Electronic Supplementary Material (ESI) for Chemical Communications.

Electronic Supporting Information

Unified biogenesis of ambiguine, fischerindole, hapalindole and welwitindolinone: Identification of a monogeranylated indolenine as a cryptic common biosynthetic intermediate by an unusual magnesium-dependent aromatic prenyltransferase

Xinyu Liu*[a], Matthew L. Hillwig[b], Leonardus M. I. Koharudin[b] and Angela M. Gronenborn[b]

a. Department of Chemistry, University of Pittsburgh, 219 Parkman Avenue, Pittsburgh, PA 15260, USA.

b. Department of Structural Biology, University of Pittsburgh School of Medicine, 3501 Fifth Avenue, Pittsburgh, PA 15260, USA

*To whom correspondence should be addressed: xinyuliu@pitt.edu

ESI Materials and Methods

General methods. All polymerase chain reactions (PCRs) were carried out on a C1000 thermal cycler (Bio-Rad). DNA sequencing was performed by Elim BioPharm Inc. Preparative-scale reverse-phase HPLC was performed using a Dionex instrument equipped with Luna C18 columns (21 x 250 mm and 4.6 x 250 mm) (Phenomenex). Analytical reverse-phase HPLC was performed using a Dionex UHPLC with a photo-diode array UV/Vis detector (Thermo Fisher Scientific) and a 4.6 x 250 mm Luna C18 column (Phenomenex). HRMS analysis was conducted using a Q Exactive Benchtop Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific) equipped with a Dionex RSLC (Thermo Fisher Scientific). NMR spectrum was recorded on a Bruker Avance III 700 MHz spectrometer equipped with a $^{1}H/^{13}C/^{15}N$ triple-resonance inverse probe (1.7mm 'microprobe').

Materials. Synthetic oligonucleotides for gene amplification by PCR were purchased from Life Technologies or Integrated DNA Technology. Kappa HiFi DNA polymerase was obtained from Kappa Biosystems. Restriction endonucleases, T4 DNA ligase and Antarctic phosphatase were purchased from New England Biolabs. LB broth and agar used for culturing *E. coli* were obtained from Teknova. All other reagents including inorganic salts and cofactors were purchased from Sigma-Aldrich or Fisher Scientific unless otherwise stated.

Strains and plasmids. *E. coli* TOP10 cell (Life Technologies) was used for routine cloning and plasmid propagation. *E. coli* C43(DE3) cell (Lucigen) was used for protein expression. pQTEV cloning plasmid was obtained from Addgene.

Protein expression. AmbP1, and its homologs WelP1 and FidP1, were cloned and expressed in *E. coli* as previously described.[1] Cell pellets were re-suspended in 10 mL of protein lysis buffer (50 mM Tris, 500 mM NaCl, 20 mM imidazole, 10 mM beta-mercatoethanol, and 0.1 % Tween 20) and

sonicated for 10 s three times on ice and transferred to 50 mL tubes and centrifuged for 30 min at 30,000*g* at 4 °C to pellet cell debris. The cleared lysate was transferred to 2 mL of pre-washed nickel-NTA bead (Qiagen, Valencia, CA) slurry and incubated for 1 h at 4 °C. Thereafter, the nickel-NTA beads were rinsed three times with 20 mL of binding buffer (50 mM Tris, 500 mM NaCl, 0.1% TWEEN20, 20 mM imidazole, 10 mM β-mercatoethanol, pH=7). The His-tagged proteins were then eluted by addition of 5 mL of elution buffer (50 mM Tris, 500 mM NaCl, 250 mM imidazole, 10 mM beta-mercatoethanol) to the bead bed. Eluted protein was subjected to dialysis (50 mM Tris, 50 mM NaCl, 10% glycerol, 0.5 mM DTT) in a 10-kDa molecular weight cutoff membrane (Spectrum Laboratory Products, Inc., Gardena, CA) to remove the imidazole and exchange buffer. The purified proteins were analyzed by SDS-PAGE to ensure homogeneity, concentrated with 30 kDa cutoff concentrator tubes, assayed, and the remainder was flash-frozen using liquid nitrogen and stored at -80 C for later assays.

AmbP1/WelP1/FidP1 in vitro assay. For a typical AmbP1 assay, 5 μM of AmbP1 was added to a 50 μL reaction containing 50 mM Tris-HCl pH 8.0, 0.5 mM (*Z*)-**1** or (*E*)-**1**, 0.5 mM GPP, 5 mM $MgCl_2$, at 30 °C. All assays were stopped at indicated time by extraction with ethyl acetate (0.5 mL x 2). The combined organic layer was dried under a stream of N_2 gas, and redissolved in methanol (100 μL). A 40-μL aliquot was used for HPLC and LC-HRMS analysis for product identification. For optimizing assay conditions, alternative buffers (MES buffer for pH <7 and Tris buffer for pH>7), metal ions ($ZnCl_2$, $CaCl_2$, $MnCl_2$) or EDTA were used as indicated. For substrate scope studies, DMAPP, NPP and FPP were used as GPP alternative. (*E*)-**1**, **4** and **5** were used as (*E*)-**1** alternative. For kinetics, AmbP1 concentration was fixed at 0.1 μM, GPP concentration was fixed at 0.5 mM and (*Z*)-**1** concentration was varied from 1 μM to 200 μM. Reactions were stopped after 2 min and conversions were estimated based on HPLC with analytical standards.

Structure determination of the AmbP1 enzymatic products. For product **2**, assays were scaled to 5 mL. Enzymatic mixtures containing 50 mM Tris buffer (pH 9.0), 20 mM $MgCl_2$, 1 mM (*Z*)-**1**, 1mM GPP with 1 μM AmbP1 were incubated overnight at 30 °C. For **3**, assays were scaled to 5 mL. Enzymatic mixtures containing 50mM MES (pH 6.0), 10 mM EDTA, 1 mM (*Z*)-**1**, 1mM GPP with 1 μM AmbP1 were incubated overnight at 30 °C. Each of the enzymatic products were extracted from the buffer with equal volume ethyl acetate twice, dried, redissolved in methanol, filtered and further purified using a 250 x 4.6 mm Luna C18 5 micron HPLC column (Phenomenex, Torrance, CA), with a gradient from 50-80% acetonitrile in 2-35 min. The dried products were re-dissolved in 40 μL of CD_3OD or CD_3Cl and analyzed in a 1 mm capillary NMR tube using a Bruker 700 MHz NMR spectrometer equipped with a capillary probe. The structure of product **5** was determined in an analogous manner.

AmbP1 expression and purification for NMR study: AmbP1 gene encoding residues 1−309 was inserted into the pET-15b(+) expression vector (Novagen), using NdeI and XhoI restriction sites at the 5' and 3' ends, respectively. For protein expression, *E. coli* Rosetta2 (DE3) cells (Novagen) were transformed with the pET-15b(+)-AmbP1 vector. Cells were grown using auto-induction medium, initially at 37^{o}C until the cell density reaches an OD_{600} value of ~1.50 and a further ~20 h at 15^{o}C for protein expression. ^{15}N-isotopic labeling of the protein for NMR studies was carried out by growth in a modified auto induction minimal medium, containing $^{15}NH_4Cl$ as sole nitrogen source and succinate for pH balance. The protein was isolated from the soluble fraction of *E. coli* after opening the cells in a microfluidizer, followed by removal of cell debris by centrifugation. Purification involved an affinity chromatography on a Ni^{2+}(HP) column (GE Healthcare) in 40 mM Tris.HCl (pH 8.0), 40 mM NaCl, 1 mM NaN_3, and 2.88 mM sodium lauroyil sarcosinate, using a linear gradient of imidazole (20-500 mM) for elution, followed by gel filtration on Superdex75 (GE Healthcare) in the same binding/washing buffer for the Ni^{2+}(HP) column. Purified protein fractions were collected and concentrated to ~10 mg/ml using centriprep devices (Millipore).

Mg^{2+} binding studies by NMR Spectroscopy: Spectra were recorded on a Bruker AVANCE900 spectrometer, equipped with a z-axis gradient, triple-resonance cryoprobe. Binding of Mg^{2+} to Orf2-1 was investigated at 25°C by ^{1}H-^{15}N 2D HSQC spectroscopy, using a 0.20 mM ^{15}N-labeled Orf2-1 sample in in 40 mM Tris-HCl buffer (pH 8.0), 40 mM NaCl, 1 mM NaN_3, 2.88 mM sodium lauroyil sarcosinate, and 90/10% H_2O/D_2O. Two 2D ^{1}H-^{15}N HSQC spectra were recorded: one in the absence of Mg^{2+} and the other in the presence of a 10 fold molar excess of Mg^{2+}. Both spectra were processed with NMRPipe[2] and analyzed using NMRView[3].

Fig. S1. Protein sequence alignment of AmbP1, FidP1 and WelP1.

