Supporting Information

Mutation of the eunicellane synthase Bnd4 alters its product profile and expands its prenylation ability

Baofu Xu,[a] Wenbo Ning,[a] Xiuting Wei,[a] Jeffrey D. Rudolf[a]*

[a]Department of Chemistry, University of Florida, Gainesville, Florida, 32611, USA

*Correspondence to E-mail: jrudolf@chem.ufl.edu

Table of Contents

Supporting Methods

Bacterial strains, plasmids, and chemicals. Strains, plasmids, and PCR primers used in this study are listed in Tables S1–S3. PCR primers were obtained from Sigma-Aldrich. Q5 high-fidelity DNA polymerase and restriction endonucleases were purchased from NEB and used according to the protocols provided by the manufacturer. DNA gel extraction and plasmid preparation kits were purchased from Omega Bio-Tek. DNA sequencing was conducted by Genewiz. Other common chemicals, biochemical, and media components were purchased from standard commercial sources.

General experimental procedures. All ^{1}H, ^{13}C, and 2D NMR experiments were run at 400 MHz for ^{1}H and 100 MHz for ^{13}C nuclei on a Bruker Ascend 400 or at 600 MHz for ^{1}H and 150 MHz for ^{13}C nuclei on a Bruker Avance III 600. Preparative HPLC was carried out on an Agilent 1260 Infinity LC equipped with an Agilent Eclipse XDB-C18 column (250 mm × 21.2 mm, 7 μm). HPLC was performed on an Agilent 1260 Infinity LC equipped with an Agilent Zorbax SB-C18 column (150 mm × 4.6 mm, 5 μm) or Agilent InfinityLab Poroshell 120 EC-C18 (50 × 4.6 mm, 2.7 μm).

Site-directed mutagenesis. Primers (Table S3) were designed for T5 exonuclease-dependent assembly (TEDA).[1] For site-directed mutagenesis of *bnd4*, overlap PCR was used with the *bnd4* gene from pJR1003 as a template with Q5 DNA polymerase. The PCR products were purified by gel extraction and cloned into pET28a, which was linearized with *Bam*HI and *Hin*dIII, using TEDA to afford plasmids pJR1048–pJR1063. Mutagenesis of *dtcycA* and *cotB2* was similarly performed affording pJR1069 and pJR1070, respectively. The sequences of all genes were confirmed by DNA sequencing.

Protein production and purification. Plasmids harboring each gene were transformed into *E. coli* BL21 Star (DE3). *E. coli* strains harboring plasmids were grown in lysogeny broth (LB) containing 50 mg mL^{-1} kanamycin for antibiotic selection. Each strain was grown in 4 × 1 L of LB at 37 °C with shaking at 200 rpm until an optical density at 600 nm (OD_{600}) of 0.6 was reached. Gene expression was induced with the addition of 0.3 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and the cells were further cultured at 16 °C for 18 h with shaking at 200 rpm. After harvesting the cells by centrifugation at 4000 *g* for 15 min at 4 °C, the pellet was resuspended in cold lysis buffer (50 mM Tris-HCl, pH 8.0, containing 150 mM NaCl). Cells were disrupted using an M-110L Microfluidizer Processor (Quadro Engineering Corp) and centrifuged at 40,000 *g* for 25 min at 4 °C. Target proteins were purified by nickel-affinity chromatography by adding the supernatant to a column packed with HisPur™ Ni-NTA Resin (Thermo Scientific) washing with wash buffer (lysis buffer containing 20 mM imidazole), and eluting with elution buffer (lysis buffer containing 500 mM imidazole). The resultant protein was immediately desalted using a PD-10 column (GE Healthcare Biosciences) and concentrated using an Amicon Ultra-15 concentrator (Millipore) in 50 mM Tris-HCl, pH 8.0, containing 150 mM NaCl. Protein purities were assessed by SDS-PAGE analysis (Figure S2) and protein concentration was determined by the Bradford assay using bovine serum albumin as the standard.[2] Individual aliquots of each protein were flash-frozen in liquid nitrogen and stored at –80 °C until use.

Enzymatic activity assays of Bnd4 or mutants. The assays were performed in 50 mM Tris-HCl, pH 7.5, containing 1 mM GGPP, 10 mM $MgCl_2$, and 5 μM Bnd4 in a total volume of 100 μL. The reactions were initiated by the addition of enzyme and incubated for 10 min at 37 °C. The reactions were quenched with 200 μL of ice-cold acetonitrile and then 50 μL of saturated NaCl solution was added to form two layers. The upper organic layer was taken for HPLC analysis.

For chromatography of enzymatic products of the Bnd4 mutants (Figure 2A–2C), an Agilent Zorbax SB-C18 column (150 mm × 4.6 mm, 5 μm) was used and chromatographic separation was carried out at 35 °C, with a flow rate of 1 mL min^{-1} and an 18 min solvent gradient from 5–95% acetonitrile in water. The linear gradient program was run as follows: 0−2 min, 5% CH_3CN; 2–20 min, 5–95% CH_3CN; 20−45 min, 95% CH_3CN. For chromatography of enzymatic products of $Bnd4^{W316A}$ and DtcycA (Figure 2D), an Agilent Poroshell 120 EC-C18 column (50 x 4.6 mm, 2.7 μm) was used and chromatographic separation was carried out at 35 °C, with a flow rate of 1 mL min^{-1} and the gradient program was run as follows: 0−0.5 min, 5% CH_3CN; 0.5–11 min, 5-95% CH_3CN; 11–11.5 min, 95-100% CH_3CN; 11.5−25 min, 100% CH_3CN. Diterpene enzyme products and prenylation products were detected by monitoring at 210 nm and 280 nm, respectively, with a photodiode array detector.

Construction of the MKI4 system for GGPP production in *E. coli*. Primers for TEDA were designed (Table S1) and genes were first amplified from *Arabidopsis thaliana*, *E. coli*, or *Streptomyces sp*. CL12-4.[3-6] Overlap PCRs were then performed to combine four genes in one DNA fragment containing a ribosome binding site (rbs) between each gene. The obtained DNA fragment was cloned into linearized pET28a (digested with *Bam*HI and *Hin*dIII) using TEDA.[1] The plasmid, pET28a-MKI4 (pJR1064), was confirmed by seqencing and tested through co-transformation with pET21a-Bnd4 (pJR1017) or pCDF-Bnd4 (pJR1065).

Isolation of products from Bnd4 mutants. *E. coli* cells harboring pET28a-MKI4 (pJR1064) and selected pCDF-Bnd4 mutants (pJR1066 and pJR1067) were prepared by picking a single transformant into LB medium containing kanamycin (50 mg L^{-1}) and streptomycin (50 mg L^{-1}). Overnight cultured cells were then inoculated into 12 × 1 L of fresh Terrific Broth (TB) media. IPTG (0.5 mM) and isoprenol (4 mM) were added when the cultures reach an OD_{600} of 0.8 and further incubated at 28 °C for 24 hours while shaking. When *E. coli* was grown in LB, 1 mM of isoprenol was added due to lower observed yields of diterpene products when higher concentrations of isoprenol was used. Cells were collected by centrifugation at 4000 g for 15 min at 4 °C and transferred to a glass beaker. The cells were lysed and diterpenes were extracted by the addition of acetone. For extraction of the supernatant, 10–30 g L^{-1} of XAD-16 resin was added and the collected resin was eluted with five column volumes of acetone or methanol. Organic extracts were combined and concentrated to dryness and fractionated with silica (preneutralized with hexanes contatining 5% Et_3N and equilibrated with 5 column volumes of hexanes.[7] Fractions containing target compounds were combined and further purified using preparatory HPLC.

Accession numbers for proteins. Bnd4 homologues (Fig. S1): *Streptomyces* sp. CL12-4, WP_239771469; *Streptomyces flaveolus*, WP_189228172; *Streptomyces* sp. SID5910, A0A7K2JR29; *Streptomyces iakyrus*, WP_033312626; *Streptomyces* sp. yr375, A0A1H9GEM7; *Streptomyces* sp. CNH287, WP_027750606; *Amycolatopsis arida*, A0A1I5P1T3; *Amycolatopsis*

taiwanensis, WP_027943377; *Nocardia* sp. SYP-A9097, A0A6I2FRA3. Bacterial cembrene synthases (Fig. S23): DtCycA, M1V9Q0; DtCycB, M1VDX3; Rx_0493, WP_011563485; CAS, WP_030430753.

Spectroscopic data for known compounds that was consistent with previously reported spectra in $CDCl_3$:

β-Springene (2). ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 6.38 (dd, J = 10.8, 17.6 Hz, 1H), 5.24 (d, J = 17.6 Hz, 1H), 5.13 (m, 3H), 5.05 (d, J = 10.9 Hz, 1H), 5.01 (d, J = 6.6 Hz, 2H), 2.23 (m, 2H), 2.20 (m, 3H), 2.08 (m, 3H), 1.99 (m, 4H), 1.68 (s, 3H), 1.60 (s, 9H) ppm; ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 146.3, 139.1, 135.6, 135.1, 131.4, 124.5, 124.4, 124.2, 115.9, 113.2, 39.88, 39.85, 31.6, 26.9, 26.78, 26.77, 25.9, 17.8, 16.20, 16.17 ppm. The NMR data, shown in Figs. S4–S8, was consistent with previously reported spectra in $CDCl_3$.[8]

Geranylgeraniol (3). ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 5.44 (tq, J = 6.9, 1.3 Hz, 1H), 5.13 (m, 3H), 4.18 (d, J = 6.4 Hz, 2H), 2.18–1.95 (m, 12H), 1.71 (s, 6H), 1.63 (m, 9H), 2.08 (m, 3H), 1.99 (m, 4H), 1.68 (s, 3H), 1.60 (s, 9H) ppm; ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 140.0, 135.5, 135.1, 131.4, 124.5, 124.3, 123.9, 123.5, 59.6, 39.9, 39.8, 39.7, 26.9, 26.8, 26.5, 25.8, 17.8, 16.4, 16.17, 16.15 ppm. The NMR data, shown in Figs. S9–S10, was consistent with previously reported spectra in $CDCl_3$.[9]

Geranyllinalool (4). ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 5.94 (dd, J = 10.7, 17.3 Hz, 1H), 5.24 (dd, J = 17.3, 1.3 Hz, 1H), 5.20–5.11 (m, 3H), 5.09 (dd, J = 10.7, 1.3 Hz, 1H), 2.13–2.05 (m, 6H), 2.05–1.97 (m, 4H), 1.70 (d, J = 1.4 Hz, 3H), 1.66–1.56 (m, 2H) 1.63 (br s, 6H) 1.62 (br s, 3H), 1.31 (s, 3H) ppm. The NMR data, shown in Fig. S11, was consistent with previously reported spectra in $CDCl_3$.[10]

3-Farnesylindole (9). ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.89 (br s, 1H), 7.59 (d, J = 7.9 Hz, 1H), 7.35 (d, J = 8.1 Hz, 1H), 7.18 (t, J = 7.5 Hz, 1H), 7.12 – 7.08 (m, 1H), 6.95 (m, 1H), 5.46 (m, 1H), 5.14 (m, 1H), 5.09 (m, 1H), 3.47 (d, J = 7.1 Hz, 2H), 2.14 (m, 2H), 2.06 (m, 4H), 1.98 (m, 2H), 1.76 (s, 3H), 1.68 (s, 3H), 1.60 (d, 6H) ppm. ^{13}C NMR (125 MHz, $CDCl_3$): δ_C 136.5, 135.7, 135.0, 131.3, 127.5, 124.4, 124.2, 122.9, 121.9, 121.1, 119.1, 119.0, 116.2, 111.0, 39.8, 39.7, 29.7, 26.8, 25.7, 24.0, 17.7, 16.1, 16.0 ppm. The NMR data, shown in Figs. S24–S27, was consistent with previously reported spectra in $CDCl_3$.[11]

Data availability. All data generated or analyzed in this study are available within the article and its Supporting Information.

Table S1. Strains used in this study.

Strain	Description	Source
E. coli NEB Turbo	Host for general cloning	New England Biolabs
E. coli BL21 Star (DE3)	Host for high-level protein production	Invitrogen

Table S2. Plasmids used in this study.

Plasmid	Description	Source (Reference)
pET28a	General plasmid for cloning and protein production	Novagen
pET21a	Plasmid for heterologous expression in *E. coli*	Novagen
pCDF-Duet	Plasmid for heterologous expression in *E. coli*	Novagen
pJBEI-2999	Plasmid harboring genes for the overproduction of FPP. Herein, co-transformed with pJR1015 to overproduce GGPP	Addgene 35152 (ref. 12)
pJR1003	pET28a harboring *bnd4*	(6 and 13)
pJR1004	pET28a harboring *cotB2*	(13)
pJR1015	pET28a harboring *ggpps*	(6 and 13)
pJR1017	pET21a harboring *bnd4* (no tag)	(6)
pJR1048	pET28a harboring *bnd4* (Y197A)	This study
pJR1049	pET28a harboring *bnd4* (Y197F)	This study
pJR1050	pET28a harboring *bnd4* (Y197W)	This study
pJR1051	pET28a harboring *bnd4* (Y197H)	This study
pJR1052	pET28a harboring *bnd4* (Y197M)	This study
pJR1053	pET28a harboring *bnd4* (Y197L)	This study
pJR1054	pET28a harboring *bnd4* (Y197E)	This study
pJR1055	pET28a harboring *bnd4* (F162A)	This study
pJR1056	pET28a harboring *bnd4* (F162Y)	This study
pJR1057	pET28a harboring *bnd4* (F162A/Y197A)	This study
pJR1058	pET28a harboring *bnd4* (F162Y/Y197F)	This study
pJR1059	pET28a harboring *bnd4* (W316A)	This study
pJR1060	pET28a harboring *bnd4* (W316H)	This study
pJR1061	pET28a harboring *bnd4* (W316F)	This study
pJR1062	pET28a harboring *bnd4* (W316Y)	This study
pJR1063	pET28a harboring *bnd4* (W67A)	This study
pJR1064	pET28a-MKI4: pET28a harboring kinases *Ec-Th*iM and *At-IPK*, *Ec-idi*, and GGPP synthase (*bnd3*). Ribosome binding sites were inserted before each gene.	This study
pJR1065	pCDF harboring *bnd4*	This study
pJR1066	pCDF-Duet harboring *bnd4* (Y197A)	This study
pJR1067	pCDF-Duet harboring *bnd4* (W316A)	This study
pJR1068	pET28a harboring codon-optimized *dtcycA*	This study
pJR1069	pET28a harboring *dtcycA* (A321W)	This study
pJR1070	pET28a harboring *cotB2* (W186A)	This study

Table S3. Primers used in this study. Mutations are shown as lowercase letters.

