string with NCBI's tBLASTn tool. The genome of *Aspergillus desertorum,* which was accessed via the Joint Genome Institute (Sequencing read coverage depth = 95.4x).[7-8] From these analyses, potential indole diterpene biosynthetic genes were identified based on sequence similarity. These genes were then annotated in SnapGene software (from Insightful Science; available at snapgene.com) based on their top homology matches. NCBI's BLASTp tool, which searches protein databases using a protein query, was used to compare hypothetical protein sequences. All NCBI BLAST analyses were carried out using default algorithm parameters. FgeneSH[9] was used to predict intron/exon boundaries and overall gene structure (based on *Aspergillus nidulans* or *Penicillium chysogenum*). In addition, we used homology of known functional IDT genes in the *PAX* cluster to compare of nucleotide sequences and support intron/exon predictions. Nucleotide sequence alignments for genes characterized in this work are shown in Figures S2-S4. The *EST* cluster can be accessed under the following accession numbers OP493542, OP493543, OP493544 and the *DES* clusters from BK062751.

The *A. desertorum* genome was subjected BLASTp searches for homologues of *paxP* and *paxQ*. No region identified by this search was considered adequate in length or similarity (relative to *paxP* or *paxQ*) to be described a homologue. A short region between *desA* and *desM* of the *DES* cluster with a residue length of 162 amino acids, did share some similarity to *paxQ* (42%), however no appropriate open reading frame (ORF) could be identified, and neither FgeneSH[9] or Augustus[10] online tools could generate reliable ORF predictions as the sequence is highly disrupted. Alongside this observation, a gene relic for an *idtB* has previously been identified in an IDT cluster of *A. flavus*;[11] and similarly speculate that an evolutionary relic of an *idtQ* may be present between *desA* and *desM,* that is highly unlikely to be a functional gene.

In the manuscript, Figure 2 was generated using Clinker and clustermap.js version 0.0.25.7.[12] Gene coloration corresponds to demonstrated or predicted function as described in the legend. The weight of shaded bands between homologous genes represents amino acid percentage identity and identity below 30% is not shown.

gDNA extraction

gDNA extraction of *A. striatus* and *P. paxilli* was performed using the Zymo Research Quick-DNA™ Fungal/Bacterial Miniprep Kit (Cat. No D6005). The standard protocol was used, with the exception of using Milli-Q® Water in place of DNA Elution buffer (to prevent interference from elution buffer components when the DNA is used for downstream applications). DNA concentration was quantified using an Invitrogen™ Qubit™ 4 Fluorometer using an Invitrogen™ Qubit™ dsDNA BR Assay Kit.

Polymerase chain reaction (PCR)

All PCRs were performed in an Applied Biosystems SimpliAmp™ Thermal Cycler. Phusion® High-Fidelity 2x PCR Master Mix (HF Buffer) (New England Biolabs) was used, unless otherwise stated. The Phusion® Master Mix contains a Phusion® High-

Fidelity DNA polymerase, deoxynucleotides and a reaction buffer optimized for PCR. Gene-specific primers are listed in **Table S3**. Annealing temperatures (T_A) were calculated for each primer set based on the melting temperature (T_m) of the primers in the specific PCR mixture using the New England BioLabs T_m Calculator version 1.12.0 (https://tmcalculator.neb.com/#!/main). PCR clean-up was carried out using Macherey-Nagel NucleoSpin® Gel and PCR Clean-up kit and the protocol followed was as per the manufacturer's instructions. **Table S1** and **Table S2** show the standard PCR components and a general protocol, respectively. Following clean-up, DNA concentration was quantified using an Invitrogen™ Qubit™ 4 Fluorometer using an Invitrogen™ Qubit™ dsDNA BR Assay Kit.

MIDAS cloning

Cloning reactions and protocols were carried out as described in van Dolleweerd *et al.* 2018[6]

Media/reagents for fungal protocols

All media was made up in in Milli-Q® Water and sterilized by autoclaving. Antibiotics and supplements were added to cool media post sterilization.

LB Broth

25 g/L LB Broth (Miller).

LB agar

25 g/L LB Broth (Miller), 20 g/L Select Agar (Invitrogen).

Trace elements (TE) solution

1.7 mM $FeSO_4.7H_2O$, 1.73 mM $ZnSO_4.7H_2O$, 0.59 mM $MnSO_4.H_2O$, 0.2 mM $CuSO_4.5H_2O$, 0.17 mM $CoCl_2.6H_2O$, 21.9 mL conc. HCl (made up in 0.6 N HCl).

Czapex-Dox Yeast Extract (CDYE) broth

34 g/L Czapex-Dox (Oxoid Ltd), 5 g/L Yeast Extract (Oxoid Ltd), 5 mL/L TE solution.

Czapex-Dox Yeast Extract Agar (CDYE-A)

34 g/L Czapex-Dox (Oxoid Ltd), 5 g/L Yeast Extract (Oxoid Ltd), 5 mL/L TE solution, 15 g/L Select Agar (Invitrogen).

YEPGA medium

10 g/L Yeast Extract (Oxoid Ltd), 20 g/L Bactopeptone (Oxoid Ltd), 20 g/L D-(+)-glucose anhydrous (VWR International BVBA), 20 mL/L TE solution, 4 g/L Select Agar (Invitrogen), pH 6.0.

1.5% RG Agar

20 g/L Malt extract (Oxoid Ltd), 20 g/L D-(+)-glucose anhydrous (VWR International BVBA), 10 g/L Mycological peptone (Oxoid Ltd), and 273.8 g/L Sucrose (ECP Ltd.) 15 g/L Select Agar (Invitrogen).

0.8% RG Agar

20 g/L Malt extract (Oxoid Ltd), 20 g/L D-(+)-glucose anhydrous (VWR International BVBA), 10 g/L Mycological peptone (Oxoid Ltd), and 273.8 g/L sucrose (ECP Ltd) 8 g/L Select Agar (Invitrogen).

OM buffer

10 mM Na_2HPO_4, 1.2 M $MgSO_4.7H_2O$, brought to pH 5.8 with 100 mM $NaH_2PO_4.2H_2O$.

Lysing enzymes solution

Lysing enzymes solution was prepared by resuspending Lysing Enzymes from *Trichoderma harzianum* (Sigma L1412) at 10 mg/mL in OM buffer and filter sterilized.

ST buffer

0.6 M D-Sorbitol (Sigma) and 0.1 M Tris-HCl at pH 8.0.

STC buffer

1 M D-Sorbitol (Sigma), 50 mM Tris-HCl at pH 8.0, and 50 mM $CaCl_2$.

40% Polyethylene glycol (PEG) solution

400 g/L PEG 4000 platelets (Sigma Aldrich), 5.6 g/L $CaCl_2$, 50 mL/L 1 M Tris-HCl (pH 8.0), 182.1 g/L D-sorbitol (Sigma).

Antibiotics

Bacterial work: Spectinomycin (1000x), 10 mg/mL (Gold Biotechnology); Kanamycin (1000x), 50 mg/mL (PanReac AppliChem)
Fungal work: Geneticin, G418 (333x), 50 mg/mL (Sigma); Nourseothricin (1000x), 100 mg/mL (GoldBio).

Protoplast preparation

The preparation of fungal protoplasts for transformation was according to Yelton *et al*. 1984 with modifications.[13] 5 × 25 mL (or 2 × 50 mL) aliquots of CDYE + TE in sterile Erlenmeyer flasks, were inoculated with 5×10^6 spores and incubated for 28+ h at 28 °C with shaking (200 rpm). The mycelia were separated from the medium using a sterile nappy liner in a sterile glass funnel. Mycelia were then rinsed three times with Milli-Q® Water and once with OM buffer. Mycelia were weighed, resuspended in 10 mL of filter-sterilized Lysing Enzymes solution per gram of mycelia, and incubated for 16+ h at 30 °C with shaking at 80 rpm. Protoplasts were filtered through a sterile nappy liner into an Erlenmeyer flask. Aliquots (5 mL) of filtered protoplasts were transferred into sterile 15 mL centrifuge tubes and gently overlaid with 2 mL of ST buffer. Tubes were centrifuged at 4300 rpm for 15 min at 4 °C, with a reduced deceleration speed to facilitate the formation of a protoplast layer. The protoplast layer was transferred (in 2 mL aliquots) into sterile 15 mL centrifuge tubes, gently washed by pipette resuspension in 5 mL of STC buffer and centrifuged at 4300 rpm for 5 min at 4 °C. The supernatant was decanted off and pelleted protoplasts from multiple tubes were combined by resuspension in 5 mL aliquots of STC buffer. The STC buffer wash was repeated three times until protoplasts were pooled into a single 15 mL centrifuge tube. The final protoplast pellet was resuspended in 500 µL of STC buffer (or less if pellet was small) and protoplast concentration was estimated with a hemocytometer. Protoplast stocks were diluted to give a final concentration of 1.25×10^8 protoplasts per mL of STC buffer. Aliquots of protoplasts (100 µL) were used immediately for fungal transformations and excess protoplasts were preserved as 80 µL aliquots with 20 µL of 40% PEG (w/v) solution in 1.7 mL micro-centrifuge tubes and stored at −80°C.

Transformation of *P. paxilli*

The following method is modified from van Dolleweerd *et al.* 2018,[1] which was originally modified from Vollmer and Yanofsky 1986[14] and Oliver *et al*. 1987.[15] Fungal transformations were carried out in 1.7 mL micro-centrifuge tubes containing 100 µL (1.25×10^7) protoplasts. 5–15 µg of plasmid DNA, 2 µL of spermidine (50 mM in H_2O) and 5 µL heparin (5 mg/mL in STC buffer) was added to the protoplasts. When plasmid volumes above 25 µL were required to achieve the desired concentration an equal volume of 2× STC buffer was added, and volumes of spermidine and heparin were adjusted appropriately. Protoplasts were incubated on ice for 30 min then 900 µL of 40% PEG solution was added and incubated on ice for a further 20 min. The protoplast mixture was gently transferred to 20 mL of 0.8% RGA medium (pre-warmed to 50 °C) in sterile 50 mL falcon tubes, mixed by inversion, and 3.5 mL aliquots were dispensed onto 1.5% RGA plates (5×) which were subsequently incubated overnight at 28 °C. The following day, 3.5 mL of 0.8% RGA (containing the appropriate antibiotic) was overlaid onto each plate. Plates were incubated at 28 °C until single colonies had formed, typically 3–5 days. Transformants were selected (up to ten per transformation) and spores were streaked onto CDYE agar plates supplemented with suitable antibiotic. Streaked plates were incubated at 28 °C until single colonies had formed, typically 3-5 days. Spores from isolated colonies (i.e., one colony from each plate/transformant) were suspended in 40 µL of 0.01% (v/v) Triton-X-100 and transferred/spotted onto either new CDYE agar plates, or 1.5% RGA plates, both supplemented with the appropriate antibiotic. These sporulation plates were incubated at 28 °C for 3–5 days. The spores from these plates were removed by washing colonies in 2 mL 0.01% (v/v) Triton X-100. Suspended spores (800 µL) were mixed with 200 µL of 50% (w/v) glycerol in a 2 mL screw cap tube. The concentration of these spore stocks was measured either by counting using a hemocytometer and Leica DM500 microscope, or by measuring OD_{600} using an Implen NanoPhotometer® NP80-Touch UV/Vis spectrophotometer. Dilutions of spore stocks were made (if necessary) so all samples have approximately the same amount of spores. Spore stocks were used to inoculate CDYE media for secondary metabolite growth, then stored at −80 °C. Fungal strains or transformants (from spores stocks, 5×10^6 spores) were grown in 25 mL cultures of CDYE medium with trace elements in 125 mL Erlenmeyer flasks capped with cotton wool for 7 days at 28 °C at 200 rpm. Mycelia from fungal cultures were separated from the media by filtration through nappy liners and excess liquid was removed by

squeezing. Total wet biomass from each culture was weighed and imaged, and the mycelia was divided into three tubes: an Eppendorf containing 500 µL RNAlater® (for DNA/RNA analysis), a 2 mL tube containing 0.85 g of mycelia and 0.1 g 1.6 mm stainless steel beads (for chemical extraction), and a 2 mL tube of mycelia only (an extra sample in case further work is required). For chemical extraction, 0.5 mL of ethyl acetate was added to 0.85 g of mycelia, and samples were homogenized for 40 s at 6.0 m/s using MP Biomedicals FastPrep-24™ 5G. The tubes were then centrifuged at 17,000 × g for 10–15 min, and the resultant supernatant was used analysis by TLC and LC-MS.

estB3 transcript analysis

RNA extraction and cDNA synthesis

Prior to RNA extraction, the surfaces of benches, pipettes, and any other potential source of RNases were cleaned thoroughly using RNaseZAP™ (Sigma, Cat. No. R2020). RNase-free 2 mL screw cap tubes were filled with one scoop of 0.5 mm RNase-Free Glass Beads (Next Advance, Cat. No. GB05-RNA) and 600 µL of TRIzol™ Reagent (ThermoFisher Scientific, Cat. No. 15596026). Frozen mycelia were allowed to thaw prior to being dried on a KimWipe™ before being transferred into the tube. Mycelia were then homogenized using an MP Biomedicals FastPrep-24™ 5G instrument in two cycles of speed 6.0 m/s for 30 s with a pause time of 5 min between each cycle. Following homogenization, the mycelia were centrifuged at 12,000 × g (with the temperature maintained at 4 °C) for 10 min to pellet the cell debris. The supernatant containing RNA (approximately 400 µL) was transferred to a fresh RNase-free 1.7 mL centrifuge tube followed by the addition of 80 µL of chloroform which was mixed by inversion. Tubes were further centrifuged at 12,000 × g (maintained at 4 °C) for 15 min.

Approximately 200 µL of the top layer was used as the starting sample material for RNA clean-up using the Invitrogen™ PureLink™ RNA Mini Kit. The manufacturer's instructions relating to clean up using TRIzol™ Reagent were followed with the exception of the final step where Milli-Q water was used instead of the RNase-Free water for eluting the RNA from the column. Instead of the optional On-column PureLink® DNase Treatment, DNA contamination was removed using the Invitrogen™ TURBO DNA-*free*™ Kit which was performed as per the manufacturer's instructions.

The concentration and quality of the obtained RNA were determined using the Invitrogen™ Qubit™ RNA BR Assay Kit and the Invitrogen™ Qubit™ RNA IQ Assay Kit, respectively. Quality scores above 7 were deemed acceptable for cDNA synthesis.

Synthesis of cDNA was performed using Applied Biosystems™ High-Capacity cDNA Reverse Transcription Kit as per the manufacturer's instructions. Typically, each reaction was set up to utilize 1 or 2 µg of RNA in a total volume of 20 µL. Controls lacking the reverse transcriptase (−RT controls) were generated at the same time as the sample. Once synthesized, the cDNA was diluted 5× with Milli-Q water and stored at −80 °C.

TLC and LC-MS screening of transformant metabolites

The EtOAc supernatant from extracted mycelia was spotted on thin layer chromatography (TLC) plates (Merck, TLC Aluminium sheets, silica gel 60 matrix, F_{254} fluorescent indicator) and the solvent system used was typically 9:1 $CHCl_3$:MeCN. TLC plates were stained with Ehrlich's reagent [1% (w/v) *p*-dimethylaminobenzaldehyde (DMAB) in 24% (v/v) HCl and 50% ethanol] for visualization of indole diterpenes (DMAB binds to two indole moieties to form resonance stabilized carbenium ion compounds that have distinctive colours: for IDTs, this colour is typically green). The remaining EtOAc extract was then transferred to a new tube for overnight evaporation or was concentrated using a vacuum concentrator. The dried extracts were resuspended in 150 µL MeCN and syringe filtered through a Millex®-LG, 4mm, 0.2 µm PTFE membrane into amber LC-MS vials (with glass inserts). LC-MS analysis of transformants was performed on an Agilent 1260 Infinity II LC-MS system with a diode array detector and electrospray ionization, coupled with a Phenomenex Kinetex®, 2.6 µm C18 (50 × 2.1 mm) column equipped with a Phenomenex C18 guard cartridge and maintained at 40 C. Compounds were eluted with a mobile phase of A: H_2O and B: MeCN, both containing 0.1% formic acid. An injection volume of 10 µL and flow rate of 0.4 mL/min were used. The gradient was as follows: 0–1 min 40% B, 1–5 min 40–60% B, 5–24 min 60–90% B, 24–25 min 90–100% B, 25–27 min 100% B. Potential indole diterpenes were identified based on their UV profiles (absorption maxima at 230 and 280 nm)[16-19] as well as the fragment ion of *m/z* 130.1.[20-22] Data was exported from Agilent OpenLab ChemStation software as .csv files and processed with Microsoft Excel.

HR-ESI-MS

High-resolution (ESI) mass spectrometric data was obtained with an Agilent 6530 Accurate Mass Q-TOF LC-MS (positive ion mode) equipped with a 1260 Infinity binary pump. Samples were directly injected at 0.3 mL/min, 50:50 A:B where A = H_2O + 0.1% formic acid and B = MeCN + 0.1% formic acid. The mass spectrometer parameters used were: mass range 100-1000 Da, fragmentor voltage 175 V, skimmer voltage 65 V, gas temperature 300 °C, nebulizer pressure 35 psi, drying gas flow rate 8 L/min, acquisition rate 2 scans/s. Calibrated to reference masses of 121.050873 and 922.009798. Data analysis was conducted using Agilent MassHunter Workstation Software, Qualitative Analysis, Version B.05.00, or Version B.08.00.

Compound isolation from *P. paxilli* strains and structural elucidation

Isolation of emindole SA (10)

Transformant #3 of pRB14:ΔpaxB5 was chosen for growth and spores from this strain were used to generate a seed culture by inoculating 50 mL of YEPGA medium in a 250 mL Erlenmeyer flask with approximately 5 × 10^6 spores. This culture was incubated at 28 °C with shaking at 200 rpm for 24 h and 4% (v/v) was used to inoculate 4 × 400 mL CDYE+TE flasks which were left to incubate at 28 °C with shaking at 200 rpm for 7 days.

Mycelia from these fungal cultures were separated from the liquid medium using a Büchner funnel and pooled together (120 g). The mycelia were then extracted with 250 mL of ethyl acetate by adding a stir bar and stirring overnight. The ethyl acetate was separated from the mycelia using a Büchner funnel and another extraction step with 250 mL of ethyl acetate was performed. The ethyl acetate from the two extractions were pooled together and were concentrated *in vacuo* (402.3 mg).

The crude extracts were fractionated on a Reveleris® HP silica flash cartridge using the Reveleris® X2 Flash Chromatography system using a petroleum ether: ethyl acetate gradient. Fractions containing UV active compounds were then further purified using semi-preparative reversed phase high-performance liquid chromatography on an Agilent 1260 Infinity II LC system coupled with a Phenomenex Luna® 5 µm C18 100 Å column (250 × 50 mm) using an isocratic solvent system of 85% MeCN in H_2O (with 0.1% formic acid), with a flow rate of 3 mL/min, giving 0.47 mg of emindole SA (**10**).

Emindole SA (**8**): colourless oil; HR-ESI-MS *m/z* 406.3107 $[M + H]^+$ (calcd. for $C_{28}H_{39}NO$, 406.3032).

Isolation of emindole DA (11) and emindole DB (12)

Transformant #8 of pRB66:CY2 was chosen for growth and spores from this strain were used to generate a seed culture by inoculating 50 mL of YEPGA medium in a 250 mL Erlenmeyer flask with approximately 5 × 10^6 spores. This culture was incubated at 28 °C with shaking at 200 rpm for 24 h and 4% (v/v) was used to inoculate 4 × 400 mL CDYE+TE flasks which were left to incubate at 28 °C with shaking at 200 rpm for 7 days.

Mycelia from these fungal cultures were separated from the liquid medium using a Büchner funnel and pooled together (140 g). The mycelia were then extracted with 250 mL of ethyl acetate by adding a stir bar and stirring overnight. The ethyl acetate was separated from the mycelia using a Büchner funnel and another extraction step with 250 mL of ethyl acetate was performed . The ethyl acetate from the two extractions were pooled together and were concentrated *in vacuo* (1.5 g).

The crude extracts were fractionated on a Reveleris® HP silica flash cartridge using the Reveleris® X2 Flash Chromatography system using a gradient of chloroform:methanol (1% methanol over 5 CV, followed by an increase to 15% over 12 CV, ending with 100% methanol for 4 CV). Fractions containing UV active compounds were then further purified using semi-preparative reversed phase high-performance liquid chromatography on an Agilent 1260 Infinity II LC system coupled with a Phenomenex Luna® 5 µm C18 100 Å column (250 × 50 mm) using an isocratic solvent system of 90% MeCN in H_2O (with 0.1% formic acid), with a flow rate of 3 mL/min, giving 2.25 mg of emindole DA (**9**) and 2.12 mg of emindole DB (**12**).

Emindole DA (**11**): colourless oil; HR-ESI-MS *m/z* 406.3112 $[M + H]^+$ (calcd. for $C_{28}H_{39}NO$, 406.3032).

Emindole DB (**12**): colourless oil; HR-ESI-MS *m/z* 422.3056 $[M + H]^+$ (calcd. for $C_{28}H_{39}NO_2$, 422.2981).

NMR

NMR spectra were recorded on a JEOL JNM-ECZ600R with a nitrogen cooled 5 mm SuperCOOL cryogenic probe (600 MHz for ^{1}H nuclei and 150 MHz for ^{13}C nuclei). The residual solvent peak was used as an internal chemical shift reference ($CDCl_3$: δ_C 77.16; δ_H 7.26). Spectra were acquired with Delta^{TM} 5.2 acquisition software and processed with MestReNova.

Supplementary Tables

Table S1. Standard PCR components

Standard PCR components	
Component	**Final concentration**
DNA template	1–10 ng DNA
Forward primer	0.5 µM
Reverse primer	0.5 µM
2x Master Mix	1x
Milli-Q® H_2O	To final volume

Table S2. Standard PCR protocol

Standard PCR protocol		
Temperature	**Time**	
98 °C	30 s	
95 °C	20 s	35 cycles
T_A °C	20 s	
72 °C	20 s/kB	
72 °C	5 min	

Table S3. Primers used in this study

ORF	*Primer Name*	*Primer Sequence 5'-3'*
estC1	estC1 Frag1 F	cgatgtacgtctcaCTCGAATGCAGGGTGAGCACAG
	estC1 Frag1 R	actgctCGTCTCATGACCGCTGAGAAAACCATCC
	estC1 Frag2 F	actgctCGTCTCAGTCACGAATGATAGGACCTAAGCAAC
	estC1 Frag2 R correct	actgctCGTCTCAAGTCTCTGATGGACCGTATAATATATGTGC
	estC1 Frag3 F correct	actgctCGTCTCaGACTCTCAATCGCGCTTACTA
	estC1 Frag3 R	gacctttcgtctctGAGTGACCCCGTCTTCAACGATGC
	estC1 Frag4 F	actgctCGTCTCCACTCTTCCGACTTCTTGGACAC
	estC1 Frag4 R	gacctttcgtctctGTCTcaAAGCTCAGCGCCCCCCG
estM1	estM1 Frag1 F	cgatgtacgtctcaCTCGAATGGGGGAGCAGACCTT
	estM1 Frag1 R	TATGCGcgtctcgGAGGCCGAACACTTCCTTCGCC
	estM1 Frag2 F	GGAAGTcgtctcGCCTCGTAATTATCATTGGATTAGCGT
	estM1 Frag2 R	gacctttcgtctctGTCTcaAAGCTCAAGTCAAACGCGATGAAGCA
estB1	esB_4gene F	cgatgtacgtctcaCTCGAATGGACGGCTACGATGTTTCTC
	esB_4gene R	gacctttcgtctctGTCTcaAAGCTTATCTTTCCTTTCGCTGGGCAAGA
estC2	estC2 F	cgatgtacgtctcaCTCGAATGTCTACGGCCACGTCC
	estC2 R	gacctttcgtctctGTCTcaAAGCTTAGCTCTGCTGGTCAAGCTTCT
estM2	esM_6gene Frag1 F	cgatgtacgtctcaCTCGAATGGAGAAATCGACCTTCAGAG
	esM_6gene Frag1 R	gacctttcgtctctATTGGATGCTAACACTGATACCACTTGAGATCGG
	esM_6gene Frag2 F	cgatgtacgtctcaCAATGACGGGGTCACTGTGACCA
	esM_6gene Frag2 R	gacctttcgtctctGTCTcaAAGCCTAATGATATTGAGCGGAAAAGAGG
estB2	esB_6gene F	cgatgtacgtctcaCTCGAATGGACGGCTTCGACGTTTCTCAG
	esB_6gene F	gacctttcgtctctGTCTcaAAGCTTACTTTGCCTTCTGTTGAGAGAGA
estB3	estB3 Frag1 F	cgatgtacgtctcaCTCGAATGGAAGAAGGTTGGGATTTCGA
	estB3 Frag1 R	gacctttcgtctctGACTCCGGATGTGCGATCTCG
	estB3 Frag2 F	actgacCGTCTCGAGTCTCCTTAATAGCACTGAGT
	estB3 Frag2 R	gacctttcgtctctGTCTcaAAGCCTAAGCAAGCTTCTGTCTCC
estB3 (RT-PCR)	estB3_qPCR_F	CTCTCCGTGCTGCTTAGTCC
	estB3_qPCR_R	TCCAACACAAAGAAAGCCCCT

Table S4. Fungal strains used in this study

Strain name	***Description***	***IDT phenotype***	***Source/Reference***
Aspergillus striatus NHL 80-NE-22 (ATCC® 64988™)	*Aspergillus striata* Wild type	Paxilline, Emindole SA	ATCC® (https://www.atcc.org/)
Penicillium paxilli PN2013 (ATCC®26601™)	*Penicillium paxilli* Wild type	Paxilline	Barry Scott, Massey University[23]
Penicillium paxilli PN2250 (CY2)	PN2013/Deletion of entire *PAX* locus (*ΔPAX*); Hyg^R	-	Barry Scott, Massey University[24]
Penicillium paxilli PN2290	PN2013/*ΔpaxC*::P_{trpC}-*hph*-T_{trpC}; Hyg^R	-	Barry Scott, Massey University[25]
Penicillium paxilli PN2257	PN2013/*ΔpaxM*::P_{glcC}-*hph*-T_{trpC}; Hyg^R	-	Barry Scott, Massey University[25]
Penicillium paxilli ΔpaxB5	PN2013/*ΔpaxB*::P_{trpC}-*nptII*-T_{trpC}; Gen^R	-	Rosannah Cameron, Ferrier Research Institute[26]

Table S5. Summary of bioinformatic analysis/gene prediction of the *EST* cluster of *Aspergillus striatus*

Gene	Contig ID	Amino Acids	Protein homologue,[a] origin, accession number	Identity (%)	Proposed function[b]
estG1	15542	364	JanG, *Penicillium janthinellum*, A0A0E3D8M9.2	57.70	Geranylgeranyl pyrophosphate synthase
estC1	15542	342	PenC, *Penicillium crustosum*, A0A0E3D8N1.1	60.90	Prenyltransferase
estM1	15542	475	JanM, *Penicillium janthinellum*, A0A0E3D8L5.1	58.92	FAD-dependent monooxygenase
estB1	15542	243	PenB, *Penicillium crustosum*, A0A0E3D8M2.1	65.37	Cyclase
estG2	15201	376	JanG, *Penicillium janthinellum*, A0A0E3D8M9.2	53.79	Geranylgeranyl pyrophosphate synthase
estC2	15201	342	PenC, *Penicillium crustosum*, A0A0E3D8N1.1	66.25	Prenyltransferase
estM2	15201	465	PenM, *Penicillium crustosum*, A0A0E3D8L6.1	61.40	FAD-dependent monooxygenase
estB2	15201	243	JanB, *Penicillium janthinellum*, A0A0E3D8Q2.1	78.01	Cyclase
estP2	15201	541	PenP, *Penicillium crustosum*, A0A0E3D8K9.1	75.10	Cytochrome P450 monooxygenase
estQ2	15201	514	PtmQ, *Penicillium simplicissimum*, A0A140JWT0.1	64.69	Cytochrome P450 monooxygenase
estA2	15201	368	PenA, *Penicillium crustosum*, A0A0E3D8L0.1	39.24	Membrane protein
estV2	15201	479	PtmV, *Penicillium simplicissimum*, A0A140JWS4.1	50.20	Acyltransferase
estB3	15582	242	Cle7, *Aspergillus versicolor*, BBG28477.1	76.13	Cyclase

[a] If available, closest characterized homologue is shown. Uncharacterized homologues referred to by locus tag.