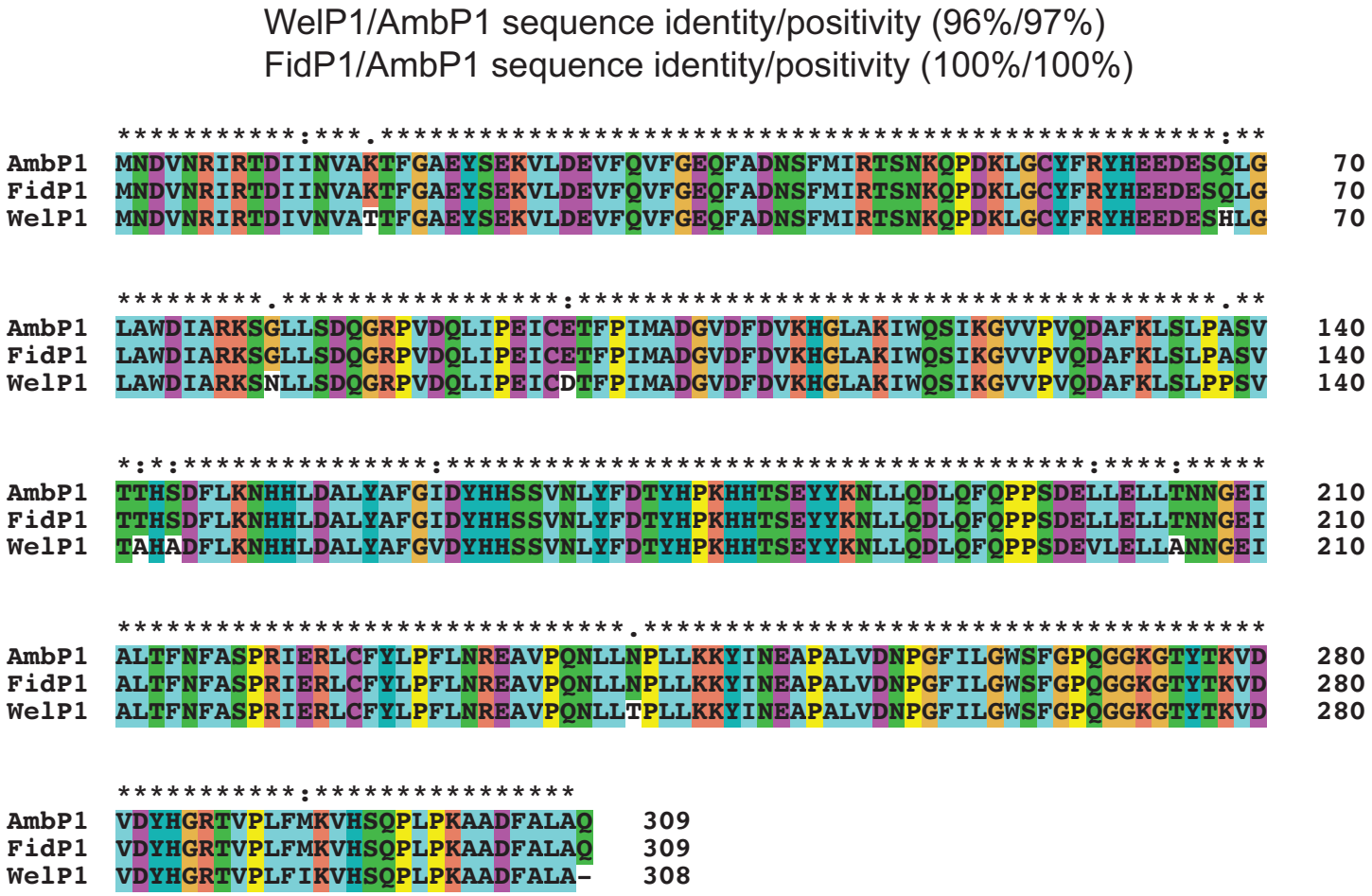


Fig. S2. UV absorption spectra of selected tri- or tetracyclic hapalindoles and fischerindoles.

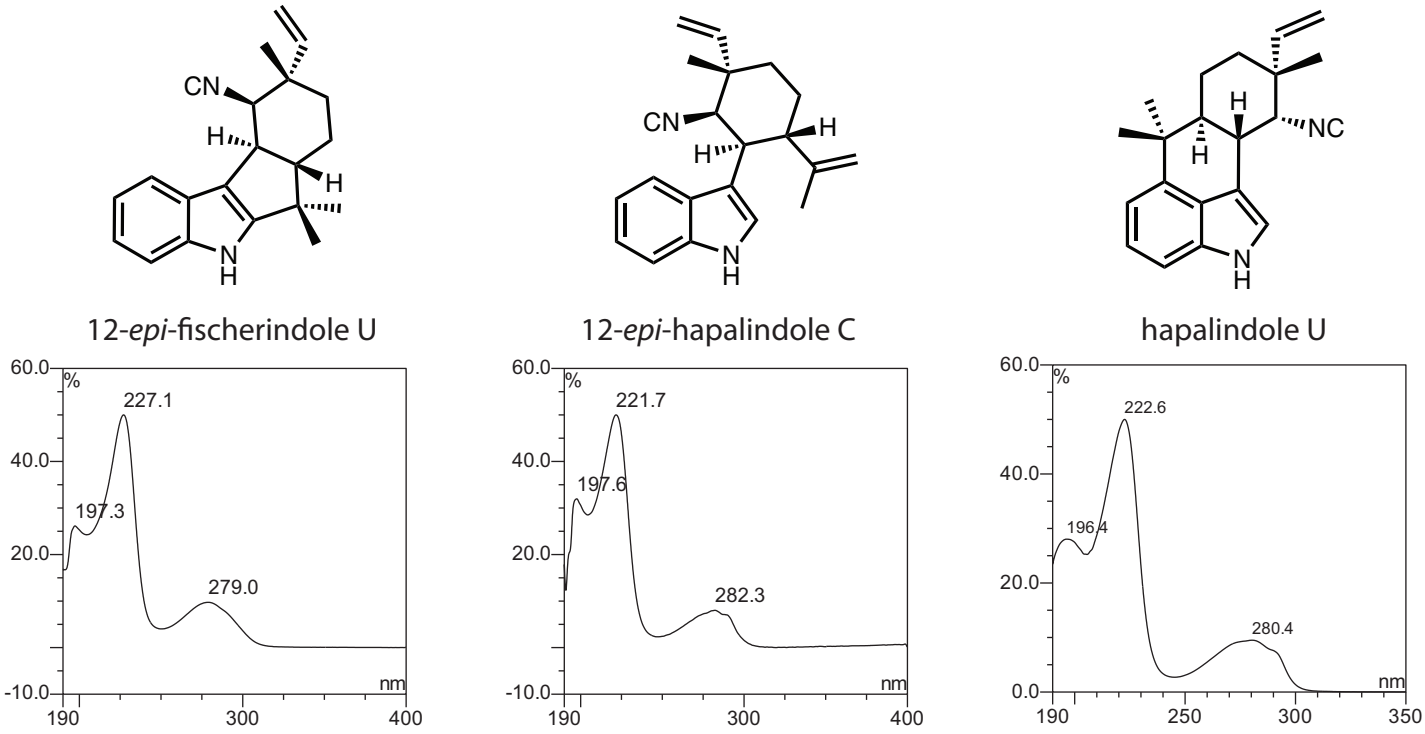


Fig. S3. Evaluation of the roles of varied pH and metal cofactors in AmbP1-mediated conversion of (*Z*)-**1** to **2** and **3**.

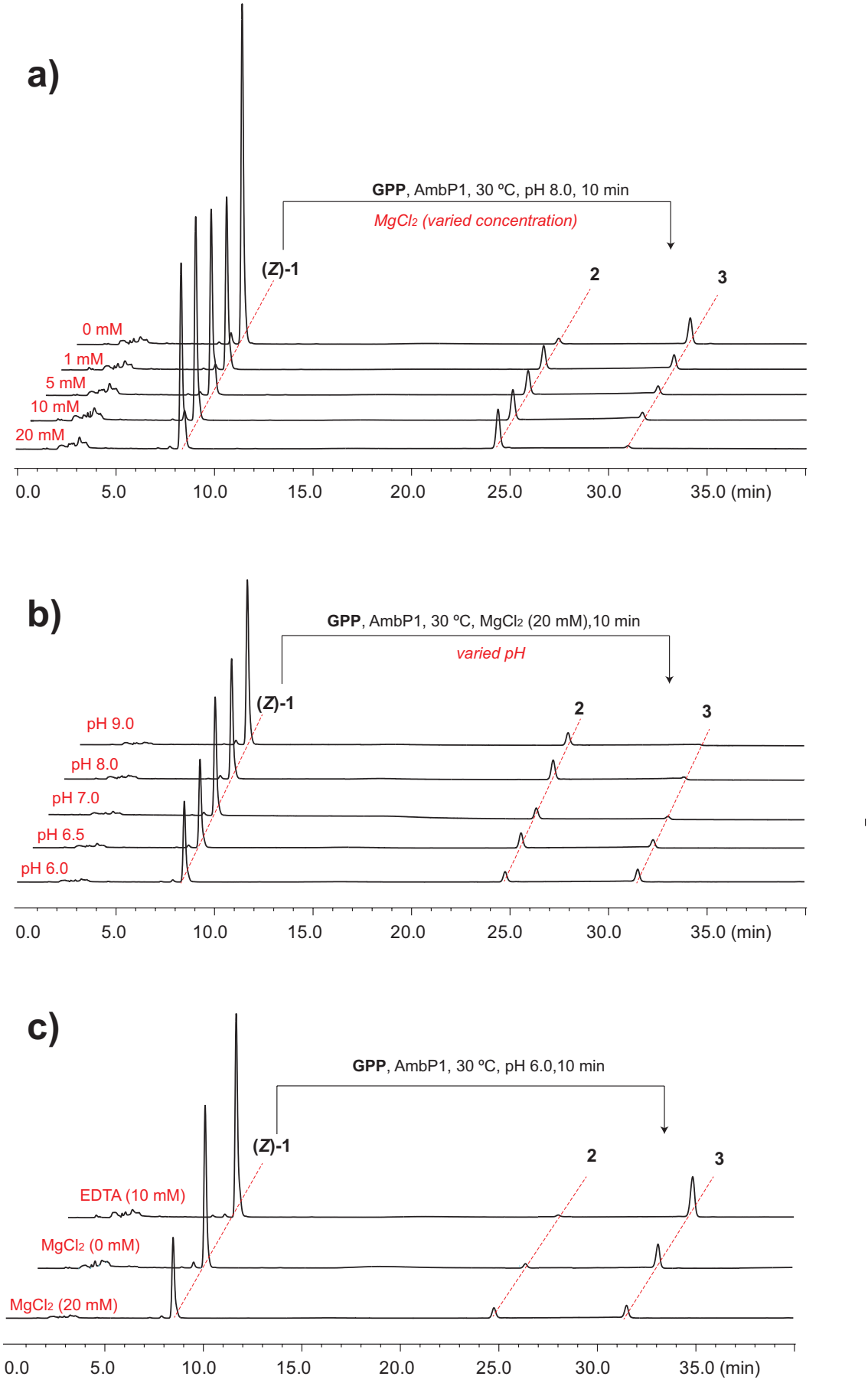
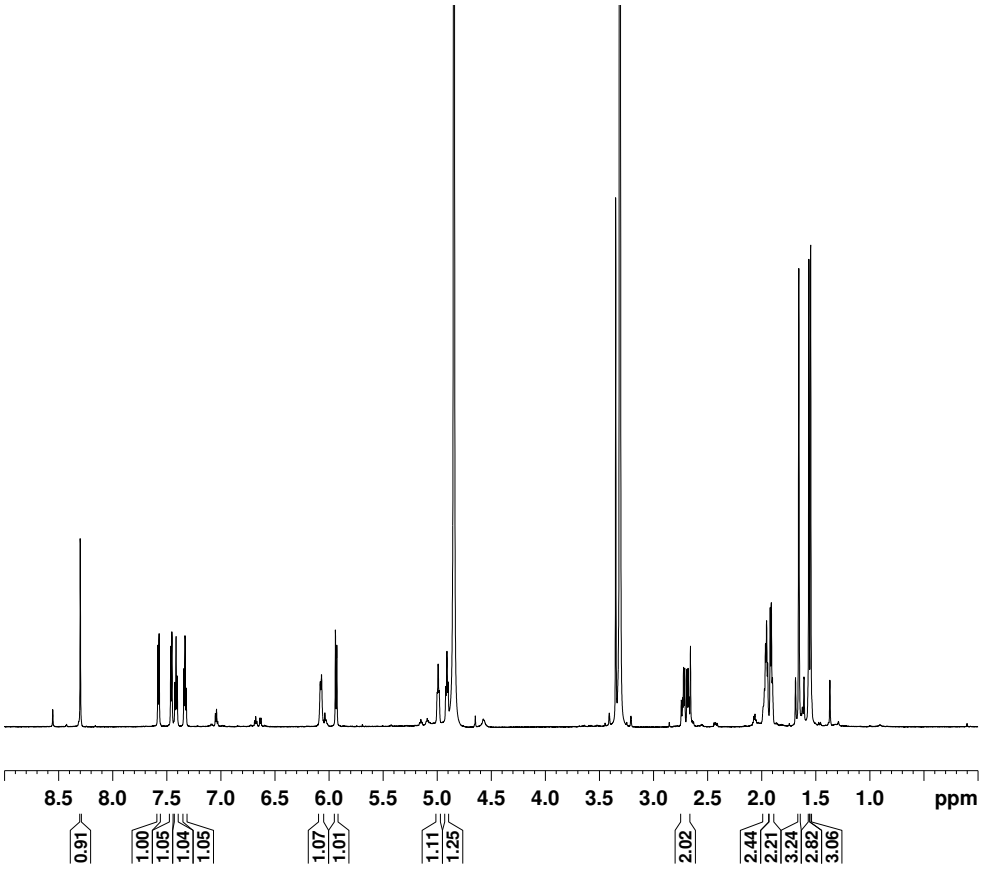


