Supplementary Information for:

Terpene biosynthesis in marine sponge animals

Kayla Wilson[1], Tristan de Rond[1,2], Immo Burkhardt[1], Taylor S. Steele[1,3], Rebecca J. B. Schäfer[1], Sheila Podell[1], Eric E. Allen[1], Bradley S. Moore[1,4, 5]

[1]Scripps Institution of Oceanography, University of California San Diego; 9500 Gilman Dr., La Jolla, CA 92093, USA.
[2]School of Chemical Sciences, University of Auckland, Auckland, New Zealand
[3]Department of Chemistry & Biochemistry, University of California, San Diego, La Jolla, California 92093, United States
[4]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego; 9500 Gilman Dr., La Jolla, CA 92093, USA.
[5]Corresponding author, email: bsmoore@ucsd.edu

Table of Contents

Supplementary Methods

Collection of marine sponges
Undescribed Bubarida sponges were collected by SCUBA from the Pacific Ocean near San Diego, California (N32° 51.344', W117° 16.621'), United States of America at approximately 15 meters depth and 15 °C temperature. The undescribed Bubarida sponges were dark orange to brown and lumpy with a shiny appearance at the surface. They tend to be found on rocky substrates, sometimes in kelp forests, often surrounded by other invertebrates and algae (Fig. S3). All specimens were collected under California Fish and Wildlife Scientific Collection Permits SC-13980 and SC-192800001. The freshly collected sponge was cut into approximately 8 cm^3 pieces and separated into groups based on sample usage. Sponge tissue for chemical extraction was flash frozen in 50 mL Falcon tubes submerged in liquid nitrogen and stored at -80 °C. Sponge tissue for morphological preservation and spicule analysis was flash frozen with liquid nitrogen in 50 mL Falcon tubes filled with calcium and magnesium free artificial sea water with glycerol (449 mM sodium chloride, 33 mM sodium sulfate, 9 mM potassium chloride, 2.5 mM sodium bicarbonate, 1 mM disodium EDTA, 15% v/v glycerol, 0.2 μM filtered) and stored at -80°C. Sponge tissue for DNA and RNA sequencing was flash frozen in liquid nitrogen in 50mL Falcon tubes filled with RNAlater (25 mM sodium citrate, 10 mM disodium EDTA, 700 g/L ammonium sulfate, pH 5.2) and stored at -80°C.
Agelas sponges were collected by SCUBA divers from Sirenas Marine Discovery (San Diego, CA, USA) off St. Thomas, Virgin Islands (N18°18.571', W64°57.681') at a depth of approximately 12 meters, and processed for metagenomic DNA extraction, Illumina paired-end sequencing, and assembly as previously described (1).

DNA extraction
For Illumina sequencing: Sponge tissue in RNAlater was thawed at room temperature and rinsed with 10 mL sterile calcium- and magnesium-free artificial sea water with glycerol (449 mM sodium chloride, 33 mM sodium sulfate, 9 mM potassium chloride, 2.5 mM sodium bicarbonate, 1 mM disodium EDTA, 15% v/v glycerol, 0.2 μM filtered). The rinsed sponge tissue was thoroughly homogenized with liquid nitrogen and a mortar and pestle. The homogenized sponge powder was transferred to a 2mL microcentrifuge tube and resuspended in 200μL chaos buffer (5M guanidine thiocyanate, 2% Sarkosyl, 50mM EDTA, 0.2μM filtered). The resuspended sponge powder was transferred to a Qiagen 0.1 mm glass bead-filled screw top plastic tube. 5μL of Proteinase K (20mg/mL, Zymo Research cat. No. D3001-2-20) and 50μL 2-mercaptoethanol was added to the tube and the tube was incubated at 55°C for 1 hour, with inversion by hand every 15 minutes. The tube was centrifuged at maximum speed for 3 minutes and the supernatant was pipetted off and saved in a separate microcentrifuge tube. 600μL of solid tissue buffer blue (Zymo Research cat. no. D4068-2-10) was added to the glass bead-filled tube containing the tissue pellet. The glass bead filled tube was homogenized on setting 4.0 for 5 seconds. The glass bead-filled tube was centrifuged at maximum speed for 30 seconds and the reserved supernatant was added back into the glass bead-filled tube. 500μL of phenol:chloroform:isoamyl alcohol (25:24:1) was added to the glass bead-filled tube, rotated by hand very gently for 5 minutes, and centrifuged at maximum speed for 5 minutes. The top aqueous layer was removed to a new microcentrifuge tube. A Zymo Research Quick-DNA Miniprep Kit (cat. no. D3024) was used for the following steps. 0.33 volumes of isopropanol was added to the aqueous layer and loaded onto a Zymo-Spin IICR column. The column was centrifuged at 10,000g for 1 minute. The outside of the column was rinsed with 500μL ethanol and the column was transferred to a new microcentrifuge tube. 200μL of Zymo PreWash Buffer was added to the column and the column was centrifuged at 10,000g for 1 minute. 5μL RNAse A (10mg/mL, Zymo Research cat. no. E1008) in 95 μL Zymo Wash Buffer was added to the column. The column in the tube was sealed with Parafilm and incubated at 37°C for 45 minutes. The Parafilm was removed, and the column was centrifuged at 10,000 g for 2 minutes. The column was air dried with the lid open for 5 minutes then eluted with 2 x 50μL of 40°C Zymo Elution Buffer. The concentration and purity of the DNA was evaluated via Qubit and Nanodrop. The DNA was stored at -80°C.
For Oxford Nanopore sequencing: A QIAGEN Genomic Tip 500/G Blood & Cell Culture DNA Maxi kit (cat. no. 13362) was used for high molecular weight DNA extraction. 400mg of frozen sponge stored in RNAlater (25mM sodium citrate, 10mM disodium EDTA, 700g/L ammonium sulfate, pH 5.2) was thawed and removed from excess RNAlater. The sponge tissue was chopped up with a sterile razor blade then

thoroughly homogenized with liquid nitrogen and a mortar and pestle. The QIAGEN Protocol: Preparation of Tissue Samples (page 31-36, June 2015 QIAGEN Genomic DNA Handbook) was used to prepare the sponge tissue for Genomic-tip loading. The Protocol: Isolation of Genomic DNA from Blood, Cultured Cells, Tissue, Yeast, of Bacteria using Genomic-tips was used to isolate DNA from the prepared sponge tissue sample (page 49-52, June 2015 QIAGEN Genomic DNA Handbook). The high molecular weight DNA was size selected using a Sage Science BluePippin and High Pass Plus >15kb Gel Cassette (Sage Science part no. BPLUS10). The size selected DNA was cleaned up using AMPure XP beads (Beckman Coulter, Item no. A63880) following the PCR Purification Process Procedure. The concentration and purity of the DNA was evaluated via Qubit and Nanodrop. The high molecular weight DNA was stored in elution buffer (10mM Tris-Cl, pH 8.5) at 4°C.

28S rRNA Barcoding

PCR amplification of 28S rDNA was performed using the 28S-C2-fwd (GAA AAG AAC TTT GRA RAG AGA GT) and 28S-D2-rev (TCC GTG TTT CAA GAC GGG) primers (2). The PCR mix consisted of 10 µL 5PRIME HotMasterMix (QuantaBio, Beverly, MA, USA), 1 µL each primer (10 µM), 1 µL template DNA, 12 µL water. The thermocycler program was as follows: initial denaturation phase of 94 °C for 2 min followed by 30 cycles of 30 s denaturation at 94 °C, 30 s annealing at 48 °C, 1 min elongation at 65 °C, and a final elongation at 65 °C for 5 min. The 28S sequence is identical to that of a Californian sponge belonging to the order Bubarida soon to be formally described by Turner and Lonhart, 2022 (3). The 28S amplicon sequence from the San Diego undescribed Bubarida is as follows:
AACGAATGCAGCCCGATGGGCTTCGGTCGGGTTCAGGGGCGGTGGTGCCGTCGGAGCGGGCGGG
CGGACCCGTGAGGGTCGCCTTTCAGTTCTCGAGCGGTCCATCGCCCTGCACTCCCGGTCGATAGCC
GGCCAACGTCGGTCCGGTTCGGCTCACGCGGATGGTCGGGCAGGTGCCTCAATGCGCCGCTTCGG
CGTGCGCAGGGAGAACTTACAGCCGATTGTTCGGCAGTCGGACGTTGCGGACCGAGGAGATGTGT
GACGCTTCACACCGGCAGTGCAGGCTCGCCGTTCTCGCGGTGGGCCGGGTCTCTTCTGATCGTTG
GGTAGGCGCTACTGGGACTGCCCGTCAGTGCCAGTCGACGTCCCGTCCCCTCGGTCGGGAGGTTG
GTCACGCCTTGCGCTGTAGTCGCTGGCTGAAGGAAGGCTGCATCCGACCCGTCTTGAAACACGGA

Illumina Sequencing

For metagenome sequencing, 375 ng of DNA from each sample was fragmented by Adaptive Focused Acoustics (E220 Focused Ultrasonicator, Covaris, Woburn, Massachusetts) to produce an average fragment size of 500 basepairs (bp). Sequencing libraries were generated using the KAPA Hyper Prep Kit (KAPA Biosystems, Wilmington, MA, USA) following manufacturer's instructions using 5 cycles of amplification at the IGM Genomics Center, University of California, San Diego, La Jolla, CA. Resulting libraries were multiplexed with another sponge sequencing project and sequenced with 250 basepair (bp) Paired End reads (PE250). 101.3 Gbp of sequence with was generated using an Illumina NovaSeq Sequencing System (Illumina, San Diego, CA, USA) at the Center for Advanced Technology, University of California, San Francisco, San Francisco, CA. NCBI BioProject number: PRJNA907134. NCBI Sequence Read Archive accession: SRR22527615

For RNA sequencing, total RNA was assessed for quality using an Agilent Tapestation 4200, and RNA sequencing libraries were generated using 500 nanograms of RNA and two methods 1) TruSeq Stranded mRNA Prep (Illumina, San Diego, CA) and 2) TruSeq Stranded Total RNA Library Prep Gold supplemented with Illumina ribodepletion probes for Bacteria. Samples were processed following manufacturer's instructions with the shearing time modified to five minutes. Resulting libraries were multiplexed with another sponge sequencing project and sequenced with 100 basepair (bp) Paired End reads (PE100) on the HiSeq 4000 Sequencing System (Illumina, San Diego, CA, USA) at the IGM Genomics Center, University of California, San Diego, La Jolla, CA. Samples were demultiplexed using bcl2fastq Conversion Software (Illumina, San Diego, CA). 51.3 Gbp and 8.1 Gbp of sequence was generated for the stranded mRNA library and ribodepleted libraries respectively. NCBI BioProject number: PRJNA907134. NCBI Sequence Read Archive accession: SRR22527612 (ribodepleted RNA) and SRR22527613 (mRNA)

Oxford Nanopore Sequencing

Size-selected, high-molecular weight (HMW) DNA was sequenced using the Oxford Nanopore MinION platform (Oxford Nanopore Technologies (ONT), Oxford, UK). A one-dimensional (1d) library was prepared with 600 ng HMW DNA using the Ligation Sequencing Kit (SQK-LSK109) (ONT, Oxford, UK).