[b] Proposed function based on homologues or analysis of NCBI Conserved Domain Search.

Table S6. Similarity/identity matrices for *EST* and *DES* gene products with known IDT homologues

Protein	AtS2B	AfB	AtS5B1	EstB1	DesB	EstB2	NodB	AtmB	PaxB	PenB	PtmB	TerB	LtmB
AtS2B	100	69.54	47.32	52.47	51.86	53.71	47.73	53.9	51.85	53.49	53.49	50.62	43.03
AfB	56.79	100	44.85	50	49.79	50.82	44.26	50.61	48.77	49.59	48.77	45.64	41.35
AtS5B1	34.97	30.86	100	56.61	58.92	61.57	60.49	60.08	62.96	61.72	61.72	57.67	54
EstB1	38.01	37.6	43.8	100	78.83	80.16	63.22	70.66	69	72.31	71.9	59.33	54.85
DesB	38.17	39.41	47.71	73.02	100	90.04	69.29	71.78	72.61	75.93	75.93	63.48	60.75
EstB2	39.25	40.49	49.17	75.61	86.72	100	73.14	73.96	76.44	80.16	80.16	68.46	63.29
NodB	35.39	33.6	48.14	52.89	60.58	64.04	100	72.83	72.95	74.59	74.18	63.07	61.6
AtmB	41.56	39.5	48.55	60.33	63.07	64.46	64.19	100	72.83	79.01	78.6	64.73	62.86
PaxB	38.27	38.11	50.61	59.09	66.39	69.83	63.11	62.13	100	82.78	81.96	63.48	60.75
PenB	42.38	40.16	47.73	62.8	68.87	72.31	68.03	70.78	72.54	100	98.36	66.39	62.86
PtmB	42.38	40.16	47.73	62.39	69.7	73.14	67.62	70.37	72.95	97.54	100	65.97	62.44
TerB	36.09	32.36	46.88	48.54	53.52	58.5	53.52	53.52	56.43	56.43	56.84	100	74.68
LtmB	33.33	31.64	40.08	45.14	51.47	53.58	50.21	50.63	51.05	53.58	53.58	67.08	100

Green = amino acid identity (%); orange = amino acid similarity (%)

Protein	NodM	AtmM	EstM1	DesM	EstM2	PtmM	PenM	JanM	PaxM
NodM	100	61.77	60.25	61.33	61.12	65.87	61.33	62.41	61.55
AtmM	51.4	100	58.57	59.7	59.29	66.11	60	59.78	60.25
EstM1	50.97	49.16	100	79.49	76.15	72.74	69.24	64	61.5
DesM	51.83	49.89	72.8	100	85.5	75.35	71.82	65.47	62.97
EstM2	51.4	49.06	69.66	79.71	100	75.11	70.53	65.05	63.8
PtmM	56.16	56.39	63.98	66.11	65.87	100	92.65	75.82	72.98
PenM	51.18	51.39	58.27	61.07	60.64	90.99	100	72.25	69.89
JanM	52.69	50.1	57.47	58.1	59.78	66.82	62.15	100	76
PaxM	50.53	50.83	53.97	54.6	55.64	64.92	61.72	68	100

Green = amino acid identity (%); orange = amino acid similarity (%)

Protein	NodC	EstC2	EstC1	AtmC	PtmC	PenC	JanC	PaxC	TerC	LtmC
NodC	100.00	65.53	65.64	69.32	67.79	68.71	66.87	65.29	55.82	56.13
EstC2	54.76	100.00	78.46	69.84	69.84	71.07	69.23	65.93	53.84	52.00
EstC1	54.60	72.92	100.00	65.26	63.90	64.79	63.60	62.77	51.63	50.88
AtmC	57.97	60.00	57.18	100.00	73.65	74.55	69.72	67.50	53.89	51.79
PtmC	56.44	63.07	58.87	66.76	100.00	96.78	76.75	70.97	52.22	49.12
PenC	57.66	64.30	59.46	66.46	95.61	100.00	77.37	72.23	52.52	50.00
JanC	56.13	61.53	55.35	59.63	69.11	69.11	100.00	77.60	51.68	51.98
PaxC	54.25	58.04	53.62	57.72	63.09	64.98	72.87	100.00	50.78	48.26
TerC	44.47	44.30	40.94	44.61	43.62	44.21	41.89	44.16	100.00	64.09
LtmC	42.33	40.30	39.34	41.01	39.18	40.05	39.44	38.80	55.48	100.00

Green = amino acid identity (%); orange = amino acid similarity (%)

Protein	AtmG	EstG2	EstG1	PtmG	PenG	JanG	PaxG	LtmG
AtmG	100.00	57.18	61.58	68.32	66.86	64.80	62.75	65.26
EstG2	48.09	100.00	70.08	54.77	56.59	54.67	54.56	55.98
EstG1	50.14	62.67	100.00	58.11	60.41	60.39	58.40	60.17
PtmG	58.35	45.50	49.57	100.00	95.60	64.88	61.51	59.58
PenG	57.77	47.50	51.02	94.72	100.00	68.32	64.80	59.88
JanG	53.66	44.23	50.14	56.46	59.23	100.00	71.70	61.37
PaxG	51.90	41.66	47.86	51.40	53.66	65.10	100.00	59.28
LtmG	52.69	43.11	48.50	49.70	49.40	49.70	48.20	100.00

Green = amino acid identity (%); orange = amino acid similarity (%)

Table S7. Summary of bioinformatic analysis/gene prediction of the *DES* cluster of *Aspergillus desertorum*

Gene	Contig ID	Amino Acids	Protein homologue,[a] origin, accession number	Identity (%)	Proposed function[b]
desG	135	340	AtmG, *Aspergillus flavus*, Q672V6.1	60.68	Geranylgeranyl pyrophosphate synthase
desA	135	348	PenA, Penicillium crustosum	40.88	Membrane protein
desC	135	342	PtmC, *Penicillium simplicissimum*, A0A140JWT3.1	64.22	Prenyltransferase
desM	135	422	PtmM, *Penicillium simplicissimum*, A0A140JWT1.1	63.21	FAD-dependent monooxygenase
desB	135	243	JanB, *Penicillium janthinellum*, A0A0E3D8Q2.1	73.44	Cyclase

[a] If available, closest characterized homologue is shown. Uncharacterized homologues referred to by locus tag.

[b] Proposed function based on homologues or analysis of NCBI Conserved Domain Search.

Table S8. MIDAS Level 1 plasmids[a]

Plasmid name	MIDAS Level 1 module (promoter, coding sequence or terminator)[b]
pRB4	*estB1*
pRB5	*estB2*
pRB6	*estB3*
pRB7	*estC2*
pRB8	*estC1*
pRB29	*estM2*
pRB32	*estM1*
pRB39	*desB*
pRB41	*desM*
pKV28	*PpaxC2*[b]
pSK12	*TpaxC*[b]
pSK4	*PpaxM*[b]
pSK6	*TpaxM*[b]
pSK7	*PpaxB*[b]
pSK9	*TpaxB*[b]
pSK11	*paxC*
pSK1	*PpaxG*[b]
pSK2	*paxG*
pSK3	*TpaxG*[b]
pSK17	*PtrpC*[b]
pSK15	*TtrpC*[b]
pRC1	*natR*

[a] All MIDAS Level 1 plasmids assembled in pML1 vector.[1]

[b] "P" at the start indicates promoter region, T at the start indicates terminator region

Table S9. MIDAS Level 2 plasmids

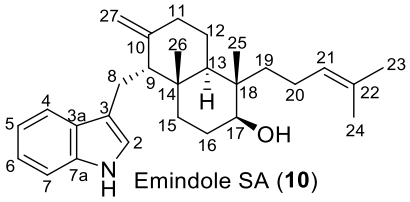
Plasmid name	Description	Level 1 entry clones (promotor, CDS, terminator)	Vector
pRB9	PpaxB_estB1_TpaxB	pSK7, pRB4, pSK9	pML2(+)BF
pRB10	PpaxB_estB2_TpaxB	pSK7, pRB5, pSK9	pML2(+)BF
pRB11	PpaxB_estB3_TpaxB	pSK7, pRB6, pSK9	pML2(+)BF
pRB12	PpaxC2_estC1_TpaxC	pKV28, pRB8, pSK12	pML2(+)BF
pRB13	PpaxC2_estC2_TpaxC	pKV28, pRB7, pSK12	pML2(+)BF
pRB30	PpaxM_estM2_TpaxM	pSK4, pRB29, pSK6	pML2(+)BF
pRB33	PpaxM_estM1_TpaxM	pSK4, pRB32, pSK6	pML2(+)BF
pRB47	PpaxM_desM_TpaxM	pSK4, pRB41, pSK6	pML2(+)BF
pRB59	PpaxB_estB1_TpaxB	pSK7, pRB4, pSK9	pML2(+)WF
pRB63	PpaxB_desB _TpaxB	pSK7, pRB39, pSK9	pML2(+)WF
pSK23	PpaxB_paxB_TpaxB	pSK7, pSK8, pSK9	pML2(+)BR
pSK59	PpaxC_paxC_TpaxC	pKV28, pSK11, pSK12	pML2(+)WF
pSK82	PpaxG_paxG _TpaxG	pSK1, pSK2, pSK3	pML2(+)BF
pRC2	PtrpC_natR_TtrpC	pSK17, pRC1, pSK15	pML2(+)WF

Table S10. MIDAS Level 3 plasmids[a]

Plasmid Name	Description	Level 2 vector	Level 3 entry vector
pRC3	PtrpC_natR_TtrpC	pRC2	pML3.1
pRB14	PtrpC_natR_TtrpC_PpaxB_estB1_TpaxB	pRB9	pRC3
pRB15	PtrpC_natR_TtrpC_PpaxB_estB2_TpaxB	pRB10	pRC3
pRB16	PtrpC_natR_TtrpC_PpaxB_estB3_TpaxB	pRB11	pRC3
pRB17	PtrpC_natR_TtrpC_PpaxC2_estC1_TpaxC	pRB12	pRC3
pRB18	PtrpC_natR_TtrpC_PpaxC2_estC2_TpaxC	pRB13	pRC3
pRB31	PtrpC_natR_TtrpC_PpaxM_estM2_TpaxM	pRB30	pRC3
pRB34	PtrpC_natR_TtrpC_PpaxM_estM1_TpaxM	pRB33	pRC3
pRB49	PtrpC_natR_TtrpC_PpaxB_paxB_TpaxB	pSK23	pRC3
pRB52	PtrpC_natR_TtrpC_PpaxM_DesM_TpaxM	pRB47	pRC3
pRB54	PtrpC_natR_TtrpC_PpaxG_paxG_TpaxG	pSK82	pRC3
pRB55	PtrpC_natR_TtrpC_PpaxG_paxG_TpaxG_PpaxC_paxC_TpaxC	pSK59	pRB54
pRB58	PtrpC_natR_TtrpC_PpaxG_paxG_TpaxG_PpaxC_paxC_TpaxC_PpaxM_desM_TpaxM	pRB47	pRB55
pRB66	PtrpC_natR_TtrpC_PpaxG_paxG_TpaxG_PpaxC_paxC_TpaxC_PpaxM_estM1_TpaxM	pRB63	pRB58

[a] MIDAS Level-3 vectors are built from pRC3[27]

Table S11. Summary of the chemical shifts (δ_C and δ_H) for emindole SA (**10**) in $CDCl_3$

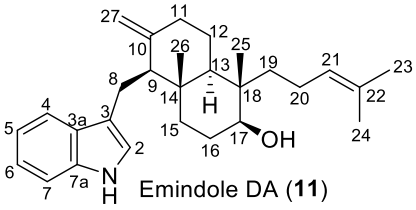


Emindole SA (**10**)

	Emindole SA (10) Nozawa *et al.*[28]		Emindole SA (10) This work	
Position	**δ_C (ppm)**	**δ_H (ppm)**	**δ_C (ppm)**	**δ_H (ppm)**
1 (NH)		7.88		7.88
2	121.7	6.89	121.8	6.91
3	116.2		116.2	
3a	127.7		127.8	
4	118.7	7.62	118.7	7.63
5	119.1	7.11	119.1	7.13
6	121.8	7.17	121.8	7.18
7	111.0	7.31	111.1	7.33
7a	147.9		136.0*	
8	19.6	2.98	19.6	2.99
		2.82		2.82
9	56.8	2.20	56.7	2.19
10	136.1		147.9*	
11	37.9	1.99	37.9	1.98
		2.40		2.39
12	23.7	1.59–1.66	23.7	1.59–1.66
		1.36–1.47		1.37–1.47
13	48.9	1.35	48.7	1.36
14	39.6		39.6	
15	37.4	2.05	37.3	2.06
		1.35–1.43		1.35–1.43
16	28.0	1.63–1.79	27.9	1.63–1.79
		1.63–1.79		1.63–1.79
17	73.3	3.60	73.2	3.61
18	41.4		41.3	
19	37.5	1.24–1.33	37.4	1.24–1.33
		1.47–1.56		1.47–1.56
20	21.6	1.80–1.95	21.5	1.80–1.95
		1.80–1.95		1.80–1.95
21	124.8	5.09	124.7	5.10
22	131.2		131.4	-
23	17.7	1.61	17.7	1.61
24	25.7	1.68	25.8	1.68
25	16.9	0.81	17.1	0.82
26	15.0	0.87	15.0	0.87
27	107.9	4.84	108.0	4.84
		4.72		4.72

*The δ_C values for positions 7a and 10 have been reversed from the assignment reported by Nozawa *et al.*[28] based on our observed HMBC correlations.

Table S12. Summary of the chemical shifts (δ_C and δ_H) of emindole DA (**11**) in $CDCl_3$



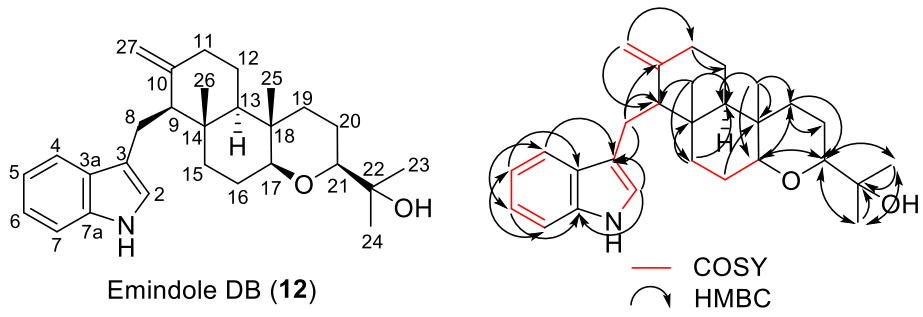
Emindole DA (**11**)

	Emindole DA (11) Nozawa *et al.*[29]		Emindole DA (11) This work	
Position	**δ_C (ppm)**	**δ_H (ppm)**	**δ_C (ppm)**	**δ_H (ppm)**
1 (NH)		7.89		7.87
2	121.9	6.89	121.9	6.89
3	115.6		115.6	
3a	127.6		127.7	
4	118.9	7.56	118.9	7.57
5	118.9	7.09	118.9	7.09
6	121.6	7.16	121.7	7.17
7	111.0	7.33	111.0	7.33
7a	148.1		136.2*	
8	23.3	3.13	23.2[a]	3.13
		2.70		2.70
9	58.6	2.06	58.5	2.06
10	136.3		148.1*	
11	30.8	2.28	30.8	2.28
		2.18		2.18
12	23.0	1.43	23.0	1.44
		1.58–1.65		1.58–1.65
13	39.1	1.73–1.76	38.9	1.73–1.76
14	37.9		37.9	
15	34.6	1.33	34.5	1.32
		1.94–2.02		1.94–2.02
16	27.8	1.73–1.80	27.7	1.73–1.80
		1.73–1.80		1.73–1.80
17	73.9	3.62	73.9	3.62
18	41.1		41.1	
19	37.5	1.34–1.40	37.4	1.34–1.40
		1.51–1.59		1.51–1.59
20	21.9	1.96–2.06	21.8	1.96–2.06
		1.96–2.06		1.96–2.06
21	124.9	5.16	124.8	5.16
22	131.2		131.4	
23	17.7	1.67	17.7	1.67
24	25.7	1.72	25.8	1.71
25	16.9	0.83	17.0	0.82
26	23.2	0.98	23.2[a]	0.98
27	110.2	4.51	110.2	4.51
		4.16		4.15

*The δ_C values for positions 7a and 10 have been reversed from the assignment reported by Nozawa *et al.*[29] based on our observed HMBC correlations.

[a]Interchangeable.

Table S13. Summary of NMR spectra obtained for emindole DB (**12**) in $CDCl_3$.



Position	δ_C (ppm)	Reported δ_C (ppm)[22]	Type	δ_H (ppm)	mult.	*J* (Hz)	COSY	HMBC
1 (NH)				7.88	br s		2	
2	122.0	121.9	CH	6.9	d	2.3	1, 8	8
3	115.8	115.6	C					2, 4, 8
3a	127.7	127.7	C					2, 5, 7, 8
4	119.1*	118.9	CH	7.57	dt	7.9, 1.0	5	6
5	119.1*	118.9	CH	7.09	ddd	8.0, 7.0, 1.0	4, 6	7
6	121.8	121.6	CH	7.17	ddd	8.1, 7, 1.2	5, 7	4
7	111.2	111.0	CH	7.33	dt	8.1, 1.0	6	5
7a	136.3	136.3	C			-		2, 4, 6
8	23.3	23.3	CH_2	3.13	dd	14.1, 3.9	2, 9	
				2.72	dd	14.4, 10.2		
9	58.9	58.9	CH	1.99–2.08	m		8	8, 11, 26, 27
10	148.2	148.0	C					8, 11
11	30.8	30.7	CH_2	2.18	m		27	27
				2.28	m			
12	22.3	22.2	CH_2	1.38–1.49	m			11
				1.56–1.67	m			
13	45.5	45.4	CH	1.49–1.39	m			11, 15, 25, 26
14	38.6	38.5	C					8, 15, 26
15	34.9	34.8	CH_2	1.99–2.08	m		16	17, 26
				1.36	dt	13.1, 3.3	-	
16	24.9	24.9	CH_2	1.72	qd	12.7, 3.4	15, 17	15, 17
				1.56–1.67	m			
17	85.9	85.9	CH	3.07	dd	11.8, 3.8	15	15, 19, 21, 25
18	36.4	36.3	C					16, 19
19	37.9	37.8	CH_2	1.86	ddd	12.7, 4.7, 2.6		17, 21, 25
				1.19	m			
20	22.1	22.0	CH_2	1.38–1.49	m			19
				1.56–1.67	m			
21	84.8	84.8	CH	3.23	dd	12.0, 2.9		17, 19, 23, 24
22	72.1	72.0	C					23, 24
23	23.8	23.8	CH_3	1.18	s			24
24	26.3	26.1	CH_3	1.20	s			23
25	13.6	13.5	CH_3	0.86	s			17, 19
26	24.0	23.9	CH_3	1.00	s			
27	110.3	110.2	CH_2	4.51	t	2.3	11	11
				4.17	t	2.1		

*Interchangeable.

Supplementary Figures

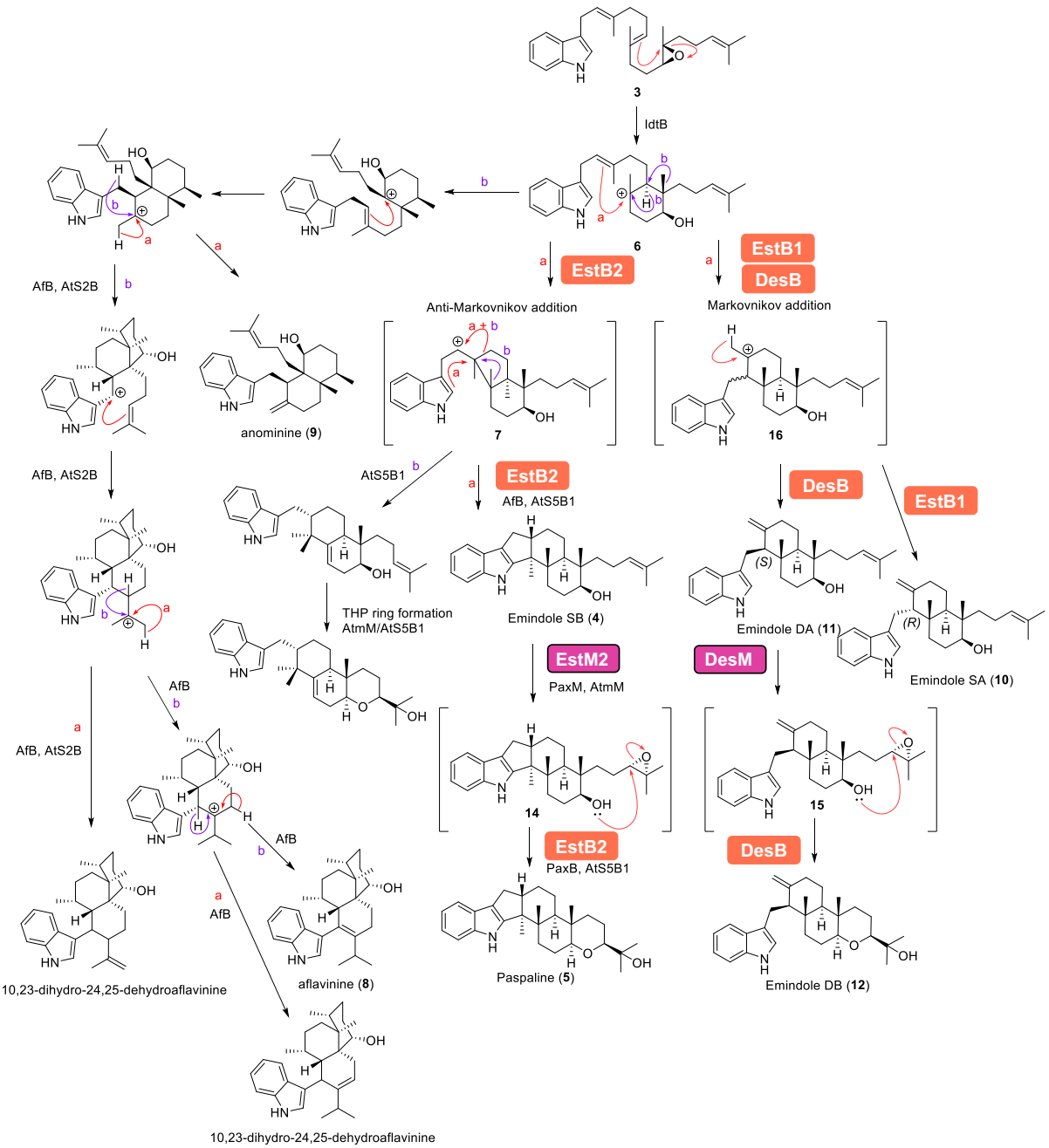


Figure S1. Mechanistic scheme summarizing IDT cyclization cascades in this work (DesB, EstB1, EstB2) and Tang *et al.* 2015.[30]