Fig. S4. ^{1}H NMR (CD_3OD, 700 MHz) of compound **2**



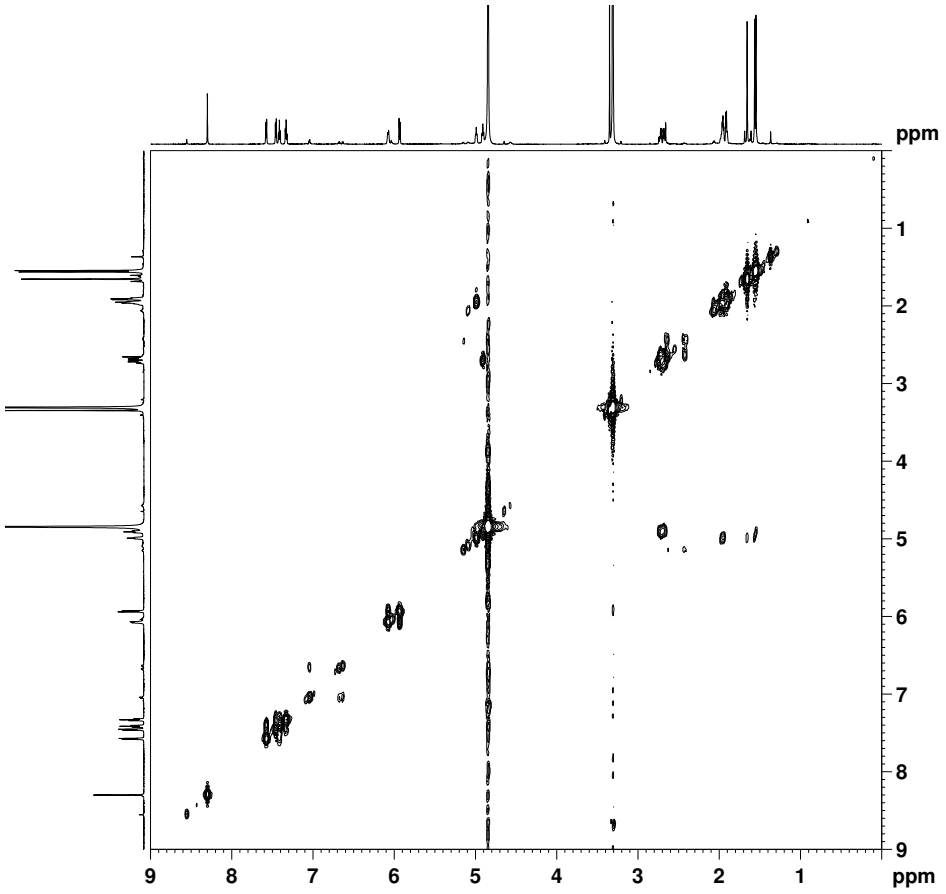
Current Data Parameters
NAME 082912_AmbP1_3–geranyl–indole isonitrile
EXPNO 1
PROCNO 1

F2 – Acquisition Parameters
Date_ 20120829
Time 20.55
INSTRUM spect
PROBHD 1.7 mm PATXI 1
PULPROG zg30
TD 65536
SOLVENT MeOD
NS 2048
DS 2
SWH 14423.077 Hz
FIDRES 0.220079 Hz
AQ 2.2719147 sec
RG 1290
DW 34.667 usec
DE 6.50 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1

======== CHANNEL f1 ========
NUC1 1H
P1 7.73 usec
PL1 7.50 dB
PL1W 4.03355265 W
SFO1 700.0543231 MHz

F2 – Processing parameters
SI 32768
SF 700.0500140 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

Fig. S5. ^{1}H-^{1}H COSY NMR (CD_3OD, 700 MHz) of compound **2**



Current Data Parameters
NAME 082912_AmbP1_3-geranyl-indole isonitrile
EXPNO 2
PROCNO 1

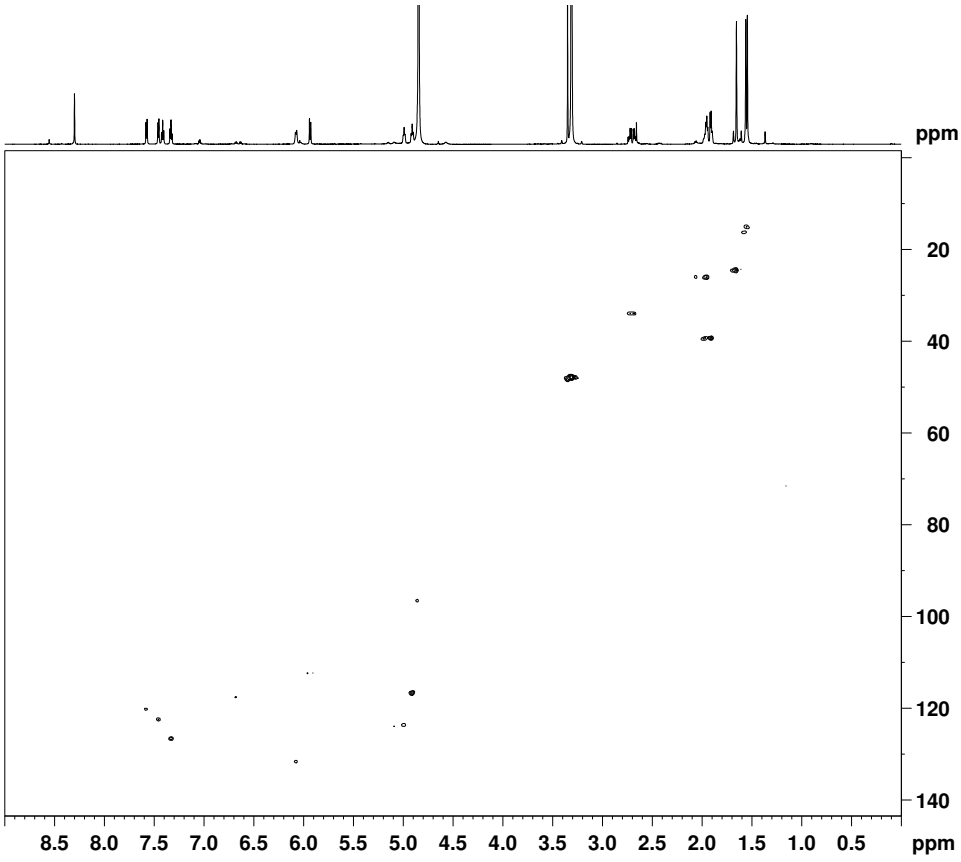
F2 - Acquisition Parameters
Date_ 20120829
Time 20.57
INSTRUM spect
PROBHD 1.7 mm PATXI 1
PULPROG cosygpqf
TD 2048
SOLVENT MeOD
NS 64
DS 8
SWH 9328.358 Hz
FIDRES 4.554862 Hz
AQ 0.1097728 sec
RG 64
DW 53.600 usec
DE 6.50 usec
TE 298.0 K
d0 0.00000300 sec
D1 1.48689198 sec
d13 0.00000400 sec
D16 0.00020000 sec
in0 0 sec
ST1CNT 128
d0orig 0.00000300 sec
ph1loop 0
t1loop 0
SFO1 700.0542088 MHz
NUC1 1H
P0 7.73 usec
P1 7.73 usec
PLW1 -1.00000000 W
GPNAM[1] SINE.100
GPZ1 10.00 %
P16 1000.00 usec

F1 - Acquisition parameters
TD 128
SFO1 700.0542 MHz
FIDRES 73.015007 Hz
SW 13.350 ppm
FnMODE QF

F2 - Processing parameters
SI 1024
SF 700.0500157 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0
PC 1.40

F1 - Processing parameters
SI 1024
MC2 QF
SF 700.0500110 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0