This library was loaded onto a R9.4 flowcell and run for 48 hr, resulting in 960 Mb of sequence with a read length N50 of 22 kb. The data from this was further processed and used for assembly as described below. NCBI BioProject number: PRJNA907134. NCBI Sequence Read Archive accession: SRR22527614

Metagenome and transcriptome assembly
Raw nanopore data was basecalled with Guppy v. 5.0.11+2b6dbffa5 using parameters --flowcell FLO-MIN111 and --kit SQK-LSK109, and trimmed with Porechop 0.2.4 using default parameters. Raw Illumina metagenome data was trimmed using Trimmomatic (4) 0.39 using parameters ILLUMINACLIP:adapters.fa:2:30:10:3:true HEADCROP:10 SLIDINGWINDOW:8:20 LEADING:10 TRAILING:10 MINLEN:40. A hybrid genome assembly was performed using Masurca v3.4.2 (5) using parameters NANOPORE= nanopore_data.fastq; PE= ns 410 60 novaseq_data_1.fastq novaseq_data_2.fastq; PE= hs 410 60 hiseq_data_1.fastq hiseq_data_2.fastq; FLYE_ASSEMBLY = 0; CA_PARAMETERS = cgwErrorRate=0.15; LIMIT_JUMP_COVERAGE = 300; MEGA_READS_ONE_PASS=0; USE_LINKING_MATES = 1; EXTEND_JUMP_READS=0. The raw Illumina data were then mapped to this assembly using bwa-mem v2.2.1 using default parameters, the mapped data was used to polish the Masurca assembly with Pilon v1.24 (6) using default parameters, and the resulting polished assembly was polished again in the same fashion. The hybrid assembly features 5.5k contigs larger than 5k bp, totaling 180M bp. Eukaryotic genome assembly completeness as estimated by BUSCO v. 5.1.2 (7) using lineage dataset metazoa_odb10 and Augustus as gene predictor was 73.6% (C:73.6%[S:69.2%,D:4.4%],F:5.0%,M:21.4%,n:954)
Genome assemblies using Flye, Metaspades and HASLR were also attempted but yielded lower quality assemblies than Masurca.
Transcriptome assembly from raw Illumina RNASeq data was performed using Trinity v2.8.4 using default parameters.

Assembly visualization using Blobtools
Blobtools v2.6.1 (8) was used to generate %GC vs coverage and snail plots as follows: Two Diamond databases, one generated from the NCBI nr dataset and one custom-built from two publicly-available *Halichondria panicea* transcriptomes (HBWD00000000 and GIFJ00000000) (9, 10), were queried for the assembly's contigs using Diamond v2.0.13(11) blastx using parameters --evalue 1e-25 --long-reads, the results concatenated and used to assign taxonomy to assembly contigs using blobtools.
Independently, trimmed reads were mapped to the assembly using bwa-mem2, position-wise coverage calculated using bedtools v2.27.1 genomecov, contig-wise coverage modes and medians calculated using awk 'BEGIN{OFS="\t";current=$1;median_cov=-1};$1!=current{print current,bestcov,median_cov;current=$1;bestcovfreq=-1;bestcov=-1;portion_along=0;median_cov=-1};{portion_along+=$5};median_cov==-1&&portion_along>=0.5{median_cov=$2};$3>bestcovfreq{bestcovfreq=$3;bestcov=$2}' and added to the blobdir using blobtools.

Chemical Extraction
Methylene Chloride Extractions for GCMS: 5 g of frozen undescribed Bubarida sponge was homogenized with liquid nitrogen and mortar and pestle. The sponge powder was extracted with 200 mL methylene chloride. The extract was vacuum filtered and carefully concentrated using a rotary evaporator to 20 mL final volume. Extract was stored at -20 °C.

Methanol Extractions for LCMS: 500 mg of frozen, dry undescribed Bubarida sponge was homogenized with liquid nitrogen and mortar and pestle. The sponge powder was extracted with 4 mL methanol for 1.5 hrs. The extract was filtered, concentrated to dryness under nitrogen gas, and redissolved in 2:1 methylene chloride:methanol. The extract in methylene chloride and methanol was concentrated to dryness under nitrogen gas, filtered, redissolved in methanol, and stored at -20 °C.

Chemical analysis
LCMS
LC-MS analysis (Fig. S50) was performed with an Agilent Technologies 1260 Infinity series HPLC equipped with a degasser, binary pump, autosampler, and diode array detector coupled to an Agilent

Technologies 6530 Accurate-Mass Q-TOF LC/MS. The instrument was calibrated using the Agilent Reference Calibration Mix®. Separations were performed with a Phenomenex Luna 5 µm C18(2) 100 Å LC Column 250 x 4.6 mm. The solvent system contained water + 0.1% formic acid as solvent A, and acetonitrile + 0.1% formic acid as solvent B. Solvents were of mass spectrometry-grade. Data were collected and analyzed using MassHunter Workstation Software version B.05.01. 0.7 mL min^{-1} flow rate; gradient: 0-5 min 5% B, 5-15 min 5-100% B, 15-20 min 100% B, 20-21 min 100-5% B, 21-26 min 5% B. Dual ESI ion source, positive polarity. T = 300 °C; Acquisition mass range: 100-1700 m/z; Sample was prepared in MeOH; V_{inj} : 10 µL.

GCMS

GCMS analysis was performed on an Agilent 7890A gas chromatograph with Agilent 5975C mass spectrometer, using a HP-5MS 30 m × 0.25 mm, 0.25 µ column. Oven program was as follows: hold 70 °C for 3 min, 10 °C/min to 325 °C, hold at 325 °C for 3 min. Flow rate: 0.44 mL/min, Injection volume: 1 µL, splitless injection, inlet temperature: 200 °C, MS transfer line temperature: 250 °C, MS source temperature 230 °C, MS quad temperature: 150 °C.

Terpene synthase discovery

TSHMM1, a terpene synthase profile HMM created from biochemically characterized terpene synthase (TS) genes (Table S3) using HMMER 3.1b2, was used to query the undescribed Bubarida transcriptome and the *Agelas clathrodes and Agelas tubulata* metagenomes (NCBI BioProject PRJNA824609) for potential TSs, as previously described (12, 13).

Searching for homologs of undescribed Bubarida terpene synthases using NCBI Blast in the publicly available sponge nuclear genomes, only gene fragments could be detected in *Stylissa carteri* (http://sc.reefgenomics.org/) and *Stylissa massa* holobiont metagenomes (JGI Integrated Microbial Genomes and Metagenomes Database id numbers Ga0072505, Ga0072501, Ga0070431) (14).The synthesized stylissa was built from two overlapping partial TS sequences from Guamanian 2 stylissa massa sequencing experiments (Ga0070431_10856351 (C-term) and Ga0072505_11956581 (N-term) (Fig. S14)(15). No homologs could be detected in *Amphimedon queenslandica*, *Ephydatia muelleri*, *Halichondria panicea*, or *Xestospongia testudinaria*(15–17). Finally, all raw sequencing data, except for amplicon sequencing data, in the NCBI Sequence Read Archives associated with taxons within Porifera (taxid 6040) and “Sponge metagenome” (taxid 1163772) was downloaded, translated *in silico*, and queried for undescribed Bubarida terpene synthase homologs using the custom HMM described above (TSHMM1). 2469 reads from *Phakellia ventilabrum* RNASeq experiment SRX6964817 were detected, prompting us to assemble this raw data using Trinity with default parameters(18). Upon querying the resulting *P. ventilabrum* transcriptome assembly we identified 1 full-length terpene synthase homolog PhTS-1.

Homology modeling

Amino acid sequences of sponge TSs were submitted to Robetta (https://robetta.bakerlab.org/) for structure prediction using *Eleutherobia rubra* cembrane A synthase (PDB entry 7S5L) as a template. Sponge TS homology models were visualized using PyMOL (https://pymol.org/).

Phylogenetic analysis

Sequences of type I TSs were obtained from Burkhardt et al(13). All sequence alignments were generated with Kalign v.2.04 The phylogenetic analysis was performed using IQ-TREE multicore version 1.6.12 substitution model LG+R6 (19, 20). Trees were visualized using IToL and Adobe Illustrator. Plant TS sequences were manually truncated to their alpha domains.

Heterologous expression for analytical scale

Genes for heterologous expression were purchased from TWIST Biosciences as codon optimized (Twist algorithm), N-terminal hexahistidine-tagged sequences cloned into pET 28a (+) vectors (cut sites XhoI and NdeI). The plasmids were chemically transformed into *E. coli* BL21 (DE3). 15 mL LB medium (Fisher) with 50 mg/L kanamycin was inoculated with transformed bacteria and grown for 18 h at 37 °C and 180 rpm. 1 mL of these cultures was transferred to 100 mL terrific broth (Fisher, supplemented with 50 mg/L kanamycin) and grown to OD600=0.6-0.8 at 37 °C and 180 rpm. Cultures were subsequently cooled to 18

°C, induced with IPTG (400 µM) and shaken at 18 °C and 180 rpm for 18h. The cells were pelleted by centrifugation (8000 g, 15 min, 4 °C) and resuspended in TS-binding buffer (20 mM Na_2HPO_4, 0.5 M NaCl, 20 mM imidazole, 1 mM $MgCl_2$, pH 7.4, 5 mL buffer per 100 mL culture). The cells were lysed by sonication (1 sec on, 1 sec off pulses for 1 min, 30 sec off, 5x) and the debris was removed by centrifugation (11000 g, 20 min, 4 °C). The supernatant was filtered through a 0.22 µM membrane sterile filter (Merck Millipore) and loaded onto a pre equilibrated (TS-binding buffer) gravity flow column containing Chelating Sepharose Fast Flow resin (GE Healthcare, 0.5 mL) loaded with Ni cations, washed with TS binding buffer (2x 2 mL) and eluted with 2.5 mL TS-elution buffer (20 mM Na_2HPO_4, 0.5 M NaCl, 0.5 M imidazole, 1 mM $MgCl_2$, pH 7.4). The elution fractions were loaded onto PD-10 columns (GE Healthcare) that were pre equilibrated with 25 mL TS reaction buffer (50 mM HEPES, 5 mM $MgCl_2$, 250 mM NaCl, 10% glycerol, pH 7.8) and eluted with 3.5 mL TS reaction buffer (Fig. S51). The protein was either used in assays immediately or flash frozen in liquid nitrogen and stored at -80 °C until use.