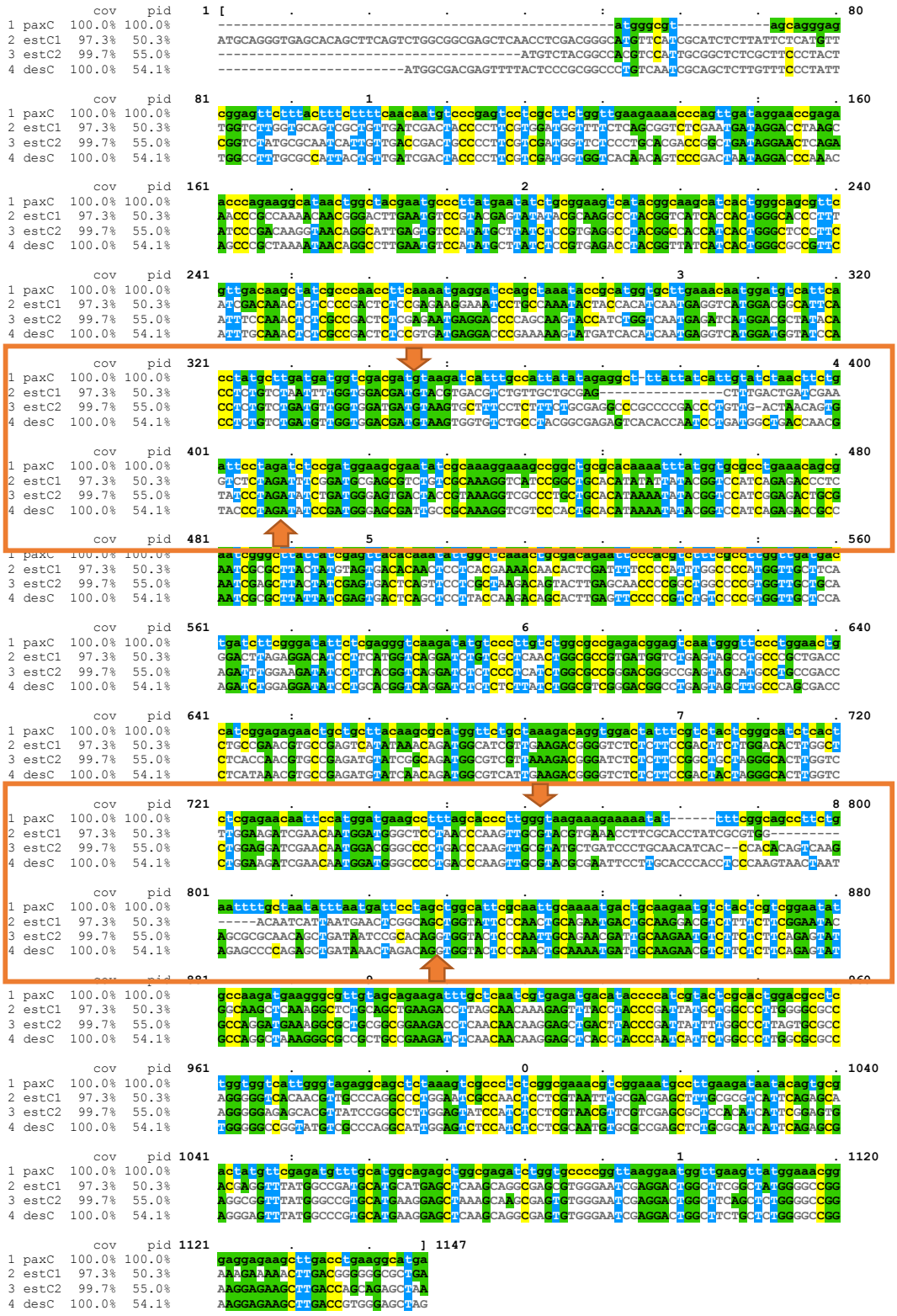
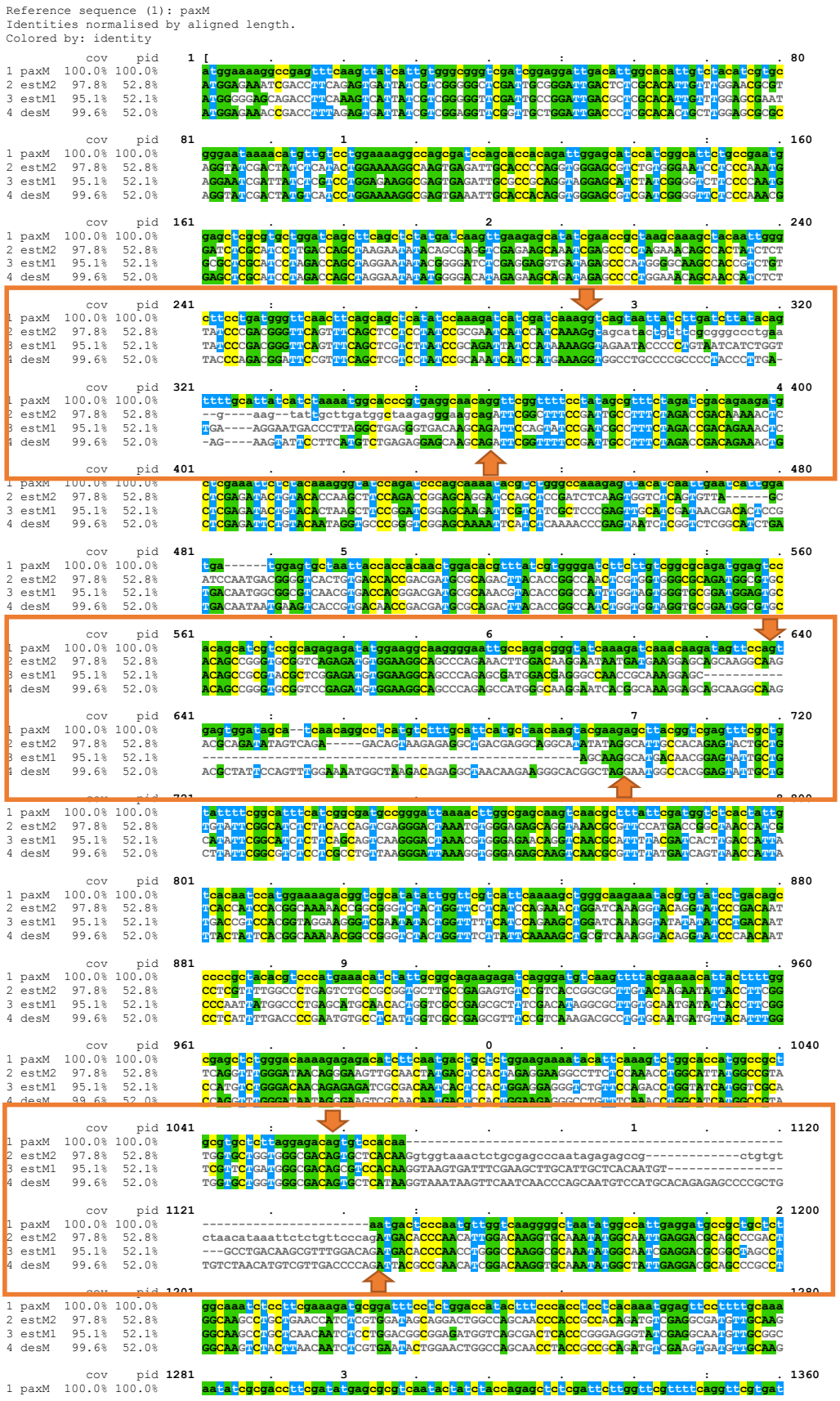


Figure S2. Nucleotide alignment for *idtCs*

Alignment of *paxC*, *estC1*, *estC2* and *desC*; demonstrating the high conservation of exon regions and low conservation of intron regions. Introns indicated by orange box and arrows. Alignment generated with ClustalW[31] and visualized with MView.[32]


Reference sequence (1): paxM
Identities normalised by aligned length.
Colored by: identity
cov pid
1 paxM 100.0% 100.0%
2 estM2 97.8% 52.8%
3 estM1 95.1% 52.1%
4 desM 99.6% 52.0%

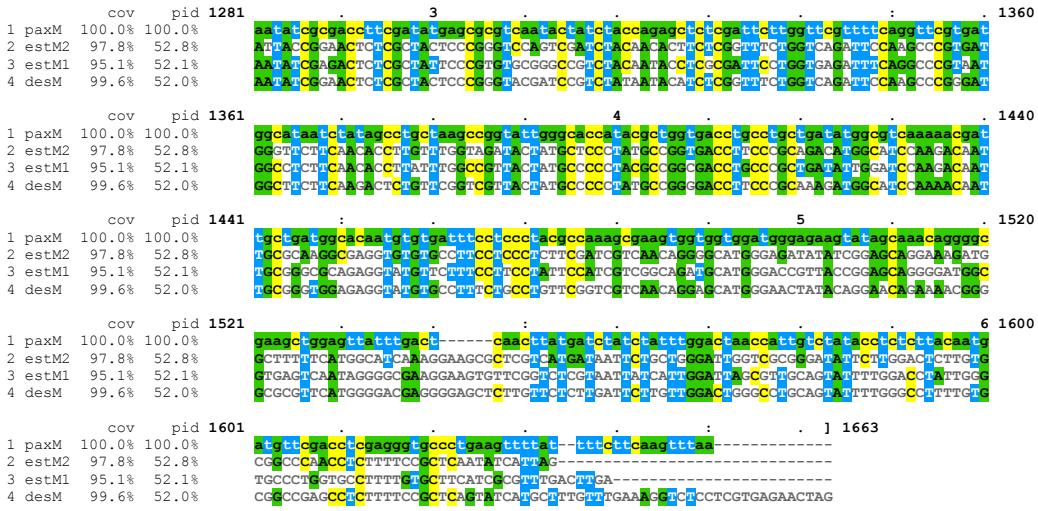


Figure S3. Nucleotide alignment for *idtMs*

Alignment of *paxM*, *estM1*, *estM2* and *desM*; demonstrating the high conservation of exon regions and low conservation of intron regions. Introns indicated by orange box and arrows. Alignment generated with ClustalW[31] and visualized with MView.[32]

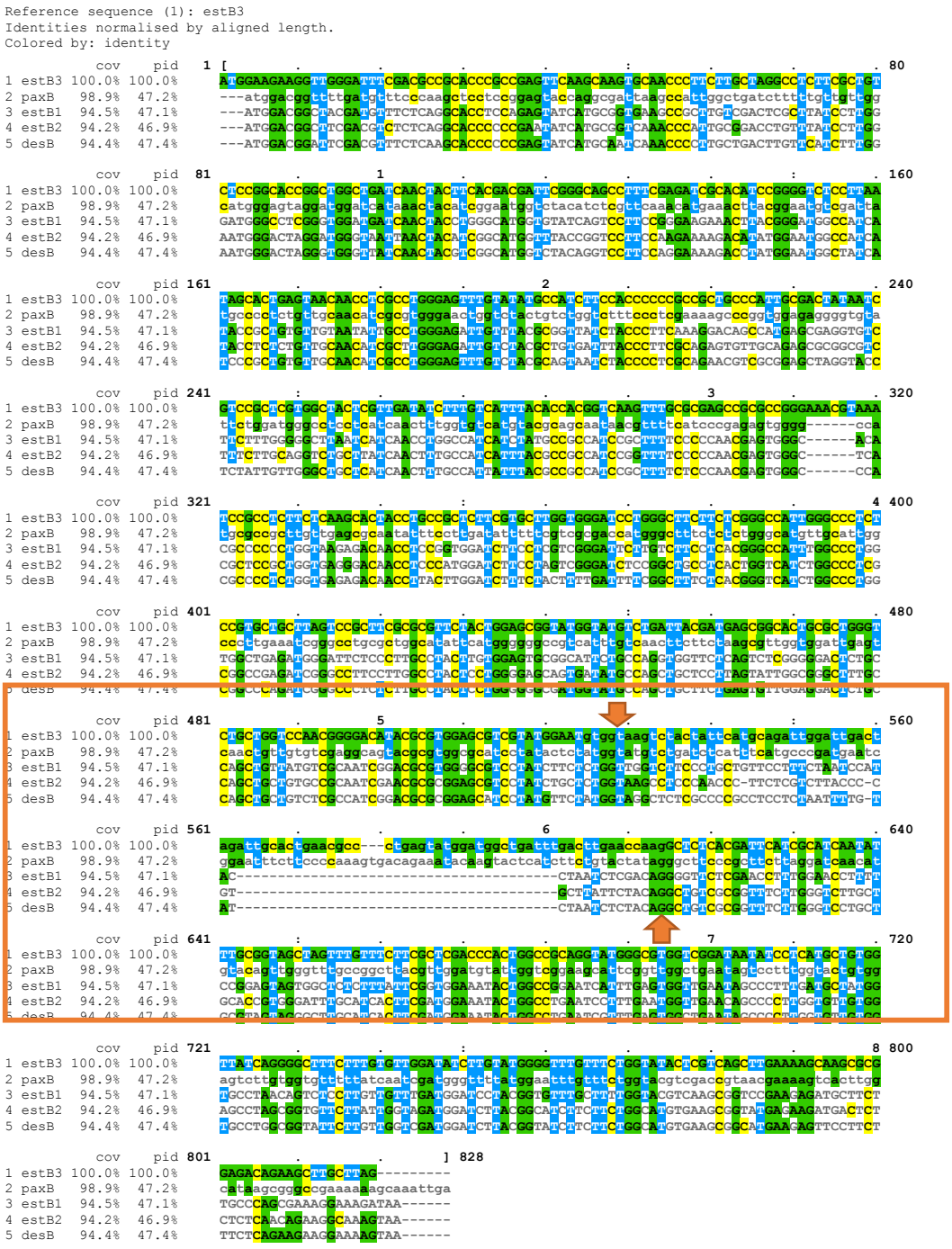


Figure S4. Nucleotide alignment for *idtBs*

Alignment of *paxB*, *estB1*, *estB2, estB3* and *desB*; demonstrating the high conservation of exon regions and low conservation of intron region. Intron indicated by orange box and arrows. Alignment generated with ClustalW[31] and visualized with MView.[32]

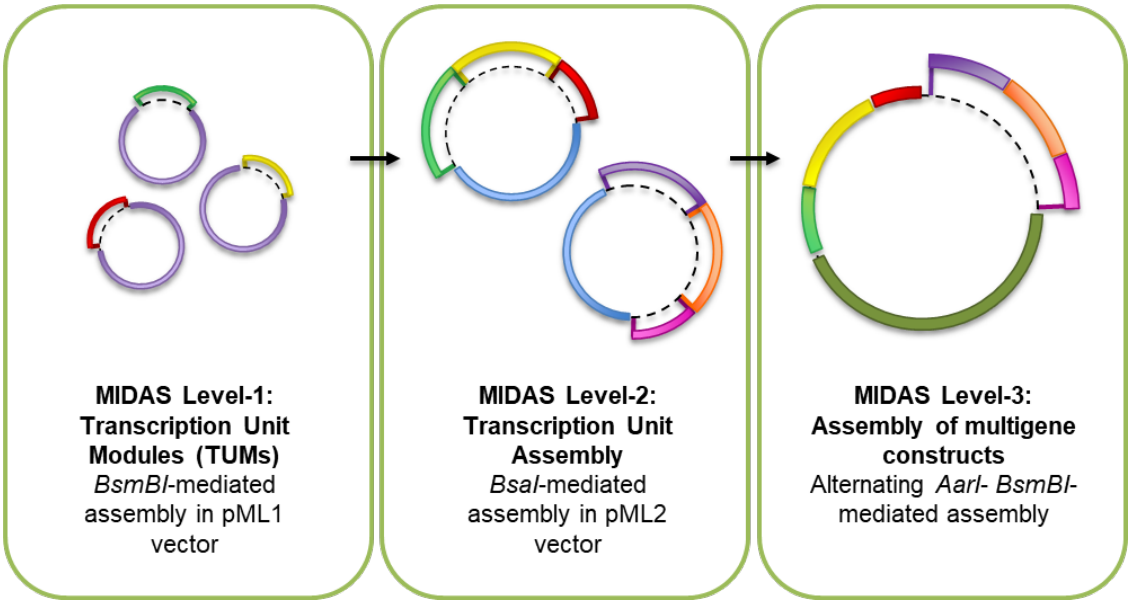


Figure S5. Overview of the MIDAS DNA assembly platform[1]

The MIDAS system is composed of three assembly levels. Level-1 is the production of libraries of DNA building blocks by cloning transcription unit (TU) modules (promoters, coding sequences, and terminators etc.) into a 'Level-1 source vector' through BsmBI-mediated reactions. Level-2 involves the assembly of TU modules in a plasmid, by combining the cloned and sequence verified Level-1 modules in a BsaI-mediated Golden-Gate reaction. Level-3 is the assembly of multigene constructs using the TU plasmids from Level-2, which are released from their plasmid and then, using alternating AarI- and BsmBI-mediated Golden Gate reactions, assembled into a Level-3 vector. This final assembly is the multigene construct that can be transformed into the desired expression host.

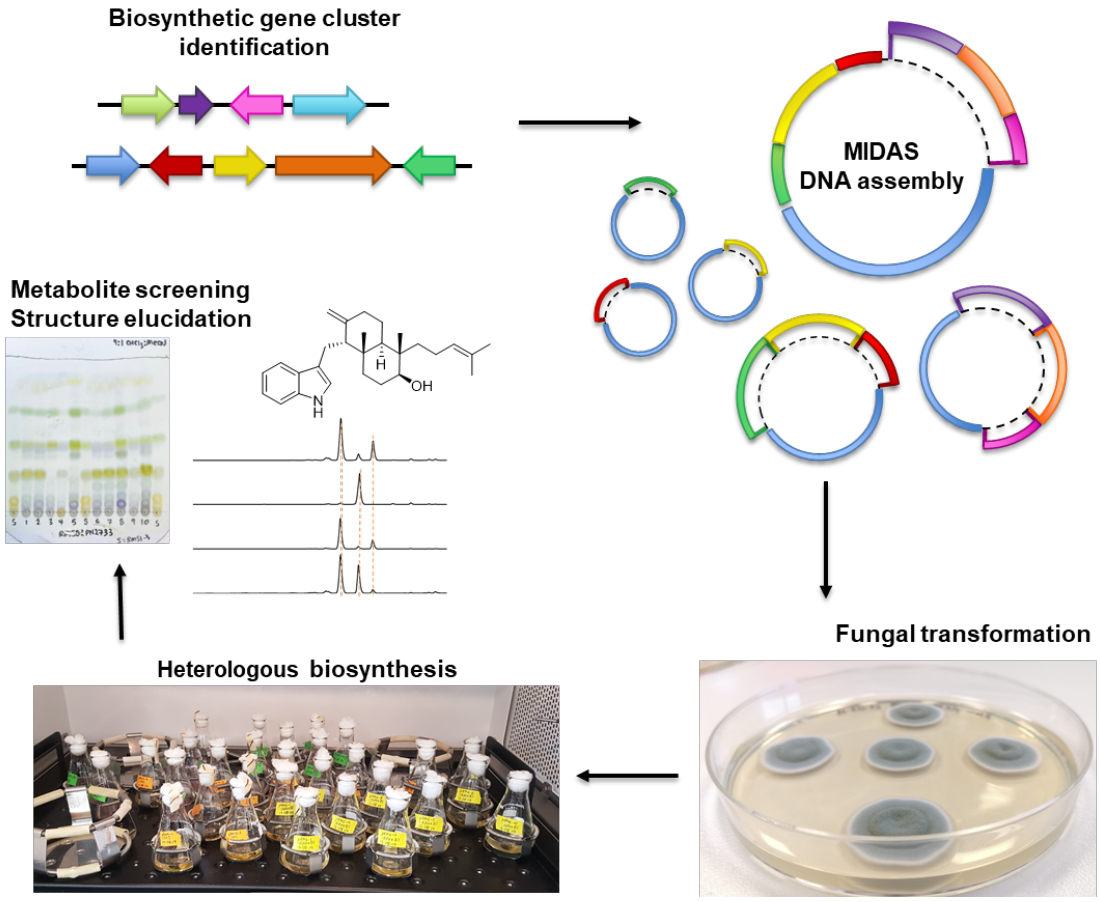


Figure S6. Overview of the heterologous expression system

This system begins identification of putative biosynthetic gene clusters in target organisms, using various bioinformatic tools and comparisons to known functional genes, to form hypotheses around putative gene function. The second is generating multi-gene plasmids using the MIDAS DNA assembly platform (**Figure S5**).[1] These plasmids are then transformed into *P. paxilli* protoplasts of appropriate strains. Transformant colonies (typically ten colonies) are then selected and grown in liquid culture for heterologous expression of indole diterpenes. Finally, extracts of fungal biomass for each transformant are screened for target secondary metabolites: first by TLC screening with Ehrlich's reagent, which is then followed up with LC-MS analysis. Strains producing compounds of interest are grown on a larger scale for isolation and NMR characterization. All steps are discussed in greater detail in the earlier experimental sections.

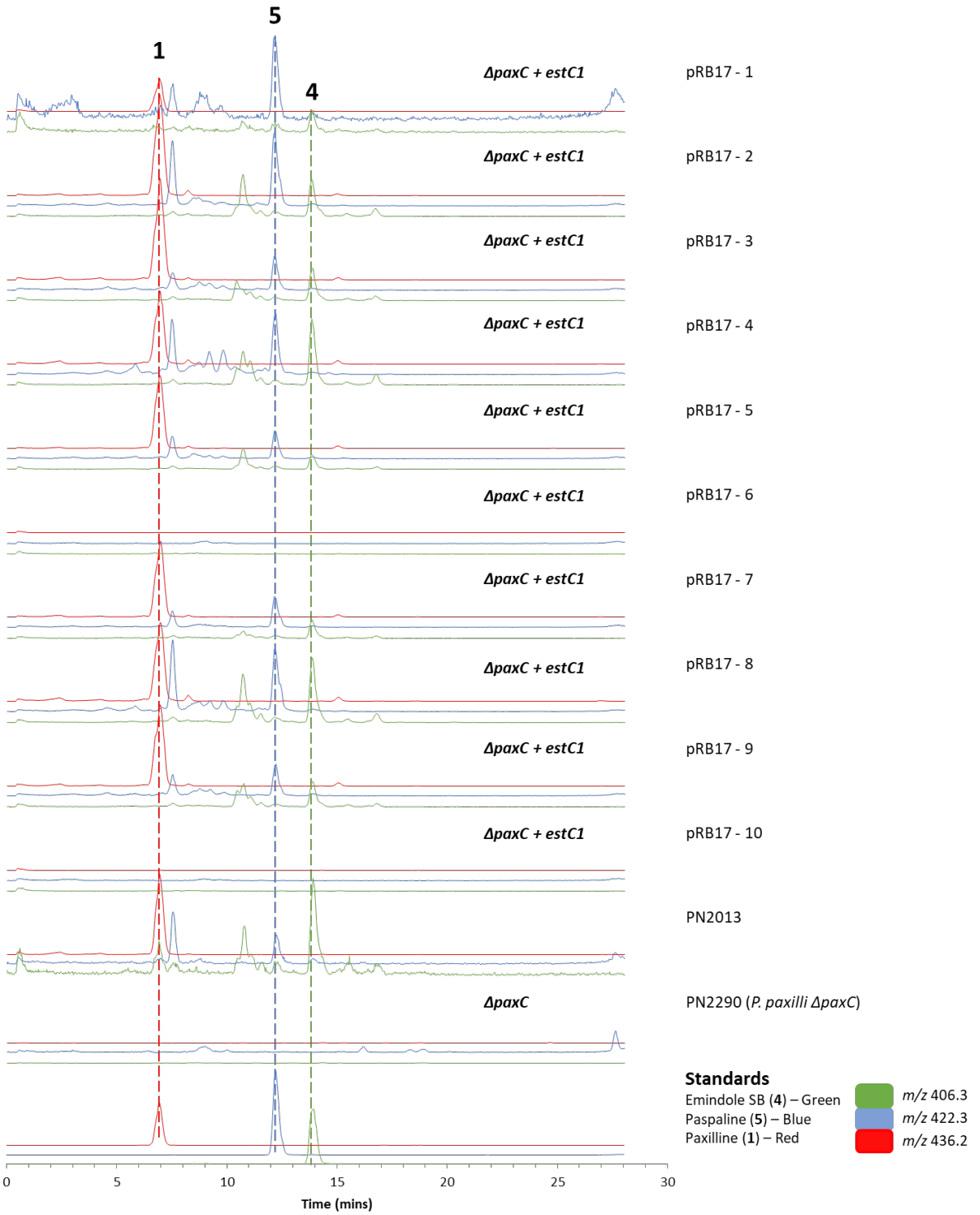


Figure S7. LC-MS analysis of ten *ΔpaxC P. paxilli* (PN2290) strains transformed with pRB17 (*estC1* expression construct).

Extracted ion chromatograms shown for the $[M + H]^+$ *m/z* of paxilline (**1**) (436.2, red), paspaline (**5**) (422.3, blue), and emindole SB (**4**) (406.3, green). The Y-axis represents counts with scales arbitrarily chosen to ensure that the key peaks are visible. PN2013 represents wild-type *P. paxilli*.