Fig. S6. ^{1}H-^{13}C HSQC NMR (CD_3OD, 700 MHz) of compound **2**



BRUKER

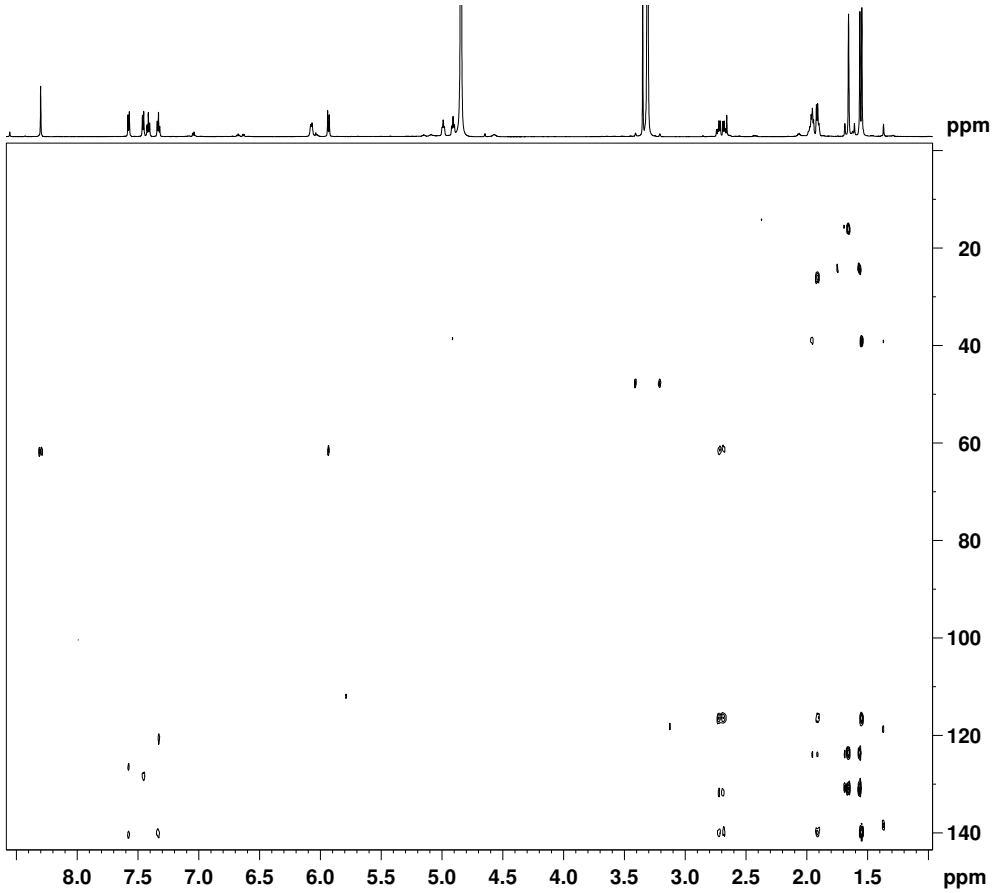
Current Data Parameters
NAME 083012_AmbP1_3–geranyl–indole isonitrile
EXPNO 1
PROCNO 1

F2 – Acquisition Parameters
Date_ 20120830
Time 17.22
INSTRUM spect
PROBHD 1.7 mm PATXI 1
PULPROG hsqcetgpsi2
TD 1024
SOLVENT MeOD
NS 64
DS 16
SWH 9328.358 Hz
FIDRES 9.109725 Hz
AQ 0.0548864 sec
RG 2050
DW 53.600 usec
DE 6.50 usec
TE 298.0 K
CNST2 145.0000000
d0 0.00000300 sec
D1 1.50000000 sec
d4 0.00172414 sec
d11 0.03000000 sec
D16 0.00020000 sec
D24 0.00086207 sec
DELTA 0.00127146 sec
DELTA1 0.00120847 sec
DELTA2 0.00006207 sec
DELTA3 0.00052414 sec
in0 0 sec
ST1CNT 128
ZGOPTNS
d0orig 0.00000300 sec
ph1loop 0
t1loop 0
SFO1 700.0542088 MHz
NUC1 1H
P1 7.73 usec
p2 15.46 usec
P28 1000.00 usec
PLW1 −1.00000000 W
SFO2 176.0409095 MHz
NUC2 13C
CPDPRG[2 garp
P3 8.65 usec
p4 17.30 usec
PCPD2 100.00 usec
PLW2 −1.00000000 W
PLW12 −1.00000000 W
GPNAM[1] SINE.100
GPNAM[2] SINE.100
GPNAM[3] SINE.100
GPNAM[4] SINE.100
GPZ1 80.00 %
GPZ2 20.10 %
GPZ3 11.00 %
GPZ4 −5.00 %
P16 1000.00 usec
P19 600.00 usec

F1 – Acquisition parameters
TD 256
SFO1 176.0409 MHz
FIDRES 113.910851 Hz
SW 165.650 ppm
FnMODE Echo–Antiecho

F2 – Processing parameters
SI 1024
SF 700.0500145 MHz
WDW QSINE
SSB 2

Fig. S7. ^{1}H-^{13}C HMBC NMR (CD_3OD, 700 MHz) of compound **2**



BRUKER

Current Data Parameters
NAME 082912_AmbP1_3-geranyl-indole isonitrile
EXPNO 3
PROCNO 1

F2 – Acquisition Parameters
Date_ 20120830
Time 0.43
INSTRUM spect
PROBHD 1.7 mm PATXI 1
PULPROG hmbcgplpndqf
TD 4096
SOLVENT MeOD
NS 256
DS 16
SWH 9090.909 Hz
FIDRES 2.219460 Hz
AQ 0.2252800 sec
RG 2050
DW 55.000 usec
DE 6.50 usec
TE 298.0 K
CNST2 145.0000000
CNST13 10.0000000
d0 0.00000300 sec
D1 1.50000000 sec
d2 0.00344828 sec
d6 0.05000000 sec
D16 0.00020000 sec
in0 0 sec
ST1CNT 128
d0orig 0.00000300 sec
ph1loop 0
t1loop 0
SFO1 700.0544103 MHz
NUC1 1H
P1 7.73 usec
p2 15.46 usec
PLW1 −1.00000000 W
SFO2 176.0452859 MHz
NUC2 13C
P3 8.65 usec
PLW2 −1.00000000 W
GPNAM[1] SINE.100
GPNAM[2] SINE.100
GPNAM[3] SINE.100
GPZ1 50.00 %
GPZ2 30.00 %
GPZ3 40.10 %
P16 1000.00 usec

F1 – Acquisition parameters
TD 128
SFO1 176.0453 MHz
FIDRES 305.459290 Hz
SW 222.095 ppm
FnMODE QF

F2 – Processing parameters
SI 2048
SF 700.0500116 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0
PC 1.40

F1 – Processing parameters
SI 1024
MC2 QF
SF 176.0277130 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0

Fig. S8. ^{1}H NMR ($CDCl_3$, 700 MHz) of compound **3**

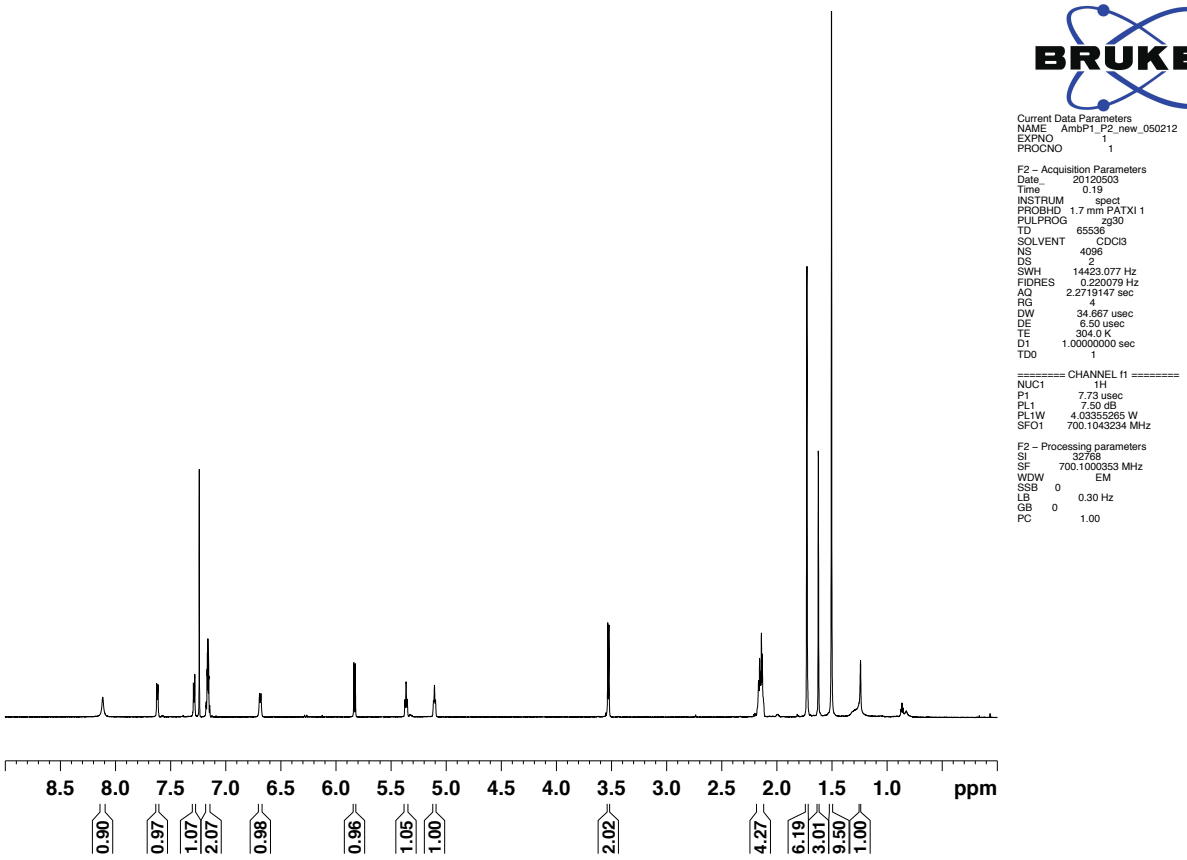
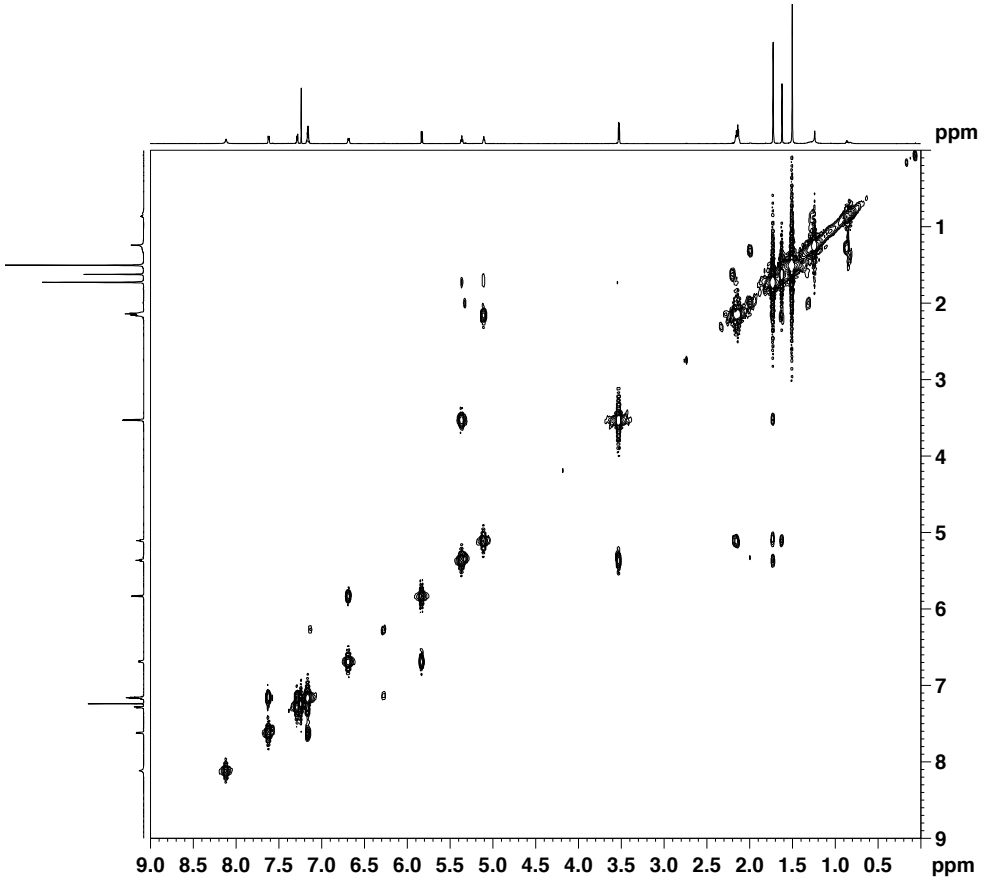