Heterologous expression for preparative scale
3x 1 L TB media were inoculated with 1 vol% of an *E. coli* overnight culture carrying the expression plasmid for the respective TS. Cultures were grown to OD600 = 0.6-0.8 at 37 °C and 180 rpm, cooled to 18 °C, induced with IPTG (400 µM) and shaken at 18 °C and 180 rpm for 18 h. The cells were pelleted by centrifugation (8000 g, 15 min) and resuspended in TS-binding buffer (20 mM Na_2HPO_4, 0.5 M NaCl, 20 mM imidazole, 1 mM $MgCl_2$, pH 7.4, 20 mL buffer per 1L culture). The cells were lysed by sonication (15 sec on, 45 sec off pulses for a total on-time of 7 min) and the debris was removed by centrifugation (11000 g, 20 min, 4 °C), transferred to a new tube and centrifuged again (11000 g, 20 min, 4 °C) after DNaseI (Roche Diagnostics, 1 mg per 20mL supernatant) was added. The supernatant was filtered through a 0.22 µM membrane sterile filter (Merck Millipore) and loaded onto a 5-mL HisTrap-FF Ni-column (GE). The column was washed with TS-binding buffer (40 mL), eluted with TS-elution buffer (20 mM Na_2HPO_4, 0.5 M NaCl, 0.5 M imidazole, 1 mM $MgCl_2$, pH 7.4) and the elution was collected in 2 mL fractions. Enzyme containing fractions were pooled and the buffer was exchanged in 2.5 mL batches using PD-10 columns equilibrated with TS-reaction buffer (50 mM HEPES, 5 mM $MgCl_2$, 250 mM NaCl, 10% glycerol, pH 7.8) and eluted with 3.5 mL TS-reaction buffer per column used. The protein was either used in assays immediately or flash frozen in liquid nitrogen and stored at -80 °C until use.

Analytical scale enzyme assays
Analytical scale assays were carried out in 1.5 mL TS reaction buffer (50 mM HEPES, 5 mM $MgCl_2$, 250 mM NaCl, 10% glycerol, pH 7.8) with an enzyme concentration of 3 µM and 0.25 mg/mL GPP, FPP or GGPP to test mono-, sesqui-, and diterpene synthase activity. The assays were run for 18 h at room temperature and extracted with hexanes (200 µL). The extracts were analyzed by GCMS.

Preparative scale enzyme assays
For product isolation, FPP (80 mg, for sesquiterpene production) were dissolved in ammonium bicarbonate solution (15 mL, 50 mM) and added dropwise over 5 h to 600 mL TS reaction buffer (50 mM HEPES, 5 mM $MgCl_2$, 250 mM NaCl, 10% glycerol, pH 7.8) containing 1-6 µM enzyme at room temperature. The assay was run for 18 h at room temperature and extracted with pentane (3x 200 mL). The organic layer was dried over $MgSO_4$, filtered, and concentrated in vacuo to 0.5 mL. The crude product was purified by silica gel column chromatography. Fractions were checked by TLC (stained with phosphomolybdic acid stain) and product containing fractions were pooled, concentrated to dryness, and analyzed by NMR spectroscopy and polarimetry.

Analytical scale enzyme assays with potassium cyanide
Assays were carried out in 1.5mL of either pH5 reaction buffer (100 mM acetate, 5 mM $MgCl_2$, 250 mM NaCl, 10% glycerol, pH 5.0), pH6 reaction buffer (100 mM Na_2HPO_4, 5 mM $MgCl_2$, 250 mM NaCl, 10% glycerol, pH 6.0), or pH7 reaction buffer (100 mM HEPES, 5 mM $MgCl_2$, 250 mM NaCl, 10% glycerol, pH 7.0). Enzyme concentration was 3 µM, FPP concentration was 0.25 mg/mL, and potassium cyanide concentration was 10mM. The assays were run for 18 h at room temperature and extracted with hexanes (200 µL). The extracts were analyzed by GCMS.

Analytical scale IDS enzyme assays

Assays were carried out in 1.5mL of TS reaction buffer (50 mM HEPES, 5 mM $MgCl_2$, 250 mM NaCl, 10% glycerol, pH 7.8). IDS and TS enzymes were added to a final concentration of 3 μM. DMAPP was added to a final concentration of 0.25 mg/mL and IPP was added to a final concentration of 0.5 mg/mL (2 eq.) or 0.75 mg/mL (3 eq.). The assay was run at room temperature for 18 h and extracted with hexanes (200 μL). The extracts were analyzed by GCMS.

NMR spectroscopy
NMR spectra were recorded on a JEOL ECZ spectrometer (500 MHz) and were referenced against deuterated chloroform (δ = 7.26 ppm) for ^{1}H-NMR and deuterated chloroform (δ = 77.2 ppm) for ^{13}C-NMR.

Optical Rotation
Optical rotation measurements were taken with a Jasco P-2000 polarimeter.

Supplementary Tables

Molecule	Molecule Class	Gene Cluster	Producer Organism	Host Sponge	Citation
Onnamide	PKS	*onn*	Uncultured bacterial symbiont	*Theonella swinhoei*	(21)
Psymberin	PKS	*psy*	Uncultured symbiont	*Psammocinia* aff *bulbosa*	(22)
Theonellamide	PKS-NRPS	*tna*	*Candidatus* Entotheonella serta	*Theonella swinhoei*	(23)
Polybrominated diphenyl ethers	PDBE	*bmp*	*Hormoscilla spongeliae*	Dysideidae sponges	(24)
Kasumigamide	PKS-NRPS	*kas*	Uncultured 'Entotheonella' bacteria	*Theonella swinhoei*	(25)
Polytheonamide	RiPP	*poy*	Uncultured 'Entotheonella' bacteria	*Theonella swinhoei*	(26)
Misakinolide	PKS	*mis*	*Candidatus* Entotheonella serta TSWA1	*Theonella swinhoei*	(27)

Table S1. Natural product biosynthetic gene clusters from the sponge microbiome.

Sequencing dataset	Total sequence	Number of reads
Oxford Nanopore Minion metagenome	960 Mbp	76,640
Illumina Novaseq metagenome	101.3 Gbp	202,618,666
Illumina HiSeq4000 stranded mRNA-seq	51.3 Gbp	256,503,457
Illumina HiSeq4000 ribodepleted RNA-seq	8.1 Gbp	40,717,030

Table S2. Sequencing data summary.

E2E2N7	B6SCF5	I6RAQ6	J7LQ09	Q5SBP3
P59287	P0CJ43	O48935	Q4KSH9	Q84UU4
B3TPQ6	Q41594	B5A434	Q9FXY7	H8ZM70
C7E5V7	O24475	Q6PWU2	A4FVP2	C5YHH7
O49853	Q84LF0	J7LH11	B1B1U3	O81191
O64404	G1JUH1	Q6Q3H2	B9S9Z3	Q94G53
Q8GUE4	B9RXW0	I6QPS5	G5CV45	B0FGA9
E2E2P0	O64961	Q29VN2	C7ASI9	B2KSJ5
D9XDR8	B1W019	D2B747	A0A1L9UKS1.1	A0A1B4XBG5.1
Q96WT2.1	Q9UR08	A0A0P0ZEM1.1	A0A0P0ZD79.1	C9K2Q3
A2PZA5	EW16201	A0A2L0VXR0	P0DL13	Q55012
M1VDX3	B2DBF0	P13513	WP_020663213	WP_011030632
XP_041688448.1	XP_002557473.1	XP_001836356.1	WP_012119179.1	WP_010981512.1
XP_031085895.1	XP_014539580.1	WP_014133196.1	WP_003956090.1	WP_011030119.1
XP_023427124.1	XP_044701492.1	WP_015102836.1	WP_003963519.1	WP_012410187.1
XP_001836356.1	XP_001276070.1	WP_003954606.1	WP_003952918.1	WP_143632835.1
XP_007824857.1	XP_024552383.1	WP_012792334.1	WP_011874125.1	WP_010984429.1
XP_001276070.1	XP_001832573.2	WP_010314578.1	WP_053126184.1	WP_005320742.1
XP_001832925.1	XP_001836556.1	WP_014153723.1	WP_011031839.1	WP_005317515.1
XP_038933630.1	XP_001832925.1	WP_012241161.1	WP_003994861.1	WP_030261827.1
WP_006348711.1	XP_645958.1	XP_644874.1	XP_640697.1	XP_642261.1
XP_642260.1	XP_645125.1	XP_642676.1	XP_638489.1	XP_629084.1
EIM83755.1	KAG5727529.1	BAJ27126.1	KID95099.1	KFG77771.1
EIM91001.1	KAH0583476.1	BAU98235.1	BAX76657.1	AB267396.1
EIM88705.1	KAG5741072.1	AQY56778.1	AQY56777.1	ASV63466.1
EIM82223.1	EAU89322	GU123140.1	QGA30878.1	ASV63465.1
EIM91236.1	LC228601.1	BBF45518	BBD74517.1	AB448947.1
LC484924	QJQ03973.1	LC228602	BBD74518.1	ASV63464.1

>Paramuricea_biscaya_TPS-1
MSCSKEIHAPRRWVDRHKQIPVLPQNAVEKLISMNELIELVIECGLCDKTSINKMYDKINTYQFMWCIVD
TVPASQYAEEIFKSSLHFLCALFLVDDAVESYSANEMQDLSRSYDTLEKEVCKTFPNFPSINEMKESLM
HLRNPFDRSSITFCMQYVNKITAILLEEGNTPHHVVYNLRRRTSNAISIAFQAVLIKSKCGSIITSHEMLW
RRVFDGLVILFYQFGELISGATETAQQHIAVVTELRMLGCLYCIVINDLYSYQRDKLASSDNMIKTWLLE
KTVSSLSEATARCSQILDAIMKYMYQRVEQCIQSNPGCPQLESLLETTIYTTVGWIRSHTTVVPRYSES
QLKVALVEVEEKELPKWLAEKDEYGWNVVEKFVETLNDEKHKGLLDALQGIADGRDQLLKTQLDIS*

>D_gigantea_TPS-1
MWCGKEVRVPRQWAVLEQEVLKEKPDPELVDIDGMIKWITECDIVDENVVRAYVKLVKPYYFIRLIYPIL
PKDKLCVEACKIFLHFVISLYCSDDRTELECDLDDIVKMCDAYENLKEPVFEMFPKFPTIKEMQSSLKFL
SDSKLISPVVLCMDFVNKVTSCVIEYGEYSETAVLDFRRRFLNTIEFYLRGLEMEKKMHRKDTENKTLW
RRIFDGGPLFFLPYVEISSFSLGKVEGHIPTISEMYIVSAFVCIIVNDLYSYKRETMESINCDSVIKHWLSD
EKISSMEEACEKVSRILNATVQYMFEKVQRVKLDYPNSPEAHVLYEYIAHVTIGWLYMHEEGNDRYKD
SPWRVSLTEVEEDKIQEWLSCKDSYGFEALNQFLASNPKAEKIIDALGGDEIIHGGLINESS