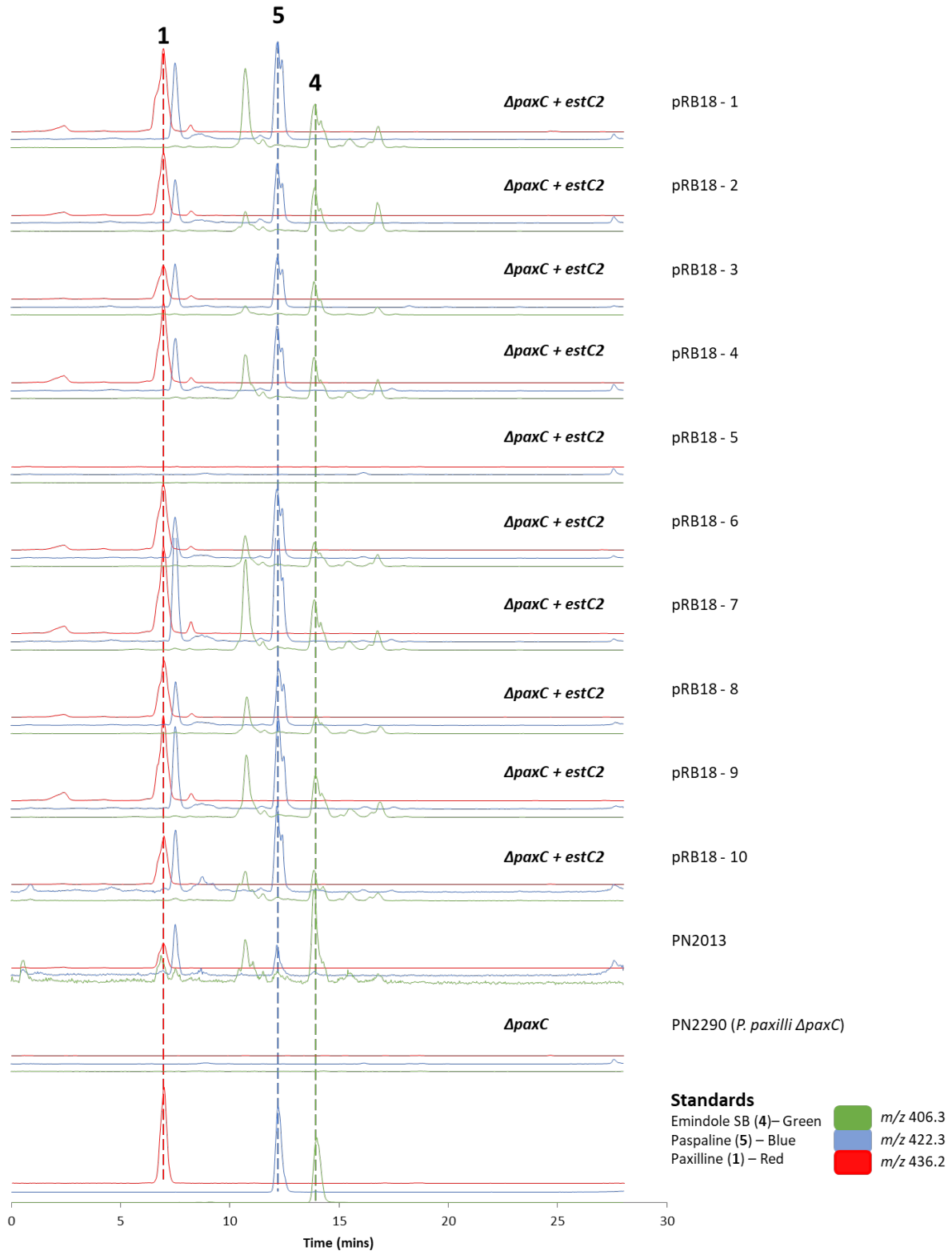


Figure S8. LC-MS analysis of ten *ΔpaxC P. paxilli* (PN2290) strains transformed with pRB18 (*estC2* expression construct)

Extracted ion chromatograms shown for the $[M + H]^+$ *m/z* of paxilline (**1**) (436.2, red), paspaline (**5**) (422.3, blue), and emindole SB (**4**) (406.3, green). The Y-axis represents counts with scales arbitrarily chosen to ensure that the key peaks are visible. PN2013 represents wild-type *P. paxilli*.

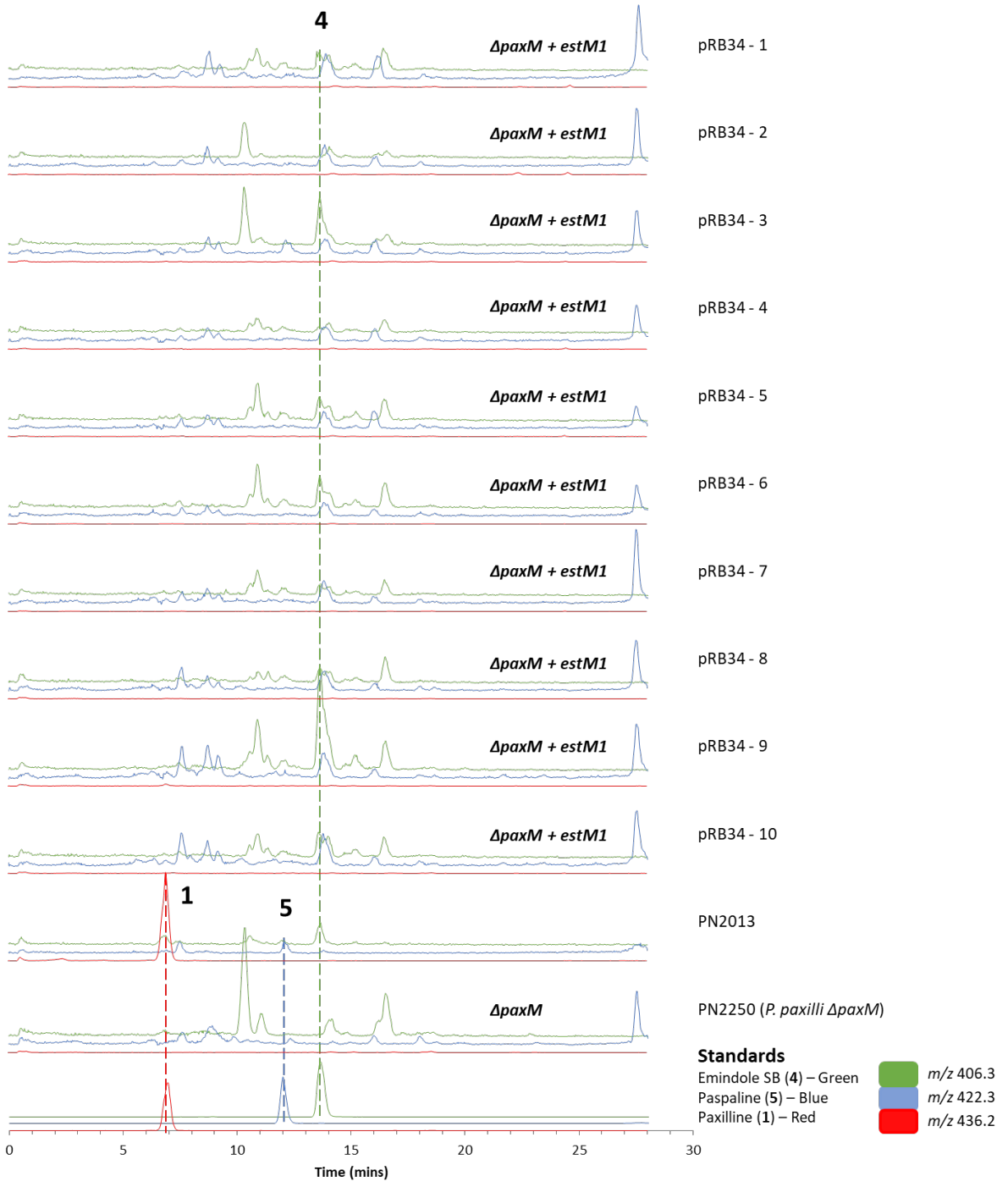


Figure S9. LC-MS analysis of ten *ΔpaxM P. paxilli* (PN2250) strains transformed with pRB34 (*estM1* expression construct).

Extracted ion chromatograms shown for the $[M + H]^+$ *m/z* of paxilline (**1**) (436.2, red), paspaline (**5**) (422.3, blue), and emindole SB (**4**) (406.3, green). The Y-axis represents counts with scales arbitrarily chosen to ensure that the key peaks are visible. PN2013 represents wild-type *P. paxilli*.

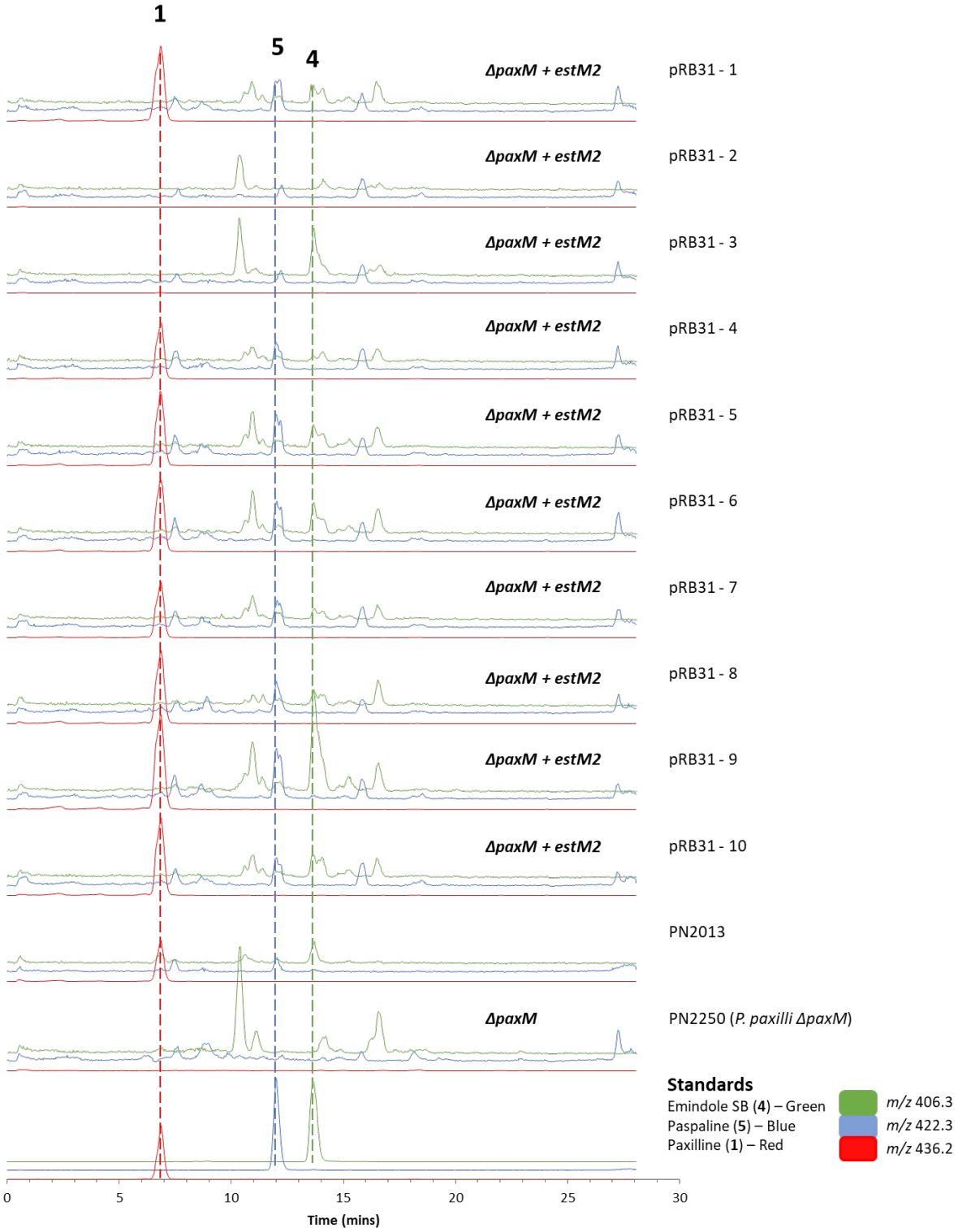


Figure S10. LC-MS analysis of ten *ΔpaxM P.paxilli* (PN2250) strains transformed with pRB31 (*estM2* expression construct). Extracted ion chromatograms shown for the $[M + H]^+$ *m/z* of paxilline (**1**) (436.2, red), paspaline (**5**) (422.3, blue), and emindole SB (**4**) (406.3, green). The Y-axis represents counts with scales arbitrarily chosen to ensure that the key peaks are visible. PN2013 represents wild-type *P. paxilli*.

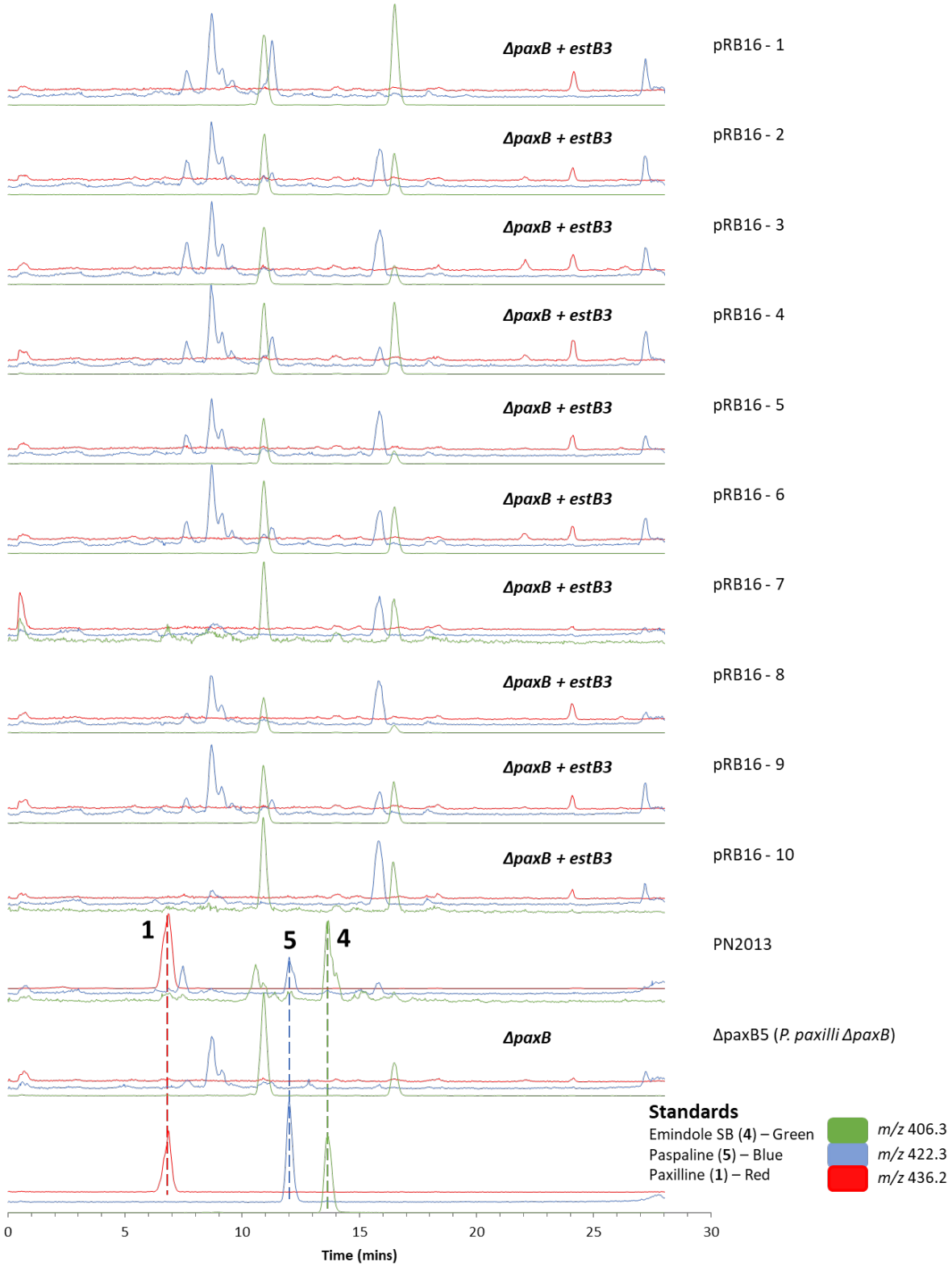


Figure S11. LC-MS analysis of ten *ΔpaxB P. paxilli* (*ΔpaxB5*) strains transformed with pRB16 (*estB3* expression construct) Extracted ion chromatograms shown for the $[M + H]^+$ *m/z* of paxilline (**1**) (436.2, red), paspaline (**5**) (422.3, blue), and emindole SB (**4**) (406.3, green). The Y-axis represents counts with scales arbitrarily chosen to ensure that the key peaks are visible. PN2013 represents wild-type *P. paxilli*.

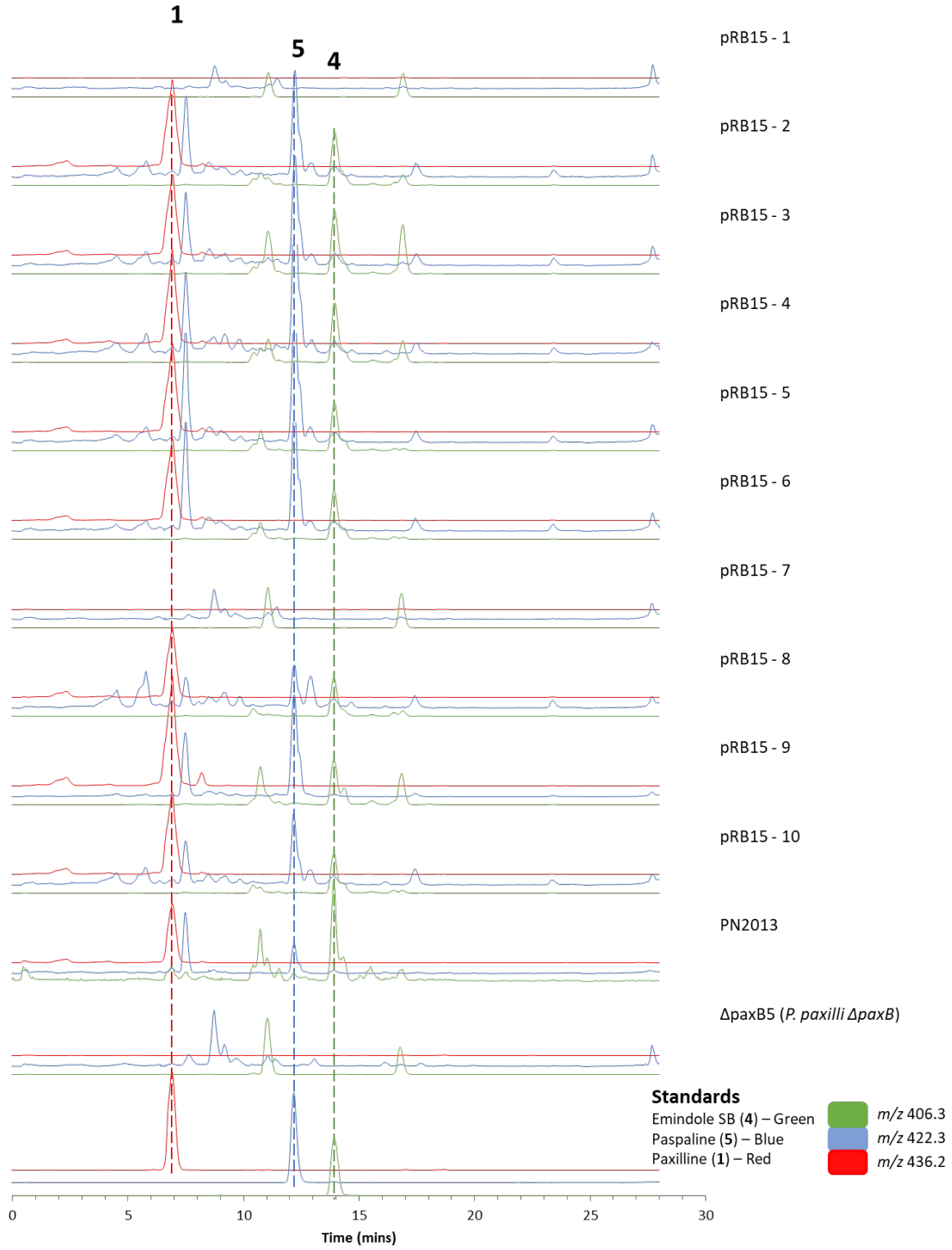


Figure S12. LC-MS analysis of ten *ΔpaxB P. paxilli* (*ΔpaxB5*) strains transformed with pRB15 (*estB2* expression construct).

Extracted ion chromatograms shown for the $[M + H]^+$ *m/z* of paxilline (**1**) (436.2, red), paspaline (**5**) (422.3, blue), and emindole SB (**4**) (406.3, green). The Y-axis represents counts with scales arbitrarily chosen to ensure that the key peaks are visible. PN2013 represents wild-type *P. paxilli*.

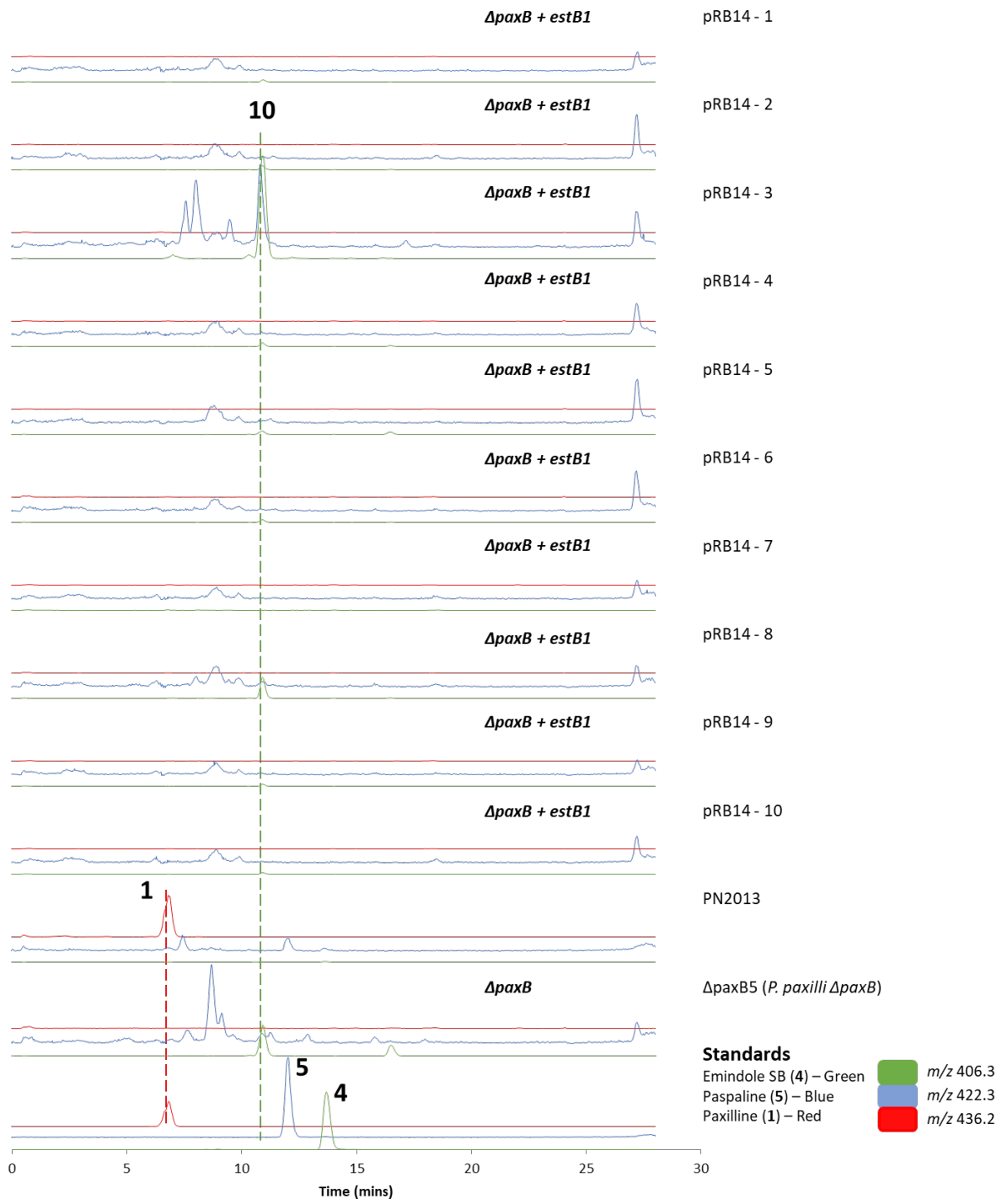


Figure S13. LC-MS analysis of ten *ΔpaxB P. paxilli* (*ΔpaxB5*) strains transformed with pRB14 (*estB1* expression construct). Extracted ion chromatograms shown for the $[M + H]^+$ *m/z* of paxilline (**1**) (436.2, red), paspaline (**5**) (422.3, blue), and emindole SB (**4**) (406.3, green). The Y-axis represents counts with scales arbitrarily chosen to ensure that the key peaks are visible except for the green (406.3) traces which are scaled to the trace shown for ΔpaxB5. PN2013 represents wild-type *P. paxilli*.

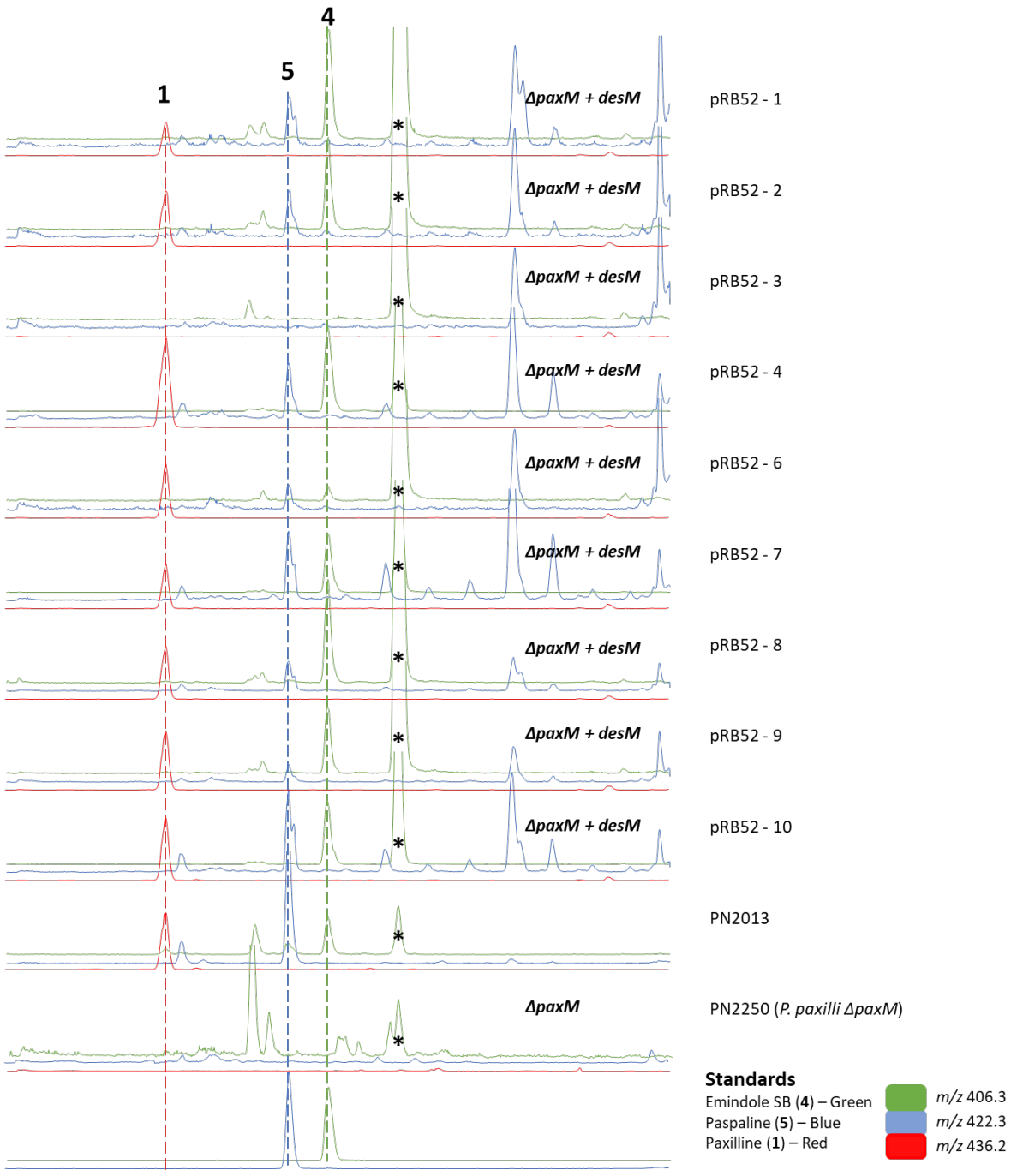


Figure S14. LC-MS analysis of ten *ΔpaxM P. paxilli* (PN2250) strains transformed with pRB52 (*desM* expression construct).