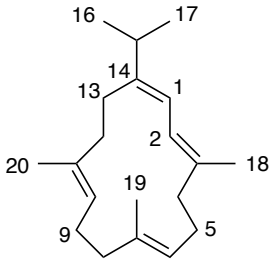
Primer	Nucleotide Sequence (5′–3′)	Purpose	Reference
Bnd4-F	CAGCAAATGGGTCGCGGATCCA TGTCGACCATCCCCAAGCC	*bnd4* or mutant amplification for expression in *E. coli*	(6 and 13)
Bnd4-R	CTCGAGTGCGGCCGCAAGCTTTC ACGCGGGGACCTCCTCGG		
Y197A-R	GAGCACGAAGgcCATGCCCATGC CGACGCTGTCCTTGCGA	Bnd4 mutagenesis for Y197A	
Y197A-F	AGCGTCGGCATGGGCATGgcCTT CGTGCTCGGCGAGTACGGACTG		
Y197F-R	GAGCACGAAGaaCATGCCCATGC CGACGCTGTCCTTGCGA	Bnd4 mutagenesis for Y197F	
Y197F-F	AGCGTCGGCATGGGCATGttCTTC GTGCTCGGCGAGTACGGACTG		
Y197W-R	GAGCACGAAccaCATGCCCATGC CGACGCTGTCCTTGCGA	Bnd4 mutagenesis for Y197W	
Y197W-F	AGCGTCGGCATGGGCATGtggTTC GTGCTCGGCGAGTACGGACTG		
Y197H-R	GAGCACGAAgtgCATGCCCATGCC GACGCTGTCCTTGCGA	Bnd4 mutagenesis for Y197H	
Y197H-F	AGCGTCGGCATGGGCATGcacTTC GTGCTCGGCGAGTACGGACTG		
Y197M-R	GAGCACGAAcatCATGCCCATGCC GACGCTGTCCTTGCGA	Bnd4 mutagenesis for Y197M	
Y197M-F	AGCGTCGGCATGGGCATGatgTTC GTGCTCGGCGAGTACGGACTG		
Y197L-R	GAGCACGAAgagCATGCCCATGC CGACGCTGTCCTTGCGA	Bnd4 mutagenesis for Y197L	
Y197L-F	AGCGTCGGCATGGGCATGctcTTC GTGCTCGGCGAGTACGGACTG		
Y197E-R	GAGCACGAAcTcCATGCCCATGC CGACGCTGTCCTTGCGA	Bnd4 mutagenesis for Y197E	
Y197E-F	AGCGTCGGCATGGGCATGgAgTT CGTGCTCGGCGAGTACGGACTG		
F162A-R	ACGCAGCCGGCGAGGgcGCGCCG GACCTCCCCCATGAACC	Bnd4 mutagenesis for F162A	
F162A-F	GGAGGTCCGGCGCgcCCTCGCCG GCTGCGTCCACGAGATC		
F162Y-R	ACGCAGCCGGCGAGGtAGCGCCG GACCTCCCCCATGAACC	Bnd4 mutagenesis for F162Y	
F162Y-F	GGAGGTCCGGCGCTaCCTCGCCG GCTGCGTCCACGAGATC		
W316A-R	CGTCAGGTACGAggcCTGCAGGTT GCCGGCCATCATGTGC	Bnd4 mutagenesis for W316A	
W316A-F	GCCGGCAACCTGCAGgccTCGTAC CTGACGTCCCGGTACA		
W316H-R	CGTCAGGTACGAgtgCTGCAGGTT GCCGGCCATCATGTGC	Bnd4 mutagenesis for W316H	
W316H-F	GCCGGCAACCTGCAGcacTCGTAC CTGACGTCCCGGTACA		

W316F-R	CGTCAGGTACGAgaaCTGCAGGTTGCCGGCCATCATGTGC	Bnd4 mutagenesis for W316F
W316F-F	GCCGGCAACCTGCAGttcTCGTACCTGACGTCCCGGTACA	
W316Y-R	CGTCAGGTACGAgtaCTGCAGGTTGCCGGCCATCATGTGC	Bnd4 mutagenesis for W316Y
W316Y-F	GCCGGCAACCTGCAGtacTCGTACCTGACGTCCCGGTACA	
W67A-R	TCGGCAGCACCAGACAGGTCgcCAGGGAGGCGGTCTCCTC	Bnd4 mutagenesis for W67A
W67A-F	CCTCCCTGgcGACCTGTCTGGTGCTGCCGACGGCCCGCGA	
28-DtcycA-F	CAGCAAATGGGTCGCGGATCCATGACCGATCCGGCGGTCACCCCATTAGC	
A321W-R	CGAGAACTCGTTGCATGccaATCGAGATTACCAGAAATTA	DtcycA mutagenesis for A321W
A321W-F	TAATTTCTGGTAATCTCGATtggCATGCAACGAGTTCTCG	
28-DtcycA-R	ctcgagtgcggccgcaagcttCTACTGGTCTAACTGTTCCCACCAC	
28CT-F	CAGCAAATGGGTCGCGGATCCATGACGACAGGACTTTCCAC	
ct-W186A-R	CGACATCTTCATCgcGAAGTCGACGCCGATGTCGGTGACC	CotB2 mutagenesis for W186A
c-W186A-F	ATCGGCGTCGACTTCgcGATGAAGATGTCGTATCCGATCT	
28CT-R	ctcgagtgcggccgcaagcttTCACTGGATGCGAGAGTTGA	

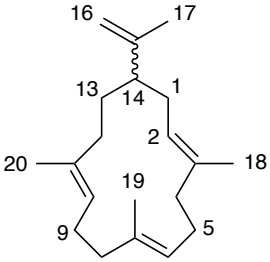
Table S4. ^{13}C NMR (100 MHz) and ^{1}H NMR (400 MHz) spectroscopic data for cembrene C (**6**) and cembrene A (**7**) in $CDCl_3$ (δ in ppm, *J* in Hz)[a]

	Cembrene C (6)		**Cembrene A (7)**	
No.	δ_C	δ_H[b]	δ_C	δ_H
1	118.59, CH	6.01 (d, 11.2)	32.58, CH_2	1.98 (m)
2	122.07, CH	5.93 (dq, 11.3, 1.3)	122, CH	5.06 (m)
3	134.68, qC		133.9, qC	
4	39.31, CH_2	2.13 (m)	39.1, CH_2	2.14 (m)*
5	25.44, CH_2	2.18 (m)	25.04, CH_2	2.26, 2.17 (m)*
6	125.18, CH	5.01 (m)	124.21, CH	5.19 (m)
7	134.39, qC		133.59, qC	
8	39.08, CH_2	2.13 (m)	39.57, CH_2	2.05 (m)
9	24.63, CH_2	2.16 (m)	23.91, CH_2	2.12 (m)
10	124.7, CH	5.01 (m)	126.06, CH	4.98 (m)
11	134.85, qC		134.96, qC	
12	38.7, CH_2	2.13 (m)	34.13, CH_2	1.94, 1.78 (m)*
13	28.15, CH_2	2.32 (m)	28.34, CH_2	1.67, 1.37 (m)*
14	147.14, qC		46.13, CH	2.02 (m)
15	33.89, CH	2.32 (m)	149.45, qC	
16	22.46, CH_3	1.05 (s)	110.27, CH	4.68 (m)
17	21.35, CH_3	1.03 (s)	19.47, CH_3	1.66 (s)
18	17.14, CH_3	1.74 (d, 1.3)	18.16, CH_3	1.56 (s)
19	15.8, CH_3	1.51 (s)	15.46, CH_3	1.59 (s)
20	17.3, CH_3	1.58 (s)	15.68, CH_3	1.57 (s)

[a] Assignments are based on 1D and 2D NMR experiments.
[b] All ^{1}H chemical shifts for **6** were shifted by ≅0.05 ppm compared with reported literature values[14]
* Chemical shifts slightly differed from reported literature values[15]



Cembrene C (**6**)



Cembrene A (**7**)

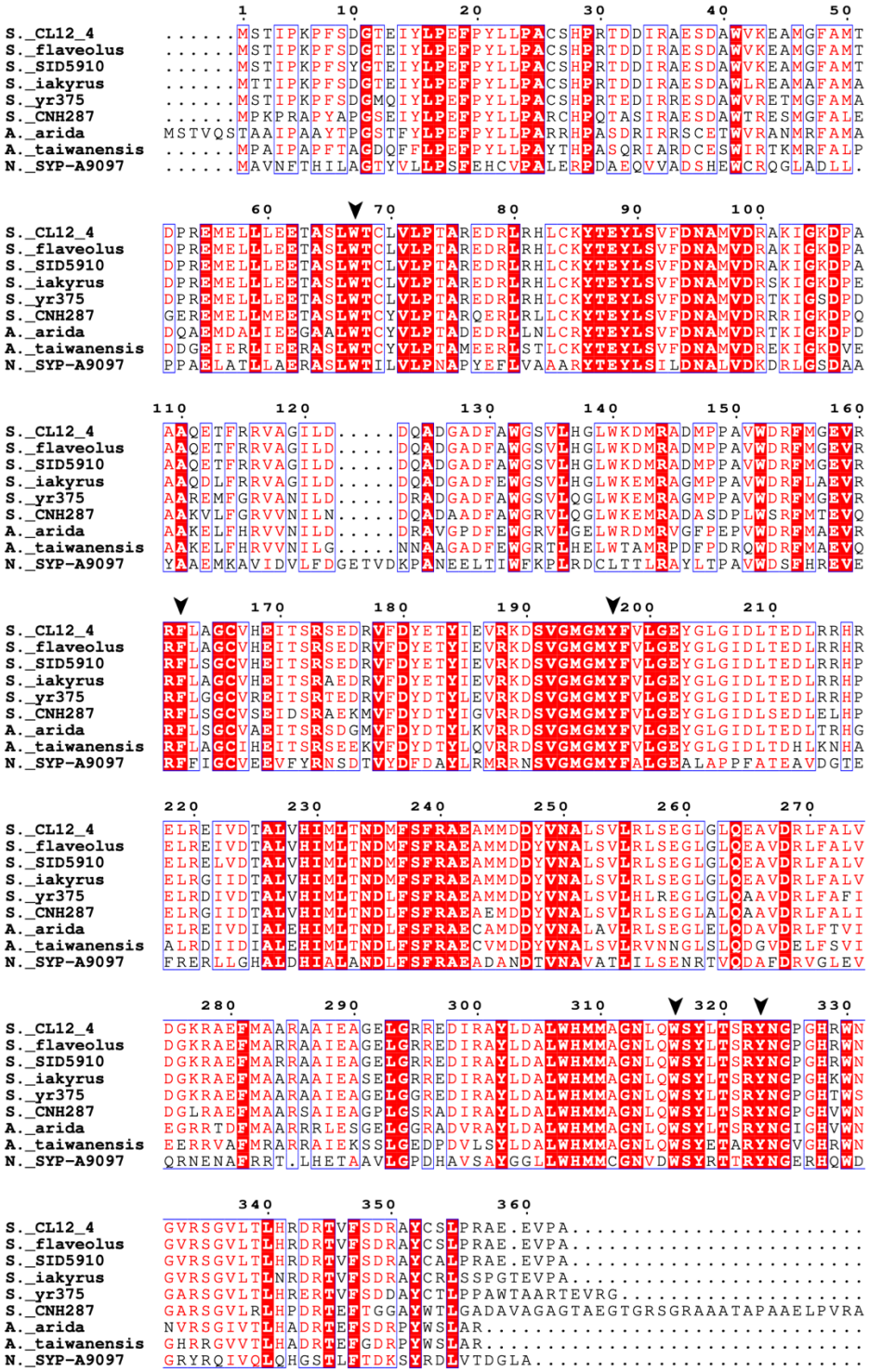


Figure S1. Sequence alignment of Bnd4 and selected homologues. Residues are colored based on the level of conservation (red box, red character, and blue frame show identity, similarity, and similarity across groups, respectively). Conserved aromatic residues found in the active site are marked with arrowheads. Clustal Omega[16] and ESPript[17] were used to generate and render the alignment.