>D_gigantea_TPS-2
MFSDNVRVPEKWTISHKKMLAEHEEDKELIALDKLLLWVKETGLTTEAASKSVFDNLNAYRYLRCLFPIL
PEDQTGKKLFQLNLHFLVMGYVIDDTIEKYSEKEMQELIDGYNFLQNQVSKTLPKLPSISKVWKLLGEA
RNKFSRSAIVTVVDYVNRTTMILLEGEICEESVLNYRNRLSNAIAVYFNAILSKTKTGCKISENEMLYRRC
FDALSIAVYMSTEVFSKTLVRNQMLPTSELYKFYLLSTLFCVVINDLYSYERDKLDDTDSVIKVWYAQNE
VTDMKAATTKVSKILDRIIQQMYLFVKEGKSCHPELSEWFERIAAMTVGWIYIHKYVVFRYVSSPFQIAV
VEVKDEMIPDWLTEKDAYGETIVQNFLDGLEYPHQKDVLDSLYGCNRL

>E_rubra-TPS-1
MSISQMVFSKELRVPKEWSKYHFDIVKEPIDPELFSEDELFDWVEDLGLTDDKSVVTKYAQRTGGYLFL
RLHMVFFPKNSLFKKLFKLWAHLVPSLFVTDDFLEATSEIEMRQVCDGFEFLAGQIRGQLPQFPTIAEM
KQSLLLQKIDEKLIPHTIHLMDFANNVAKSIIQHSSIPIESVNEYWRRLVVVCSLYYAGVAFEVKHSVRSY
SEDVWTGVFTVAYLLWLPCLEILSGAVGKTTEHLSLINELSFLATFSARVVNDIYSYNRELVLEHKPVSM
VARIAESKEVTTADEALIKNVEILGAIVKVMYQRIEKAKQENQTNKELCKWLDNIGVGTIGVHFWHHYIP
RYTSAPCRLSFVGVEEDELKEWRKCSEEEPPKELFPVLLDHSQIAKRINDAIISGVVPTQVKNVLCTD*

>E_rubra-TPS-2_7S5L_A
MSGKIVRAPSNWTTPHKKMLKEDEDQELIAFDKLLSWVSETGVGTEEQAKIVFKKINAYFYLRCLYPVL
PRDPKSMKIFQLNLHFLILGYIIDDAIEKYNENEMKELIAGYNLLQSQVSETFPNFPSIFEMKQLLGNMKN
DFSKSAITTMFDYVNKTTLILLEEGEIAEHNVLNYRKRLSNAVAVYFDALLSKTKTGCEISDGAMLWRRC
FDALAIAVYMLTEVFSKTLVKNHTLPVSEFYKFYLLSILFCVVINDLHSYERDKLDDTDSVVKVWFKEGS
VANMEAATSKVSKILDAIIQQMYLFVEEGKARHPELSEWFERIAYMTVGWIYIHKTVVPRYVSSPFQIEV
VEIHENMISNWLLKKDAYGQRVVQQFLKNLNDPQKKNIMLYDE*

>B.asbestinum_TPS-1
MVFSKPIHAPKKWVEKGKGLLDEEVDEQLEGLQEIIDTFIQCDLCDETTVRKYRGKLRPYNFTRSLNPII
PETKLGREVFKALAHFTMGIFIADDIMEKQPEEEIRKMCTTFAHLDAKSRQSFPSFPAIDDIKEVVDASS
VSKNVAGPIVFLEQFLNKQASLLLSHGDFSLSDVKEFRLRGYNMMAIYFLRVLEETEIQKKPGSIEVIWR
RAFGGCVPFYLLTAEIFSGAIGKLKEHYTALTELYVLGGIYVATLNDVFSYYREAHTVADNVVKAMVSQ
KEVETLEEAAANVAEFLDSIMQYLYEKVEELQRKDPNNRELHVCLDYVGRMTVGSVYAHLLPRYKDTT
LSYYLEEVKEDQLSSWLSKKGEYGTRMLKNILEVIHQRMADGQMEALAGAFPYCGKFLGVVIPDTICCT
SSIV*

>B.asbestinum_TPS-2
MTKPFEVTAPTRWVKFHARVLSQPPIENFIRMEQLINWGVACGVTNEAGIRKSFQKLNAYVYLRCMYP
LVADDAWSAEMFRINLHFTTAGYIVDDRIERYTMEEMNELCDGYDMLEKEVSKLFPKCPSIEEMRCRL
QHLKNKFSIAAITMVMDFVNQSALFFLRQGKTSRDRVDNFRKRLSNAVTIYYQAVRNKVKTGCKITEGE
MLWRRCFDVLAVPSYLAPESFTHAVEKQEWPLEMLYELYMLGILYSTVINDLYSYHREKLDDCDNVIKV
WMQENSVSSIEEANEKICMILDAILQIMYEKIEKAKAQYSKCPELQRLLDYTGIVSAGWIFVHNTAAPRY
LESPYQVVLREIEDKDIQGWLENKNEFGWRVVRHFMETMKSEKGKPVMDAMCGFTDARNILI*

>Xenia_TPS-1_trans
MSEKNVVRIPMKWGRIEREILTQNTIPELVDTNRLISWVKECNLADEALVTKYMNVIRPYHFSRLVFPILP
DKDVCREAFAVFIEFLIVLYLCDDELEAKCNLNELEIVCSAYDFLDEKLKQCFPRIPSVEELRSFLPHVKK
ERLLALVVSLIDSVSRVVSPLLKYSVFPQESVFDFRCRFSHSLGYNLKGVLHEKKMLGGVSENKLLWR
RIFEGAPITFLMYLEISSLSVTGSNKHIPIATEMYILSSLCCMVTNDVYSYHRECNEGLKVDNIISLWLHNK
QIASVAEGVSRISRILNSAVKYMLAKVKSMKSEYPNNFNVQTLTEFIALSSVGWLYMHDQGVPRYSDS
PWRLNLVDIEESDITSWLAEKDPYADGVIDQFKYSNLDAKKFIDSLCEKTSVSEEQWND

>Heliopora_coerulea_TPS-1
MAFSKPIYAPKKWAEKAKGVTSMNEEKQLVGMDELINWVAECELAGEATVRKYWDTVRPYHFMRCL
YPVFPDNILCKEFFKVLLHASISIYIADDQLEKQSREEAKIACAAHARIDQQSTKRFPNLPTIQEMKQILST
FSTPSIVGPTTMFADFGNRLAKVLLLHGNWNGNVVADFRLRNSNLVSMYFQAIQAEKTPNKKNTTLET
LWRRIFGGVAAPYSSLAEIASGAIENSKQHIAVITDMHLLSGLFSTTINDLYSYFLEKSSICDNVVAALLHE
KTAENVSEAVDKVAQILDAILKLMYKKAEQIKLQFPDNRELRDFFDYVGQVTVGWYYLHECALPRYQG
SPWRVHLEEIDEDEIPKWLSEKGKYGSKVMQELMELVQERMADGKMDAVHGKFPINEKYVAEKEDGL
YKII*

>Heliopora_coerulea_TPS-2
MACSKEMYIPKEWTKYHHDIVKEVKRADLEADDEVAKWVVGLGLSDKPATEKYIDSTRPYHFMRCIAV
LVPSNPLCRNIFKTWQIYVTGLFVSDDYLEQTSLAEVVQVSDSFDMLNEQISERFPVIPTHAQIMRSLKI
NEKLIPHVLVQIDWLNTIANDLLQHGDFSEDDVWDFRCRMSAWIRAYFDGLKSQVGPKVKVTTDDFM
WLRIIDGGVGPFLLINEILVGELGKCKGHVKLITQLYILSSCVCVLINDCYSYHRETVVTIYNTIKVILENKE
VSEVPDAVSKLLQTINAIIKYMYQTIEKTKHEYPDNLMLNKLLDAIGGATAGWFFVHDKVLPRYQATSW
RFSLVEVEEMELKEWRKSADEKPSDLVQPLLYRFNAKGKEIIDALAYVPKRSRRIVPE*

>Heliopora_coerulea_TPS-3
MSCSKEVRAPRKWVQRQREMQKVKVDDNLISMDELIELVVECGLCDEASVHKMYERTNTYQFMWTLI
PTVPSDQWQFEIFKTSLHFLCALFLVDDAVESYSEKQMKEMSDAYDMLEKQVCENFPCFPSINEMKQ
SLKHLKNPFDIASITFCMHYVNKIASVLVNEGSTPQNVVYNFRRRTSNAISISFQAVLIKAKLGSNVTGDE
MLWRRVYDGLVILFYQFGELVLGLTKNTEKYMTVITELRLLGCLYCIVINDLYSYHRDKLVSSDNIIKTWL
LGKAVSNLSEASSRCCEILDAVMQYMHERVERIRQSHPGCPELEALLETTVYTTVGWIFAHTAVVPRY
SESPLKVHLVEVEESELHTWFCEKNHYGRSVIEKFLKAVNDEKHKGILDALSGFVDGRGQLLKTQL*

>Renilla_reniformis_TPS-1
MLRTEVKIPEVWCIPCDLKCSTPPQLKKDIIDWCVRTGITADRRTAEKALDRLKPYMYMKILFPKLKPDN
PLIASLYELNASVIVTGFTMDDVLETYTLSAIEELDHTFRETQNWLSPLPPKDYPDLDAILDHIPSTKPPY
SRSVVAMFTDYFNRYCAAIHTAYKPSETRLGAFRRRLAAAVTAYLEMAKRKRRRDGNLHEEEFLWQR
SADNLSFPVLMLAEVFTGILADADVPSTTLHYYHFYSNLFAIVLNDLNSYHRDKDNDDNSLVKLWLKLG
VAKDFDDAATKIIDFLNSIMVRMYVATKELLESNPENEPLRRFVEAVGYTFNGWILVHTTAVERYKLSPF
QTVVAPVARGKEEVWLKGETAFGKRCVAMFEELMGERAEEMNKLYGLDS

>Tubipora_musica_TPS-2
MSCSKEIRIPSDWASVQRKSMQEGPDESLLDFEELANWLRECGVTDDQCKVRKYVDATRPVLSIRTIL
RVFPNNTLCRMHFKFLTQYTIGIYIDDDVLESHYPLEVLKEICYEYDQLDGKLLGQFPQIPSRKELKNFLA
HLKNEKVISTVTFHMDFVNRVTVNILQNGDFSEEAVLEFRRRLSNNILLYLQGVQYEKRGAGGKPSTVV
EALWNRVFGGAAVTWRLFGEVPSGTLGKIGEHITLISELYILSALYCIVINDLYSYHREIDAVPTAGENFV
RVMFNQKEVANLTVAAGKVASILNEITKYTYEQVREFKASYPNSSELHQLFDDIGCGTVGWYYFHECT
NPRYKESNVRISMKDVDKNEIEDWLSTKDSYGWNIVKQFLASYDAKAKRLNDAVACRIPAYAETF*