Extracted ion chromatograms shown for the $[M + H]^+$ *m/z* of paxilline (**1**) (436.2, red), paspaline (**5**) (422.3, blue), and emindole SB (**4**) (406.3, green). The Y-axis represents counts with scales arbitrarily chosen to ensure that the key peaks are visible while peaks marked with an asterisk (*) represent a known contamination associated with a brand of extraction tubes. PN2013 represents wild-type *P. paxilli*.

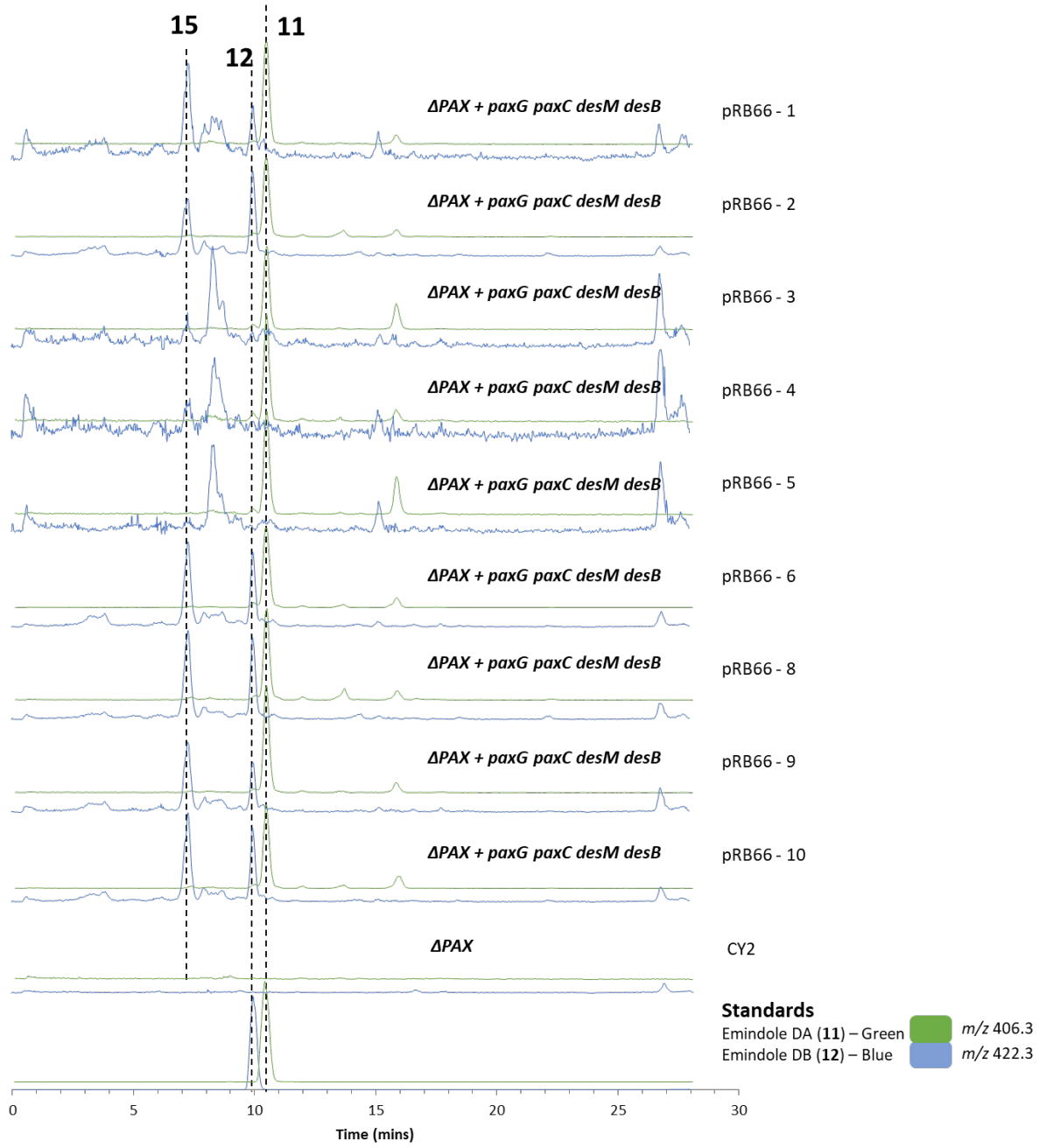


Figure S15. LC-MS analysis of ten CY2 *P. paxilli* strains transformed with pRB66 (expression construct for the IDT biosynthetic genes *paxG paxC desM* and *desB*)

Extracted ion chromatograms shown for *m/z* 422.3 (blue), and 406.3 (green). The Y-axis represents counts with scales arbitrarily chosen to ensure that the key peaks are visible.

Gene expression analysis of *estB3* using RT-PCR

For the RT-PCR an amplicon spanning the predicted intron of *estB3* was designed using the primers estB3_qPCR_F and estB3_qPCR_R. Extraction of RNA (including removal of contaminating gDNA) was carried out for each of the ten ΔpaxB5:pRB16 transformants and reverse transcribed to generate the cDNA pool. For each cDNA sample, a –RT control reaction (lacking the reverse transcriptase enzyme) was set up. These control reactions were designed to detect the presence of gDNA contamination within the cDNA samples. The *estB3* gene and a reference gene (*P. paxilli sac7*) were amplified from cDNA and –RT control samples. Amplification of *sac7* was used as an indication of the integrity of the cDNA.

The detection of bands of the expected size for the amplification of *sac7* indicated that the cDNA pool was of sufficient quality to observe amplification of a general amplicon. The amplification of *estB3* in two bands, differing in size, suggested gDNA contamination. However, no bands were observed in the –RT control which excluded this possibility.

A possible explanation of the multiple bands is that *P. paxilli* is unable to completely splice the intron of a foreign gene and results in transcripts containing the intron or that this "inefficient" processing of the *estB3* transcript is the default state in the native and the heterologous host. Nonetheless, the amplification of *estB3* transcripts from cDNA obtained from the ten pRB16:ΔpaxB5 transformants suggests that low/absent expression of *estB3* is not sufficient to explain the absence of any new IDTs in these transformants.

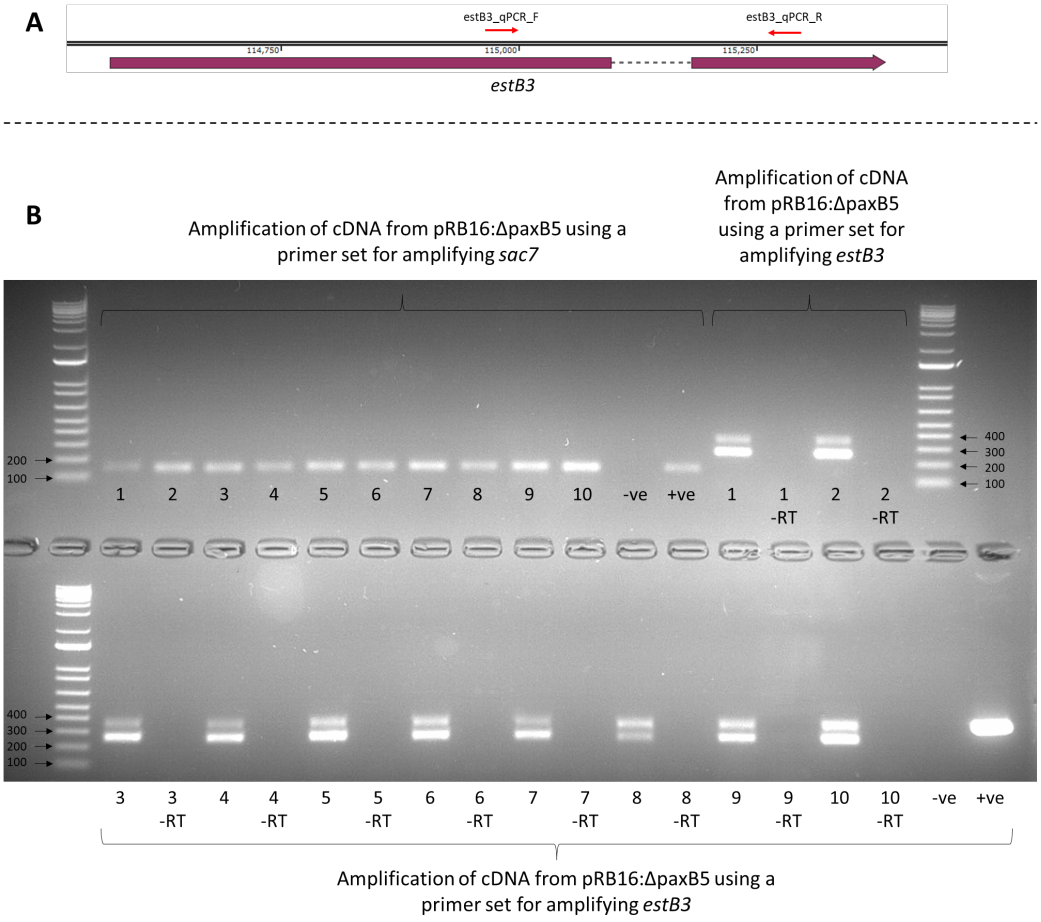


Figure S16. Analysis of the *Penicillium paxilli* pRB16:*ΔpaxB5* transformations investigating the presence of *estB3* transcripts. (**A**) Location of the primers flanking the putative intron of *estB3*. (**B**) Agarose gel (2% (w/v)) electrophoresis of the amplicons for the reference gene, *sac7*, as well as the target gene *estB3*. The positive control (marked as +ve on the gel) for *sac7* was cDNA generated from the *P. paxilli* wild-type strain and the positive control for *estB3* was the pRB16 plasmid. The negative control (marked as -ve on the gel) is water. Lanes marked as –RT were amplified from the controls lacking the reverse-transcriptase enzyme.

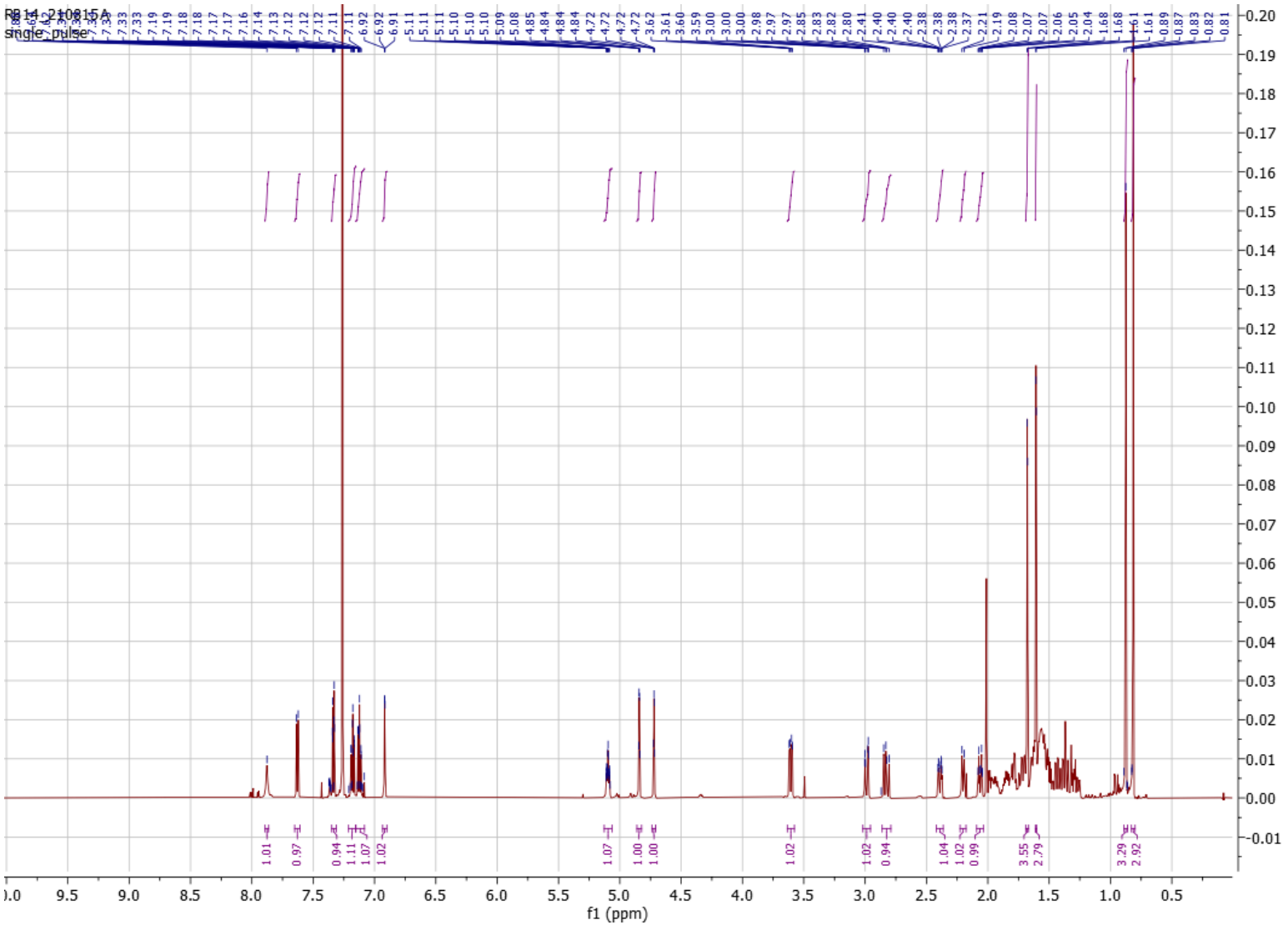


Figure S17. ^{1}H NMR spectrum of emindole SA (**10**) in $CDCl_3$ (600 MHz)

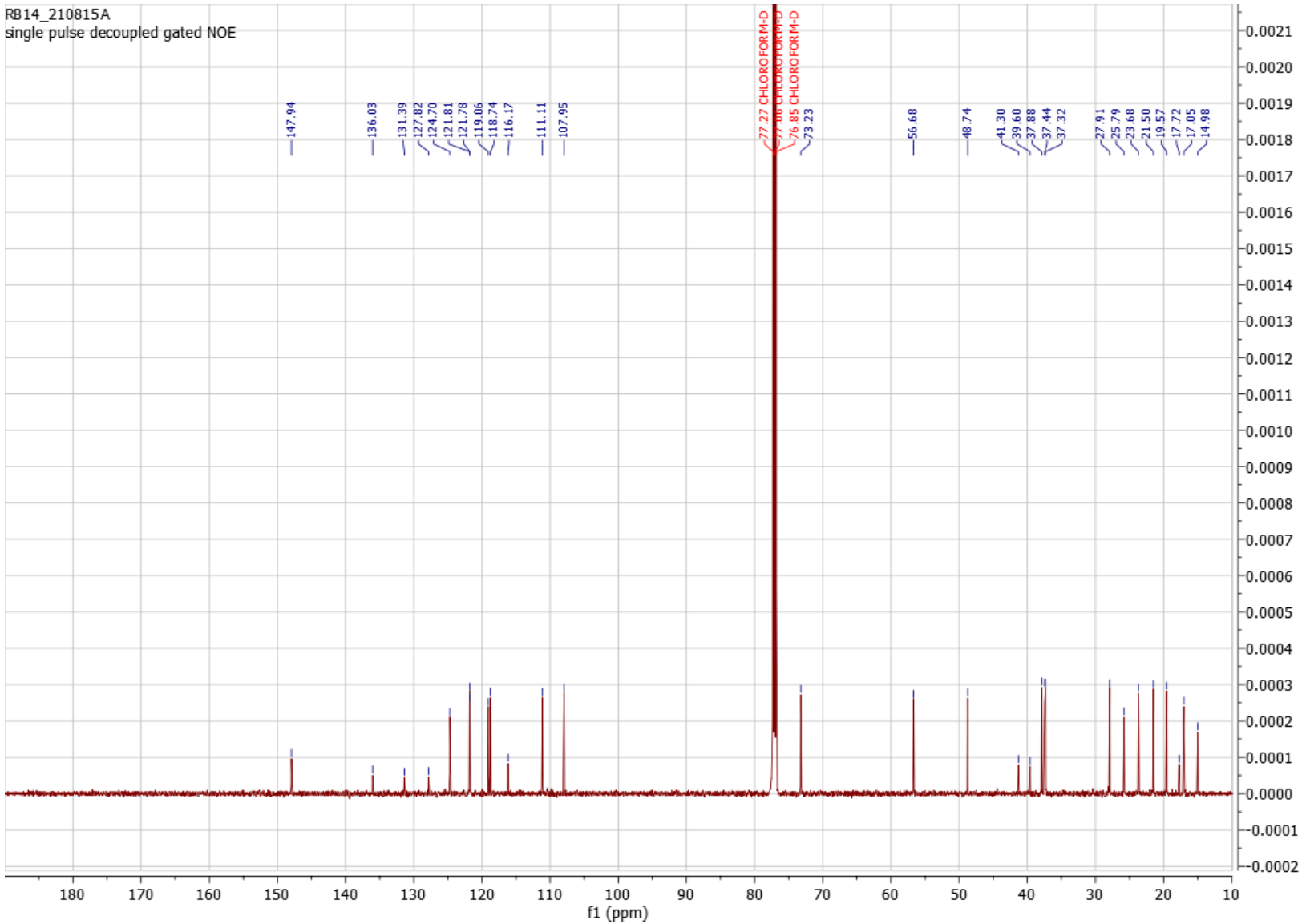


Figure S18 ^{13}C NMR spectrum of emindole SA (**10**) in $CDCl_3$ (150 MHz)

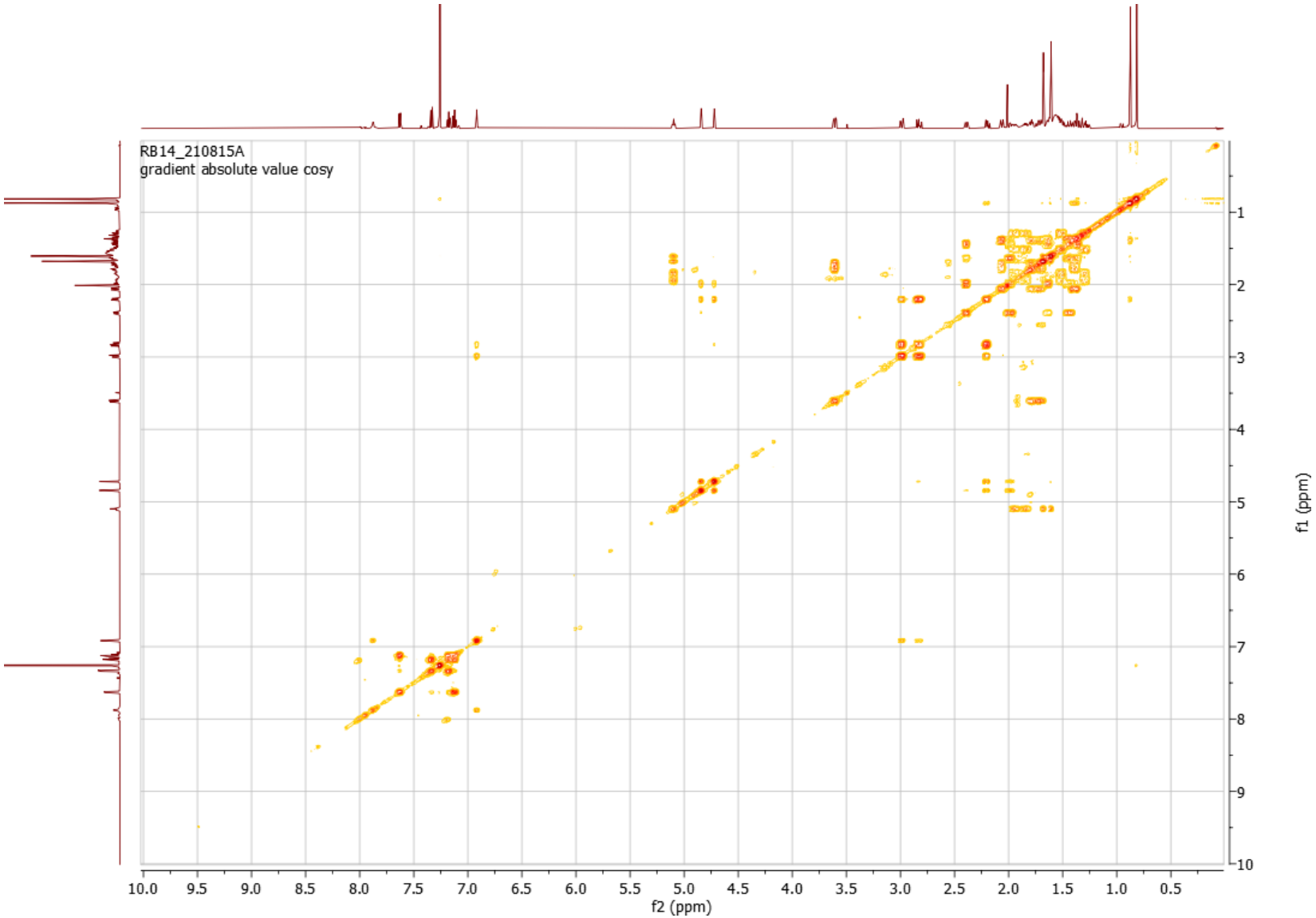


Figure S19. ^{1}H-^{1}H COSY NMR spectrum of emindole SA (**10**) in $CDCl_3$ (600 MHz)

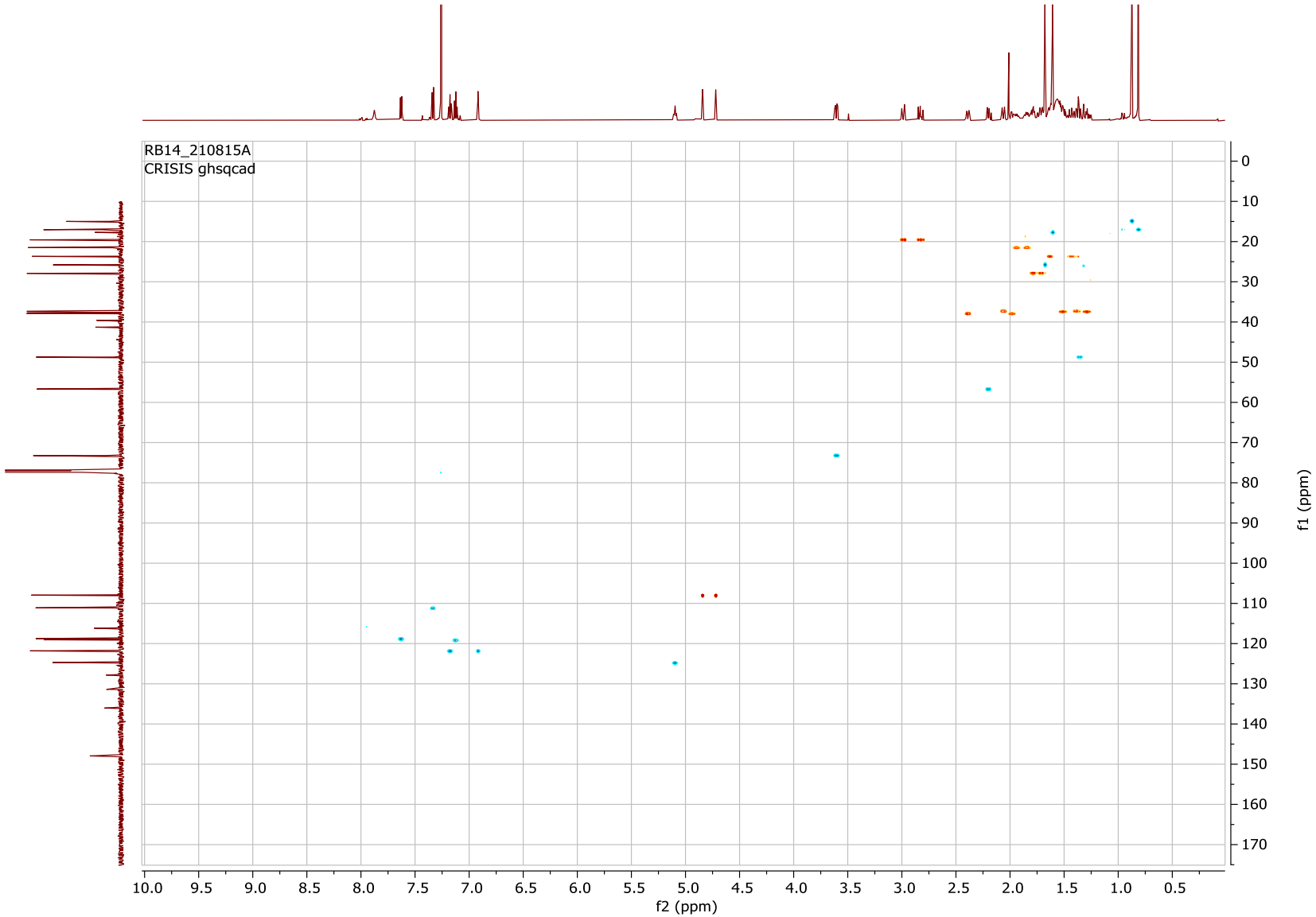


Figure S20. ^{1}H-^{13}C HSQC NMR spectrum of emindole SA (**10**) in $CDCl_3$ (600 MHz)