Fig. S9. ^{1}H-^{1}H COSY NMR ($CDCl_3$, 700 MHz) of compound **3**



BRUKER

Current Data Parameters
NAME AmbP1_P2_new_050212
EXPNO 2
PROCNO 1

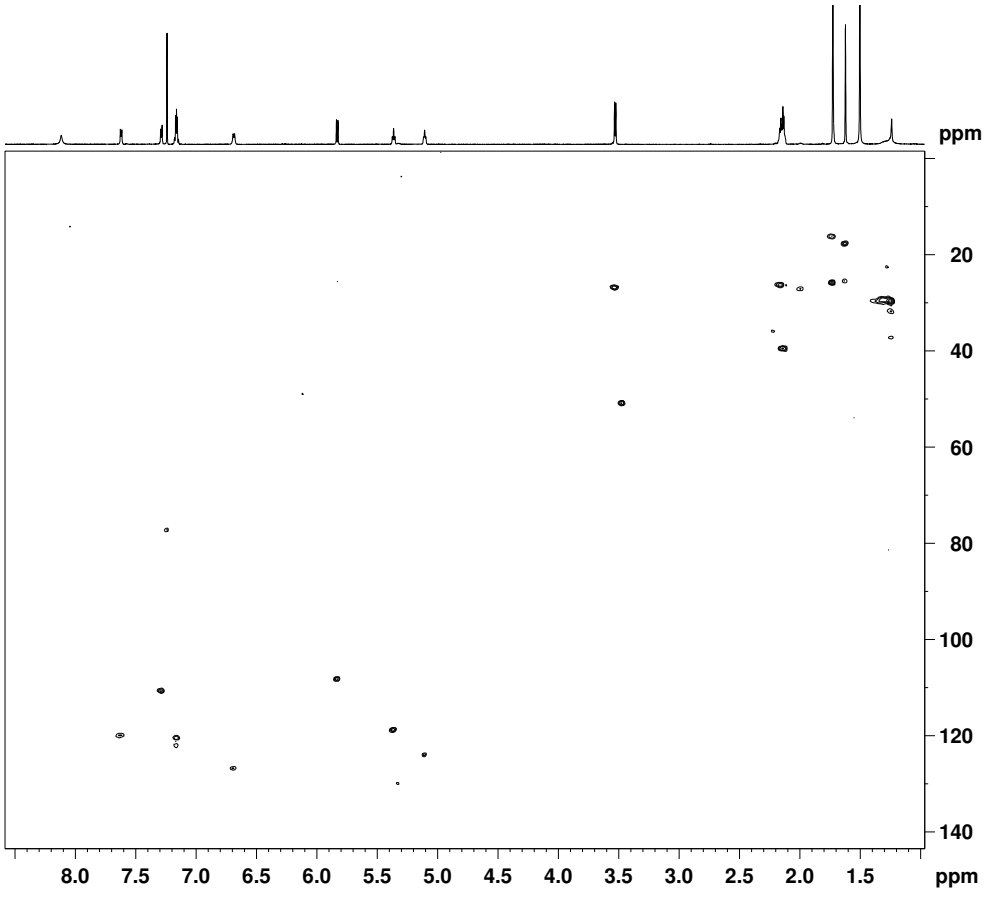
F2 – Acquisition Parameters
Date_ 20120503
Time 0.26
INSTRUM spect
PROBHD 1.7 mm PATXI 1
PULPROG cosygpqf
TD 2048
SOLVENT CDCl3
NS 256
DS 8
SWH 9328.358 Hz
FIDRES 4.554862 Hz
AQ 0.1097728 sec
RG 64
DW 53.600 usec
DE 6.50 usec
TE 304.0 K
d0 0.00000300 sec
D1 1.48689198 sec
d13 0.00000400 sec
D16 0.00020000 sec
in0 0 sec
ST1CNT 128
d0orig 0.00000300 sec
ph1loop 0
t1loop 0
SFO1 700.1042091 MHz
NUC1 1H
P0 7.73 usec
P1 7.73 usec
PLW1 −1.00000000 W
GPNAM[1] SINE.100
GPZ1 10.00 %
P16 1000.00 usec

F1 – Acquisition parameters
TD 128
SFO1 700.1042 MHz
FIDRES 73.020226 Hz
SW 13.350 ppm
FnMODE QF

F2 – Processing parameters
SI 1024
SF 700.1000304 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0
PC 1.40

F1 – Processing parameters
SI 1024
MC2 QF
SF 700.1000293 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0

Fig. S910. ^{1}H-^{13}C HSQC NMR ($CDCl_3$, 700 MHz) of compound **3**



BRUKER

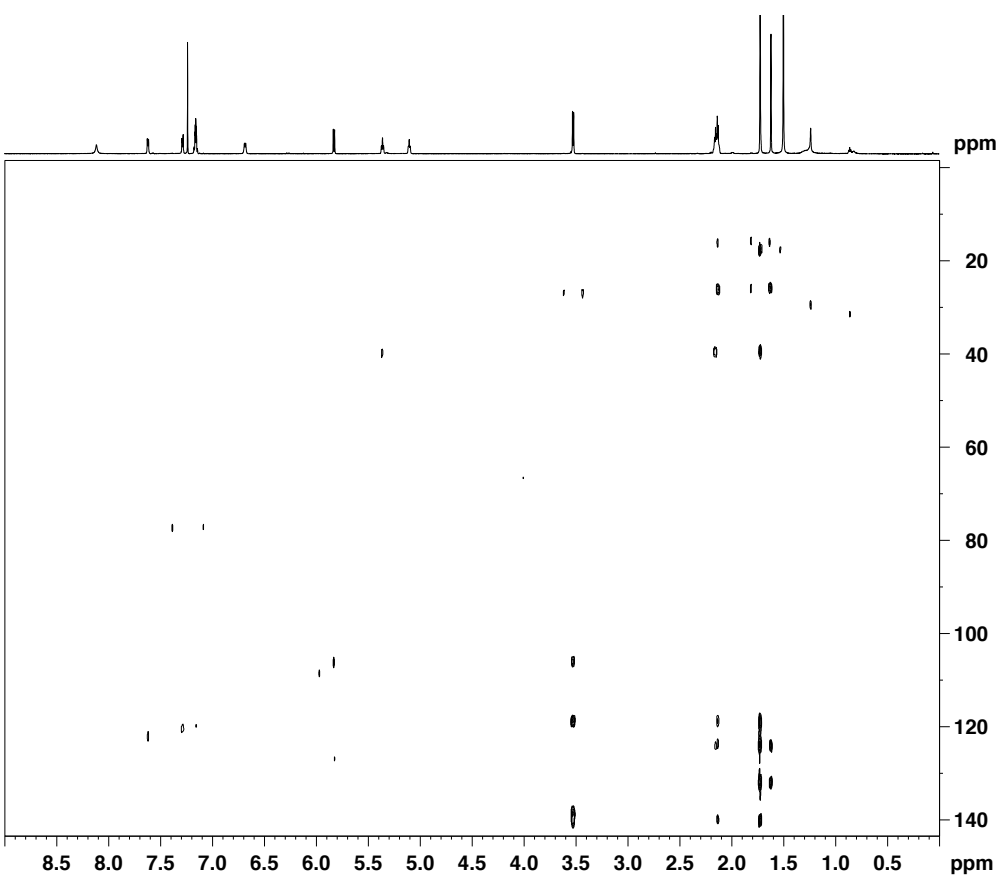
Current Data Parameters
NAME 050612_AmbP1-C2-geranyl-isonitrile-indole
EXPNO 2
PROCNO 1

F2 - Acquisition Parameters
Date_ 20120507
Time 0.01
INSTRUM spect
PROBHD 1.7 mm PATXI 1
PULPROG hsqcetgpsi2
TD 1024
SOLVENT CDCl3
NS 64
DS 16
SWH 9328.358 Hz
FIDRES 9.109725 Hz
AQ 0.0548864 sec
RG 2050
DW 53.600 usec
DE 6.50 usec
TE 304.0 K
CNST2 145.0000000
d0 0.00000300 sec
D1 1.50000000 sec
d4 0.00172414 sec
d11 0.03000000 sec
D16 0.00020000 sec
D24 0.00086207 sec
DELTA 0.00127146 sec
DELTA1 0.00120847 sec
DELTA2 0.00006207 sec
DELTA3 0.00052414 sec
in0 0 sec
ST1CNT 128
ZGOPTNS
d0orig 0.00000300 sec
ph1loop 0
t1loop 0
SFO1 700.1042091 MHz
NUC1 1H
P1 7.73 usec
p2 15.46 usec
P28 1000.00 usec
PLW1 -1.00000000 W
SFO2 176.0534829 MHz
NUC2 13C
CPDPRG[2 garp
P3 8.65 usec
p4 17.30 usec
PCPD2 100.00 usec
PLW2 -1.00000000 W
PLW12 -1.00000000 W
GPNAM[1] SINE.100
GPNAM[2] SINE.100
GPNAM[3] SINE.100
GPNAM[4] SINE.100
GPZ1 80.00 %
GPZ2 20.10 %
GPZ3 11.00 %
GPZ4 -5.00 %
P16 1000.00 usec
P19 600.00 usec