>uBuTS-1_undescribed_Bubarida
MSLSVFAKIPVPALKLENPLHYSDSAVSLAKVREPLLSGELKSRLHRHALDLKLDPKYIENLNVHDFVAK
VHDLGVSNHPTPGQNRQWLFASTLVLFIFEFDDHFDTAVRLTPENIARVSKDMRDVLRSLSRHNLCGL
QGSLEDWPAKVPCKEAYHWLLREGEDLREGTAQLIHDTFVDYCYGTEGEVIEWQPDMYRGDFTAWN
LDRYCEIRKRSIGVLFAVLGPLFIRYNWIQKEHITTCIDLLHEAAIVVALANDVAGMNRDLEDNEAIDITSL
KIAATSSEVVQYHNKKVEVLHKKVLELECNTWRFMEQVEASVVGLFLWHLNAQRYAVNSEHYTL

>uBuTS-2_undescribed_Bubarida
MSLSVFAKIPVPTLELEHPMHYVDSTVSVAKVRISSVSAGLNNRLYRRVLDLKLNPEYINHMDINDFVAK
ACNLGLPTPDQERLWIFTSVNALFIFEFDDHFDKPLLIAPEYVARRSKQMRAVLRSLSSHKLSGLQGSL
EDWPAEVPCKEAYLWLLREAEDLRKGAAELLHDVFIDFCNGVEEEVIQWQSDAHRGDFTGWNLDRYK
EARKRSVGYIFTPVVPLFIINKWIQKDYFTTCIDLLYEAAIIVALANDILGIPRDHGDDGTMTVLRVITSQDK
VVEVHNEMVEVLHKKVLELECNKRRFMEEMEAAAVGLFLWQLDSNRYAGRLPITIY

>uBuTS-3_undescribed_Bubarida
MSLSAFTKIPVPALELEYPMHYADSAVSEAKLRETSVSTELKSKLHRHILNLKLKAQYAEHLNIYDFIAKV
HSIGVSNHPTVNKSEKRQVLFASSLIVLLFQFDDHFEATDPLHLTLESIAKRSKEMRSVLSSLSAHNLSG
LQGSLEEWPAAVPGKEAYLWLLREADDLREGTAELVHDTFVDYCHGVLMEIIEWQSDMYRGDFTAW
NLDRYKEVRKRSSGVIFGTVGILYITQKWIQKEHITTCTDLLYEAAIVVSFGNDILGKNRDSVDSSVIDITS
LKIATSSNIVLKHNTMVKALHSKVLELEGDTRHFMEEMEASVVGLFLWQLYSKRYL

>uBuTS-4_undescribed_Bubarida
MSLSVFAKIPVPTLELEQPVHYVDSAVSVVKFREPLVSGDVKRKLHRHVLDLKLNPEYIQHLGIYDLVAK
VYNLGVSNHPLNQYRLWLLVSTLEILIFEFDDYFDKSLLRTPENIAKLSKEMRGVLRSLSRHNLSGLQGS
LEDWPAEVPCKQAYLWHLREAEDLREGTAQLIHDTFVDYCYGVEVEVTEWQSDVNKGDLTAWNLDR
YNDARKHSVGLAYAIVGPLYIANKWLQKEHIITCTDLVYDASILTVLGNDILGMNRDRGDSGTMTSLKLL
ASSSEVVQYHNELVEVLHKKVIELECNTRHYMEEVEAAVVAFFLWQSRAERYAV

>uBuTS-5_undescribed_Bubarida
MSLSVFTKIPIPAVELEHLMHYADSAVSEAKLRKHSVSEELKSQMHRHLVDLKLNPRYPDNLNIYDFVA
KLFNLGVSKHPTFDQNRLWLFSSILILFLFEFDDHFDDHFDNPLHLTTENIARLSKEMRAVLRSLSRHNL
SGLQGSLEDWPAEVPCKEAYRWLLREAKDLKEGAAELLHDIFVDYCFGVEEEVIEWKSDMYRGDLTA
WSLDRYKEVRKRAAGVVFAIVVPLFVTRKWIQKEHVNTCTDLLYEAALLVGLSNDILGIPRDLRDSGTM
TVLKIASTSEVVQHHNMMVEVLHKKVLALEGNTMHFMDAVETSVAGVFLWQCHSKRYTL

>uBuTS-6_undescribed_Bubarida
MSLGVFAKIPVPALELEHPVHYADSAVSVAKVRISSVSAGLNNRLYRRVLGLKMNPEYIKHMDINDFVA
KACNLGLPTPDQDRLWIFASVTGLFIFEFDDHFDKPLLIAPEYVARRSKQMRAVLRSLSSHKLSGLQGS
LEDWPAEVPCKQAYLWHLREAEDLREGAAELVQDLFVDYCYGTEEEIIEWQPDAYRGDYTAWNLDRY
NEVRKRSAGLTFVSVGPLFVAHKWIQKKHVVTCTDLLYEATIVIALANDILGVNRDRDDRLCVTSLKLSS
CSEVVQYHNELVEVLHKKVLELEGNTRRFMEEVEASVVGLAIPVAVPFKTLWCGLEG

>uBuTS-7_undescribed_Bubarida
MSLGVFAKIPVPALELEHPVHYADSAVSVAKVRISSVSAGLNNRLYRRVLDLKLNPEYINHMDINDFVAK
ACNLGLPTPDQERLWIFTSVNALFIFEFDDHFDKPLLIAPEYVARRSKQMRAVLRSLSSHKLSGLQGSL
EDWPAEVPCKEAYLWLLREAEDLRKGAAELLHDVFIDFCNGVEEEVIQWQSDAHRGDFTGWNLDRYK
EARKRSVGYIFTPVVPLFIINKWIQKDYFTTCIDLLYEAAIIVGLGNDILGMLRDSAEDGSMTVLKIITSHDK
IVEVHNKMVEVLHKKVLELECNKRRFMEEMEAAAVGLFLWQLDSNRYAGRLPITIY

>uBuTS-8_undescribed_Bubarida
MSLSTFAKIPVPELELEHQMHYADSIVSVTKLREPSVSKELKSRLHRRVLELRLNPQNIEILNSYDFVAK
VCTFGVSNHSTTCSHQNRWWIFASISVIFLFEFDDQFDKPLLLSPENIASLSKKMRGVLRSLSGHKLSG
LQGNLEDWPAEVVGKEAYLWLLREAEDLREGAAELVQDLFVDYCYGTEEEIIEWQPDAYRGDYTAWN
LDRYNEVRKRSAGLTFVSVGPLFVAHKWIQKKHVVTCTDLLYEATIVIALANDILGVNRDRDDRLCVTSL
KLSSCSEVVQYHNELVEVLHKKVLELEGNTRRFMEEVEASVVGLFLWQCRSKRYAVVFEC

>uBuTS-9_undescribed_Bubarida
MSLFAEVPIKALPLERPLQYAHSALNLARLRKASESVCSSLTKRLQKHCSALKLNPQYIEAANFSHFAAK
VYNIGVATEPGQDRLGIFVATLALFVFEFDDHFDKPCGTPENVSRLSMEMRILLRALSRHGLAGLQGVL
DDWPTGAFGREAYLWLLKEAEDLRKGAAELMHYSFLDYCFGVESEILEWAPDAYRGDTTAWNLDRY
SEVRKRSVGVTITLVAPLYIINKWTKKEHFATCNDLLYNAALIGALANDVLGMKRDQEESQSLGMTALKI
VSASEAAQYHNEKVECLKKDILDLDGDTRRLMEEVEVINVAQFLWQCNARRYN

>uBuTS-10_undescribed_Bubarida
MSLSVFAKIPVPTLELEHPMHYVDSTVSVAKVRISSVSAGLNNRLYRRVLDLKLNPEYINHMDINDFVAK
ACNLGLPTPDQERLWIFTSVNALFIFEFDDHFDKPLLIAPEYVARRSKQMRAVLRSLSSHKLSGLQGSL
EDWPAEVPCKEAYLWLLREAKDLKEGAAELLHDIFVDYCFGVEEEVIEWKSDMYRGDLTAWSLDRYK
EVRKRAAGVVFAIVVPLFVTRKWIQKEHVNTCTDLLYEAALLVGLSNDILGIPRDLRDSGTMTVLKIAST
SEVVQHHNMMVEVLHKKVLALEGNTMHFMEEVETSVAGVFLWQCHSRRYTL

>uBuTS-11_undescribed_Bubarida
FAKIPVPALELEHPVHYADSAVSVAKVRISSVSAGLNNRLYRRVLGLKMNPEYIKHMDINDFVAKACNL
GLPTPDQDRLWIFASVTGLFIFEFDDHFDKPLLIAPEYVARRSKQMRAVLRSLSGHKLSGLQGSLEDW
PTEVPCKEAYLWLLREAEDQRKGGAELLHDVFIDFCYGVEEEIIQWQSDAHRGDFTAWTFDRYNEAR
KHSVGYIFTPVVPLFIINKWIQKEYFTSCIDLLYEAAIIVGLGNDILGMLRDSAEDGSMTVLKIITSHDKIVE
VHNEMVEVLHKKVLELECNKRRFMEEMEAAAVGLFLWQLDSNRYAGRLPITIY

>uBuTS-12_undescribed_Bubarida
MSLGVFAKIPVPALELEHPVHYADSAVSVAKVRISSVSAGLNNRLYRRVLGLKMNPEYIKHMDINDFVA
KACNLGLPTPDQDRLWIFASVTGLFIFEFDDHFDKPLLIAPEYVARRSKQMRAVLRSLSGHKLSGLQGS
LEDWPTEVPCKEAYLWLLREAEDQRKGGAELLHDVFIDFCYGVEEEIIQWQSDAHRGDFTAWTFDRY
NEARKHSVGYIFTPVVPLFIINKWIQKEYFTSCIDLLYEAAIIVGLGNDILGMLRDSAEDGSMTVLKIITSHD
KIVEVHNKMVEVLHKKVLELECNKRRFMEEMEAAAVGLFLWQLDSNRYAGRLPITIY