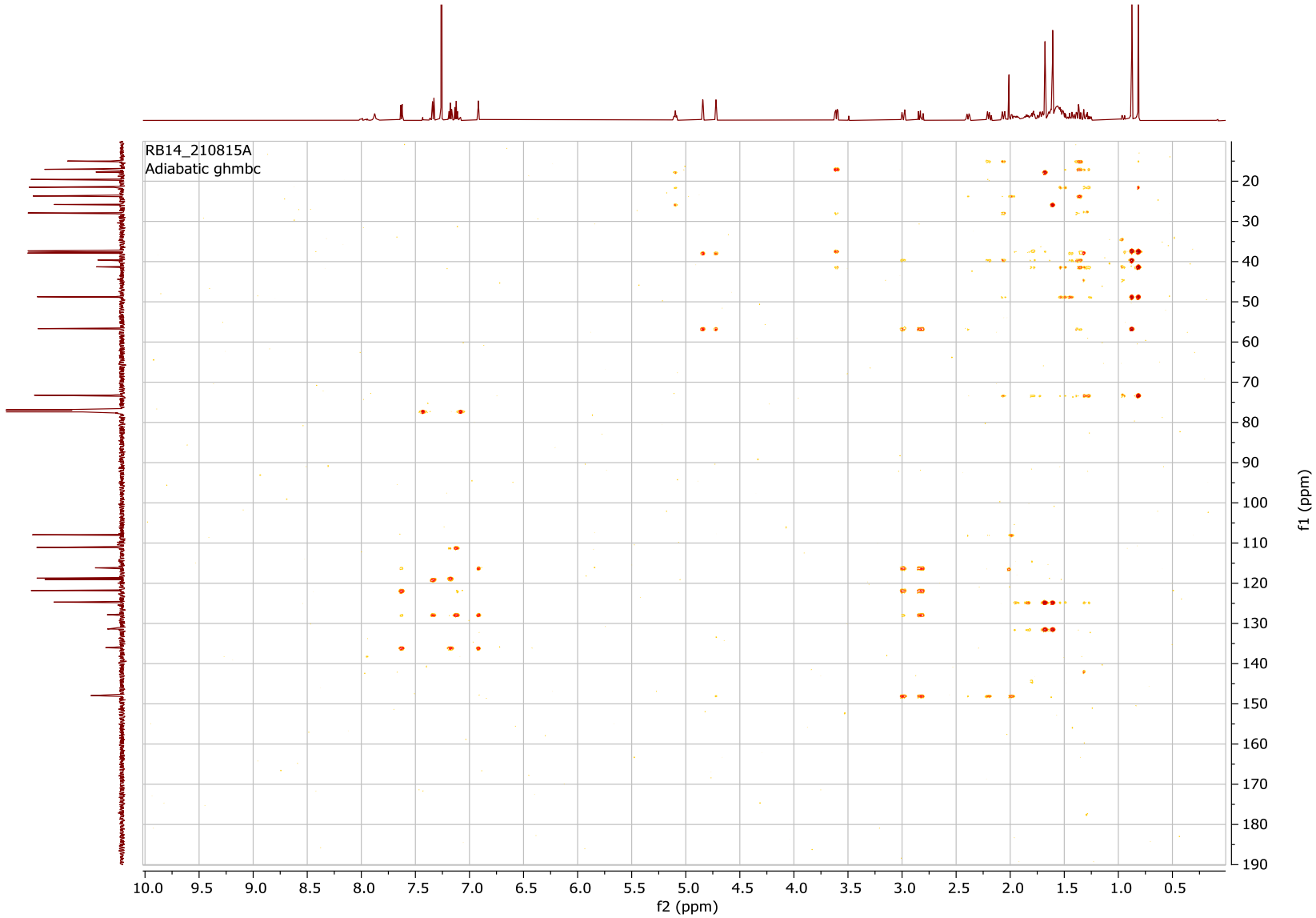


Figure S21. ^{1}H-^{13}C HMBC NMR spectrum of emindole SA (**10**) in $CDCl_3$ (600 MHz)

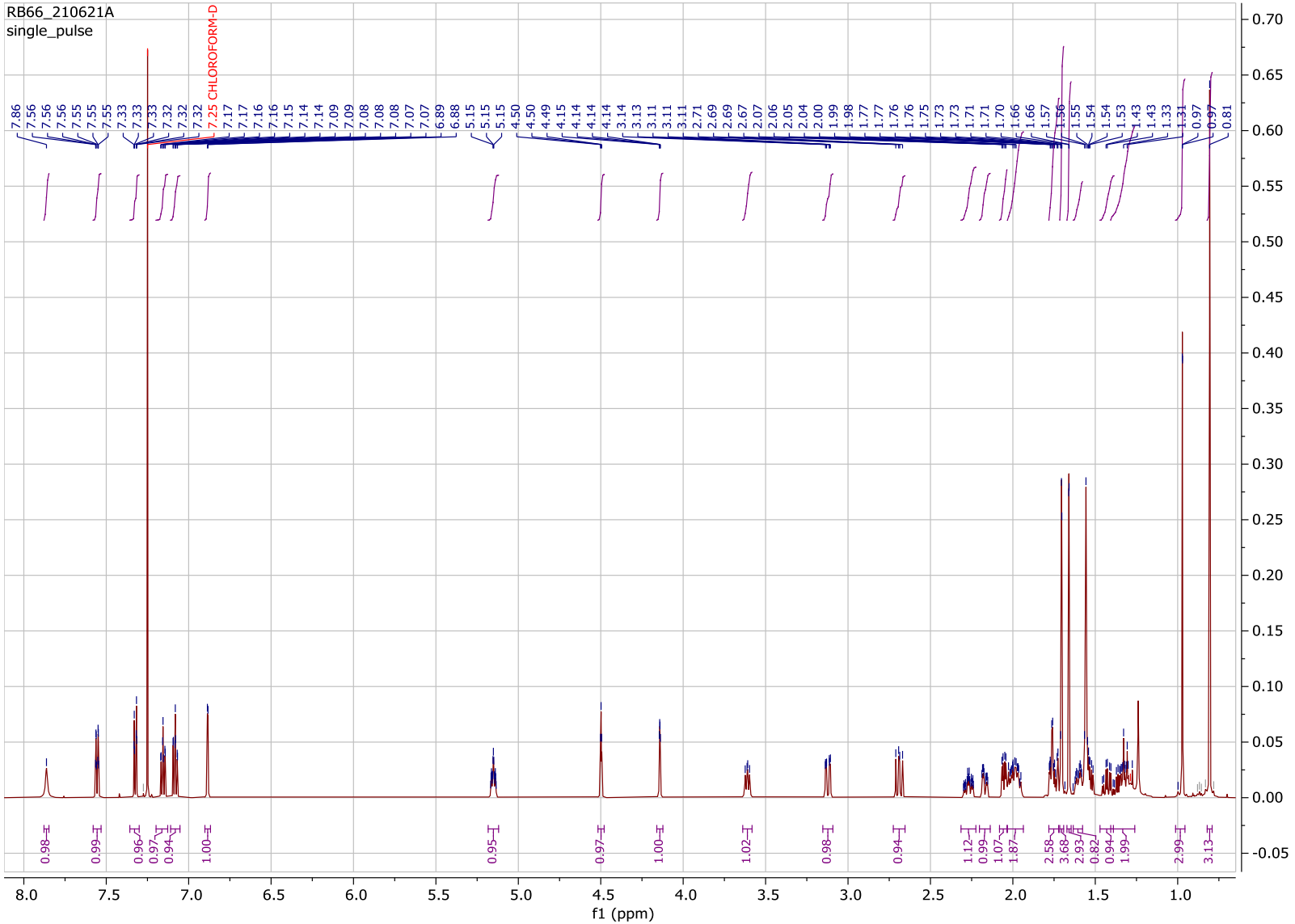


Figure S22. ^{1}H NMR spectrum of emindole DA (**11**) in $CDCl_3$ (600 MHz)

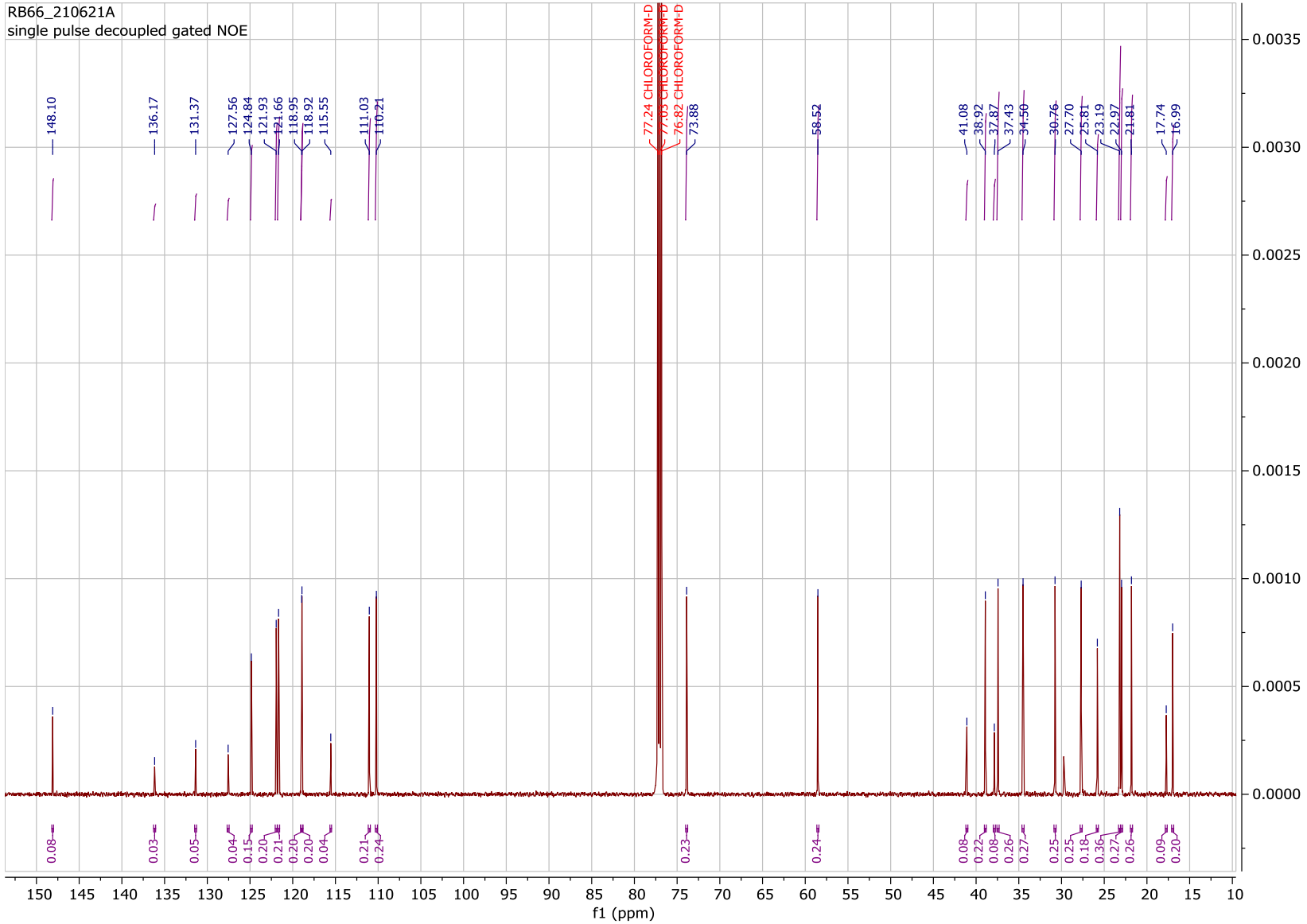


Figure S23. ^{13}C NMR spectrum of emindole DA (**11**) in $CDCl_3$ (150 MHz)

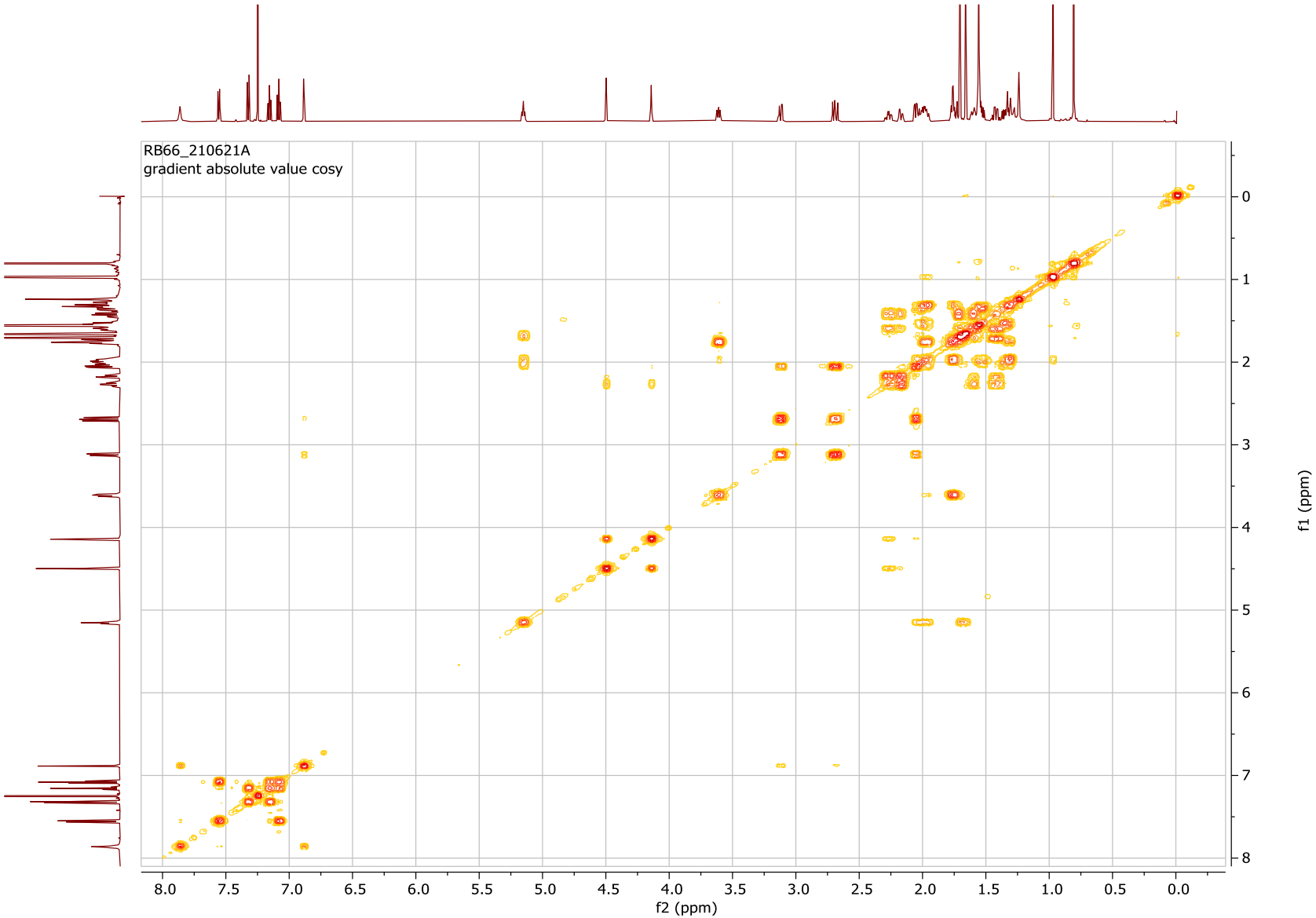


Figure S24. ^{1}H-^{1}H COSY NMR spectrum of emindole DA (**11**) in $CDCl_3$ (600 MHz)

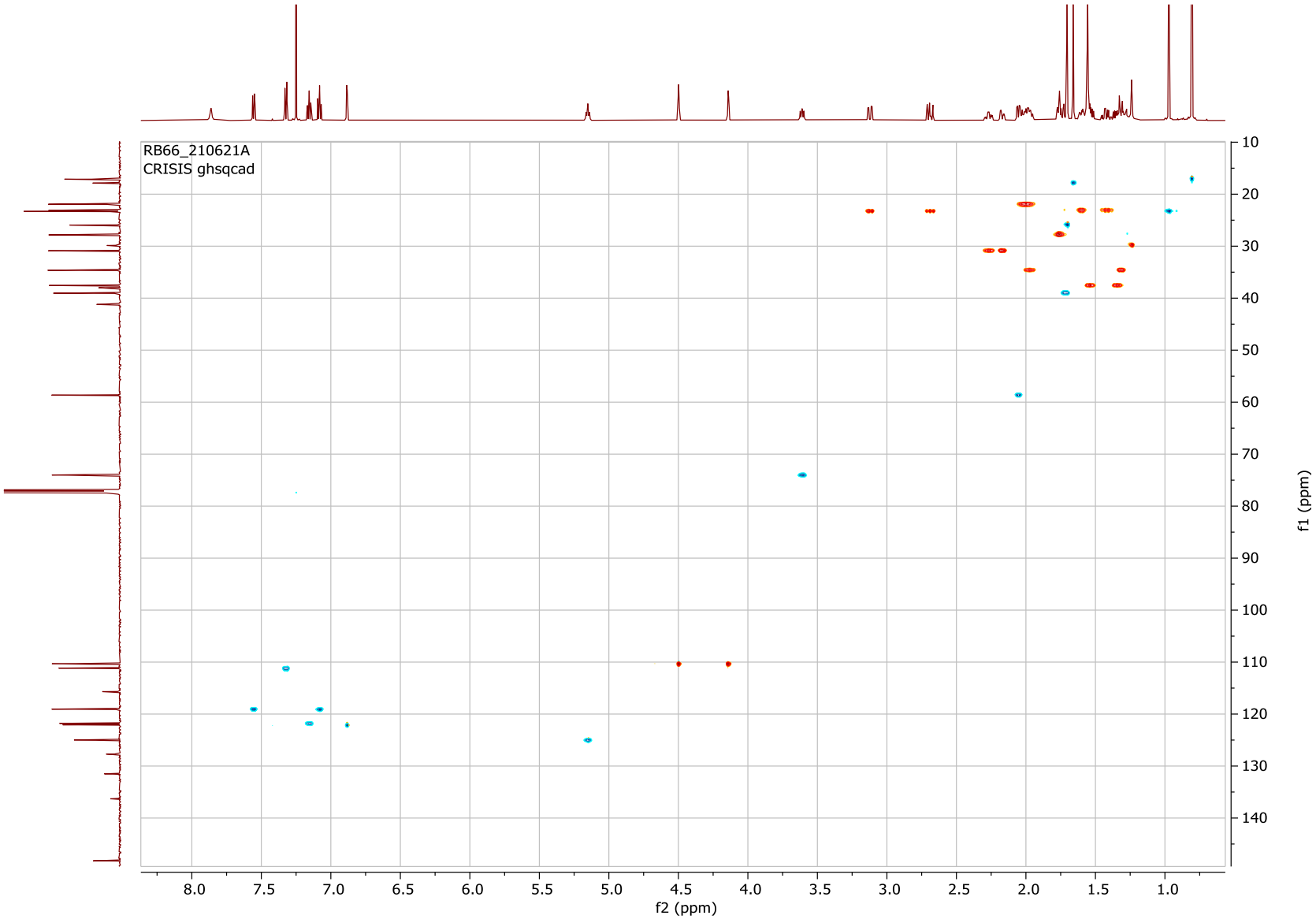


Figure S25. ^{1}H-^{13}C HSQC spectrum of emindole DA (**11**) in $CDCl_3$ (600 MHz)

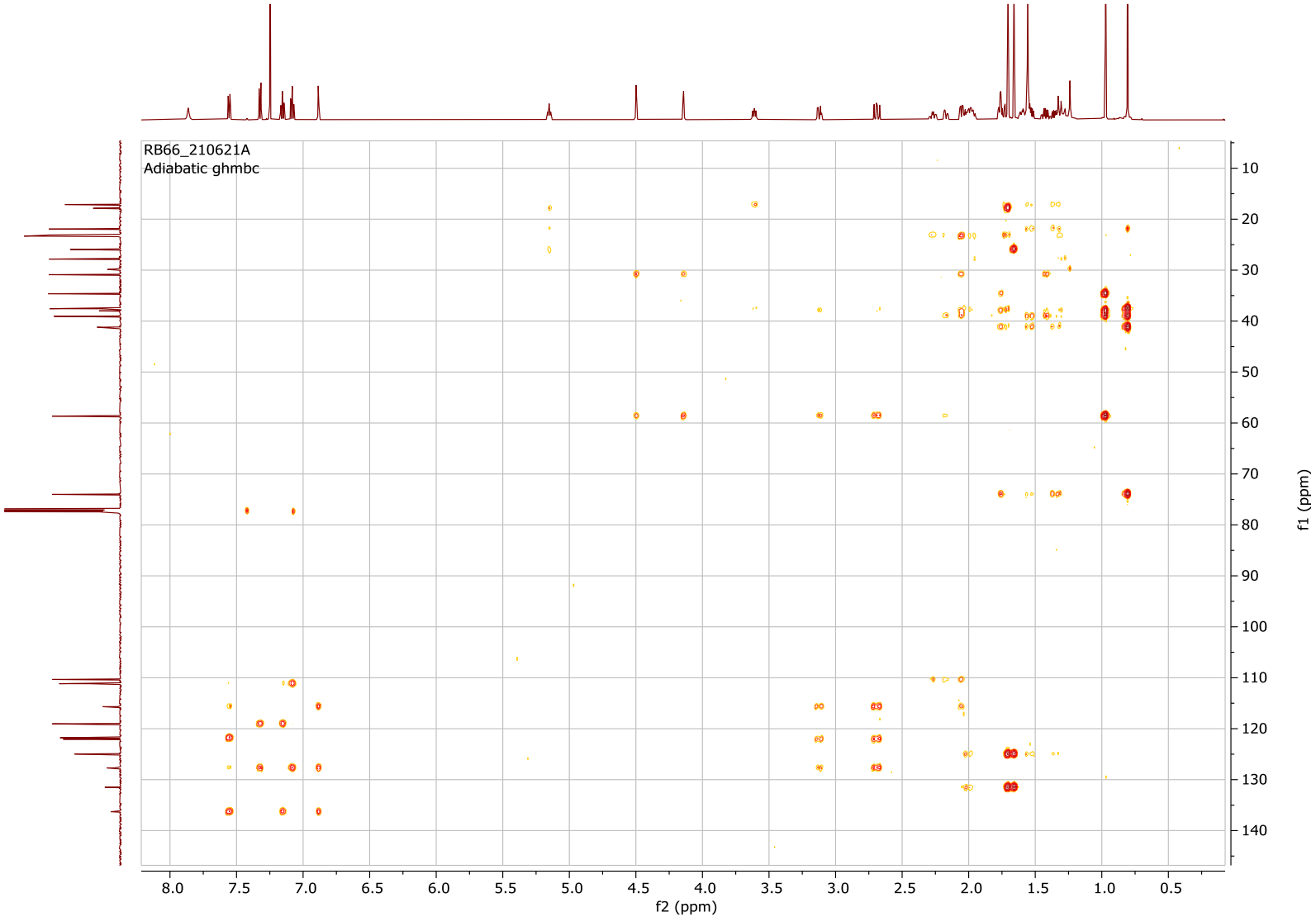


Figure S26. ^{1}H-^{13}C HMBC NMR spectrum of emindole DA (**11**) in $CDCl_3$ (600 MHz)

Figure S27. ^{1}H NMR spectrum of emindole DB (**12**) in $CDCl_3$ (600 MHz)

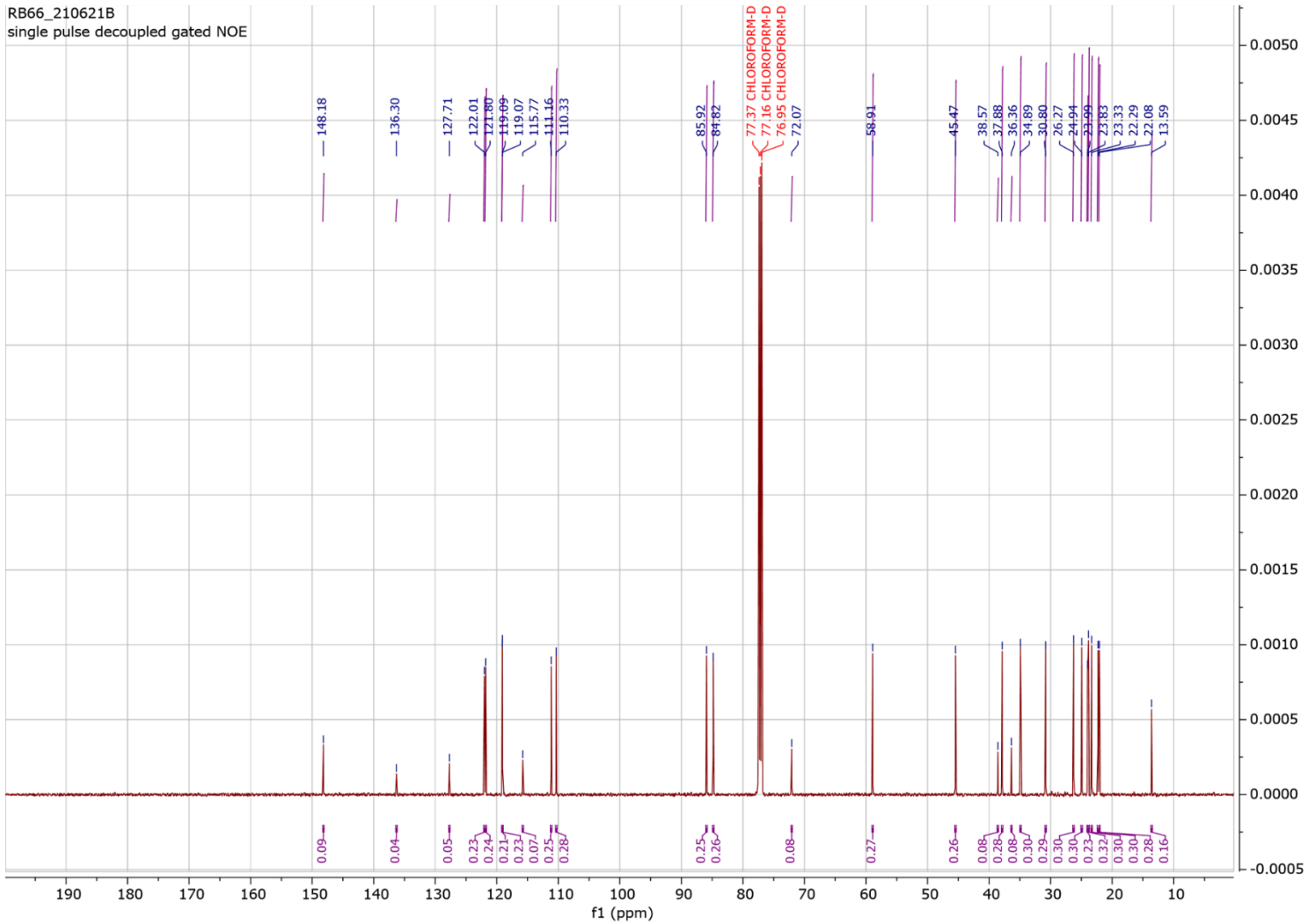


Figure S28. ^{13}C NMR spectrum of emindole DB (**12**) in $CDCl_3$ (150 MHz)