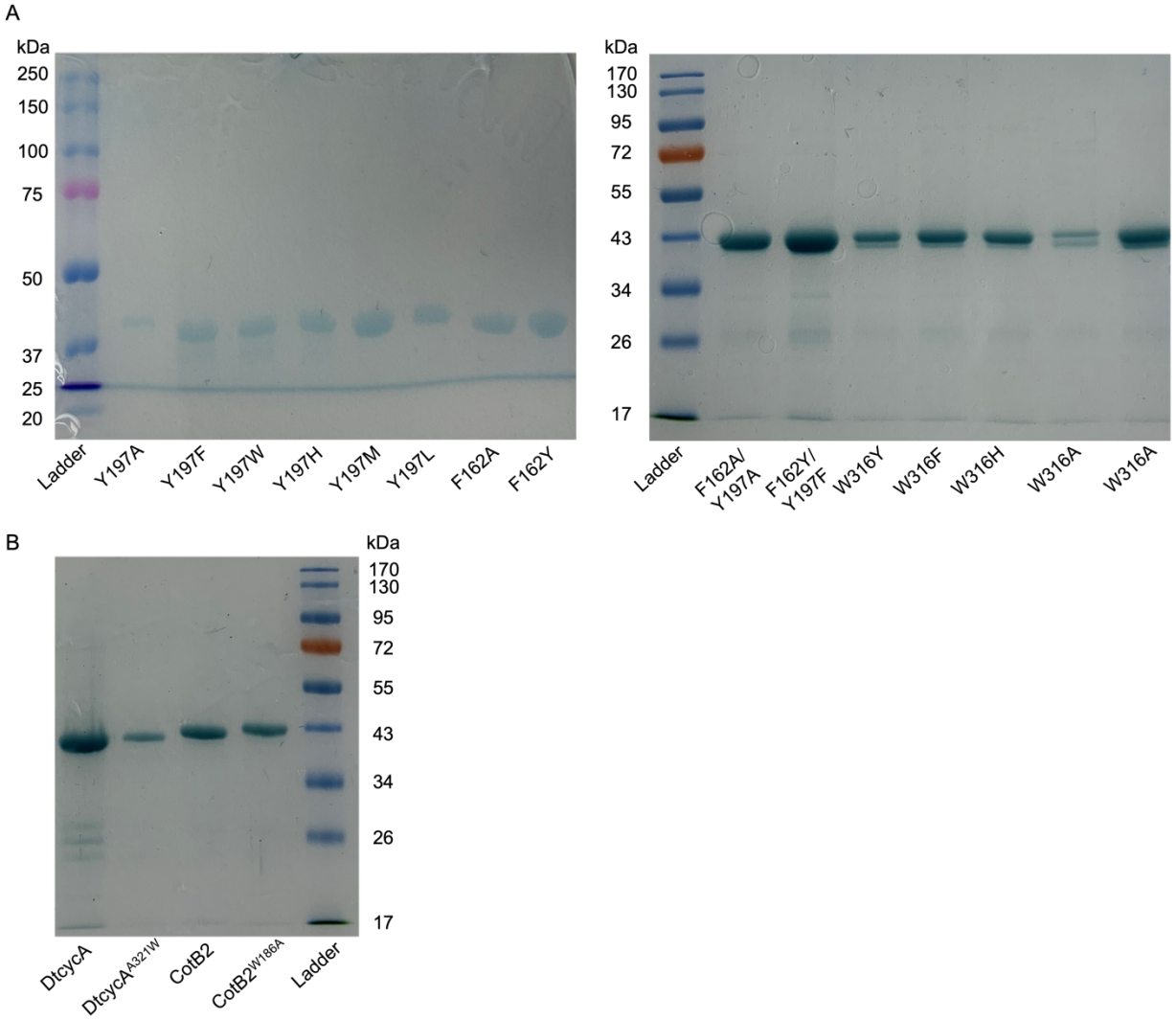


Figure S2. SDS-PAGE analysis of purified proteins.

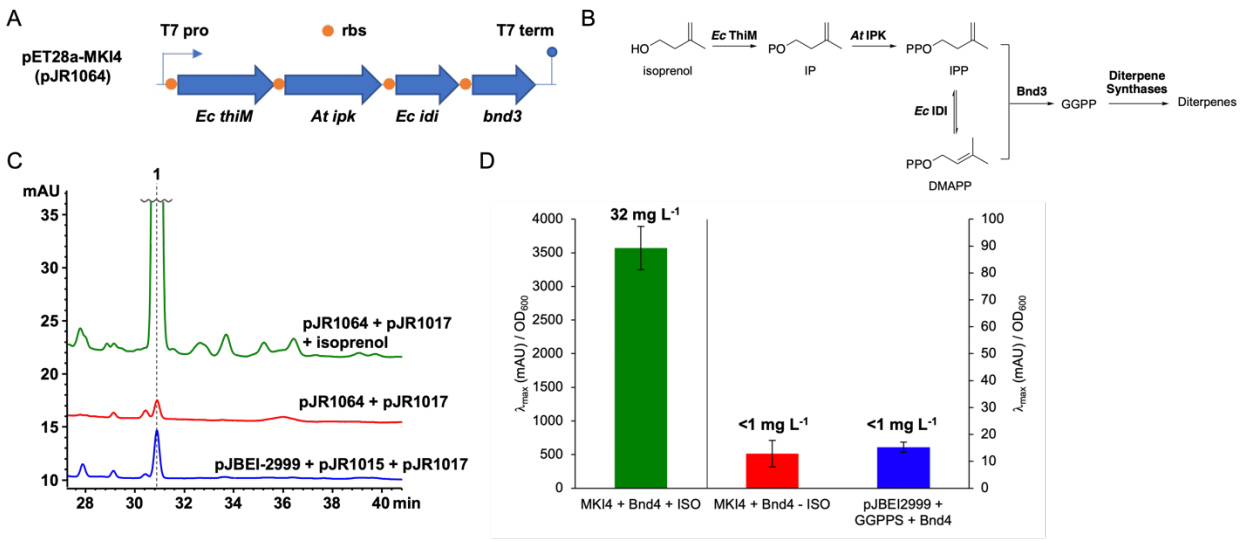


Figure S3. Diterpene overproduction system in *E. coli*. (A) Plasmid design of pET28a-MKI4 (pJR1064), a GGPP overproduction system using two kinases, hydroxyethylthiazole kinase (ThiM) from *E. coli* and isopentenyl phosphate kinase (IPK) from *Arabidopsis thaliana*, with isopentenyl diphosphate isomerase (IDI) from *E. coli* and a GGPP synthase (Bnd3) from *Streptomyces* sp. (CL12-4). All genes are under a single T7 promoter-*lacO* transcription/regulation module. Ribosome binding sites (rbs) were included before each gene to ensure maximum translation. (B) Scheme of isoprenoid production in *E. coli* harboring pET28a-MKI4 (pJR1064). (C) Diterpene overproduction was confirmed using *E. coli* harboring pJR1017 and fed isoprenol in comparison to a negative control without the addition of isoprenol and a previously employed diterpene production system (pJBEI-2999, pJR1015, and pJR1017).[6] (D) Quantification of benditerpe-2,6,15-triene (**1**) production using λ_{max} = 210 nm divided by OD_{600}. Values represent average values of four independent experiments with error bars showing the standard deviation; calculated titers of **1** are shown above each bar. Primary (left) axis represents value for MKI4 + Bnd4 + isoprenol (green); secondary (right) axis represents values for the two controls (red and blue).

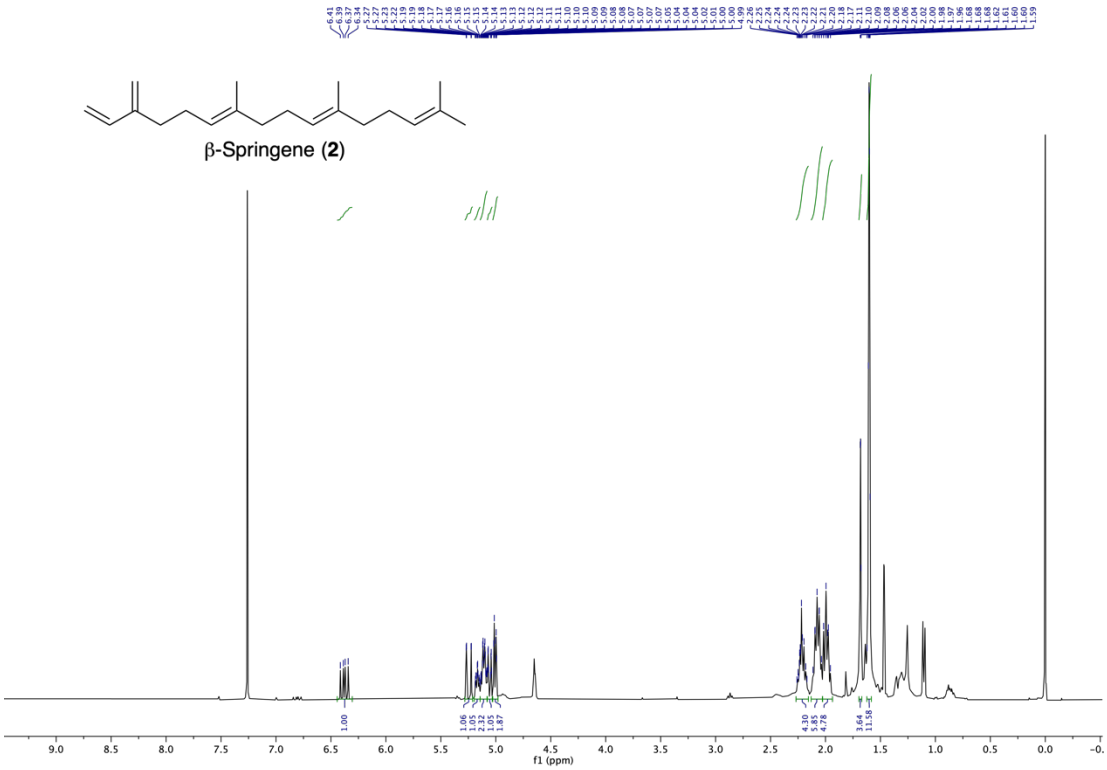


Figure S4. ^{1}H NMR spectrum of β-springene (**2**) in $CDCl_3$ (400 MHz).

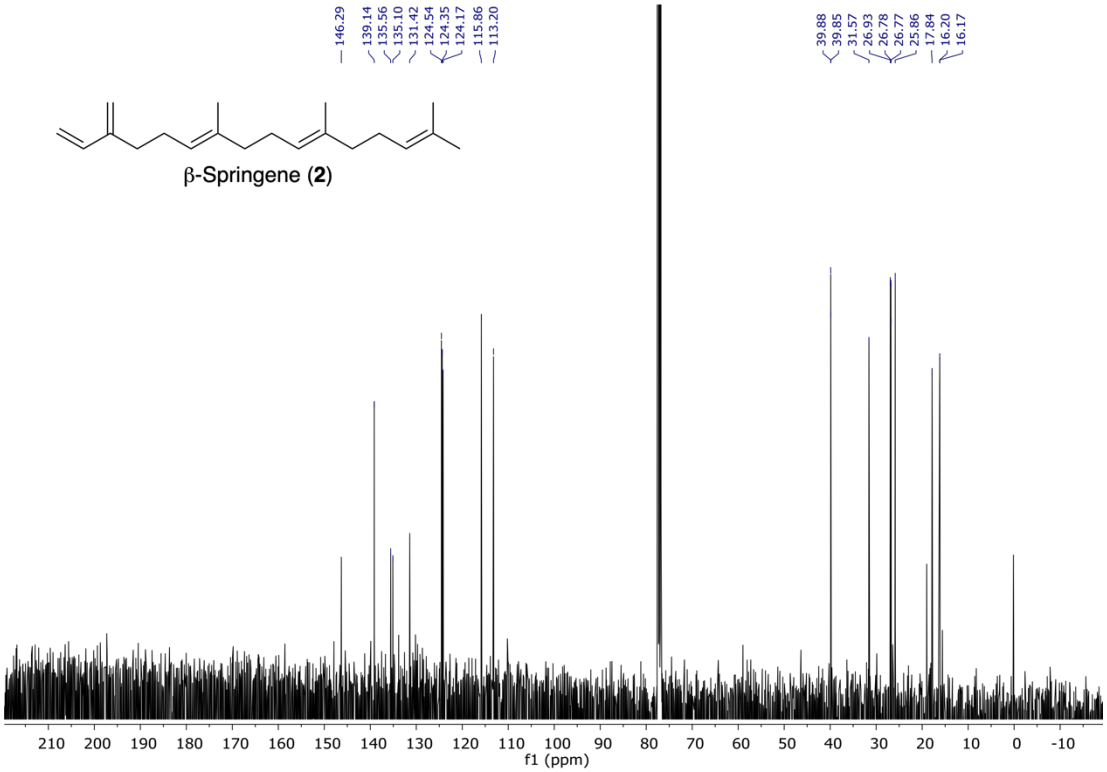


Figure S5. ^{13}C NMR spectrum of β-springene (**2**) in $CDCl_3$ (100 MHz).

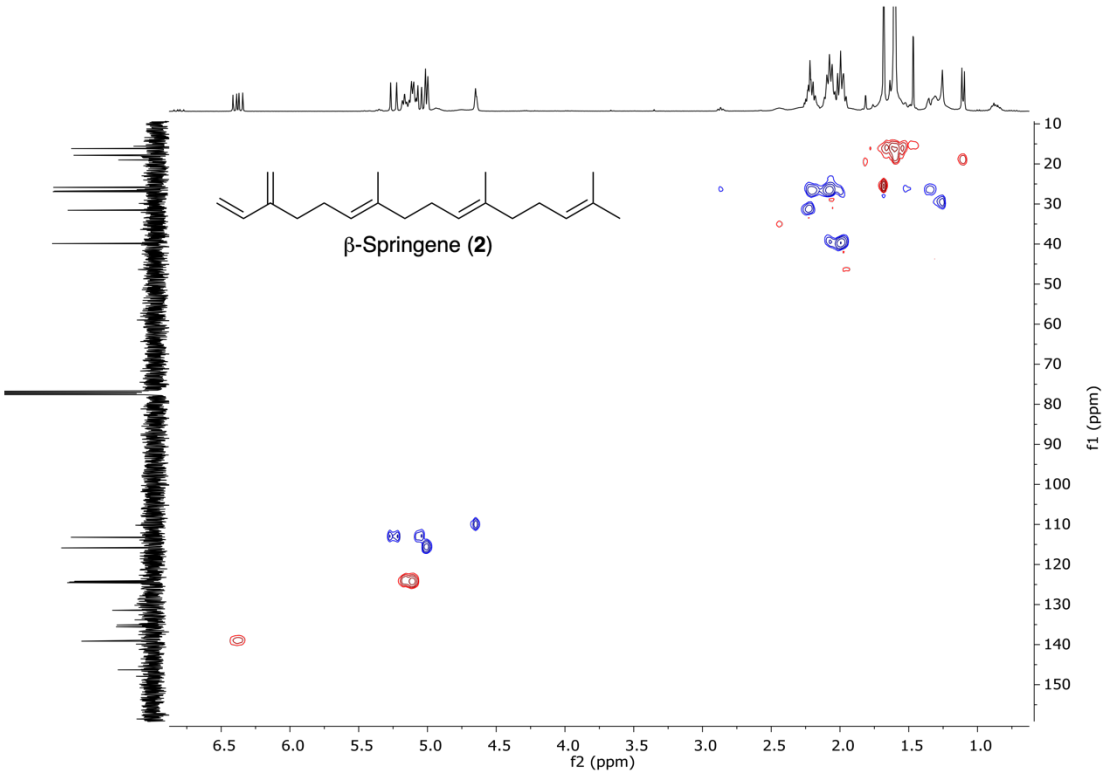


Figure S6. HSQC spectrum of β-springene (**2**) in $CDCl_3$ isolated from Bnd4^{Y197A}.