>uBuTS-13_undescribed_Bubarida MLHSQLSSLTSSLAVYKGILQALNSIIACLRPACSCECKLHCSVIMSLGVFAKIPVPALELEHPVHYADSA VSVAKVRISSVSAGLNNRLYRRVLGLKMNPEYIKHMDINDFVAKACNLGLPTPDQDRLWIFASVTGLFIF EFDDHFDKPLLIAPEYVARRSKQMRAVLRSLSGHKLSGLQGSMEDWPTEVPCKEAYLWLLREAEDLR KGAAELLHDVFIDFCYGVEEEIIQWQSDAHRGDFTAWTFDRYNEARKHSVGYIFTPVVPLFIINKWIQKE YFTSCIDLLYEAAIIVGLGNDILGMLRDSAEDGSMTVLKIITSHDKIVEVHNKMVEVLHKKVLELECNKRR FMEEMEAAAVGLFLWQLDSNRYAGRLPITIY
>uBuTS-14_undescribed_Bubarida MSLSVFTKIPIPAVELEHLMHYADSAVGETKLRKHSVSGELKSLLHRHLLDLKLNPQYLDHLNIYDCVAK CFNLGVNHHPTLDQNQLWIFASILILFLFEFDDHFDNPLHLTPENTARLSKEMRAVLRSLSRHNLSGLQ GGLEDWPAEVPGKEAYLWLLREAEDLKEGAAELTHGLFVDYCFGVEEEVIEWKSDTYRGDLTAWSLD RYKEVRKRTAAGGFAIVVPLFVTYKWIQKEHVHKCTDLLYEASLLLALSNDILGIPRDRRDSATMTALKI ASTSEVVQHHNEMVEVLHKKVLALEGNTKHFMEEVEASVVGFFLWQCHSKRYTL
>PhTS-1_Phakelia_ventilabrum MSLSVFTKIPIPTMKLEPPMYYADSAVSVAKLRESLVSEKLNNRLHRLVLDLKLNPQYIDHMNINDFVAK ACNLGAPTPDQDWLWLFTSVHGIFIFEFDDHFDKPRQIAPECVAKRAKEMRAVLRSLGSHKLSGLQGS LEHWPSEVPCKEAYLWLLRKAEELREGAAELLHDVFIDFCNGVEEEIIQWQSDAYRGDFTAWNLDHYK EARKRSVGYIYSPVVPLFIINKWIQKEYFTTCIDLLYEAAIIVALANDVLGIPRDHGDDGTMTVLRIITNLCD VIKVHNEMVEELHKKVLELECNAMYFMEEMEAGMVGLFLWQLNSNRYAGRLPITIY
>AgTS-1_Agelas_clathrodes MSQFVEVPIKSLVLQHPLRYAHSAFDSAKLQISSVSMSLKDKLHNHCSALKLNPRYIEELNIADYVAKTS NLGVEYHATEEEDRLWIFAATLGIFLFEFDDYFDKPSNTPENVARLSTEMRAVQRALSRHGLSGLQGNI DDWPTAVPYREAYCWLLREAEDLRRGAAEVVHYTFLDYCFGVEMEITEWAPDMYQHDTSSWNLDRC NEVRKRSGGTHFAFVGPLYVVNKWITKEHFNACNDLLYDAAIVITLANDLLGKKKDLQTENDSISIKTIKI ATTSEIAQHHNEKVECLRKDILQLDGDIRRLMEEWEVIVAGLFLWQYNSRRYE
>StTS-1_Stylissa_massa MSRFVEVPVKALALQHPFKYAHSALSVAKLREASVSSLLKTRLRNHLSALNLNPKYMEALNIPDFIAKIY GIECPSTEDKERQWVFSATFTNCFFVFDDHFDEQVGTPEIVAKLSLEMRDILRALSRHTLSGLQGVLDD WPTDVPCKKAYLWLLQEAEGLRKGAAELVHATFNDYCLGVESEIVEWAPDVHRGDMSAWNLDRCTE VRKRSAAGNMALAPLYVINKWMTREHYRACNDLLYDVSLIIALPNDVIQSVKRNNHSISMETTKILSADE IVQCHNKKVECLRKDILELDGDTRRFMEEIEISAVGVFLWQCNCKRYVD
>AgTS-2_Agelas_tubulata MSQFVEVHIKSLLLQHPLRYAHSALNLAKLRNSSVSSSLKDKFEQHCAALELNPRSLTDIGYSDFVAKIN NIGVKYHPAEEQDRLWIFAATLFTFLMQFDGHFDKPSITPENISRLTMEIKAILRALSRHGLSGFQGNLD DWPTSLPFREPYLWLLREAEDLRKGAAELVHYTFLDYCLGLEMEVTEWAPDVYRHDISSWSLARCYE ARKRSRGQTVDLVAPLYIANKWTKKEHFTACIDLFYDVAIITTLANDLLGEKEDLENDHIGMTTLKVANK SKIVQVHNEKVEGLRNDILGLDGDIRRLMEEVEVTTAAAFLCECTRKKLDYSESDYDS

Table S3. List of TSs included in TSHMM1 and in the TS phylogenetic tree in Figure 3. Grey shaded cells are UniProt IDs, blue shaded cells are NCBI accession numbers, and yellow shaded cells are GenBank accession numbers. White cells are previously published coral TS sequences (13). Orange shaded cells are sponge TSs described in this study, which were included in Figure 3, but not TSHMM1.

Name	Origin	Sesquiterpene synthase activity	Monoterpene synthase activity	Diterpene synthase activity	Major product(s)
uBuTS-1	undescribed Bubarida metagenome and transcriptome	Yes	No	No	(+)-bicyclogermacrene
uBuTS-2	undescribed Bubarida metagenome and transcriptome	Yes	Yes	No	(-)-germacrene D
uBuTS-3	undescribed Bubarida metagenome	Yes	Yes	No	(+)-bicyclogermacrene
uBuTS-4	undescribed Bubarida metagenome	Yes	Yes	No	mixed sesquiterpenes
uBuTS-5	undescribed Bubarida metagenome and transcriptome	Yes	No	No	mixed sesquiterpenes
uBuTS-6	undescribed Bubarida metagenome	No	No	No	None
uBuTS-7	undescribed Bubarida transcriptome	No	No	No	None
uBuTS-8	undescribed Bubarida metagenome and transcriptome	No	No	No	None
uBuTS-9	undescribed Bubarida metagenome	No	No	No	None
uBuTS-10	undescribed Bubarida transcriptome	N/A	N/A	N/A	N/A
uBuTS-11	undescribed Bubarida transcriptome	N/A	N/A	N/A	N/A
uBuTS-12	undescribed Bubarida metagenome	N/A	N/A	N/A	N/A
uBuTS-13	undescribed Bubarida transcriptome	N/A	N/A	N/A	N/A
uBuTS-14	undescribed Bubarida metagenome	N/A	N/A	N/A	N/A
AgTS-1	Agelas clathrodes metagenome	Yes	Yes	No	(+)-alloaromadendrene
AgTS-2	Agelas tubulata metagenome	No	No	No	None
PhTS-1	Phakellia ventilabrum transcriptome	Yes	Yes	No	(-)-germacrene D
StTS-1	Stylissa massa metagenome	No	No	Yes	mixed diterpenes

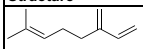
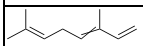
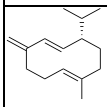
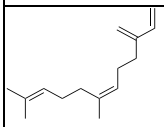
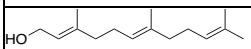
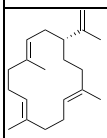
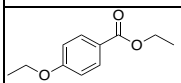
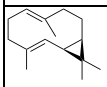
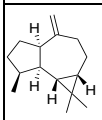
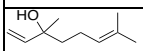
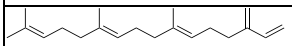
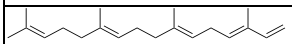
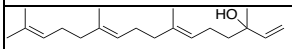
Table S4. Sponge terpene synthase summary table.