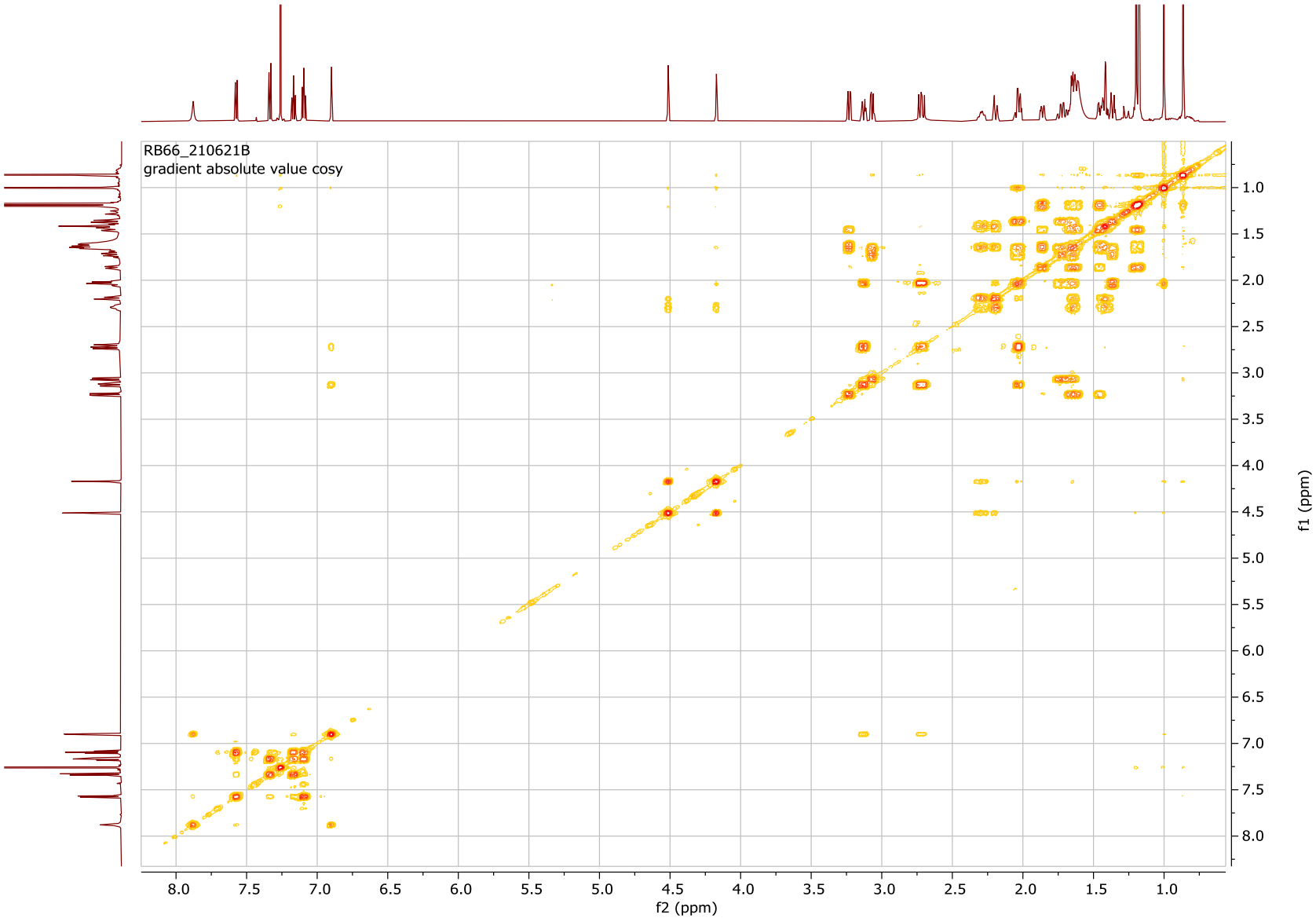


Figure S29. ^{1}H-^{1}H COSY NMR spectrum of emindole DB (**12**) in $CDCl_3$ (600 MHz)

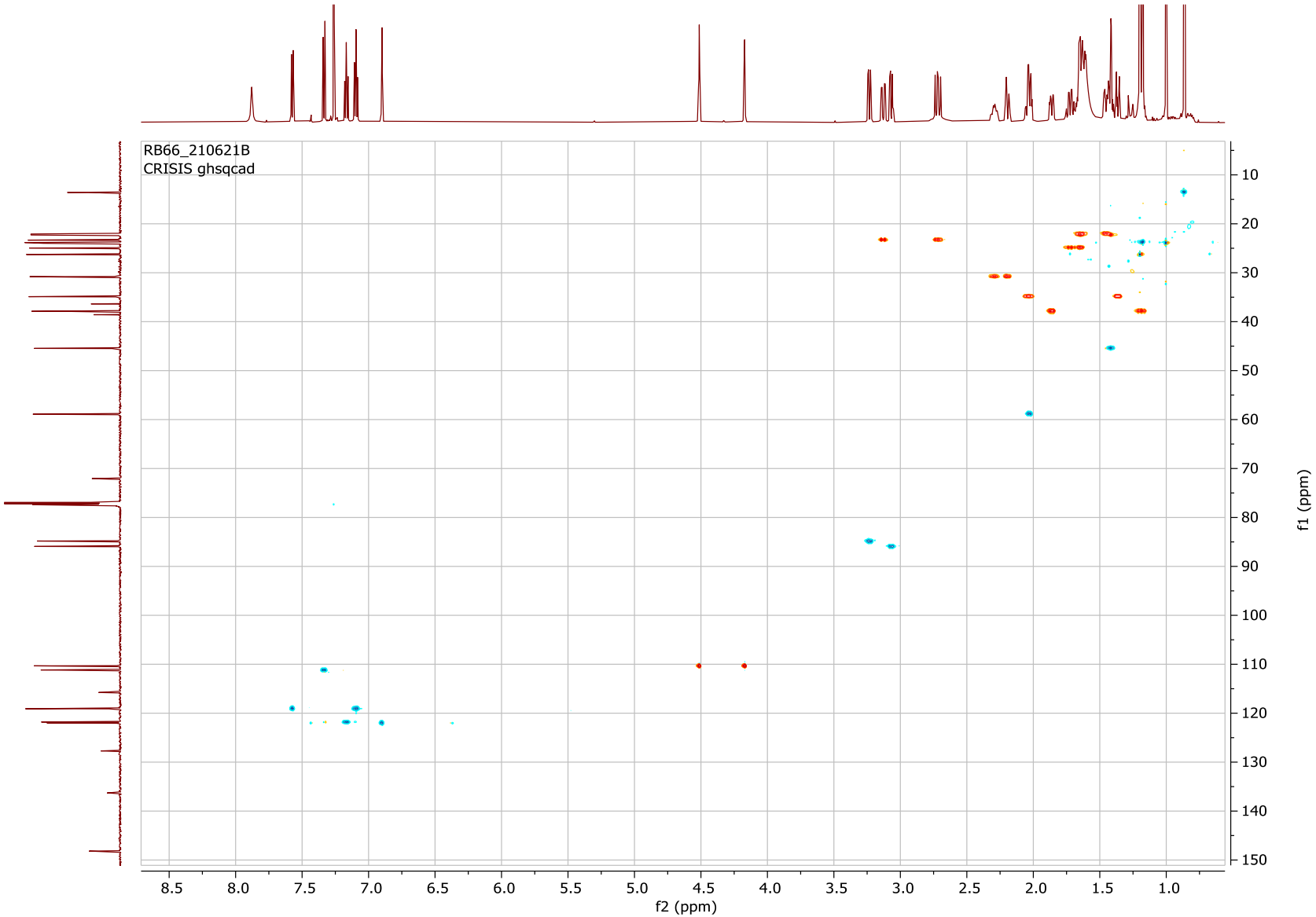


Figure S30. ^{1}H-^{13}C HSQC NMR spectrum of emindole DB (**12**) in $CDCl_3$ (600 MHz)

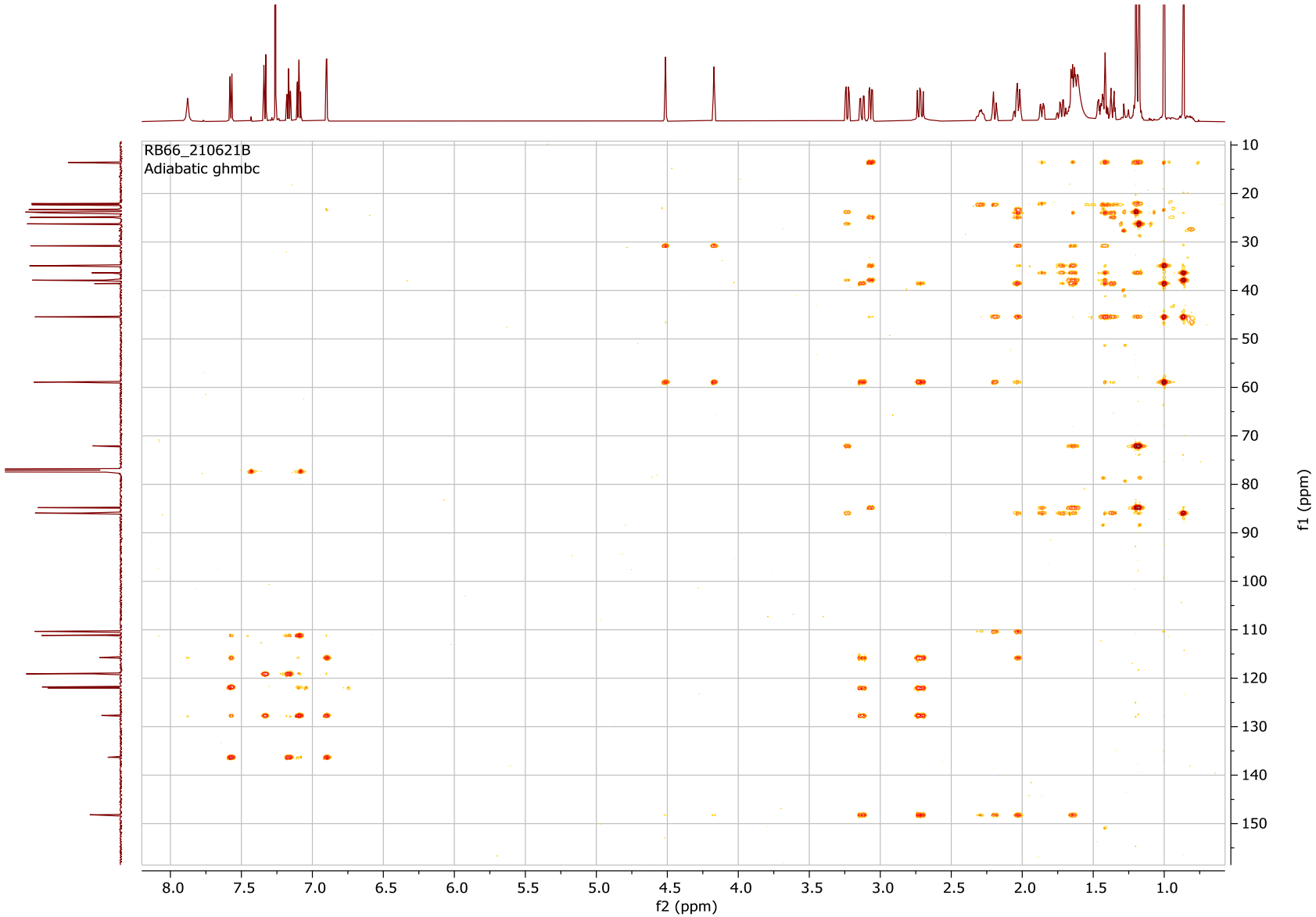


Figure S31. ^{1}H-^{13}C HMBC NMR spectrum of emindole DB (**12**) in $CDCl_3$ (600 MHz)

Phylogenetic analysis

Tang *et al.* (2015)[30] identified the first alternative IDT cyclases (*afB*, *atS5B1* and *atS2B*) through phylogenetic clustering based on sequence variation. In this study, we have identified and characterized several new IDT cyclases with new functions and we wanted to determine the phylogenetic relationship of these novel cyclases with previously characterized IDT cyclases to better understand the relationship between sequence and function. Since the representation of sequences corresponding to alternate cyclases is low, we have also included sequences corresponding to fungal meroterpenoid cyclases (delivering products such as anditomin,[33] terretonin,[34] and austinol[35] among others) to expand the pool of catalytically diverse terpene cyclase sequences of this type. These meroterpenoid cyclases catalyze similar chemistry on the terpenoid portion of the molecule as the IDT cyclases, but the terpenoid portion of the meroterpenoids are fused to a different functionality (typically the products of PKSs) instead of the indole ring in the IDTs. In our analysis we used Bayesian inference-based phylogeny[36-37] to estimate the tree of curated cyclase amino acid sequences. The Markov chain Monte Carlo sampling, used to generate the posterior probability of the hypothesis, was terminated as the model approached convergence (potential scale reduction factor approximating 1, and average standard deviation of split frequencies falling below 0.005) suggesting good support for the distribution of topology.

We observed that the IDT and meroterpenoid cyclases could be separated into 3 clades (**Figure S32**). The first clade contained a group of meroterpenoid cyclases. The second clade also contained a group of meroterpenoid cyclases as well as the majority of IDT cyclases. These clade 2 meroterpenoid cyclases catalyze the same chemistry on the terpene portion of the meroterpenoid backbone and the cyclisation outcome delivers chemistry resembling that of emindole SA (**31**). Clade 3 contained two previously characterized alternative IDT cyclases (AtS2B and AfB). The other previously characterized alternative cyclase, AtS5B1 was clustered in clade 2.[30] Three of the cyclases we characterized in this study clustered with the other IDT cyclases (EstB1, DesB and EstB2). EstB3 was basal to the other clades, indicating it could be functionally divergent from these cyclases.

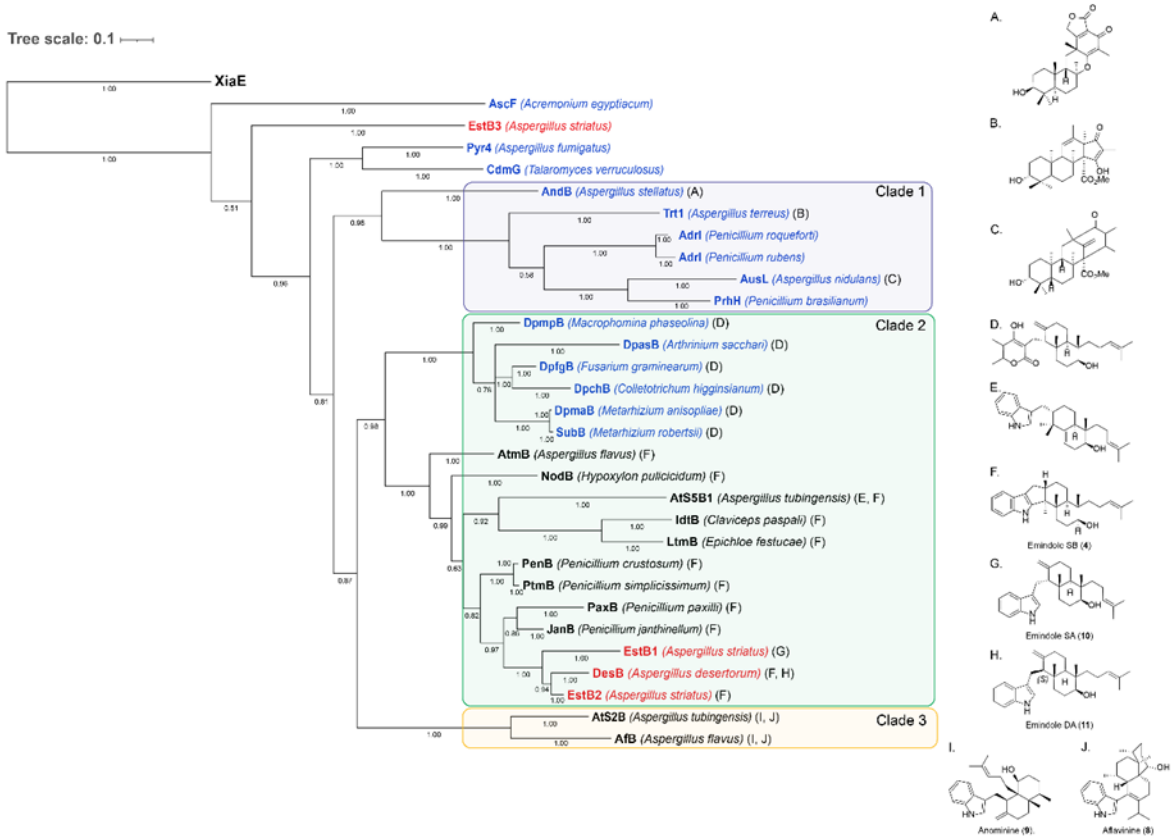


Figure S32. Phylogenetic tree of amino acid sequences of fungal IDT and meroterpenoid cyclases (blue text), as determined by Bayesian inference (the node support is displayed as posterior probabilities). XiaE, a bacterial terpene cyclase, was used as an outgroup to root the tree. The IDT cyclases investigated in this study are shown in red. Compounds produced by the cyclases are indicated with letters.

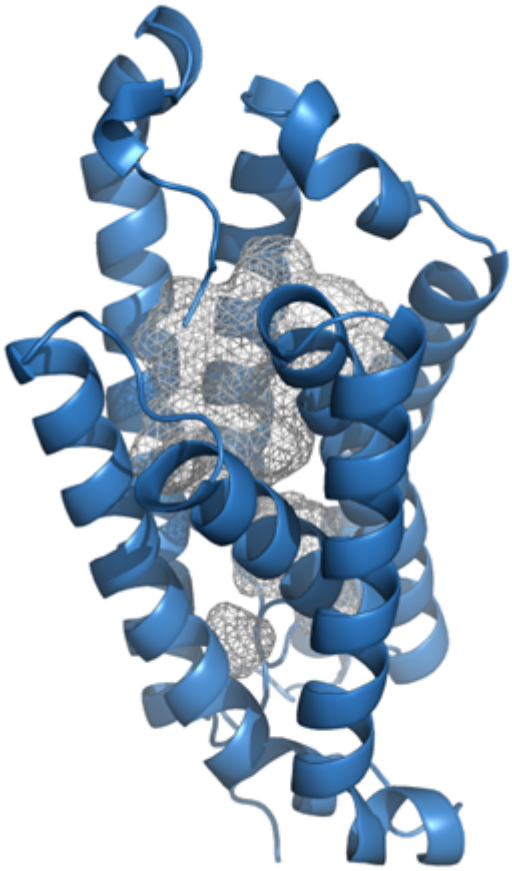


Figure S33. Structural model of PaxB with mesh showing the internal surface of the putative substrate binding pocket contained within the membrane-spanning α-helical barrel. The PaxB model (AF-E3UBL6-F1-model_v4) was obtained from the EBI AlphaFold Protein Structure Database[2-3] (alphafold.ebi.ac.uk) on 20 Dec 2022.

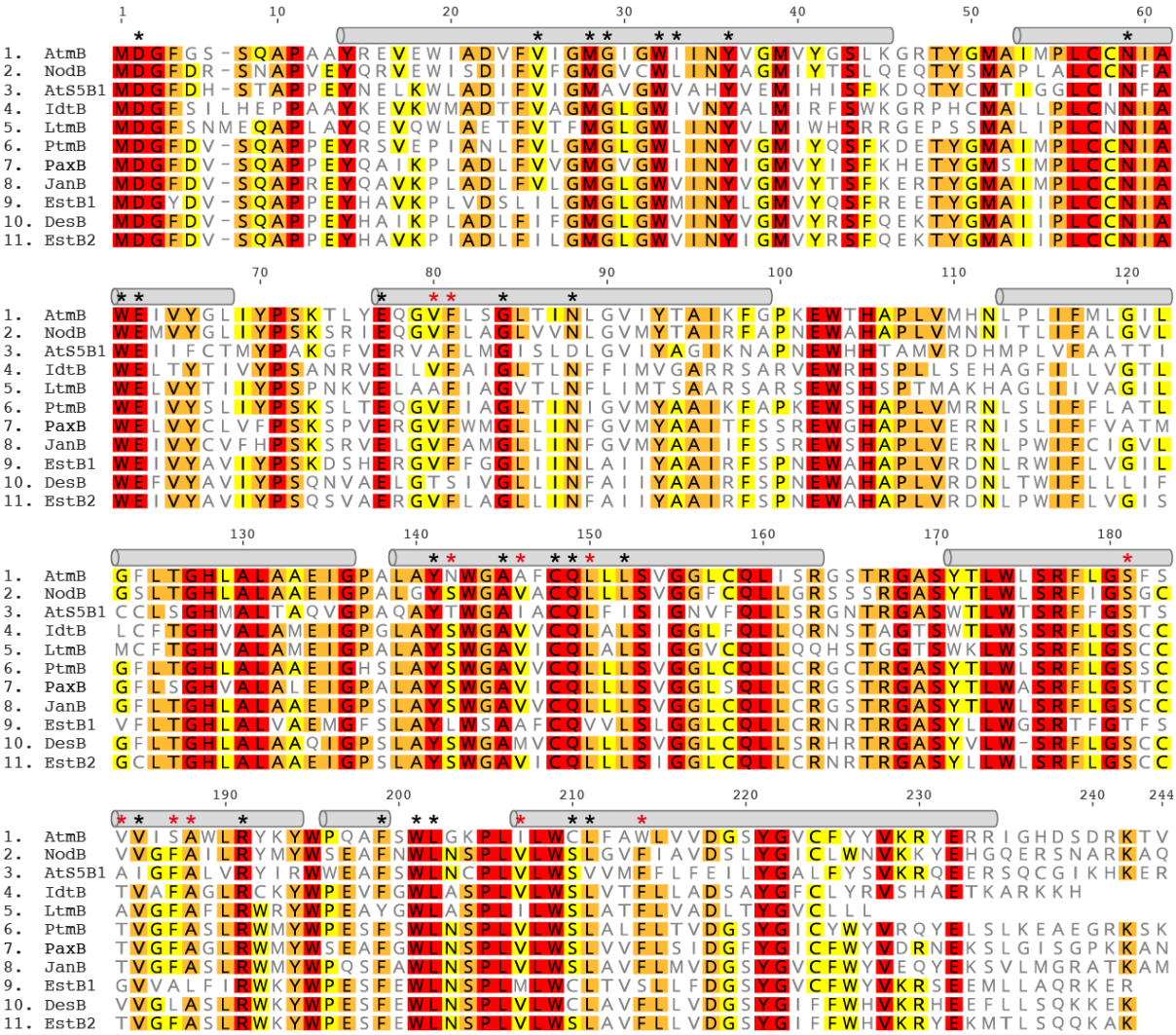


Figure S34. Multiple sequence alignment of functionally characterized fungal IDT cyclases from Clade 2 in Figure S32.