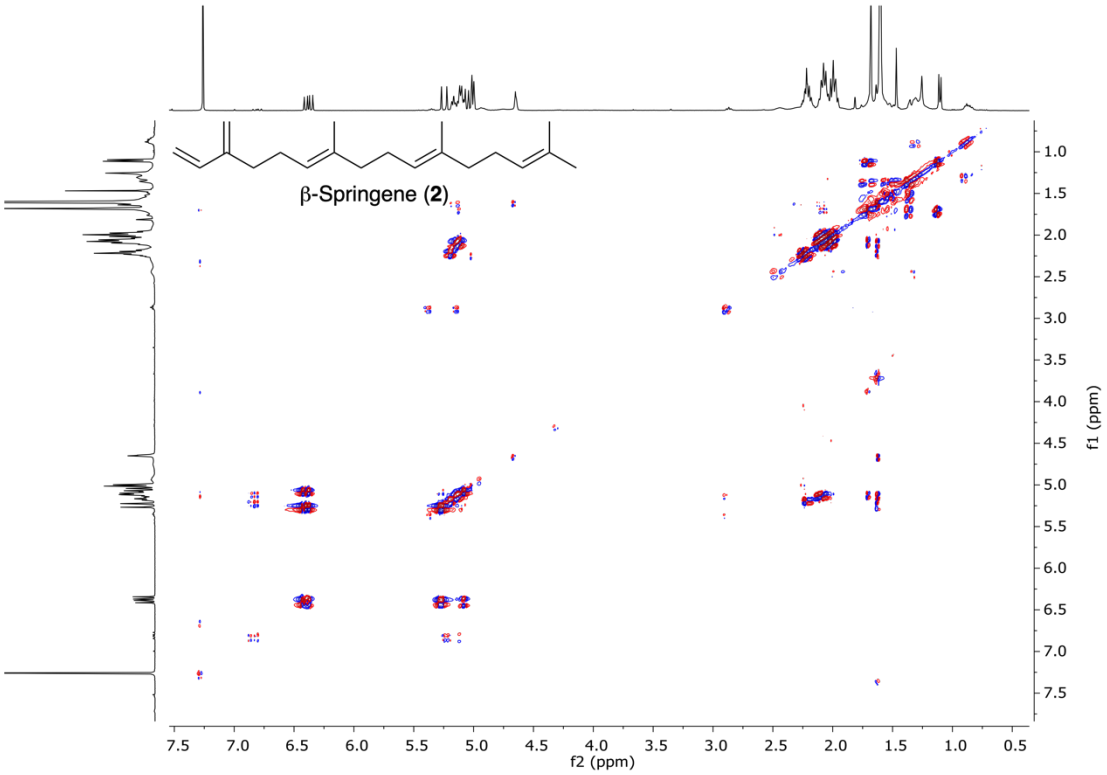


Figure S7. ^{1}H-^{1}H COSY spectrum of β-springene (**2**) in $CDCl_3$.

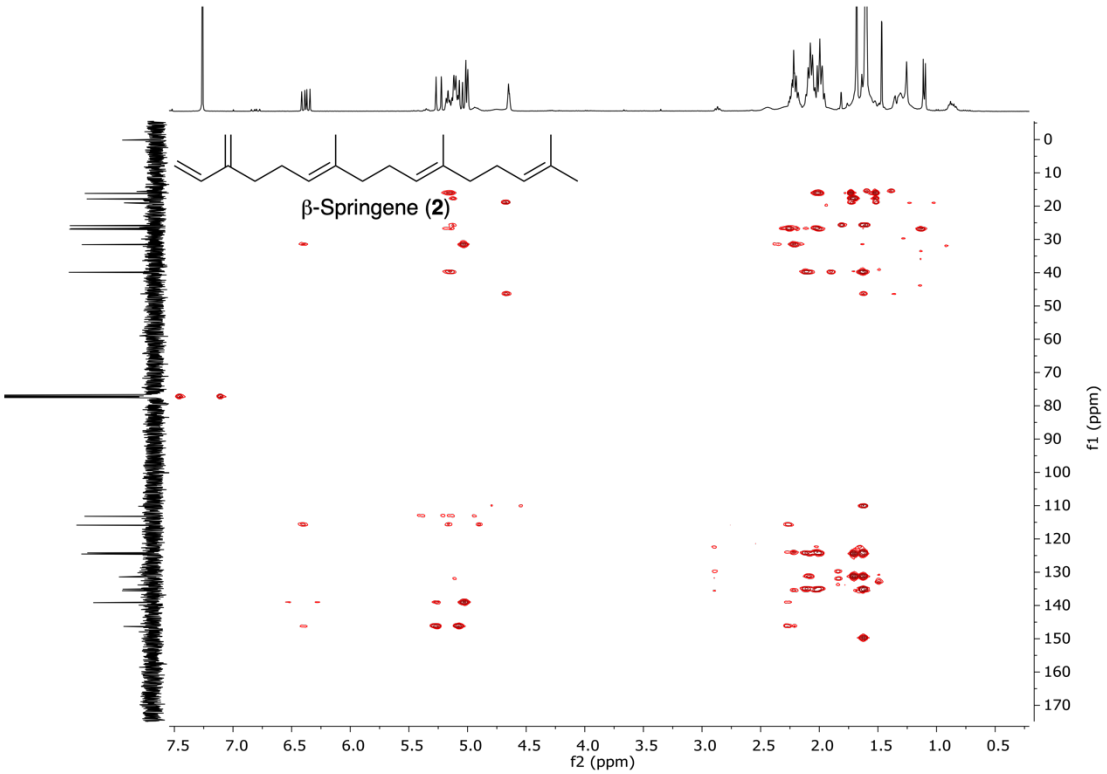


Figure S8. ^{1}H-^{13}C HMBC spectrum of β-springene (**2**) in $CDCl_3$.

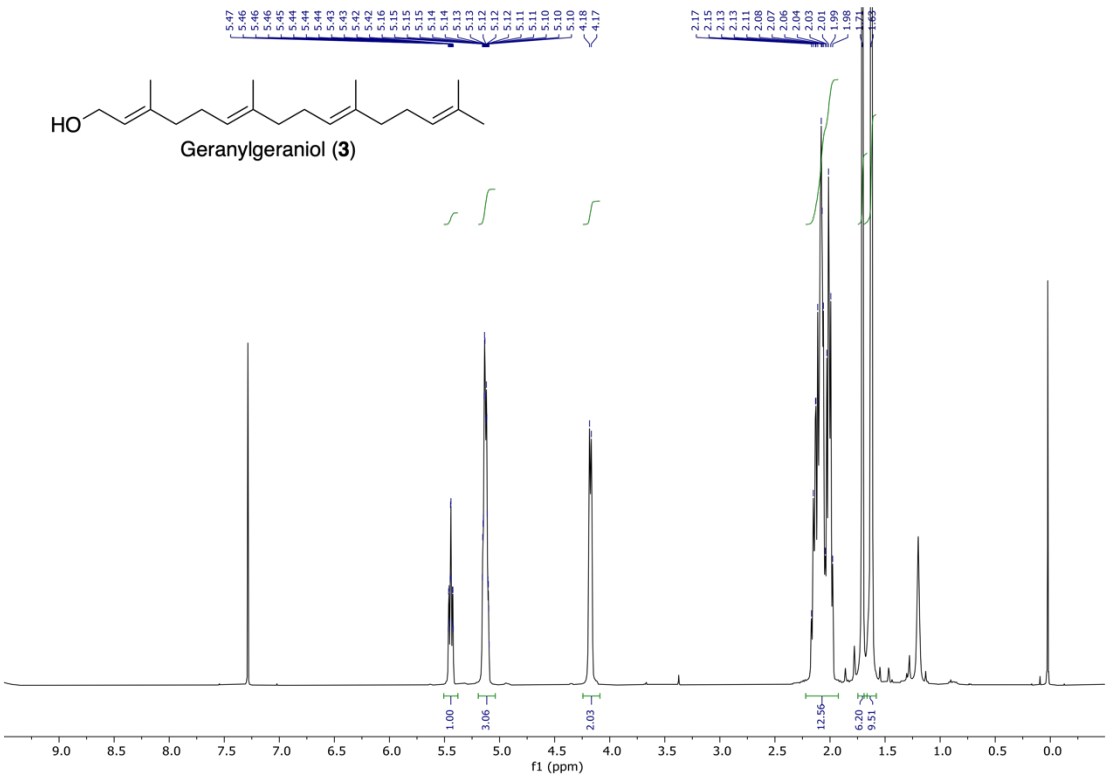


Figure S9. ^{1}H NMR spectrum of GGOH (**3**) in $CDCl_3$ (400 MHz).

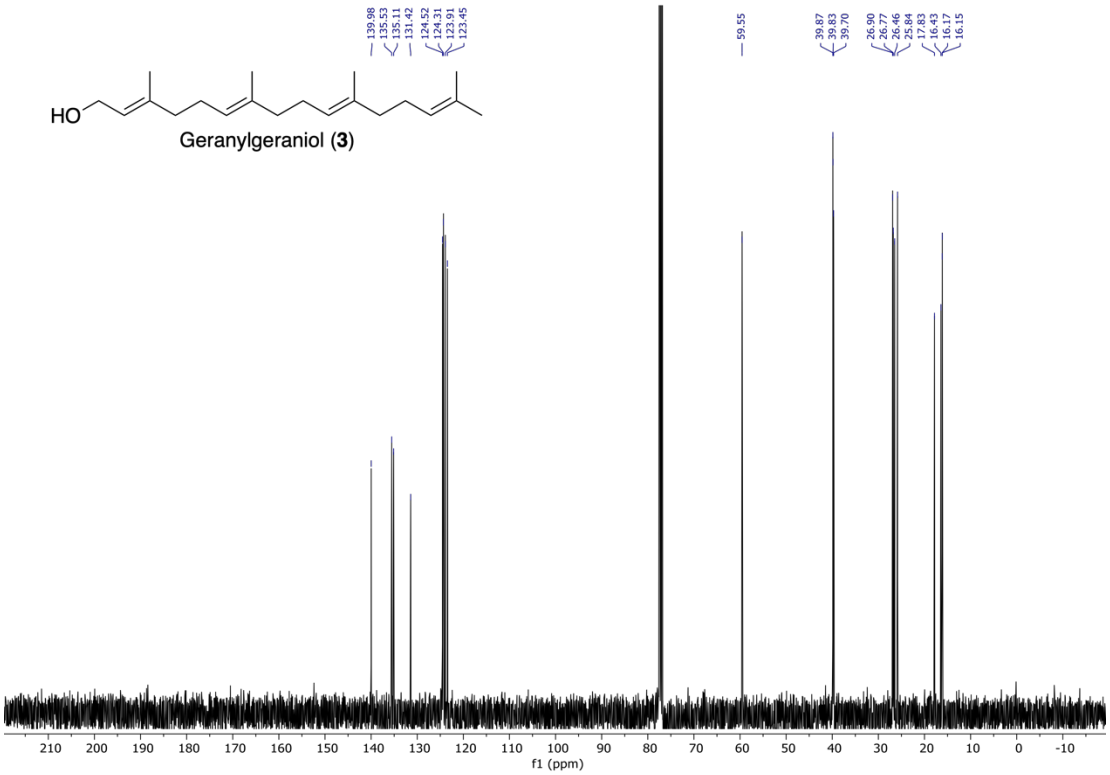


Figure S10. ^{13}C NMR spectrum of GGOH (**3**) in $CDCl_3$ (100 MHz).

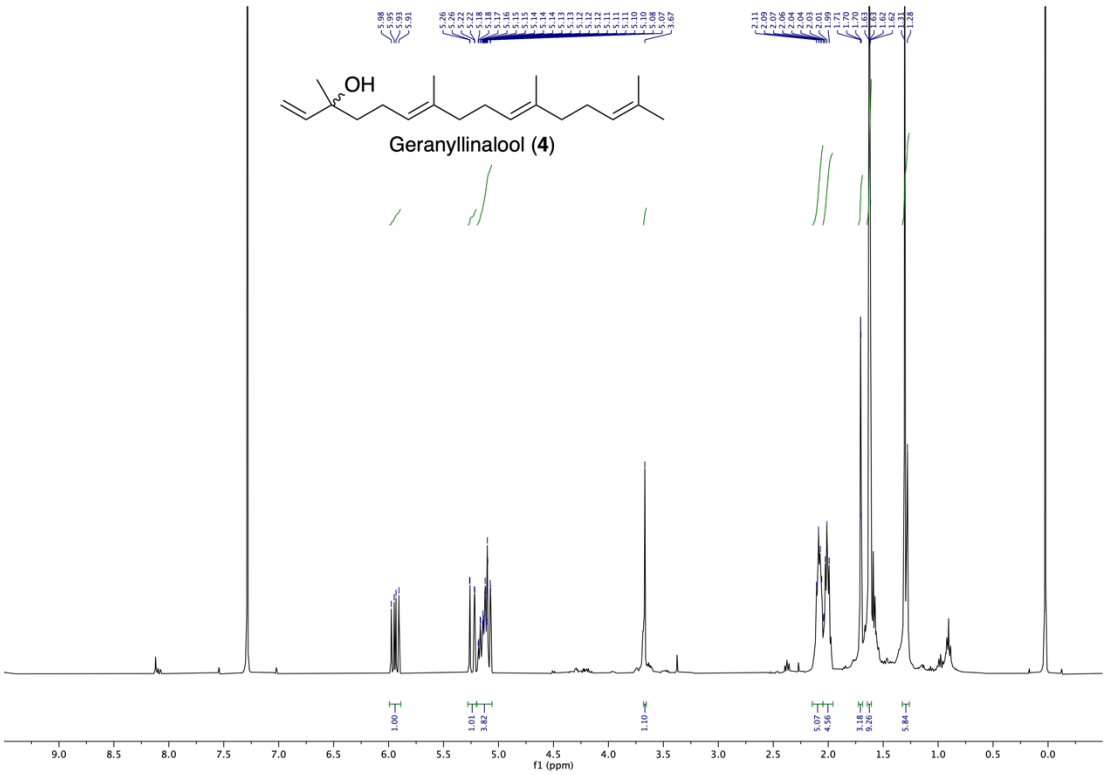


Figure S11. ^{1}H NMR spectrum GLOH (**4**) in $CDCl_3$ (400 MHz).