Start	End	Length	Introns?	Blast Annotation	Blast Annotation Organism	Blast Annotation Accession Number	Blast E value	Halichondria panicea hit	H. panicea E value
uBuTS-9 contig 1 (72kb)									
Start	End	Length	Introns?	Blast Annotation	Annotation Organism	Annotation Accession Number	Blast E value	Halichondria panicea hit	H. panicea E value
6851	11435	4585	yes	alsin-like	Stylophora pistillata	XP_022793651.1	2.00E-07	GIFJ01263445.1	3.00E-21
13656	14960	1305	no	bystin-like	Amphimedon queenslandica	XP_003388529.1	2.00E-163	HBWD01518048.1	5.00E-66
15162	15900	739	no	Branched-chain-amino-acid aminotransferase-like protein 2	Amphimedon queenslandica	XP_003383850.2	4.00E-37	GIFJ01253076.1	4.00E-16
21281	25316	4037	no	uncharacterized protein	Amphimedon queenslandica	XP_019851048.1	4.00E-93	HBWD01019842.1	7.00E-100
28635	29621	987	no	uBuTS-9					
32769	46513	13744	yes	alsin-like	Amphimedon queenslandica	XP_019854288.1	4.00E-19	HBWD01019722.1	7.00E-43
53000	61934	8934	yes	cyclin-Y-like protein 1	Amphimedon queenslandica	XP_003388010.1	4.00E-15		
64295	66087	1793	yes	no blast hit					
uBuTS-2 contig 1 (63kb)									
3008	4242	1234	yes	piggyBac transposable element derived protein 4-like	Stylophora pistillata	XP_022798311	9.00E-07		
6832	7714	882	no	no blast hit					
8023	8950	927	no	no blast hit				HBWD01105536.1	6.00E-20
11212	15012	3800	yes	no blast hit				HBWD01373978.1	2.00E-13
15451	20081	4630	yes	uncharacterized protein	Amphimedon queenslandica	XP_019855287.1	7.00E-26	HBWD01009091.1	6.00E-69
23055	23823	768	no	uncharacterized protein	Amphimedon queenslandica	XP_019854607.1	2.00E-04	GIFJ01039005.1	5.00E-07
25849	26652	803	no	uncharacterized protein	Ectocarpus sp. CCAP 13	CAB1113566.1	3.00E-23		
26853	27419	566	no	uncharacterized protein	Ectocarpus sp. CCAP 13	CAB1116628.1	3.00E-20	GIFJ01283391.1	7.00E-05
31393	32400	1007	no	uBuTS-2					
36476	50257	13781	yes	uncharacterized/retrovirus related pol polyprotein	Stylophora pistillata	PFX30308.1	4.00E-23	GIFJ01196170.1	1.00E-57
52757	63023	10266	yes	uncharacterized protein	Oopsacas minuta	KAI6647352.1	1.00E-36	GIFJ01515311.1	1.00E-12
uBuTS-1 contig 1 (74kb)									
18935	19957	1022	no	uBuTS-1					
20717	21677	960	no	hypothetical protein LOD99_12348	Oopsacas minuta	KAI6647352.1	1.00E-32		
28569	29525	956	no	hypothetical protein LOD99_12348	Oopsacas minuta	KAI6647352.1	1.00E-37		
33244	34419	1175	yes	pogo transposable element with ZNF domain-like	Saccoglossus kowalevskii	XP_006822101.1	1.00E-54	GIFJ01059185.1	3.00E-05
36033	36819	786	no	hypothetical protein LOD99_12348	Oopsacas minuta	KAI6647352.1	4.00E-33		
50879	55506	4627	yes	PREDICTED: serine/threonine-protein kinase svkA-like	Amphimedon queenslandica	XP_019855734.1	4.00E-82	HBWD01267899.1	5.00E-144
70866	73178	2312	no	PREDICTED: serine/threonine-protein kinase 10-like	Amphimedon queenslandica	XP_019859200.1	2.00E-65	GIFJ01029571.1	3.00E-134
uBuTS-8 contig 1 (61kb)									
3727	7370	3643	yes	PREDICTED: runt-related transcription factor	Amphimedon queenslandica	XP_019858567.1	2.00E-33	GIFJ01194107.1	8.00E-56
14197	19827	5630	yes	runt-related transcription factor	Amphimedon queenslandica	ACF96957.1	3.00E-49	HBWD01004530.1	5.00E-54
46761	47771	1010	no	uBuTS-8					
50497	60666	10169	yes	uncharacterized protein LOC111340013	Stylophora pistillata	XP_022802513.1	7.00E-22	HBWD01011394.1	4.00E-54
uBuTS-1 contig 2 (57kb)									
1802	2815	1013	no	uBuTS-1					
4622	7297	2675	yes	PREDICTED: 60S ribosomal protein L12-like	Amphimedon queenslandica	XP_003382502.1	2.00E-26	GIFJ01084458.1	1.00E-21
9219	11890	2671	no	Hypp4351	Branchiostoma lanceolatum	CAH1270521.1	3.00E-90	GIFJ01153800.1	2.00E-151
13646	14517	871	yes	Hypothetical predicted protein	Paramuricea clavata	CAB4002751.1	9.00E-45	HBWD01137827.1	5.00E-53
20900	26503	5603	yes	PREDICTED: uncharacterized protein LOC109582994	Amphimedon queenslandica	XP_019853667.1	2.00E-60	HBWD01255793.1	1.00E-49
41880	42758	878	yes	archaemetzincin-2-like	Stylophora pistillata	XP_022791489.1	3.00E-60	HBWD01208581.1	1.00E-31
uBuTS-14 contig 1 (47kb)									
1021	5706	4685	yes	PREDICTED: pleckstrin homology domain-containing family F member 2-like isoform X2	Amphimedon queenslandica	XP_019864174.1	2.00E-26	HBWD01207577.1	6.00E-33
13957	15744	1787	yes	PREDICTED: pleckstrin homology domain-containing family F member 2-like isoform X3	Amphimedon queenslandica	XP_003382481.1	1.00E-10	HBWD01207582.1	3.00E-15
28994	30166	1172	yes	PREDICTED: proprotein convertase subtilisin/kexin type 9-like	Amphimedon queenslandica	XP_019849385.1	3.00E-13	GIFJ01293424.1	3.00E-13
30301	36391	6090	yes	PREDICTED: uncharacterized protein LOC100638844	Amphimedon queenslandica	XP_019849384.1	9.00E-57	GIFJ01125311.1	4.00E-102
42985	44406	1421	yes	hypothetical protein MCG8623886.1	Proteobacteria bacterium	MCG8623886.1	1.00E-51	HBWD01267899.1	1.30E-43
44681	45673	992	no	uBuTS-14					
uBuTS-2 contig 2 (272 kb)									
219	3532	3313	yes	PREDICTED: vinculin-like	Amphimedon queenslandica	XP_003382439.1	7.00E-11	HBWD01162922.1	2.00E-15
11347	27492	16145	yes	PREDICTED: mucin-17-like	Amphimedon queenslandica	XP_019849308.1	4.00E-06	GIFJ01353008.1	4.00E-21
57307	72065	14758	yes	TPR and ankyrin repeat-containing protein 1-like	Haliotis rufescens	XP_046368821.2	5.00E-145	GIFJ01089742.1	0
73001	77134	4133	no	uncharacterized protein K02A2.6-like	Oreochromis aureus	XP_039459596.1	0	GIFJ01340966.1	0
78527	88037	9510	yes	interferon-stimulated 20 kDa exonuclease-like 2	Chrysemys picta bellii	XP_005310999.1	4.00E-17	GIFJ01335375.1	4.00E-17
88149	99651	11502	yes	PREDICTED: protein asteroid homolog 1-like	Amphimedon queenslandica	XP_011407600.1	3.00E-76	HBWD01172737.1	1.00E-86
109646	121637	11991	yes	TPR and ankyrin repeat-containing 1-like	Paramuricea clavata	CAB4029118.1	8.00E-160	GIFJ01089739.1	0.00E+00
123248	136125	12877	yes	TPR and ankyrin repeat-containing protein 1-like isoform X2	Orbicella faveolata	XP_020607958.1	3.00E-145	GIFJ01089738.1	0.00E+00
141266	143344	2078	yes	Hypothetical predicted protein	Mytilus galloprovincialis	VDI68231.1	3.00E-11	GIFJ01053260.1	8.00E-08
151338	160040	8702	yes	TPR and ankyrin repeat-containing protein 1	Lingula anatina	XP_013405370.1	9.00E-161	GIFJ01089740.1	0.00E+00
170415	180500	10085	yes	PREDICTED: uncharacterized protein LOC100635005	Amphimedon queenslandica	XP_019859701.1	0	HWD01149209.1	5.00E-117
193727	195100	1373	no	protein O-mannose kinase-like isoform X1	Acropora millepora	XP_029212908.2	1.00E-48	HBWD01506726.1	4.00E-66
196853	203076	6223	yes	PREDICTED: balbiani ring protein 3-like	Amphimedon queenslandica	XP_019849443.1	2.00E-10	HBWD01221237.1	1.00E-29
210324	212035	1711	no	PREDICTED: balbiani ring protein 3-like	Amphimedon queenslandica	XP_019849443.1	2.00E-37	HBD01261137.1	5.00E-46
219388	221995	2607	yes	hypothetical protein BSL78_12423	Aposticho pus japonicus	MRZV01000408.1	9.00E-20	HBWD01045466.1	8.00E-21
237060	240800	3740	no	PREDICTED: balbiani ring protein 3-like	Amphimedon queenslandica	XP_019849443.1	4.00E-109	HBWD01261143.1	1.00E-166
244030	247157	3127	yes	PREDICTED: 40S ribosomal protein S11-like	Amphimedon queenslandica	XP_003382520.1	5.00E-33	HBWD01596638.1	5.00E-32
247280	253239	5959	yes	protein kinase	Proteobacteria bacterium	MCG8623461.1	5.00E-99	GIFJ01138049.1	5.00E-112
266184	267194	1010	no	uBuTS-2					
267920	269747	1827	no	hypothetical protein LOD99_12348	Oopsacas minuta	KAI6647352.1	6.00E-32		

Table S5. Annotations for predicted genes on TS-containing contigs from the undescribed Bubarida sponge. Gene and organism annotations are based on the closest NCBI blastx hit against the non-redundant protein sequences database. The contigs were also queried against a custom BLAST database generated from two publicly available *Halichondria panicea* transcriptomes (HBWD00000000 and GIFJ00000000) using tblastx, the results of which are shown in the 2 right-most columns.

Repository	Accession	Sequence
BioProject	PRJNA907134	Undescribed Bubarida BioProject
BioProject	PRJNA824609	*Agelas clathrodes* and *Agelas tubulata* BioProject
BioSample	SAMN31952758	Undescribed Bubarida specimen SCS0036
BioSample	SAMN31952757	Undescribed Bubarida specimen SCS0044
BioSample	SAMN27408796	*Agelas clathrodes* specimen SIA219
BioSample	SAMN08501975	*Agelas tubulata* specimen SBM112
SRA	SRR22527612	Illumina undescribed Bubarida ribodepleted RNA-seq
SRA	SRR22527613	Illumina undescribed Bubarida poly-A enriched RNA-seq
SRA	SRR22527614	Nanopore undescribed Bubarida metagenomic sequencing
SRA	SRR22527615	Illumina undescribed Bubarida metagenomic sequencing
SRA	SRR18681978	*Agelas clathrodes* metagenomic sequencing
SRA	SRR23375712	*Agelas tubulata* metagenomic sequencing
GenBank	OP977975-OP977983	uBuTS genes
GenBank	OQ320741-OQ320747	uBuTS contigs

Table S6. Accession numbers of sequences from this study.

Structure	Name	IUPAC name	Molecular Formula	Method of identification	Figures appeared in
	β-Myrcene	7-methyl-3-methyleneocta-1,6-diene	C10H16	NIST spectral library comparison	S5, S24, S25-S27, S32, S34
	Ocimene	3,7-dimethylocta-1,3,6-triene	C10H16	NIST spectral library comparison	S5
	(-)-germacrene D **8**	(S,1E,6E)-8-isopropyl-1-methyl-5-methylenecyclodeca-1,6-diene	C15H24	NMR, OR	S5-S6, S16, S19-S20, S38, S41, S44, S48
	(E)-β-Farnesene	(Z)-7,11-dimethyl-3-methylenedodeca-1,6,10-triene	C15H24	NIST spectral library comparison	S7, S46
	farnesol	(2E,6E)-3,7,11-trimethyldodeca-2,6,10-trien-1-ol	C15H26O	NIST spectral library comparison	S7, S47
	cembrene A	(1E,5E,9E,12R)-1,5,9-Trimethyl-12-(prop-1-en-2-yl)cyclotetradeca-1,5,9-triene	C20H32	comparison to purified standard	S7
	4-ethoxy ethylbenzoate	4-Ethoxy ethylbenzoate	C11H14	NIST spectral library comparison	S8-S11
	(+)-bicyclogermacrene **6**	(1S,2E,6E,10R)-3,7,11,11-tetramethylbicyclo[8.1.0]undeca-2,6-diene	C15H24	NMR, OR	S15, S17, S19, S37-S38, S41, S45, S48
	(+)-alloaromadendrene **7**	(1aS,4aR,7S,7aS,7bR)-1,1,7-trimethyl-4-methylenedecahydro-1H-cyclopropa[e]azulene	C15H24	NMR, OR	S21, S45
	linalool	3,7-dimethylocta-1,6-dien-3-ol	C10H18O	NIST spectral library comparison	S22-S35
	1,2-dibromobenzene	1,2-dibromobenzene	C6H4BR2	NIST spectral library comparison	S22-S24, S27- S35
	β-springene	(6E,10E)-7,11,15-trimethyl-3-methylenehexadeca-1,6,10,14-tetraene	C20H32	NIST spectral library comparison	S22, S24-S25, S27-S28, S30-S34
	α-springene	(3E,6E,10E)-3,7,11,15-tetramethylhexadeca-1,3,6,10,14-pentaene	C20H32	NIST spectral library comparison	S22, S24-S25, S27, S30-S34
	geranyllinalool	(6E,10E)-3,7,11,15-tetramethylhexadeca-1,6,10,14-tetraen-3-ol	C20H34O	NIST spectral library comparison	S22, S27, S33