F1 - Acquisition parameters
TD 256
SFO1 176.0535 MHz
FIDRES 113.918991 Hz
SW 165.650 ppm
FnMODE Echo-Antiecho

F2 - Processing parameters
SI 1024
SF 700.1000318 MHz
WDW QSINE
SSB 2

Fig. S11. ^{1}H-^{13}C HMBC NMR ($CDCl_3$, 700 MHz) of compound **3**



BRUKER

Current Data Parameters
NAME AmbP1_P2_new_050212
EXPNO 3
PROCNO 1

F2 – Acquisition Parameters
Date_ 20120503
Time 15.06
INSTRUM spect
PROBHD 1.7 mm PATXI 1
PULPROG hmbcgplpndqf
TD 4096
SOLVENT CDCl3
NS 256
DS 16
SWH 9090.909 Hz
FIDRES 2.219460 Hz
AQ 0.2252800 sec
RG 2050
DW 55.000 usec
DE 6.50 usec
TE 304.0 K
CNST2 145.0000000
CNST13 10.0000000
d0 0.00000300 sec
D1 1.50000000 sec
d2 0.00344828 sec
d6 0.05000000 sec
D16 0.00020000 sec
in0 0 sec
ST1CNT 128
d0orig 0.00000300 sec
ph1loop 0
t1loop 0
SFO1 700.1044106 MHz
NUC1 1H
P1 7.73 usec
p2 15.46 usec
PLW1 −1.00000000 W
SFO2 176.0578596 MHz
NUC2 13C
P3 8.65 usec
PLW2 −1.00000000 W
GPNAM[1] SINE.100
GPNAM[2] SINE.100
GPNAM[3] SINE.100
GPZ1 50.00 %
GPZ2 30.00 %
GPZ3 40.10 %
P16 1000.00 usec

F1 – Acquisition parameters
TD 128
SFO1 176.0579 MHz
FIDRES 305.481110 Hz
SW 222.095 ppm
FnMODE QF

F2 – Processing parameters
SI 2048
SF 700.1000344 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0
PC 1.40

F1 – Processing parameters
SI 1024
MC2 QF
SF 176.0402850 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0

Fig. S12. Nonenzymatic conversion of **2** to **3** at pH 6.5: a time course analysis.

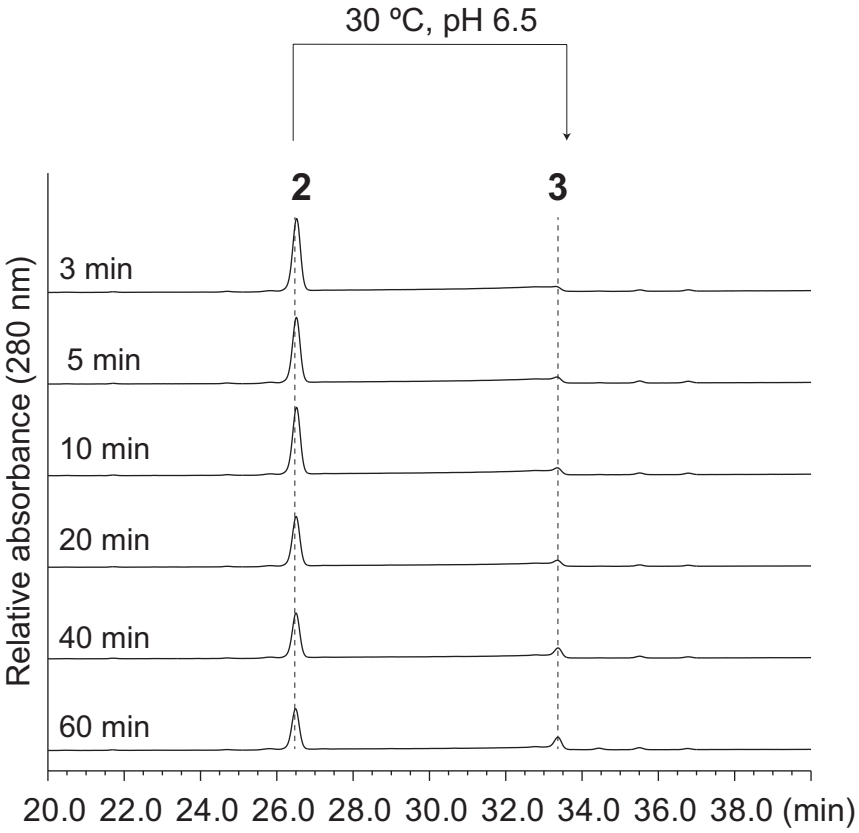


Fig. S13. UV absorption spectra of **2a** and **3a**

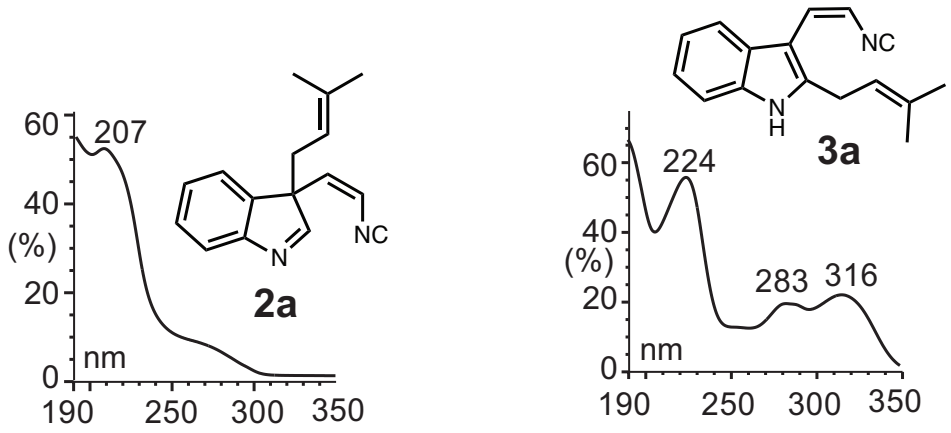


Fig. S14. HPLC chromatographs showing the enzymatic reaction outcomes of **(Z)-1** in combination of NPP or FPP in the presence of AmbP1 and $MgCl_2$.