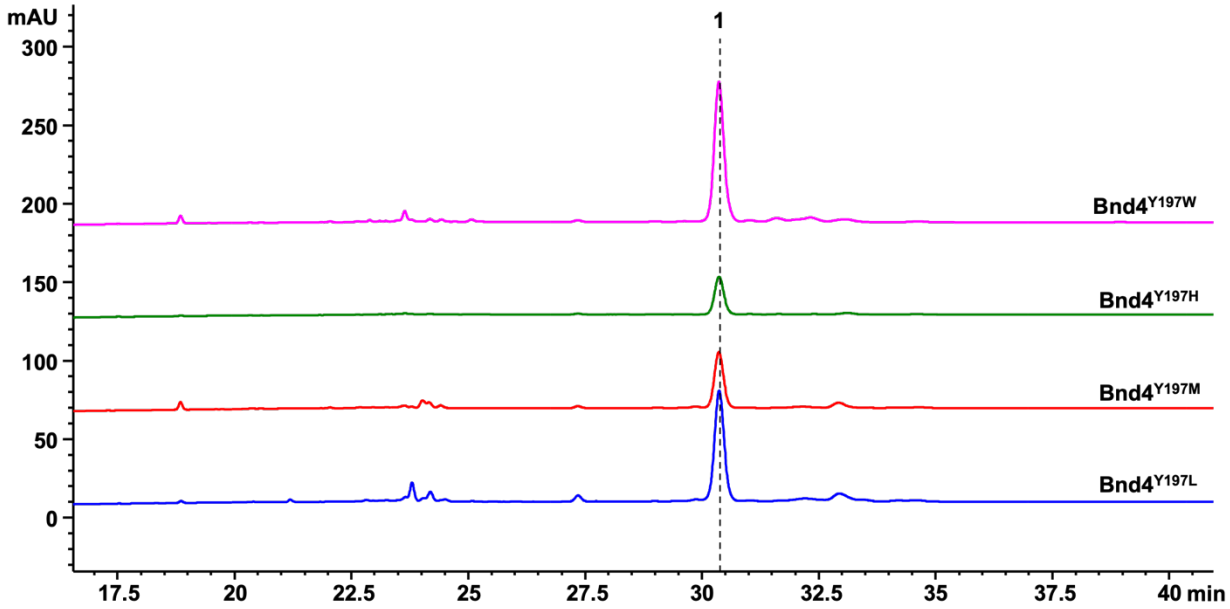


Figure S12. Mutation of Y197 in Bnd4 to Trp, His, Met, or Leu did not change its product profile. HPLC analyses of the Bnd4 mutants.

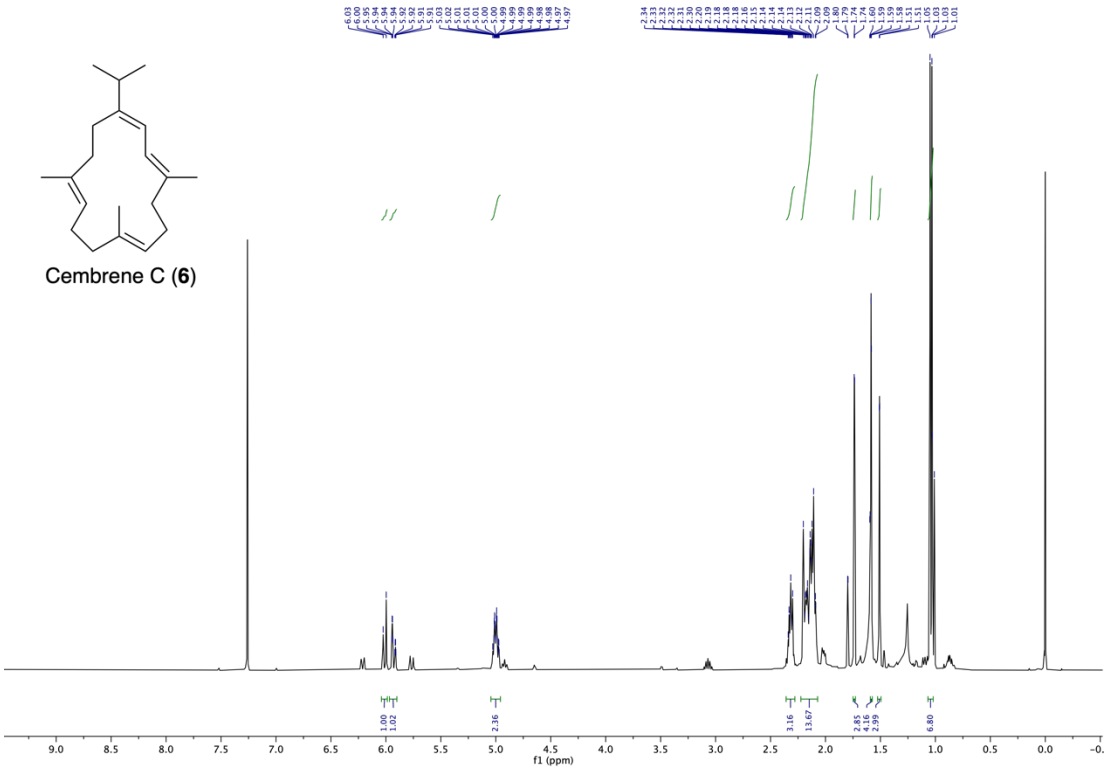


Figure S13. ^{1}H NMR spectrum of cembrene C (**6**) in $CDCl_3$ (400 MHz).

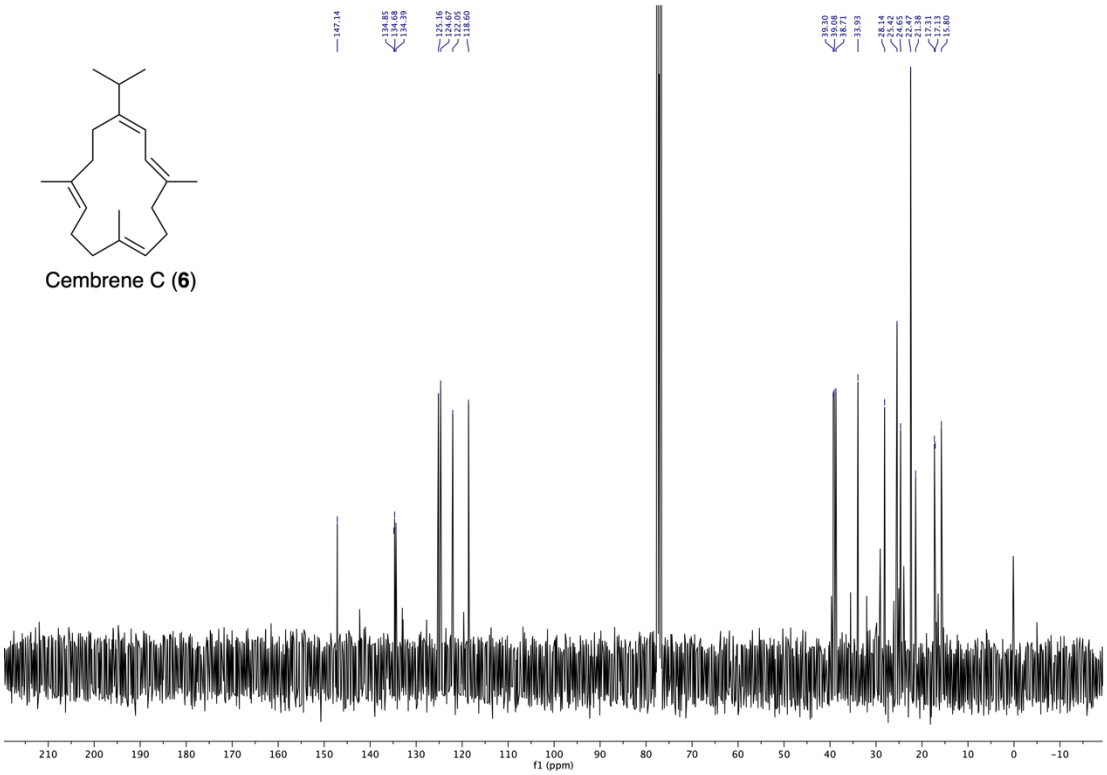


Figure S14. ^{13}C NMR spectrum of cembrene C (**6**) in $CDCl_3$ (100 MHz).

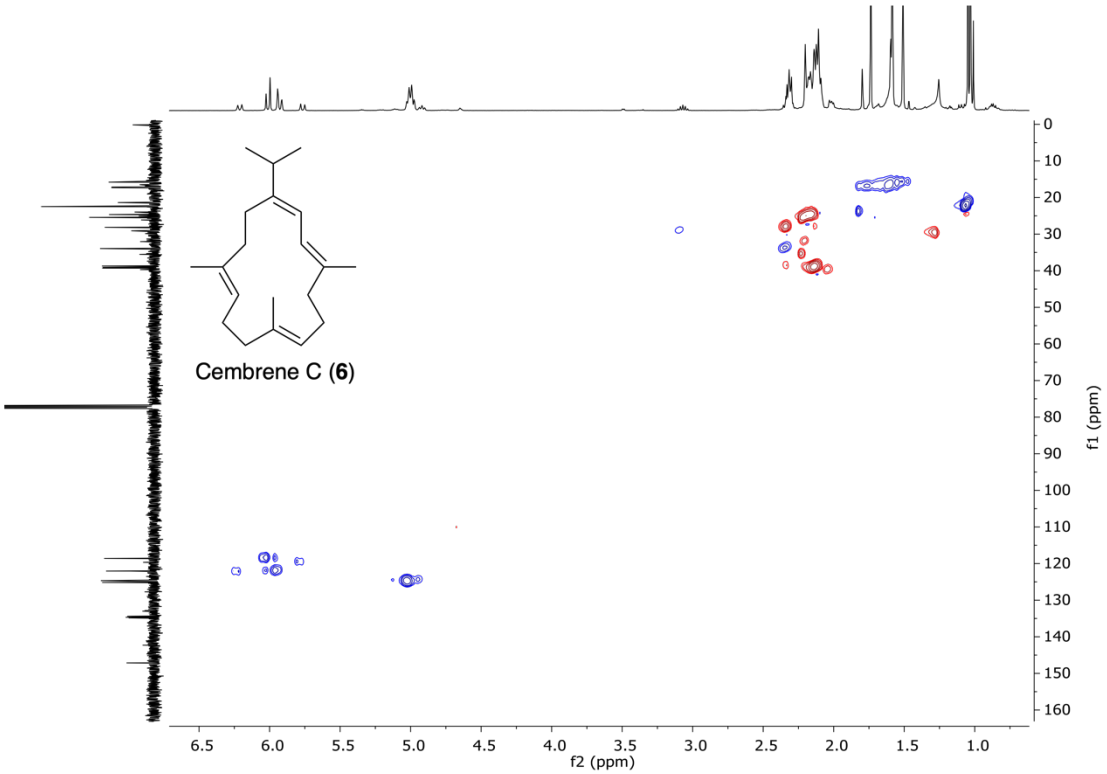


Figure S15. ^{1}H-^{13}C HSQC spectrum of cembrene C (**6**) in $CDCl_3$.

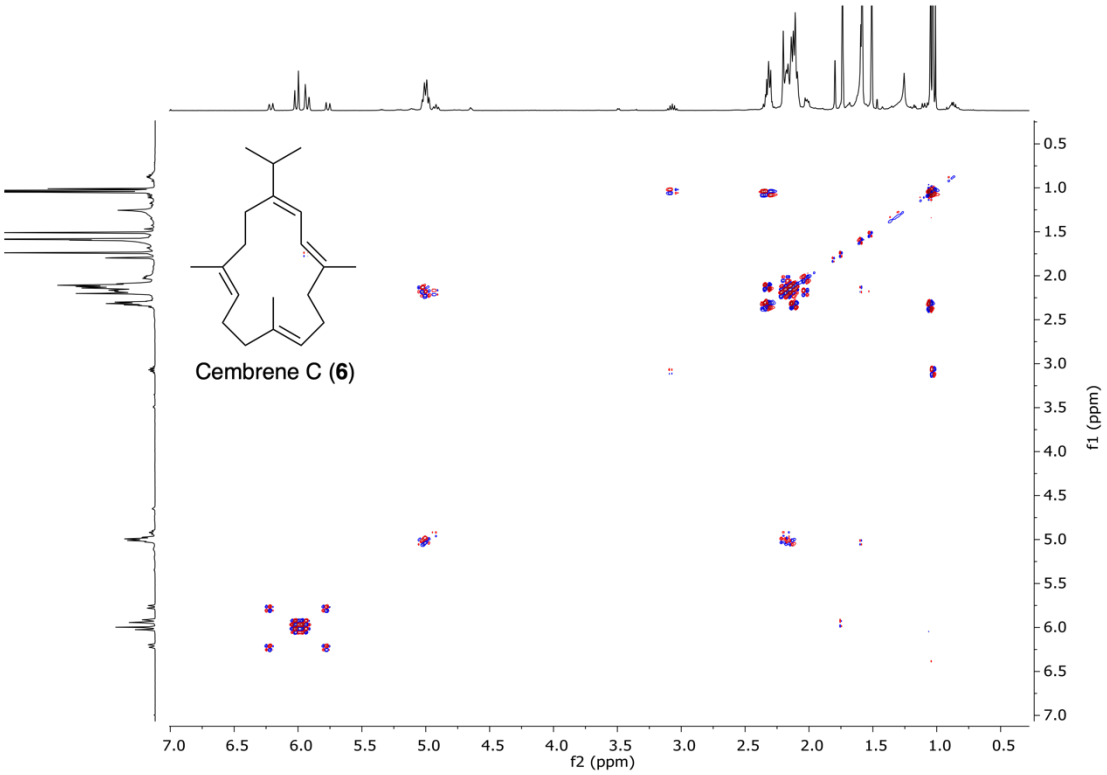


Figure S16. ^{1}H-^{1}H COSY spectrum of cembrene C (**6**) in $CDCl_3$.