Highlights show percentage sequence identify (red, 100%; orange, >80%; yellow, >60%). Grey boxes show the positions of alpha helices relative to AlphaFold 2 PaxB protein structural model shown in Figure S33. Asterisks show residues that line the putative internal substrate binding pocket. Asterisks coloured red highlight binding pocket residues that exhibit sequence variations unique to the EstB1 and/or DesB sequences, and therefore may contribute to the unique functions of these enzymes. Multiple sequence alignment generated using Clustal Omega version 1.2.3.[38]

Sequences used in phylogenetic analysis

>DpchB
MNVADISQAPEAYRDVVWIADTCKLIMGIGWTANYVGMIRKSLKDQTYAMALLPLCCNFAWELTYAIMYAFTTSLEKYVHFSGLLLNCGVMYTAVKN
APREWEHAPLVQRNLRLIFVLAVAGFASAHVVLAKQVGPELGQAWSAYACQLLLSVGGLCQLLCRGHSRGASYFLWFSRFFGSLVLVPQDIIRYTY
WKEAHEFMGSPMYIWFVTIFLILDGSYGLCLWYVRRFEQQNPAAGKLKK

>CdmG
MDYFYGTSPPPEYERYASIVDAATLVQGFLWALNYGEASYRSIKDRTYGMAIFPLCCNYAWELVYTVIYSSQNKYERIIMTTWLILNSIMMGFTIKFAP
NEWRHAPLVQRNIPFIFLAGVAAFVIAQLALAATVGPGLAMNWVAALCYLLLTIGSLCQLMTRGSSRGVSYTMWLSRFVGTYVGVICVYFRYNYWP
QNFSWVDEPIMKCFSGISLAVEIVYGVTLWHIRKQERHHIVEKSK

>Trt1
MPSIISDPQAYDIMLRLLQFSCWSLSYINTVRTTLSDQLPSVSFMSICCDVAWEFVYAFVYPIASSHWAGGIRIWFAMHCVMLFIVAKYAPNDWDHVP
LMKRFARLAYVAITIGFMAGHLALASEIGPALGFFWSGALCQITASLGSLCLLVCRGSTRGASIKTCPLCWFYIAITLTLDAIYPVFFFYFRAIEHPKKDS
ERKVE

>SubB
MNAADISRAPPGYLEVAWIADTCKLLMGLGWTTNYAGMIYKSLKDRTYGMALMPLCCNFA
WELTYAVIYPFGSRQDKFTHYFGLMLNCGVMYTAVKNAERFWTHAPLVRRNLPFIFIICI
AAWTTAHLALALQIGPSHAQAFSAYGCQLLLSVGALCQLLSRGSSRGASYFLW

>PtmB
MDGFDVSQAPPEYRSVEPIANLFVLGMGLGWLINYVGMIYQSFKDETYGMAIMPLCCNIAWEIVYSLIYPSKSLTEQGVFIAGLTINIGVMYAAIKFAP
KEWSHAPLVMRNLSLIFFLATLGFLTGHLALAAEIGHSLAYSWGAVVCQLLLSVGGLCQLLCRGCTRGASYTLWLSRFLGSSCTVGFASLRWMYWP
ESFSWLNSPLVLWSLALFLTVDGSYGICYWYVRQYELSLKEAEGRKSK

>PrhH
MEEPLTVAAIFRDPFNILAISEVLKVVAAVGWSVNYIGMVHRAWKDQIPSIGILPLCCDIGWEFVYAWMFPDFSSHWQGVVRVWFFLHSAVLLVTLKV
SPNDWVHTPLGHRHIVFIYIFVTLVFGAGQYALAAEIGPALGFHWGGALCQFLSSSCGIAQLLSRGHTRGASYLIWFARAISTFAGFIKLCIRFQHNVD
GAPWLDSPMCWFYIVTVLSFDAAYPFLYSSMRKLETPALRKESRIKNQ

>Pyr4
MDGWSDLSSAPPQYREVAGIADWALLAQGLGWSINYLAMIYHSYKDRTYGMAILPLCCNFAWEFVYSVIYPSHNSAERAVLTTWMILNLFVMYTAIK
FAPNEWQHAPLVRQCLPWIFPVAIAAFTAGHLALAATVGVSKAANWGAFLCFELLTSGAVCQLMSRGSSRGASYTIWLSRFLGSYIGGIFLHVRETH
WPQEFGWISHPFVTWHGLMCFSLDIAYVTFLWRIRRQEHRSQRKKAL

>JanB
MDGFDVSQAPREYQAVKPLADLFVLGMGLGWVINYVGMVYTSFKERTYGMAIMPLCCNIAWEIVYCVFHPSKSRVELGVFAMGLLINFGVMYAAIIF
SSREWSHAPLVERNLPWIFCIGVLGFLTGHLALAAEIGPSLAYSWGAVVCQLLLSVGGLCQLLCRGSTRGASYTLWLSRFLGSCCTVGFASLRWMY
WPQSFAWLNSPLVLWSLAVFLMVDGSYGVCFWYVEQYEKSVLMGRATKAM

>PaxB
MDGFDVSQAPPEYQAIKPLADLFVVGMGVGWIINYIGMVYISFKHETYGMSIMPLCCNIAWELVYCLVFPSKSPVERGVFWMGLLINFGVMYAAITFS
SREWGHAPLVERNISLIFFVATMGFLSGHVALALEIGPALAYSWGAVICQLLLSVGGLSQLLCRGSTRGASYTLWASRFLGSTCTVGFAGLRWMYW
SEAFGWLNSPLVLWSLVVFLSIDGFYGICFWYVDRNEKSLGISGPKKAN

>PenB
MDGFDVSQAPPEYRAVEPIANIFVLGMGLGWLINYVGMIYQSFKDETYGMAIMPLCCNIAWEIVYSLIYPSKSLIEQGVFIAGLTINIGVMYAAIKFAPK
EWSHAPLVMRNLSLIFFLATLGFLTGHLALAAEIGHSLAYSWGAVVCQLLLSVGGLCQLLCRGSTRGASYTLWLSRFLGSSCTVAFASLRWMYWPE
SFSWLNSPLVLWSLALFLTVDGSYGLCYWYVRQYELSLKEAEGRKSK

>DpfgB
MEVADPSRAPPEYKDVAWIADTCKLLMGIGWTTNYVGMIYKSLKDETYAMALMALCCNFAWELTYALIYPFGSDLEMYVHFSGLMLNCGVMYTAVK
NAHREWGHSPLVLRNLPLIFIICVSGFMSGHVALAAQVGPSLAQAWSAYGCQLLLSVGGLCQLLCRGHSRGASYFLWFSRFFGSLVLVPQDILRYK
YWRVDHEYMGSPLYIWFVCIFLLLDGSYGICLWYVRRFERQTAVAHKKKK

>DpasB
MDVHDLTRAPPEYLEVVWVTDVCKLVMAVGWLSNYIGMIAKSIKEQTYSMALMPLCCNFAWEFTYFFIYPYKVPMERNIHTLAFLLNCGVMYTAVRY
GAREWGHAPLVQRNLPVIFVVCIACWVSAHVAFAEQYGPSLAQAVSGFACQILLSAGGTCQLLCRGHSRGASYKLWLARFMGSFALILPNMLRYKY
WRDDHQYIGSPLYIWFLGMFLFLDGSYGFVLWYVRRHEREQVLVAKPKVQ

>DpmaB
MNAADISRAPPGYLEVAWIADTCKLLMGLGWTTNYAGMIYKSLKDRTYGMALMPLCCNFAWELTYAVIYPFGSRQDKFTHYFGLMLNCGVMYTAV
KNAEREWTHAPLVRRNLPFIFIICIAAWTTAHLALALQIGPSHAQAFSAYGCQLLLSVGALCQLLSRGSSRGASYFLWFCRFFGSLVLIPQDVLRYRY
WRQDHEYMGSPLYIWFVSIFLLLDGSYALCLWYVRRFESEQEEAKKAKSI

>LtmB
MDGFSNMEQAPLAYQEVQWLAETFVTFMGLGWLINYVLMIWHSRRGEPSSMALIPLCNNIAWELVYTIIYPSPNKVELAAFIAGVTLNFLIMTSAARS
ARSEWSHSPTMAKHAGLIIVAGILMCFTGHVALAMEIGPALAYSWGAVICQLALSIGGVCQLLQQHSTGGTSWKLWSSRFLGSCCAVGFAFLRWRY
WPEAYGWLASPLILWSLATFLVADLTYGVCLLL

>Adrl_Prub
MEKSTLLSAVLKHRDALASVAEFLRILAGICWTLNYFSMLRTSQKDKIPSTGIFPLCNDIGWEFIYAFIYPKASAHWEGGVRVWFLVHCIVIFFIIKNAHN
EWDYFPLIQRNLYFLYGIVTIGFAIGQYSFAREVGPDLGFFYGGVLCQTLASLGPIAQILSRNSTRGASLLLRAVATFGGFIKLTIYYLTGNAAGPWFES
PMCKFYIGLTLILDFTYPICYYVIRRQELVNDEGDKKKKTKSGKAA

>AusL
MSQLTISKIIEEPFSALSLSEMLKILAALGWSTNYLAMVYRTQADKLPAIAVLPLCCDIAWEFTYAWIYPQASGHWQGVVRVWFFLHTAVLAATLRYA
PNDWAGTPLGESRGRLVLLYAAVIAAFAAGQLCLALEMGGALGFHWGGALCQFLSSSAAVGQLLTRGHTRGASLLIWGARAISTAGGDRALIGCVV
SGAVPIDQKA

>AndB

MQPITQIPLTFDTVVNLLGSASGIGWILNYILMTYYSFRDKTYSMSMLPLCCNIAWEFVYGILCPSSTFVVRPVILSWLVLNCLVVYAAIKYSPNEWAH
APLVQRHLPLLFTVGIAACTGFHIALIRKFDPATAFLWSARSCQVLLSIGGLFQLLCRSSTKGGSYVLWLSRFLGSICGVLKMTLMWKYGESRFPWLD
DPLTAYCIALWIISDVLYGVVFYSLRSKELAGAGKAKAI
>DpmpB
MNIVPLSQAPPEFLEVAWLADACKLLMGVGWTANYIGMIYKSIKDRTYGMALMPLCCNFAWELVYALILPFDSGMEKWVHVTGLAFNCGVMYTAIK
FAPGEWAHARLVQRHLTWIFIASVAGWMSAHLALAAQLGPSLAQAWSAYGCQLLLSVGGLCQLLCRGHSRGTSYLLWFSRFFGSLVLIPQDILRYK
YWRRDHEWMKSPLYLWFVSIFLILDGSYGILLWYVRRFERETAEAENRKRR

>Adrl_Proq
MEESSLLSAILDHRDALASVAEFLRILAGICWTLNYFSMLRTSRKDKIPSTGIFPLCNDIGWEFIYAFIYPTASAHWEGGVRVWFLVHCIVIIFIIKYAHNE
WDHFPLIQRNLYFLYGVVTIGFAIGQYSFAREVGPDLGFFYGGVLCQTLASLGPIAQILSRNSTRGASLLTWLLRAIATFGGFIKLTIYYLTGNAAGPW
FESPMCKFYIGLTLVLDFTYPICYYVIQRQELANAQKEKKEKSK

>AtmB_Afla
MDGFGSSQAPAAYREVEWIADVFVIGMGIGWIINYVGMVYGSLKGRTYGMAIMPLCCNIAWEIVYGLIYPSKTLYEQGVFLSGLTINLGVIYTAIKFGP
KEWTHAPLVMHNLPLIFMLGILGFLTGHLALAAEIGPALAYNWGAAFCQLLLSVGGLCQLISRGSTRGASYTLWLSRFLGSFSVVISAWLRYKYWPQ
AFSWLGKPLILWCLFAWLVVDGSYGVCFYYVKRYERRIGHDSDRKTV

>AscF
MAFGVEPPEHVTPWFKPVYEATFQFGGVAWTLCYILIAREGMRTKSYGMPLFALANNFAWEMVYALWVVDNAFEKTAMTIWMLIDTPIIYSILKHGV
LEWQHAPMVSRNLKSILVGLIALCAAAHWSWQSWWIGNEMGKRDDLEGADLTQMAYWAVSMCQFLVSTMSLAMLCVRGHSGGVSWMIWLSRFL
GTLIGLNMNYAWAYYTWPEAHEYFMSAPAVFVWGVTTVCDIIYGFVLYHVKSNERELSDGRKVAAEADDEQVGGWSKMKTGKN

>IdtB
MDGFSILHEPPAAYKEVKWMADTFVAGMGLGWIVNYALMIRFSWKGRPHCMALLPLCNNIAWELTYTIVYPSANRVELLVFAIGLTLNFFIMVGARR
SARVEWRHSPLLSEHAGFILLVGTLLCFTGHVALAMEIGPGLAYSWGAVVCQLALSIGGLFQLLQRNSTAGTSWTLWSSRFLGSCCTVAFAGLRCK
YWPEVFGWLASPLVLWSLVTFLLADSAYGFCLYRVSHAETKARKKH

>NodB
MDGFDRSNAPVEYQRVEWISDIFVFGMGVCWLINYAGMIYTSLQEQTYSMAPLALCCNFAWEMVYGLIYPSKSRIEQGVFLAGLVVNLGVMYTAIR
FAPNEWAHAPLVMNNITLIFALGVLGSLTGHLALAAEIGPALGYSWGAVACQLLLSVGGFCQLLGRSSSRGASYTLWLSRFIGSGCVVGFAILRYMY
WSEAFNWLNSPLVLWSLGVFIAVDSLYGICLWNVKKYEHGQERSNARKAQ

>EstB1
MDGYDVSQAPPEYHAVKPLVDSLILGMGLGWMINYLGMVYQSFREETYGMAIIPLCCNIAWEIVYAVIYPSKDSHERGVFFGGLIINLAIIYAAIRFSPN
EWAHAPLVRDNLRWIFLVGILVFLTGHLALVAEMGFSLAYLWSAAFCQVVLSLGGLCQLLCRNRTRGASYLLWGSRTFGTFSGVVALFIRWKYWPE
SFEWLNSPLMLWCLTVSLLFDGSYGVCFWYVKRSEEMLLAQRKER

>EstB2
MDGFDVSQAPPEYHAVKPIADLFILGMGLGWVINYIGMVYRSFQEKTYGMAIIPLCCNIAWEIVYAVIYPSQSVAERGVFLAGLLINFAIIYAAIRFSPNE
WAHAPLVRDNLPWIFLVGISGCLTGHLALAAEIGPSLAYSWGAVICQLLLSIGGLCQLLCRNRTRGASYLLWLSRFLGSCCTVGFASLRWKYWPESF
EWLNSPLVLWSLAVFLLVDGSYGIFFWHVKRYEKMTLSQQKAK

>EstB3
MEEGWDFDAAPAEFKQVQPFLLGLFAVSGTGWLINYFTTIRAAFRDRTSGVSLIALSNNLAWEFVYAIFHPPPLPIATIIVRSWLLVDIFVIYTTVKFAR
AAPGNVNPPLLKHYLPLFVLGGILGFFSGHWALSVLLSPLRAFYWSGMVCLITMSGTALGLLVQRGHTRGASYGMWLSRFIASIFAVASLFLRSTHW
PQVWAWSDNILMLWLSGAFFVLDILYGVCFWYTRQLEKQARRQKLA

>AfB
MDSFDLANAPPEIRAYATPIILLNLYTNASWLYVYFGMVYRSVKDKSYAMPLYSQCLNIAWEITYGYIYGDDWMLFATFLVTFPTDCLVIWAAIYHGAK
EWDRSPLVQRNLLWYYVIGTGIAVALHMCAASELGVEKAFFAGAIGCQAVLSVGYLGNLIQMGSTRGFSMHLWFFRFTGSLTLVPEFYLRVKYWPE
RFGFLGQPLMLWCCAVFLGFDLVYGILFWYIRRQERETGMLLADGRKRK

>AtS5B1
MDGFDHSTAPPEYNELKWLADIFVIGMAVGWVAHYVEMIHISFKDQTYCMTIGGLCINFAWEIIFCTMYPAKGFVERVAFLMGISLDLGVIYAGIKNAP
NEWHHTAMVRDHMPLVFAATTICCLSGHMALTAQVGPAQAYTWGAIACQLFISIGNVFQLLSRGNTRGASWTLWTSRFFGSTSAIGFALVRYIRWW
EAFSWLNCPLVLWSVVMFFLFEILYGALFYSVKRQEERSQCGIKHKER

>AtS2B
MDAFDLSTAPPEFASWATTLYACNIYTNFIWLYVYYGMIYRSYKDKSFAMPLISQCLNIAWEIVFGFLFSQDHWFITLSFQAAVISNCGVIYAAIKYGAP
EWNRSPMIQRNLPWIYIGGTLLAIAGHLALATELGMVRACFQGAIVCQAILSVGYVCQLLVRGSTRGFSLNLWFFRFTGSLVMVPEFYIRVNYWPDA
FSWLGEPFMLWCCFIYLGFDLAYPVLFWYIQRREKEEALAKSIKSL

>DesB
MDGFDVSQAPPEYHAIKPLADLFIFGMGLGWVINYVGMVYRSFQEKTYGMAIIPLCCNIAWEFVYAVIYPSQNVAELGTSIVGLLINFAIIYAAIRFSPN
EWAHAPLVRDNLTWIFLLLIFGFLTGHLALAAQIGPSLAYSWGAMVCQLLLSVGGLCQLLSRHRTRGASYVLWSRFLGSCCVVGLASLRWKYWPES
FEWLNSPLVLWCLAVFLLVDGSYGIFFWHVKRHEEFLLSQKKEK

>XiaE
MVWLPPFLIPMSEVPPVTVGAADVSDLLFAAVAGPTALGWMVTYVLAIRQARRDGRTGIPAYLIAVNIAWEFSLTFLLEQTPTQRQINFLWLVFNVFL
FAQALRYGPRDYPGLSARTFRWTLAGVLVWASVVVMVGANELHDVDGMYTGMIIQVPLSAAFILMLRRRGSSAGQSMHIAVAKTVGSLFAGLTAVI
VYPSHHLLQVLVPTYVVLDVAYVVLLRRTMLREGRPLWAFRHPGAGVPGGCRLPVR

Supplementary References

1. van Dolleweerd, C. J.; Kessans, S. A.; Van de Bittner, K. C.; Bustamante, L. Y.; Bundela, R.; Scott, B.; Nicholson, M. J.; Parker, E. J., MIDAS: A Modular DNA Assembly System for Synthetic Biology. *ACS Synth. Biol.* **2018,** *7* (4), 1018-1029.
2. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Žídek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S., AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2021,** *50* (D1), D439-D444.
3. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D., Highly accurate protein structure prediction with AlphaFold. *Nature* **2021,** *596* (7873), 583-589.
4. Cox, M. P.; Peterson, D. A.; Biggs, P. J., SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinform.* **2010,** *11* (1), 485.
5. Simpson, J. T.; Wong, K.; Jackman, S. D.; Schein, J. E.; Jones, S. J. M.; Birol, I., ABySS: a parallel assembler for short read sequence data. *Genome Res.* **2009,** *19* (6), 1117-1123.
6. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J., Basic local alignment search tool. *J. Mol. Biol.* **1990,** *215* (3), 403-10.
7. Nordberg, H.; Cantor, M.; Dusheyko, S.; Hua, S.; Poliakov, A.; Shabalov, I.; Smirnova, T.; Grigoriev, I. V.; Dubchak, I., The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res.* **2014,** *42* (D1), D26-31.
8. Grigoriev, I. V.; Nordberg, H.; Shabalov, I.; Aerts, A.; Cantor, M.; Goodstein, D.; Kuo, A.; Minovitsky, S.; Nikitin, R.; Ohm, R. A.; Otillar, R.; Poliakov, A.; Ratnere, I.; Riley, R.; Smirnova, T.; Rokhsar, D.; Dubchak, I., The genome portal of the Department of Energy Joint Genome Institute. *Nucleic Acids Res.* **2012,** *40* (D1), D26-32.
9. Solovyev, V.; Kosarev, P.; Seledsov, I.; Vorobyev, D., Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol.* **2006,** *7 Suppl 1*, S10 1-12.
10. Stanke, M.; Morgenstern, B., AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* **2005,** *33* (suppl_2), W465-W467.
11. Nicholson, M. J.; Koulman, A.; Monahan, B. J.; Pritchard, B. L.; Payne, G. A.; Scott, B., Identification of two aflatrem biosynthesis gene loci in Aspergillus flavus and metabolic engineering of *Penicillium paxilli* to elucidate their function. *Appl. Environ. Microbiol.* **2009,** *75* (23), 7469-81.
12. Gilchrist, C. L. M.; Chooi, Y.-H., clinker & clustermap.js: automatic generation of gene cluster comparison figures. *Bioinformatics* **2021,** *37* (16), 2473-2475.
13. Yelton, M. M.; Hamer, J. E.; Timberlake, W. E., Transformation of *Aspergillus nidulans* by using a trpC plasmid. *Proc. Natl. Acad. Sci. U.S.A.* **1984,** *81* (5), 1470.
14. Vollmer, S. J.; Yanofsky, C., Efficient cloning of genes of Neurospora crassa. *Proc. Natl. Acad. Sci. U.S.A.* **1986,** *83* (13), 4869-73.
15. Oliver, R. P.; Roberts, I. N.; Harling, R.; Kenyon, L.; Punt, P. J.; Dingemanse, M. A.; van den Hondel, C. A. M. J. J., Transformation of *Fulvia fulva*, a fungal pathogen of tomato, to hygromycin B resistance. *Curr. Genet.* **1987,** *12*, 231-233.
16. Zhang, P.; Li, X.-M.; Li, X.; Wang, B.-G., New indole-diterpenoids from the algal-associated fungus *Aspergillus nidulans*. *Phytochem. Lett.* **2015,** *12*, 182-185.
17. Fan, Y.; Wang, Y.; Liu, P.; Fu, P.; Zhu, T.; Wang, W.; Zhu, W., Indole-Diterpenoids with Anti-H1N1 Activity from the Aciduric Fungus *Penicillium camemberti* OUCMDZ-1492. *J. Nat. Prod.* **2013,** *76* (7), 1328-1336.
18. Zhang, Y. H.; Huang, S. D.; Pan, H. Q.; Bian, X. Q.; Wang, Z. Y.; Han, A. H.; Bai, J., Structure determination of two new indole-diterpenoids from Penicillium sp. CM-7 by NMR spectroscopy. *Magn. Reson. Chem.* **2014,** *52* (6), 306-9.
19. Belofsky, G. N.; Gloer, J. B.; Wicklow, D. T.; Dowd, P. F., Antiinsectan alkaloids: Shearinines A-C and a new paxilline derivative from the ascostromata of *Eupenicillium shearii*. *Tetrahedron* **1995,** *51* (14), 3959-3968.
20. Uhlig, S.; Egge-Jacobsen, W.; Vrålstad, T.; Miles, C. O., Indole–diterpenoid profiles of Claviceps paspali and Claviceps purpurea from high-resolution Fourier transform Orbitrap mass spectrometry. *Rapid Commun. Mass Spectrom.* **2014,** *28* (14), 1621-1634.
21. Uhlig, S.; Rangel-Huerta, O. D.; Divon, H. H.; Rolén, E.; Pauchon, K.; Sumarah, M. W.; Vrålstad, T.; Renaud, J. B., Unraveling the Ergot Alkaloid and Indole Diterpenoid Metabolome in the Claviceps purpurea Species Complex Using LC–HRMS/MS Diagnostic Fragmentation Filtering. *J. Agric. Food. Chem.* **2021,** *69* (25), 7137-7148.
22. Liu, D. Q.; Sun, M., Formation of the Ions of Methylindoles in APCI Mass Spectrometry. *ISRN Spectroscopy* **2012,** *2012*, 973649.
23. Itoh, Y.; Johnson, R.; Scott, B., Integrative transformation of the mycotoxin-producing fungus, *Penicillium paxilli*. *Curr. Genet.* **1994,** *25* (6), 508-13.
24. Saikia, S.; Parker, E. J.; Koulman, A.; Scott, B., Four gene products are required for the fungal synthesis of the indole-diterpene, paspaline. *FEBS Lett.* **2006,** *580* (6), 1625-1630.
25. Scott, B.; Young, C. A.; Saikia, S.; McMillan, L. K.; Monahan, B. J.; Koulman, A.; Astin, J.; Eaton, C. J.; Bryant, A.; Wrenn, R. E.; Finch, S. C.; Tapper, B. A.; Parker, E. J.; Jameson, G. B., Deletion and gene expression analyses define the paxilline biosynthetic gene cluster in *Penicillium paxilli*. *Toxins* **2013,** *5* (8), 1422-46.
26. McLellan, R. M.; Cameron, R. C.; Nicholson, M. J.; Parker, E. J., Aminoacylation of Indole Diterpenes by Cluster-Specific Monomodular NRPS-like Enzymes. *Org. Lett.* **2022,** *24* (12), 2332-2337.
27. Van de Bittner, K. C.; Cameron, R. C.; Bustamante, L. Y.; Bundela, R.; Kessans, S. A.; Vorster, J.; Nicholson, M. J.; Parker, E. J., Nodulisporic acid E biosynthesis: in vivo characterisation of NodD1, an indole-diterpene prenyltransferase that acts on an emindole SB derived indole-diterpene scaffold. *Medchemcomm* **2019,** *10* (7), 1160-1164.
28. Nozawa, K.; Yuyama, M.; Nakajima, S.; Kawai, K.-i.; Udagawa, S.-i., Studies on fungal products. Part 19. Isolation and structure of a novel indoloditerpene, emindole SA, from Emericella striata. *J. Chem. Soc., Perkin Trans. 1* **1988,** (8), 2155-2160.
29. Nozawa, K.; Nakajima, S.; Kawai, K.-i.; Udagawa, S.-i., Isolation and structures of novel indoloditerpenes, emindoles DA and DB, from *Emericella desertorum*: X-ray molecular structure of emindole DA acetate. *J. Chem. Soc., Perkin Trans. 1* **1988,** (7), 1689-1694.

30. Tang, M. C.; Lin, H. C.; Li, D.; Zou, Y.; Li, J.; Xu, W.; Cacho, R. A.; Hillenmeyer, M. E.; Garg, N. K.; Tang, Y., Discovery of Unclustered Fungal Indole Diterpene Biosynthetic Pathways through Combinatorial Pathway Reassembly in Engineered Yeast. *J. Am. Chem. Soc.* **2015,** *137* (43), 13724-7.
31. Madeira, F.; Pearce, M.; Tivey, A. R. N.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R., Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* **2022**, gkac240.
32. Brown, N. P.; Leroy, C.; Sander, C., MView: a web-compatible database search or multiple alignment viewer. *Bioinformatics* **1998,** *14* (4), 380-381.
33. Matsuda, Y.; Wakimoto, T.; Mori, T.; Awakawa, T.; Abe, I., Complete Biosynthetic Pathway of Anditomin: Nature's Sophisticated Synthetic Route to a Complex Fungal Meroterpenoid. *J. Am. Chem. Soc.* **2014,** *136* (43), 15326-15336.
34. Guo, C.-J.; Knox, B. P.; Chiang, Y.-M.; Lo, H.-C.; Sanchez, J. F.; Lee, K.-H.; Oakley, B. R.; Bruno, K. S.; Wang, C. C. C., Molecular Genetic Characterization of a Cluster in A. terreus for Biosynthesis of the Meroterpenoid Terretonin. *Org. Lett.* **2012,** *14* (22), 5684-5687.
35. Lo, H.-C.; Entwistle, R.; Guo, C.-J.; Ahuja, M.; Szewczyk, E.; Hung, J.-H.; Chiang, Y.-M.; Oakley, B. R.; Wang, C. C. C., Two Separate Gene Clusters Encode the Biosynthetic Pathway for the Meroterpenoids Austinol and Dehydroaustinol in *Aspergillus nidulans*. *J. Am. Chem. Soc.* **2012,** *134* (10), 4709-4720.
36. Holder, M.; Lewis, P. O., Phylogeny estimation: traditional and Bayesian approaches. *Nat. Rev. Genet.* **2003,** *4* (4), 275-284.
37. Ronquist, F.; Deans, A. R., Bayesian Phylogenetics and Its Influence on Insect Systematics. *Annu. Rev. Entomol.* **2010,** *55* (1), 189-206.
38. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011,** *7* (1), 539.