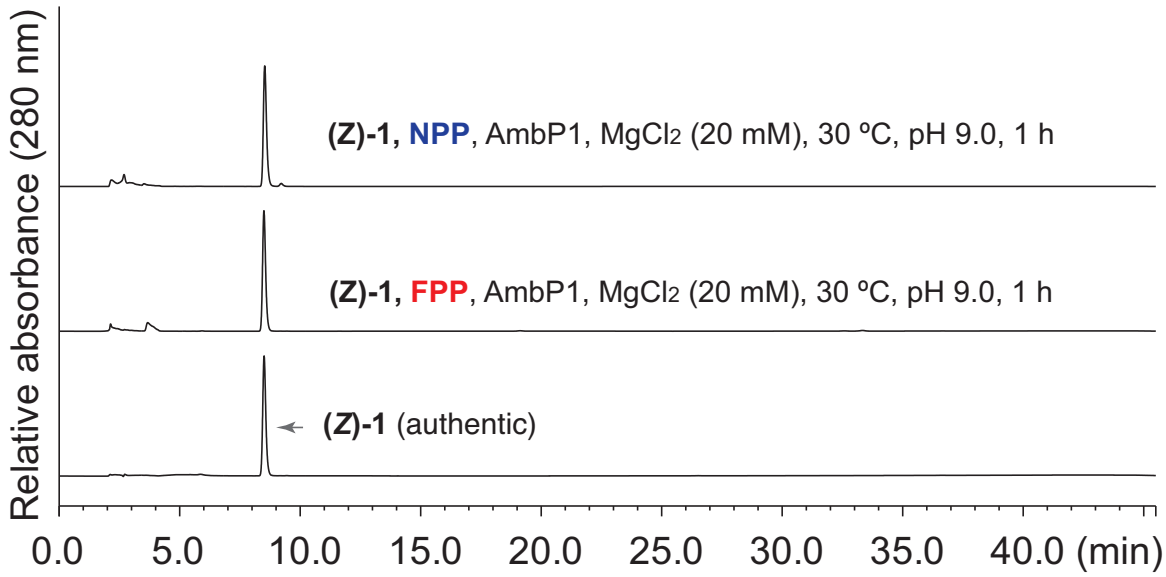


Fig. S15. ^{1}H NMR ($CDCl_3$, 700 MHz) of compound **6**

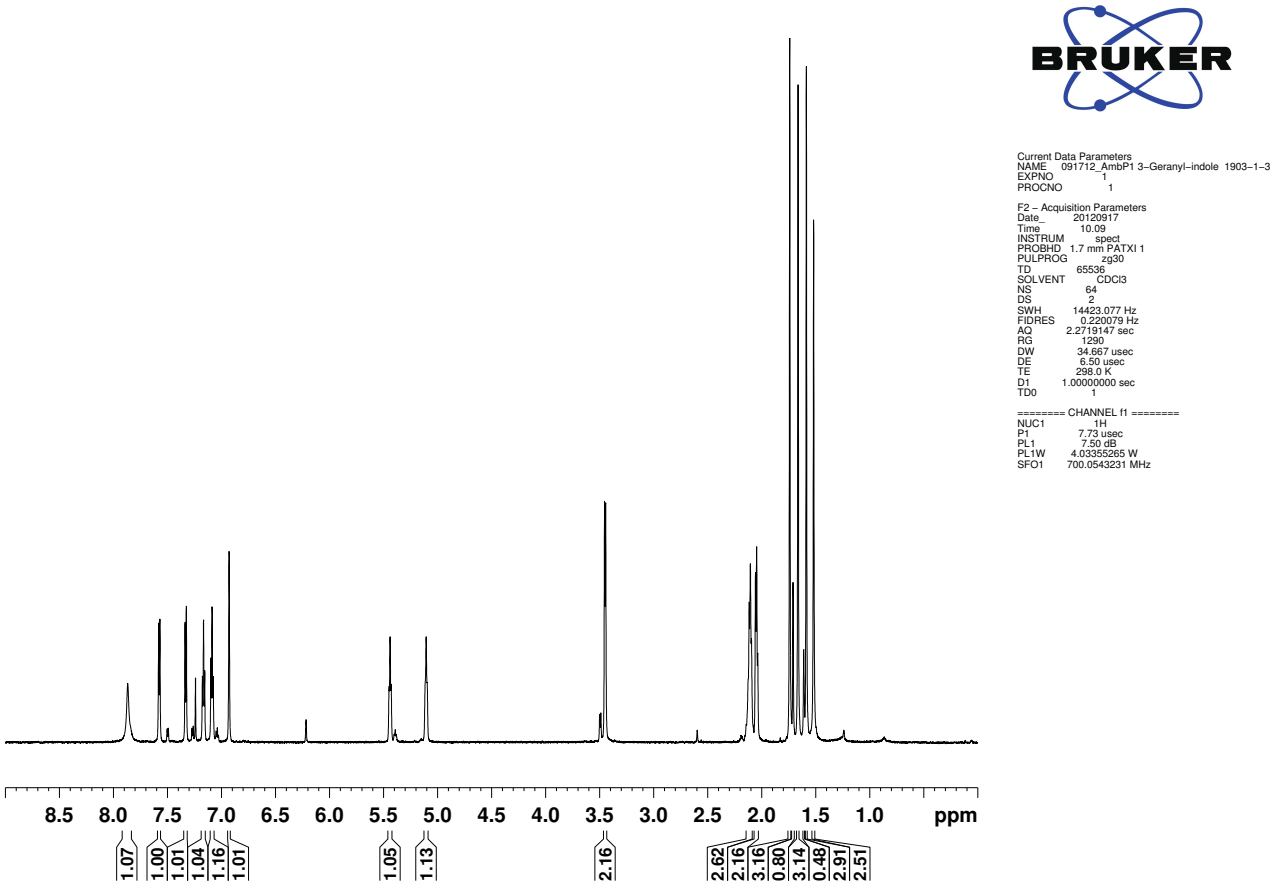
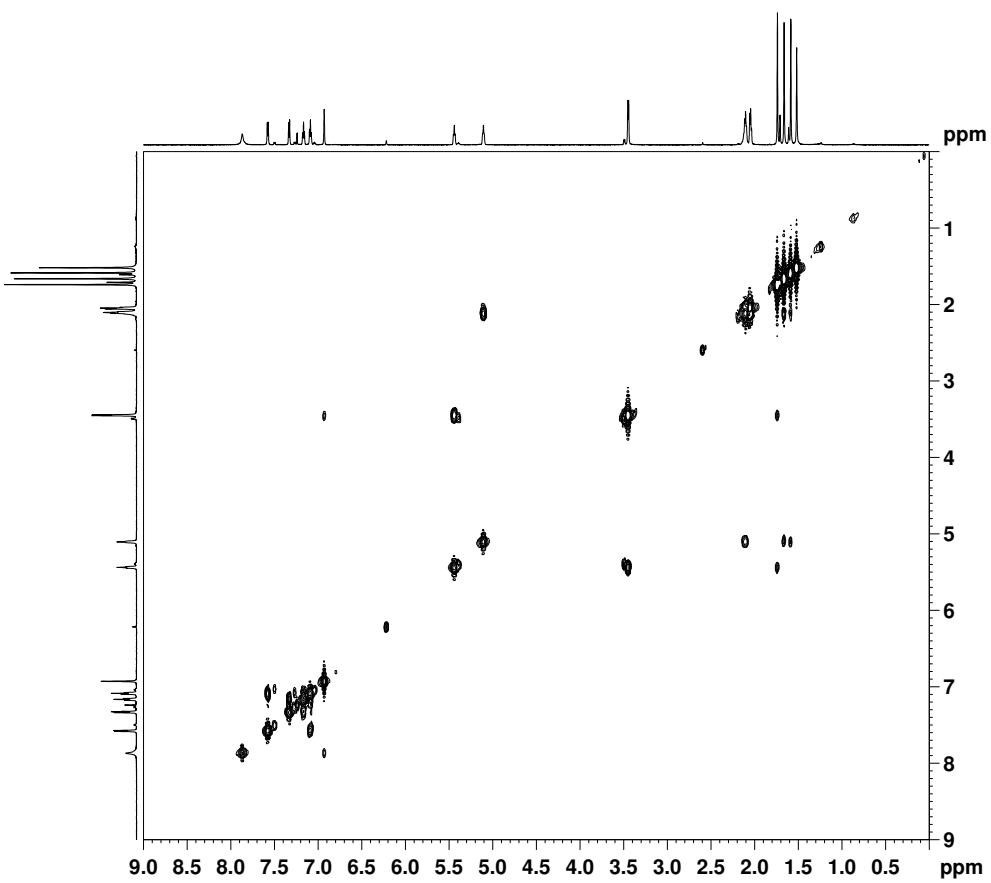


Fig. S16. ^{1}H-^{1}H COSY NMR ($CDCl_3$, 700 MHz) of compound **6**



Current Data Parameters
NAME 091712_AmbP1 3-Geranyl-indole 1903-1-3_36m
EXPNO 2
PROCNO 1

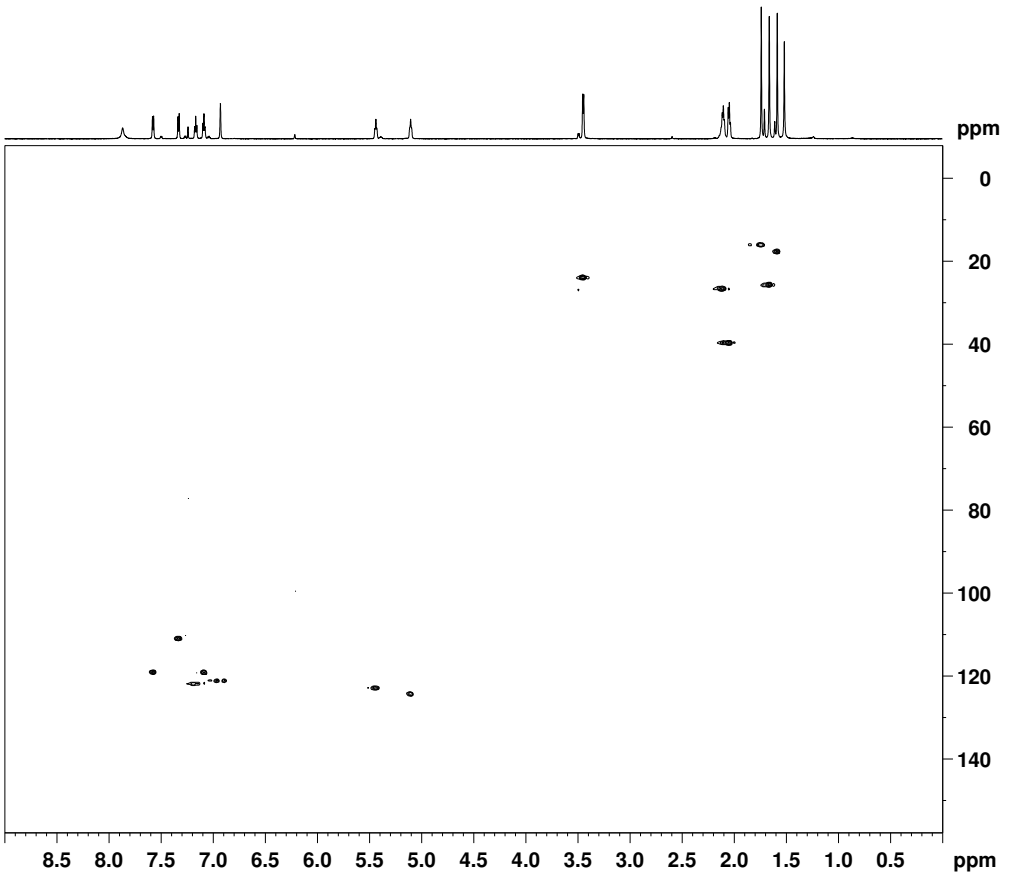
F2 - Acquisition Parameters
Date_ 20120917
Time 10.22
INSTRUM spect
PROBHD 1.7 mm PATXI 1
PULPROG cosygpqf
TD 2048
SOLVENT CDCl3
NS 16
DS 8
SWH 9328.358 Hz
FIDRES 4.554862 Hz
AQ 0.1097728 sec
RG 64
DW 53.600 usec
DE 6.50 usec
TE 298.0 K
d0 0.00000300 sec
D1 1.48689198 sec
d13 0.00000400 sec
D16 0.00020000 sec
in0 0 sec
ST1CNT 128
d0orig 0.00000300 sec
ph1loop 0
t1loop 0
SFO1 700.0542088 MHz
NUC1 1H
P0 7.73 usec
P1 7.73 usec
PLW1 -1.00000000 W
GPNAM[1] SINE.100
GPZ1 10.00 %
P16 1000.00 usec

F1 - Acquisition parameters
TD 128
SFO1 700.0542 MHz
FIDRES 73.015007 Hz
SW 13.350 ppm
FnMODE QF

F2 - Processing parameters
SI 1024
SF 700.0500264 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0
PC 1.40

F1 - Processing parameters
SI 1024
MC2 QF
SF 700.0500260 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0

Fig. S17. ^{1}H-^{13}C HSQC NMR ($CDCl_3$, 700 MHz) of compound **6**



BRUKER

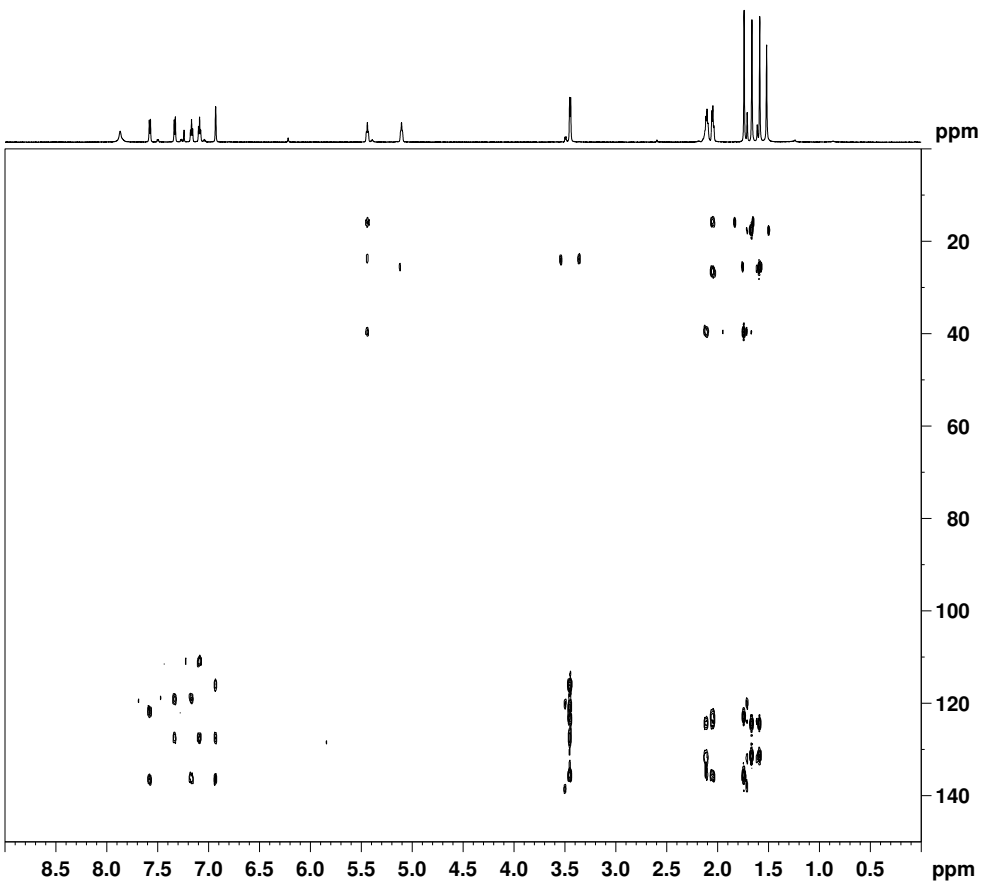
Current Data Parameters
NAME 091712_AmbP1 3-Geranyl-indole 1903-1-
EXPNO 3
PROCNO 1