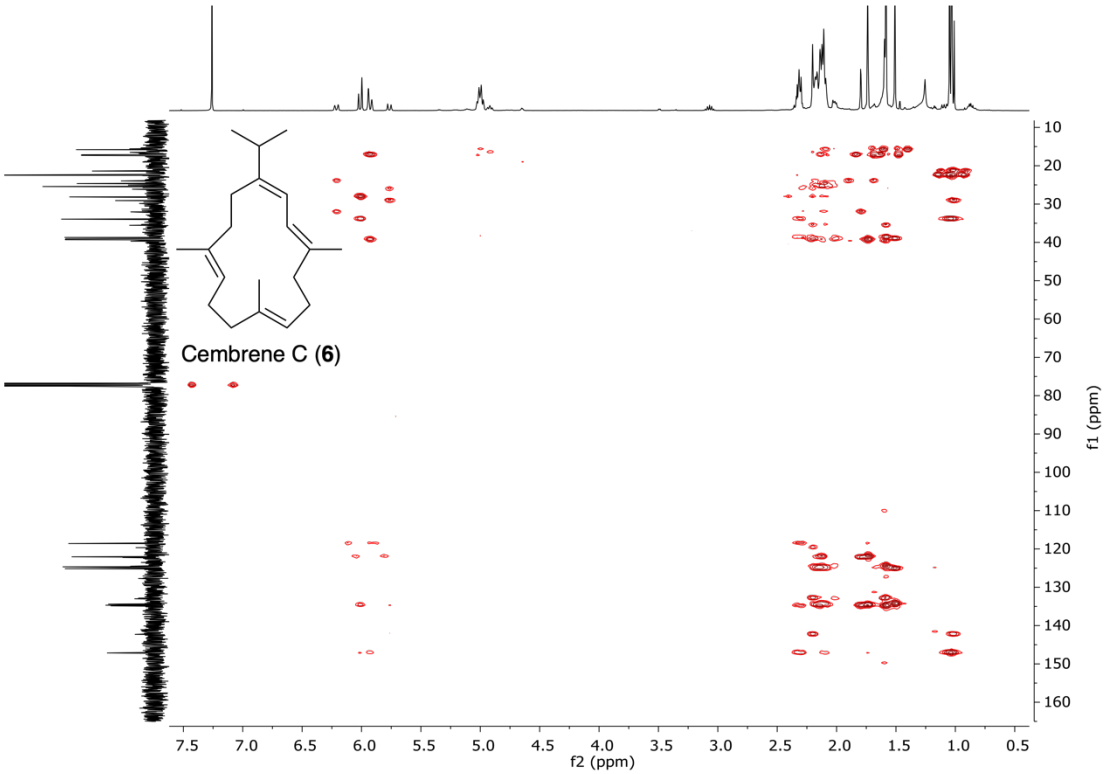


Figure S17. ^{1}H-^{13}C HMBC spectrum of cembrene C (**6**) in $CDCl_3$.

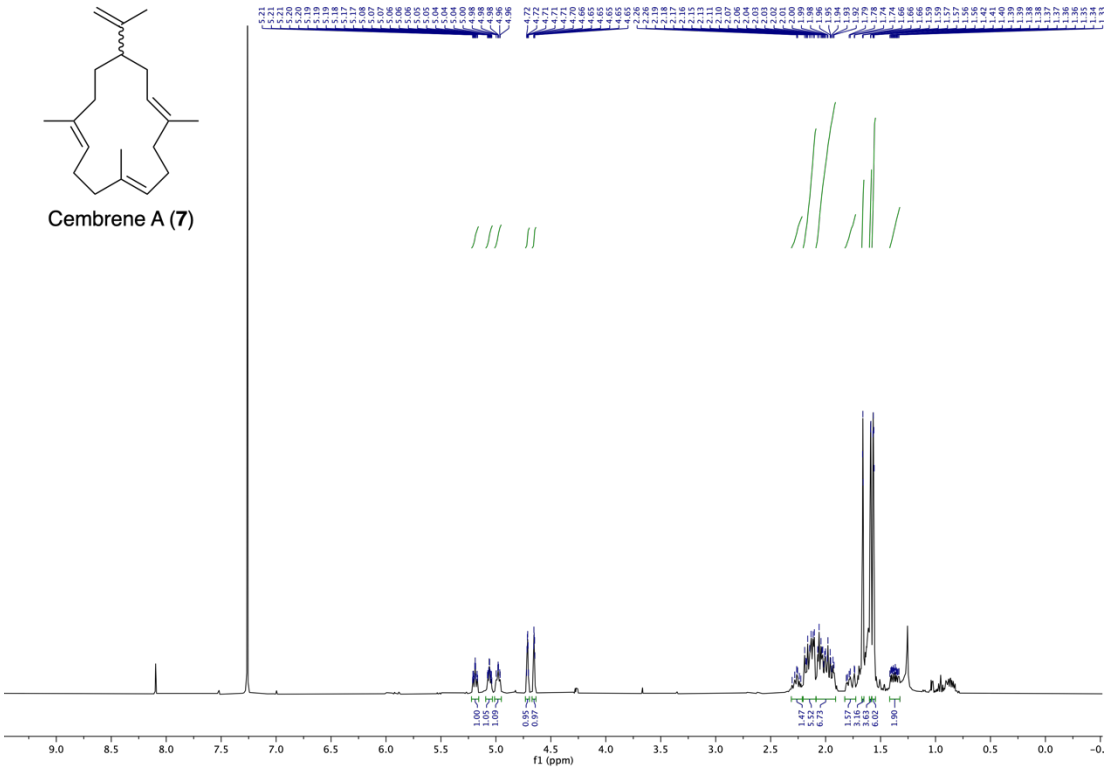


Figure S18. ^{1}H NMR spectrum of cembrene A (**7**) in $CDCl_3$ (400 MHz).

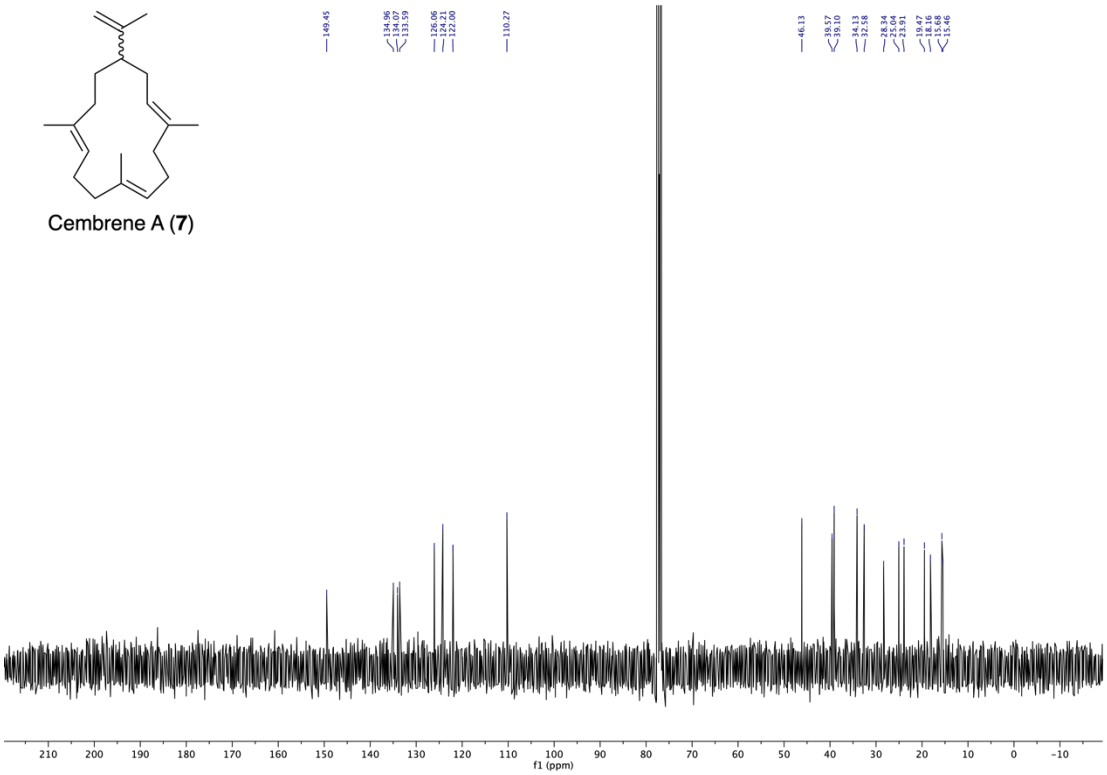


Figure S19. ^{13}C NMR spectrum of cembrene A (**7**) in $CDCl_3$ (100 MHz).

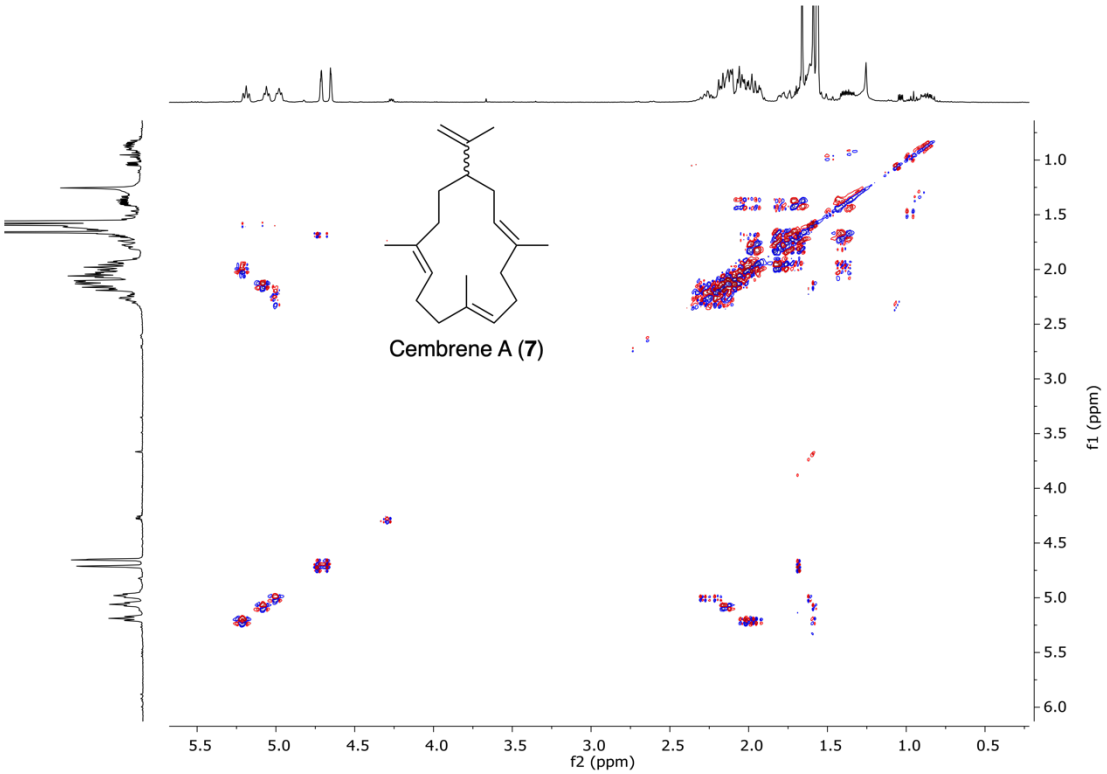


Figure S20. ^{1}H - ^{13}C HSQC spectrum of cembrene A (**7**) in $CDCl_3$.

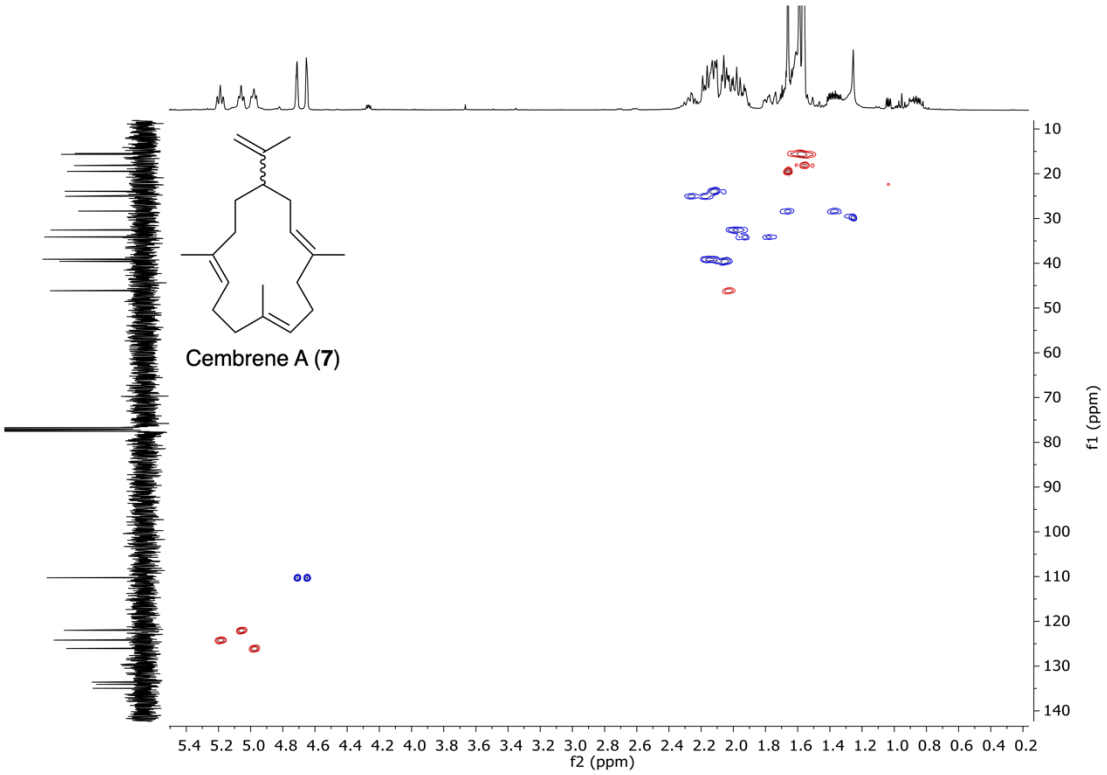


Figure S21. ^{1}H – ^{1}H COSY spectrum of cembrene A (**7**) in $CDCl_3$.

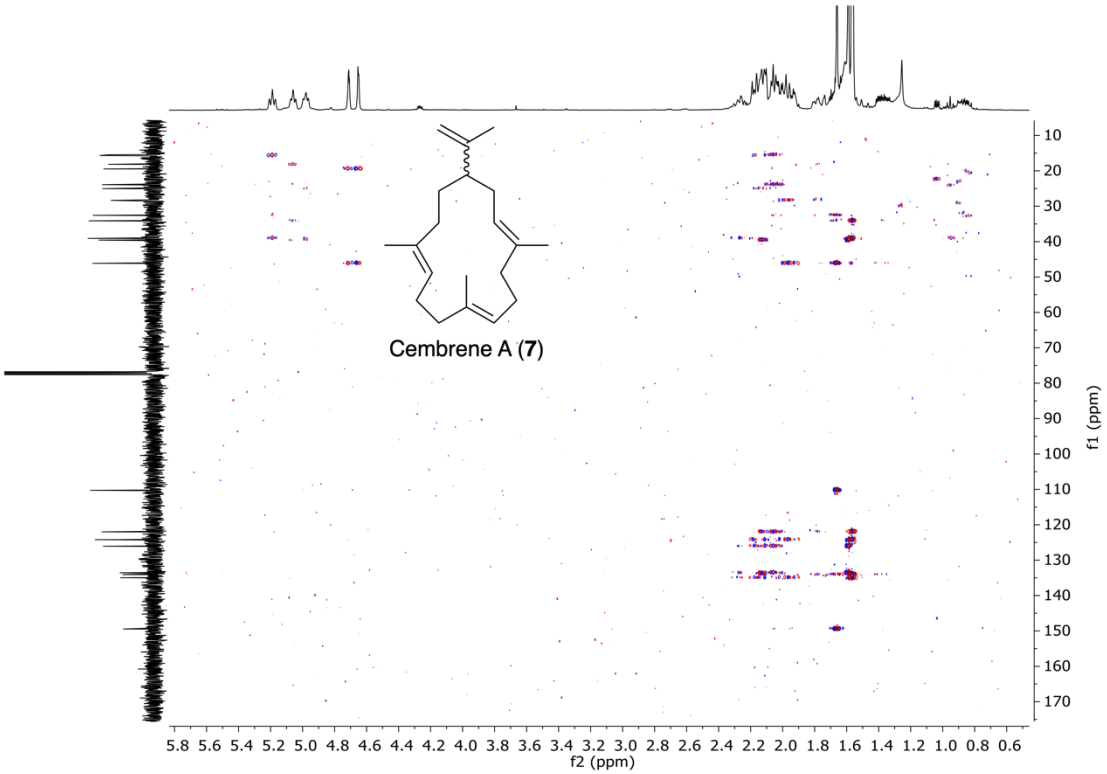


Figure S22. ^{1}H – ^{13}C HMBC spectrum of cembrene A (**7**) in $CDCl_3$.

```
           1         10         20         30         40         50
Bnd4       MSTIPKPFSDGTEIYLPEFPYLLPACSHPRTDDIRAESDAWVKEAMGFAMTDPRE
Rx_0493    .....MSTENGHRDGLGPLRCPFPAAINPHADEVHRETVEWAE...GFGLLGPGD
CAS        .......MAGTAKFSLPELTYPYPRRIHPETGEAEQHMLDWLA...GHGLIRGPE
DtcycA     ...MTDPAVTPLAFSIPQLYCPFPTAIHPEVDTLTRAGMDFMT...HHGFCNTEA
DtcycB     .........MDLPPALLSFYCPIASEVSPEHEAVAQEMYAWIH...AMSLTSDNR

                60         70         80         90        100
Bnd4       MELL.LEETASLWTCLV..LPTAREDRLRHLCKYTEYLSVFDNAMVDRAKIGKDP
Rx_0493    GHRMMRDTGIGRLAGRF..HPGAGREELRLISDWCAWMFLRDDLADAPAYFRHP.
CAS        VHEAFLRTGFAGLVGEE..YAAADSEGLRVVANFYGWMFVMDDYVTDTAAFGKDL
DtcycA     DRLVVANIDAGAIVARWYPNPDFPVDRLQMVTDFLYLYFLIDDLRFEVINSDTGL
DtcycB     QAKMLAQ..AGAGFNSYFT.PRARGELARALSKYNVCAWIANGMVQEI..RDPGT

           110        120                                130        140
Bnd4       AAAQETFRRVAGILD.DQADG..................ADFAWGSVLHGLWKD
Rx_0493    ERLAALDAAFLDILS...........GRAS........GEGCGSFGRALRDLRER
CAS        GKLSAFTA.WMRQLCDSPTDRTHLAMGQAATENMSGPARDFCHQIAEAAADLFGA
DtcycA     AGPIALFAQHLDLWEYPQAHR..................REELDLFHQAIHDLASR
DtcycB     FGAMAA..RWARIMEEPATCP..................ADGIPM.DFALADAFSH

                150        160        170        180        190
Bnd4       MRA.DMPPAVWDRFMGEVRRFLAGCVHEITSRSEDRVFDYETYIEVRKDSVGMGM
Rx_0493    LLPKVPAPLWLRRFLRSVEEHFESTLWEATNRARGVVPDLETYLRMRPITGGMHV
CAS        VAD.RATPSQYMRLVAEMSYFFQGMQWEAGHHVAGTLPTPDEYVIGRRMTSATPA
DtcycA     MAE.LTTPTKAARMRRSINGWFLALLREIALFNDDHAVMAEEYLPIRVVTVASRL
DtcycB     IRR.TLSPVKWQHFSAAQSHWMHGLAWENCLH.QVKGLTVHDYLSFRYVMSGCFA

             200        210        220        230        240
Bnd4       YFV.LGEYGLGIDLTEDLRRHRELREIVDTALVHIMLTNDMFSFRAEAMMDDY..
Rx_0493    DTD.FIEISSGVYLPPEVRRHPAVSALTGASNNVVCWANDIISLAKERSRGDV..
CAS        GLA.LQDIAAGYEVPANDYHQPRLRELRAMTANINSWCNDIFSYGKESDSPDT..
DtcycA     MIDVNGFICPA.EVPGDEWYSLKVQAAAEAAMSVCLYDNELYSAGKEQWLKSRAT
DtcycB     .AAAFAYAVPERHPSAEEWAHPKVRAAADAAMMVDALDNDRYSYLKESLTEAD..

                250        260        270        280        290
Bnd4       ....VNALSVL..RLSEGLGLQEAVDRLFALVDGKRAEFMAARAAIEAGELGRRE
Rx_0493    ....HNLVLVL..RASRRLTTREAVAEAARMYEAEVRRFVRLERELPPFGPAIDA
CAS        ....AVLNLPASLVRYHGYSEQGAIEEAARRHNDEVMNYLAAEEAV...ARDAGP
DtcycA     AHDRRPRNLVALIQAQTGGSTEHALQEVAEYRNRTVCLYLNLRSQL...EKTASP
DtcycB     ...KK.TIFAALRHENPALGREEVIVRGVQLRDRILTLYLTLRGEL...LCDASE

           300        310        320        330        340        350
Bnd4       DIRAYLDALWHMMAGNLQWSYL.TSRYNGPGHRWNGVRSGVLTLHRDRTVFSDRA
Rx_0493    NLRRYVSVLKSRMRGNLDWTYE.SARYRAGAASR.....................
CAS        EVMRFLDAMRMMQRGFYDWGLT.TSRYNVRRYFTNCPAPTAAAPH..........
DtcycA     ALLAYLSVLDGVISGNLDAHAT.SSRYHNPDGHHPHAI..AFTPLRTT......D
DtcycB     GLRSYLTGLDLIIAGNLVFCADMGLRYGLPEGSVRTDA....EPL..........

                360
Bnd4       YCSLPRAE.EVPA..........
Rx_0493    .......................
CAS        .......................
DtcycA     ECSARAHTPIAPPIAWWWEQLDQ
DtcycB     ..DRTVAPPGIGAIDHWWAQAGA
```

Figure S23. Sequence alignment of Bnd4 with bacterial cembrene synthases. Residues are colored based on the level of conservation (red box, red character, and blue frame show identity, similarity, and similarity across groups, respectively). The first residue of the WxxxxxRY motif, corresponding to W316 of Bnd4 and A321 of DtcycA, is marked with an arrowhead. Clustal Omega[16] and ESPript[17] were used to generate and render the alignment.

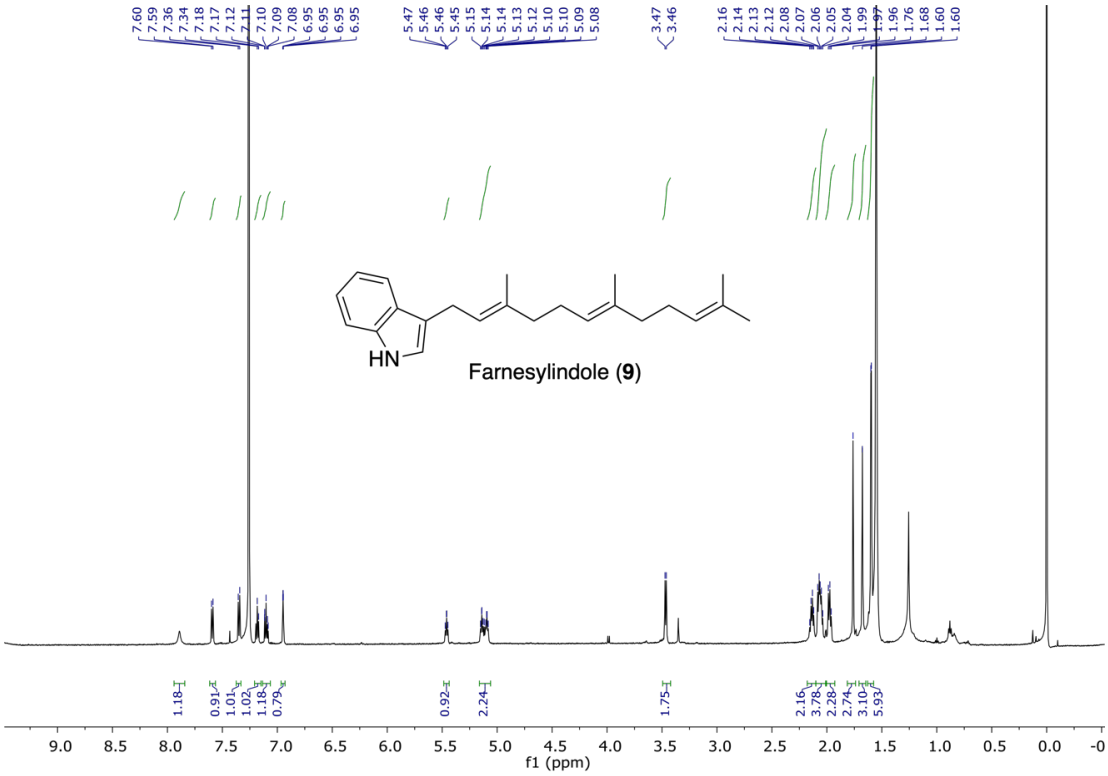


Figure S24. ^{1}H NMR spectrum of farnesylindole (**9**) in $CDCl_3$ (400 MHz).

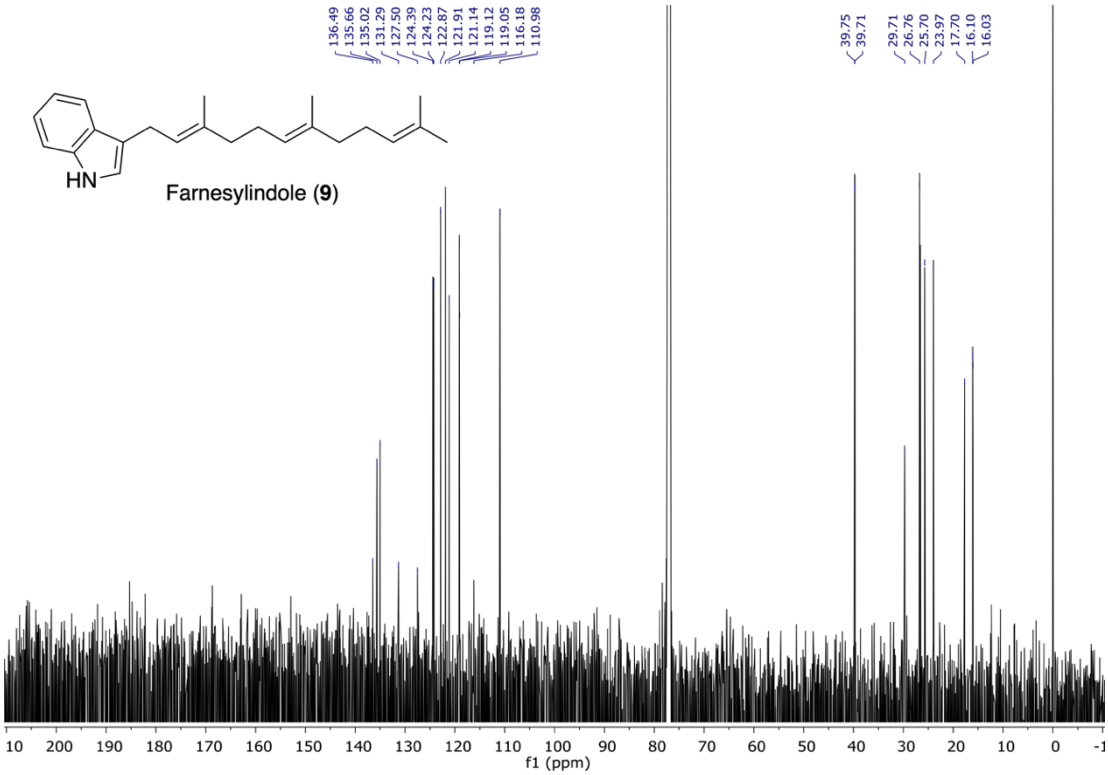


Figure S25. ^{13}C NMR spectrum of farnesylindole (**9**) in $CDCl_3$ (100 MHz).