F2 - Acquisition Parameters
Date_ 20120917
Time 16.52
INSTRUM spect
PROBHD 1.7 mm PATXI 1
PULPROG hsqcetgpsi2
TD 1024
SOLVENT CDCl3
NS 64
DS 16
SWH 9328.358 Hz
FIDRES 9.109725 Hz
AQ 0.0548864 sec
RG 2050
DW 53.600 usec
DE 6.50 usec
TE 298.0 K
CNST2 145.0000000
d0 0.00000300 sec
D1 1.50000000 sec
d4 0.00172414 sec
d11 0.03000000 sec
D16 0.00020000 sec
D24 0.00086207 sec
DELTA 0.00127146 sec
DELTA1 0.00120847 sec
DELTA2 0.00006207 sec
DELTA3 0.00052414 sec
in0 0 sec
ST1CNT 128
ZGOPTNS
d0orig 0.00000300 sec
ph1loop 0
t1loop 0
SFO1 700.0542088 MHz
NUC1 1H
P1 7.73 usec
p2 15.46 usec
P28 1000.00 usec
PLW1 -1.00000000 W
SFO2 176.0409095 MHz
NUC2 13C
CPDPRG[2 garp
P3 8.65 usec
p4 17.30 usec
PCPD2 100.00 usec
PLW2 -1.00000000 W
PLW12 -1.00000000 W
GPNAM[1] SINE.100
GPNAM[2] SINE.100
GPNAM[3] SINE.100
GPNAM[4] SINE.100
GPZ1 80.00 %
GPZ2 20.10 %
GPZ3 11.00 %
GPZ4 -5.00 %
P16 1000.00 usec
P19 600.00 usec

F1 - Acquisition parameters
TD 256
SFO1 176.0409 MHz
FIDRES 113.910851 Hz
SW 165.650 ppm
FnMODE Echo-Antiecho

F2 - Processing parameters
SI 1024
SF 700.0500280 MHz
WDW QSINE
SSB 2

Fig. S18. ^{1}H-^{13}C HMBC NMR ($CDCl_3$, 700 MHz) of compound **6**



BRUKER

Current Data Parameters
NAME 091712_AmbP1 3-Geranyl-indole 1903-1-
EXPNO 4
PROCNO 1

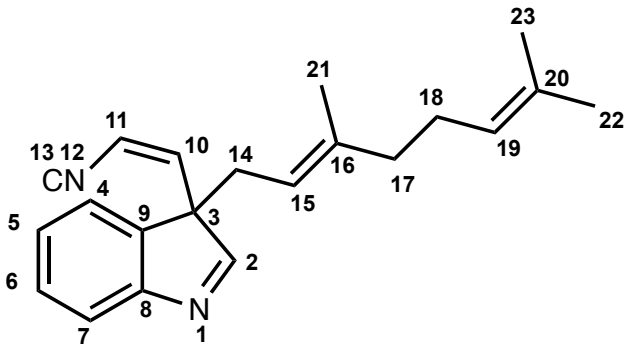
F2 - Acquisition Parameters
Date_ 20120918
Time 0.05
INSTRUM spect
PROBHD 1.7 mm PATXI 1
PULPROG hmbcgplpndqf
TD 4096
SOLVENT CDCl3
NS 128
DS 16
SWH 9090.909 Hz
FIDRES 2.219460 Hz
AQ 0.2252800 sec
RG 2050
DW 55.000 usec
DE 6.50 usec
TE 298.0 K
CNST2 145.0000000
CNST13 10.0000000
d0 0.00000300 sec
D1 1.50000000 sec
d2 0.00344828 sec
d6 0.05000000 sec
D16 0.00020000 sec
in0 0 sec
ST1CNT 128
d0orig 0.00000300 sec
ph1loop 0
t1loop 0
SFO1 700.0544103 MHz
NUC1 1H
P1 7.73 usec
p2 15.46 usec
PLW1 -1.00000000 W
SFO2 176.0452859 MHz
NUC2 13C
P3 8.65 usec
PLW2 -1.00000000 W
GPNAM[1] SINE.100
GPNAM[2] SINE.100
GPNAM[3] SINE.100
GPZ1 50.00 %
GPZ2 30.00 %
GPZ3 40.10 %
P16 1000.00 usec

F1 - Acquisition parameters
TD 128
SFO1 176.0453 MHz
FIDRES 305.459290 Hz
SW 222.095 ppm
FnMODE QF

F2 - Processing parameters
SI 2048
SF 700.0500275 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0
PC 1.40

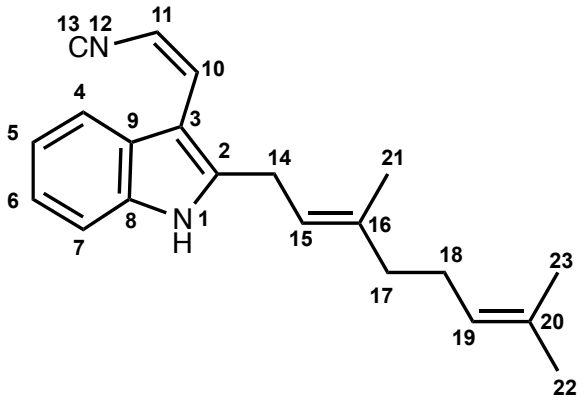
F1 - Processing parameters
SI 1024
MC2 QF
SF 176.0277130 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0

Table S1. Compound **2** NMR peak assignments



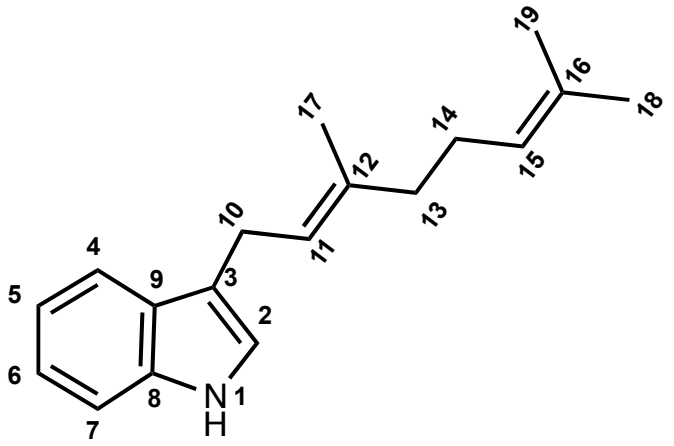
C,N	δ_C (ppm)	δ_H (ppm), multiplicity	*J* (Hz)
N1			
2	174.3	8.299(s)	
3	62.8		
4	121.5	7.577(d)	7.7
5	127.9	7.330(t)	7.6
6	121.6	7.414(t)	7.6
7	123.3	7.456(d)	7.7
8	155.4		
9	141.2		
10	132.8	6.075(d)	9.0
11	113.7	5.932(d)	9.0
N12			
13			
14	35.2	2.700(d), 2H	7.4
15	117.3	4.908(t)	7.3
16	141.7		
17	40.5	1.907(t), 1.975(t)	7.3
18	27.2	1.959(m), 2H	
19	124.9	4.989(t)	6.4
20	132.3		
21	16.3	1.546(s)	
22	25.7	1.657(s)	
23	17.5	1.562(s)	

Table S2. Compound **3** NMR peak assignments



C,N	δ_C (ppm)	δ_H (ppm), multiplicity	*J* (Hz)
N1		8.11(s)	
2	138.7		
3	105.9		
4		7.620(m)	
5	122.0	7.161(m)	
6	120.3	7.154(m)	
7	110.4	7.289(m)	
8	150.6		
9	126.5		
10	126.4	5.830(d)	8.9
11	108.2	6.687(d)	8.6
N12			
13			
14	27.1	3.528(d), 2H	7.4
15	118.7	5.364(t)	7.5
16	140.1		
17	39.3	2.132(m), 2H	
18	26.1	2.151(m), 2H	
19	124.0	5.107(t)	6.8
20	131.9		
21	16.2	1.728(s)	
22	25.7	1.727(s)	
23	17.7	1.625(s)	

Table S3. Compound **6** NMR peak assignments



C,N	δ_C (ppm)	δ_H (ppm), multiplicity	*J* (Hz)
N1		7.865(s)	
2	120.8		
3	116.1		
4	119.0	7.576(d)	7.7
5	121.8	7.171(t)	7.2
6	119.0	7.087(t)	7.26
7	110.4	7.331(d)	8
8	136.4		
9	127.5		
10	24.0	3.450(d), 2H	7.4
11	123.2	5.445(t)	7.5
12	135.5		
13	39.3	2.046(m), 2.105(m)	
14	26.6	2.110(m), 2H	
15	124.3	5.107(t)	6.6
16	131.4		
17	16.0	1.740(s)	
18	25.5	1.587(s)	
19	17.6	1.663(s)	

Reference:

1. Hillwig, M. L., Zhu, Q. & Liu, X. Biosynthesis of ambiguine indole alkaloids in cyanobacterium *Fischerella ambigua*. *ACS Chem. Biol.* 2014, **9**, 372-377.

2. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR* **1995,** 6, (3), 277-93.

3. Johnson, B. A.; Blevins, R. A. NMR View: A computer program for the visualization and analysis of NMR data. *Journal of biomolecular NMR* **1994,** 4, (5), 603-14.