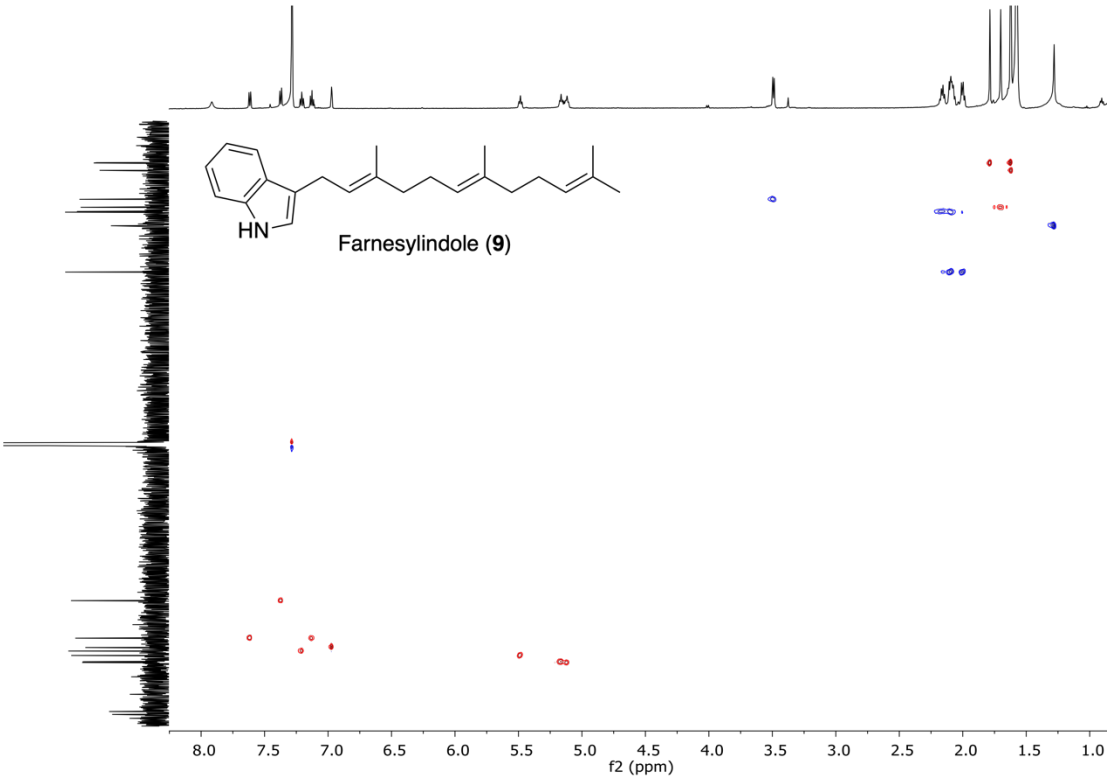


Figure S26. HSQC spectrum of farnesylindole (**9**) in $CDCl_3$.

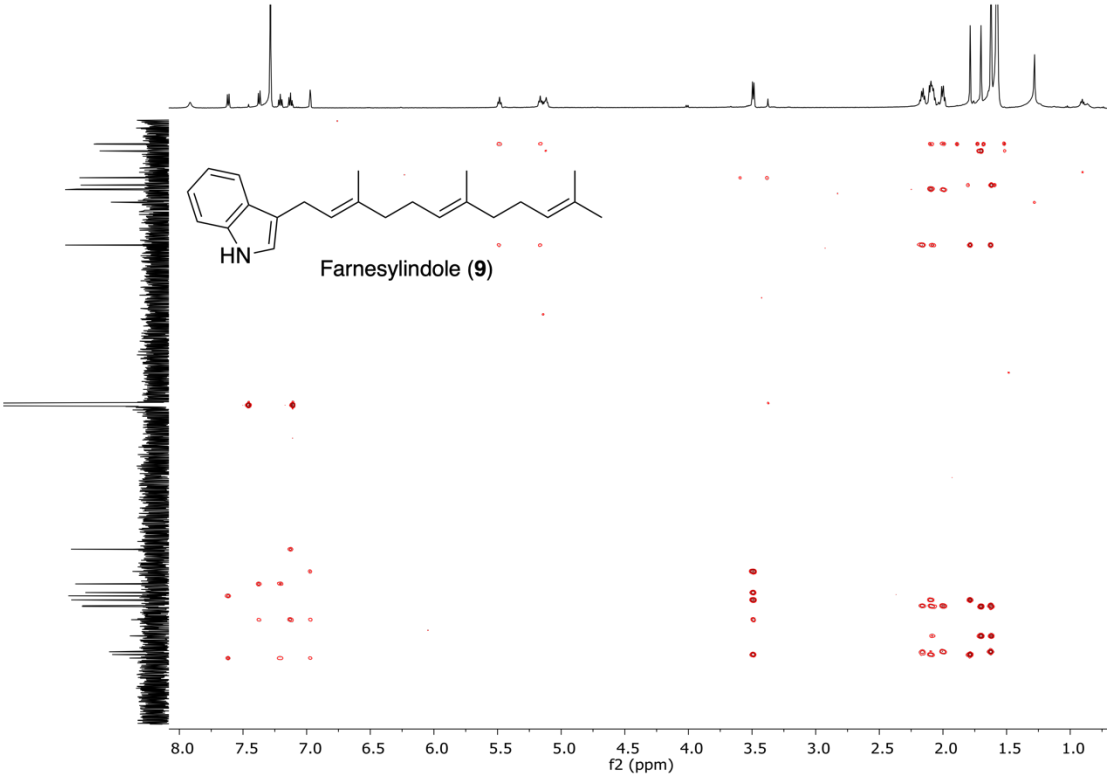


Figure S27. ^{1}H-^{13}C HMBC spectrum of farnesylindole (**9**) in $CDCl_3$.

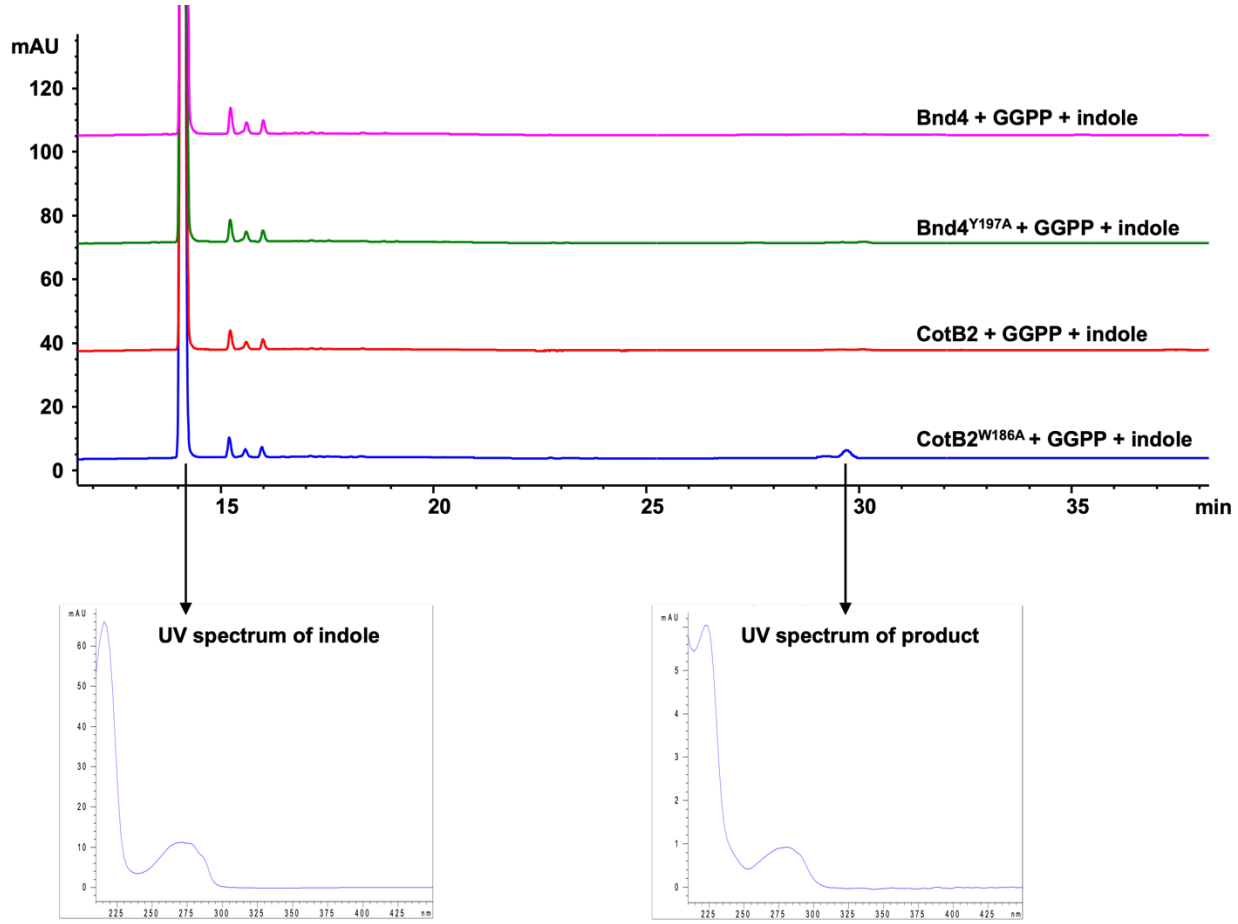


Figure S28. Prenylation test with GGPP and terpene synthase mutants. Bnd4^{Y197A} and CotB2^{W186A} did not catalyze geranylgeranylation of indole. Trace amounts of putative geranylgeranylindole were detected, although the low yield precluded confirmation by isolation and structural determination.

Supporting References

1. Xia, Y.; Li, K.; Li, J.; Wang, T.; Gu, L.; Xun, L. T5 exonuclease-dependent assembly offers a low-cost method for efficient cloning and site-directed mutagenesis. *Nucleic Acids Res.*, **2019**, *47*, e15.
2. Kruger, N. J. The Bradford method for protein quantitation. In *The Protein Protocols Handbook*; Walker, J. M., Ed., Springer, **2009**; pp 17–24.
3. Chatzivasileiou, A.O.; Ward, V.; Edgar, S. M.; Stephanopoulos, G. Two-step pathway for isoprenoid synthesis. *Proc. Natl. Acad. Sci. U. S. A.*, **2019,** *116*, 506–511.
4. Clomburg, J. M.; Qian, S.; Tan, Z.; Cheong, S; Gonzalez, R. The isoprenoid alcohol pathway, a synthetic route for isoprenoid biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.*, **2019,** *116*, 12810–12815.
5. Lund, S.; Hall, R; Williams, G. J. An artificial pathway for isoprenoid biosynthesis decoupled from native hemiterpene metabolism. *ACS Syn. Biol.*, **2019**, *8*, 232–238.
6. Zhu, C.; Xu, B.; Adpressa, D. A.; Rudolf, J. D.; Loesgen, S. Discovery and biosynthesis of a structurally dynamic antibacterial diterpenoid. *Angew. Chem., Int. Ed.*, **2021**, *60*, 14163–14170.
7. Huang, A. C.; Kautsar, S. A.; Hong, Y. J.; Medema, M. H.; Bond, A. D.; Tantillo, D. J.; Osbourn, A. Unearthing a sesterterpene biosynthetic repertoire in the Brassicaceae through genome mining reveals convergent evolution. *Proc. Natl. Acad. Sci. U. S. A.*, **2017**, *114*, E6005–E6014.
8. Baeckström, P.; Li, L. Syntheses of all-trans acyclic isoprenoid pheromone components. *Tetrahedron*, **1991**, *47*, 6533–6538.
9. Mu, Y.; Eubanks, L. M.; Poulter, C. D.; Gibbs, R. A. Coupling of isoprenoid triflates with organoboron nucleophiles: synthesis and biological evaluation of geranylgeranyl diphosphate analogues. *Bioorg. Med. Chem.*, **2002**, *10*, 1207–1219.
10. Rinkel, J.; Rabe, P.; Chen, X.; Köllner, T. G.; Chen, F.; Dickschat, J. S. Mechanisms of the diterpene cyclases β-pinacene synthase from *Dictyostelium discoideum* and hydropyrene synthase from *Streptomyces clavuligerus*. *Chem. Eur. J.*, **2017**, *23*, 10501–10505.
11. Li, H.; Sun, Y.; Zhang, Q.; Zhu, Y.; Li, S.-M.; Li, A.; Zhang, C. Elucidating the cyclization cascades in xiamycin biosynthesis by substrate synthesis and enzyme characterizations. *Org. Lett.*, **2015**, *17*, 306–309.
12. Peralta-Yahya, P. P.; Ouellet, M.; Chan, R.; Mukhopadhyay, A.; Keasling, J. D.; Lee, T. S. Identification and microbial production of a terpene-based advanced biofuel. *Nature Commun.* **2011**, *2*, 1–8.
13. Xu, B.; Li, Z.; Alsup, T.A.; Ehrenberger, M.A.; Rudolf, J.D. Bacterial diterpene synthases prenylate small molecules. *ACS Catal.*, **2021**, *11*, 5906–5915.
14. Yamada, Y.; Tomohisa, K.; Komatsu, M.; Shin-ya, K.; Omura, S.; Cane, D. E.; Ikeda, H. Terpene synthases are widely distributed in bacteria. *Proc. Natl. Acad. Sci. U. S. A.*, **2015**, *112*, 857–862.
15. Meguro, A.; Tomita, T.; Nishiyama, M.; Kuzuyama, T. Identification and characterization of bacterial diterpene cyclases that synthesize the cembrane skeleton. *ChemBioChem*, **2013**, *14*, 316–321.
16. 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*, 539.

17. Robert, X.; Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. **2014**, *42*, W320–W324.