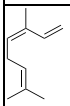
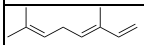
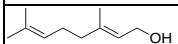
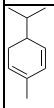
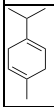
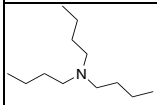
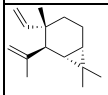
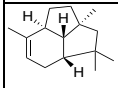
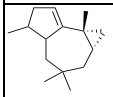
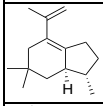
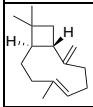
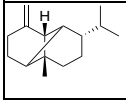
Structure	Name	IUPAC name	Molecular Formula	Method of identification	Figures appeared in
	decane	decane	C10H22	NIST spectral library comparison	S23, S28, S29
	cis-β-Ocimene	(Z)-3,7-dimethylocta-1,3,6-triene	C10H16	NIST spectral library comparison	S24, S26
	trans-β-Ocimene	(E)-3,7-dimethylocta-1,3,6-triene	C10H16	NIST spectral library comparison	S24-S26, S32, S34
	geraniol	(E)-3,7-dimethylocta-2,6-dien-1-ol	C10H18O	NIST spectral library comparison	S24
	α-phellandrene	5-isopropyl-2-methylcyclohexa-1,3-diene	C10H16	NIST spectral library comparison	S34
	γ-terpinene	1-isopropyl-4-methylcyclohexa-1,4-diene	C10H16	NIST spectral library comparison	S34
		1-methyl-5-(prop-1-en-2-yl)cyclohex-1-ene	C10H16	NIST spectral library comparison	S34
	tributylamine	tributylamine	C12H27N	NIST spectral library comparison	S37, S47
		(1S,2S,3S,6S)-3,7,7-trimethyl-2-(prop-1-en-2-yl)-3-vinylbicyclo[4.1.0]heptane	C15H24	NIST spectral library comparison	S37
		(2aS,2a1R,4aS,7aR)-1,1,2a,5-tetramethyl-2,2a,2a1,3,4,4a,7,7a-octahydro-1H-cyclopenta[cd]indene	C15H24	NIST spectral library comparison	S37
		(1aS,7bR)-3,3,5,7b-tetramethyl-1a,2,3,4,4a,5,6,7b-octahydro-1H-cyclopropa[e]azulene	C15H24	NIST spectral library comparison	S40
	brasila-5,10-diene	(3S,3aR)-3,5,5-trimethyl-7-(prop-1-en-2-yl)-2,3,3a,4,5,6-hexahydro-1H-indene	C15H24	NIST spectral library comparison	S40
	caryophyllene	(1R,9S,Z)-4,11,11-trimethyl-8-methylenebicyclo[7.2.0]undec-4-ene	C15H24	NIST spectral library comparison	S40
	β-copaene	(1R,2S,8S)-8-isopropyl-1-methyl-3-methylenetricyclo[4.4.0.02,7]decane	C15H24	NIST spectral library comparison	S41

Table S7. Summary table for compounds shown in the Supplementary Information.

Supplementary Figures

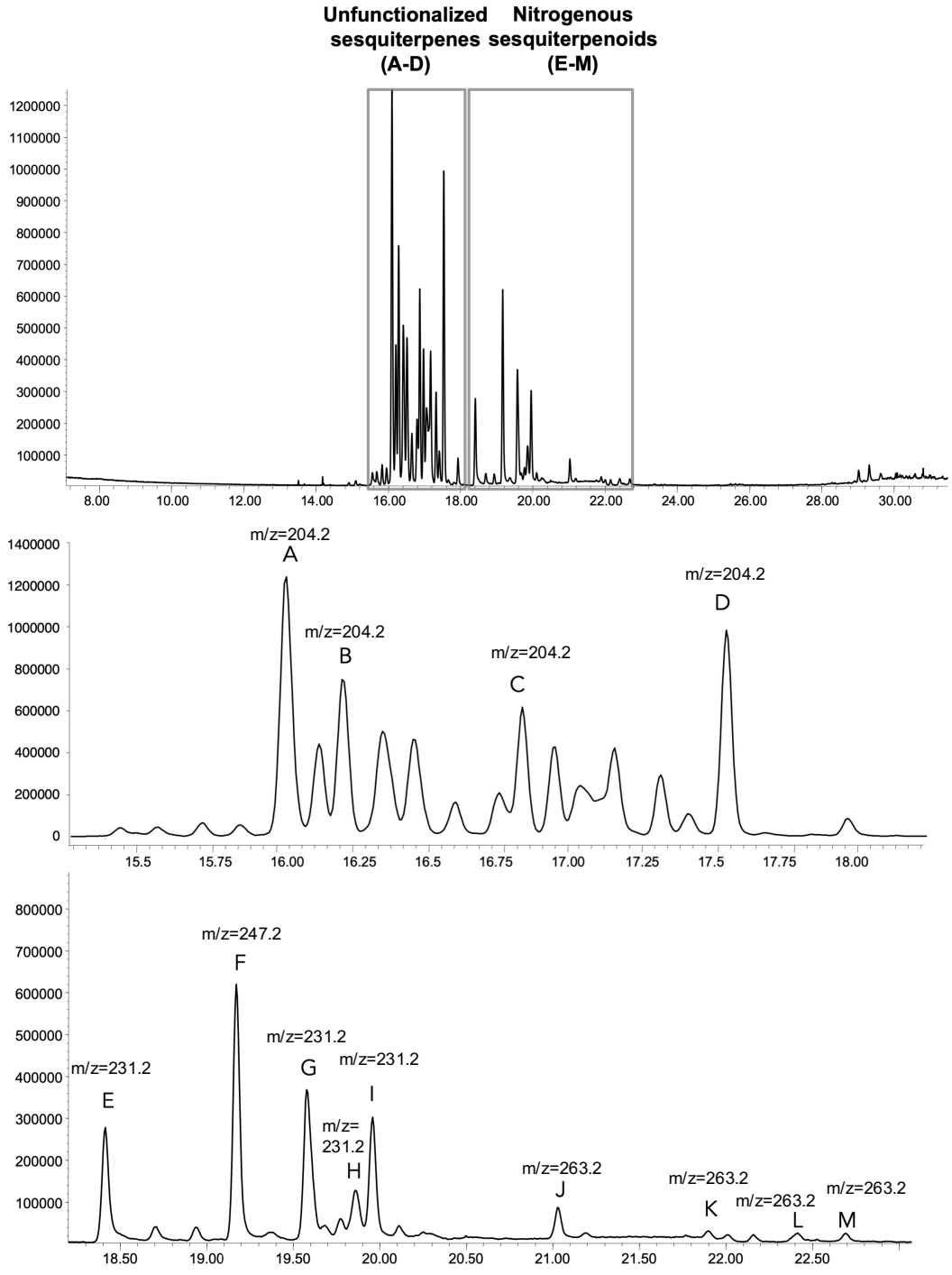


Figure S1. GCMS analysis of crude methylene chloride extracts of undescribed Bubarida sponge. Top: complete GCMS chromatogram; middle: Zoom on signals corresponding to unfunctionalized sesquiterpenes; bottom: Zoom on signals corresponding to isonitrile sesquiterpenoids. Mass spectra for major peaks are shown in Fig. S2 in alphabetized boxes on the next page.

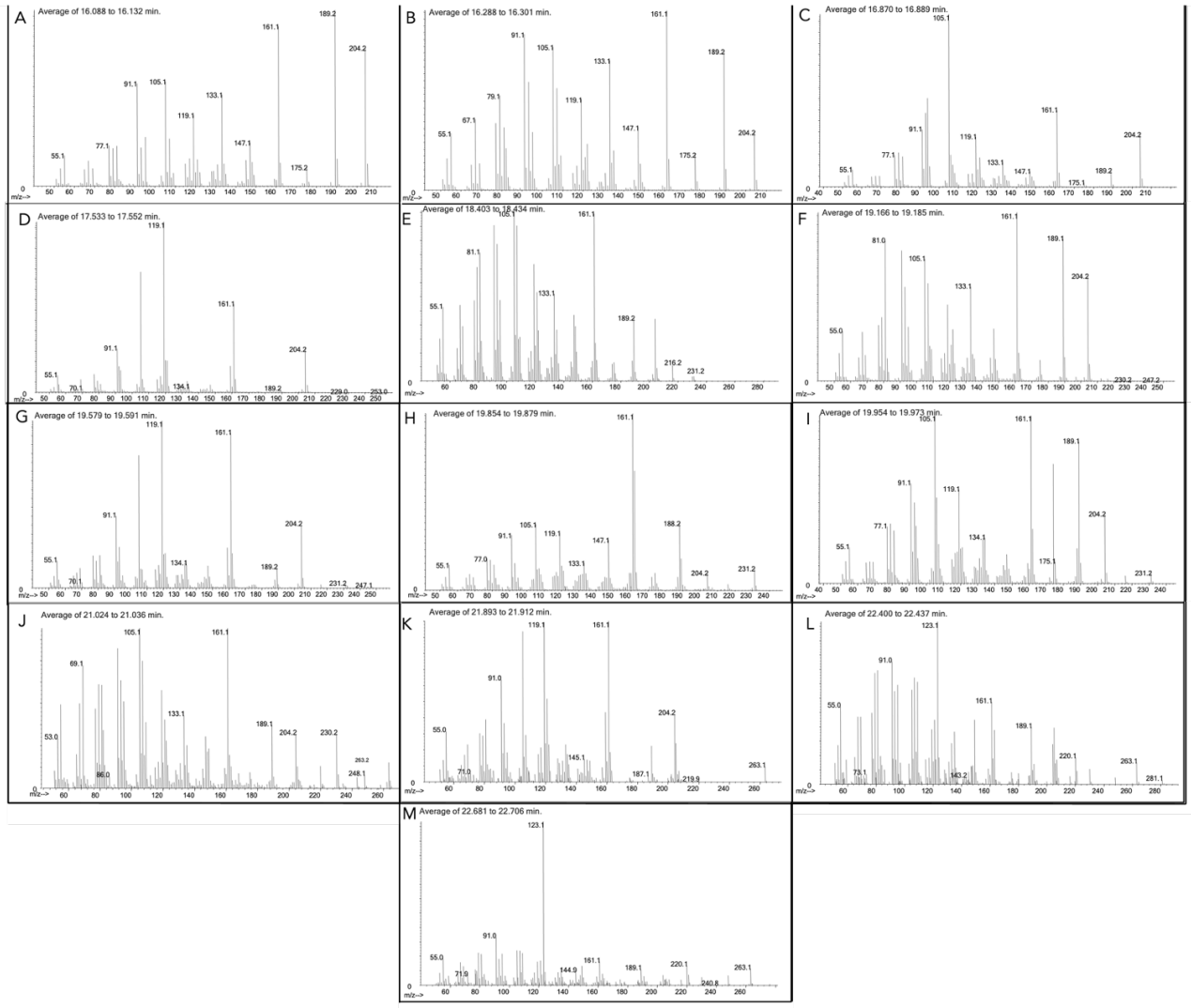


Figure S2. Mass spectra for the alphabetized peaks shown in Fig. S1.

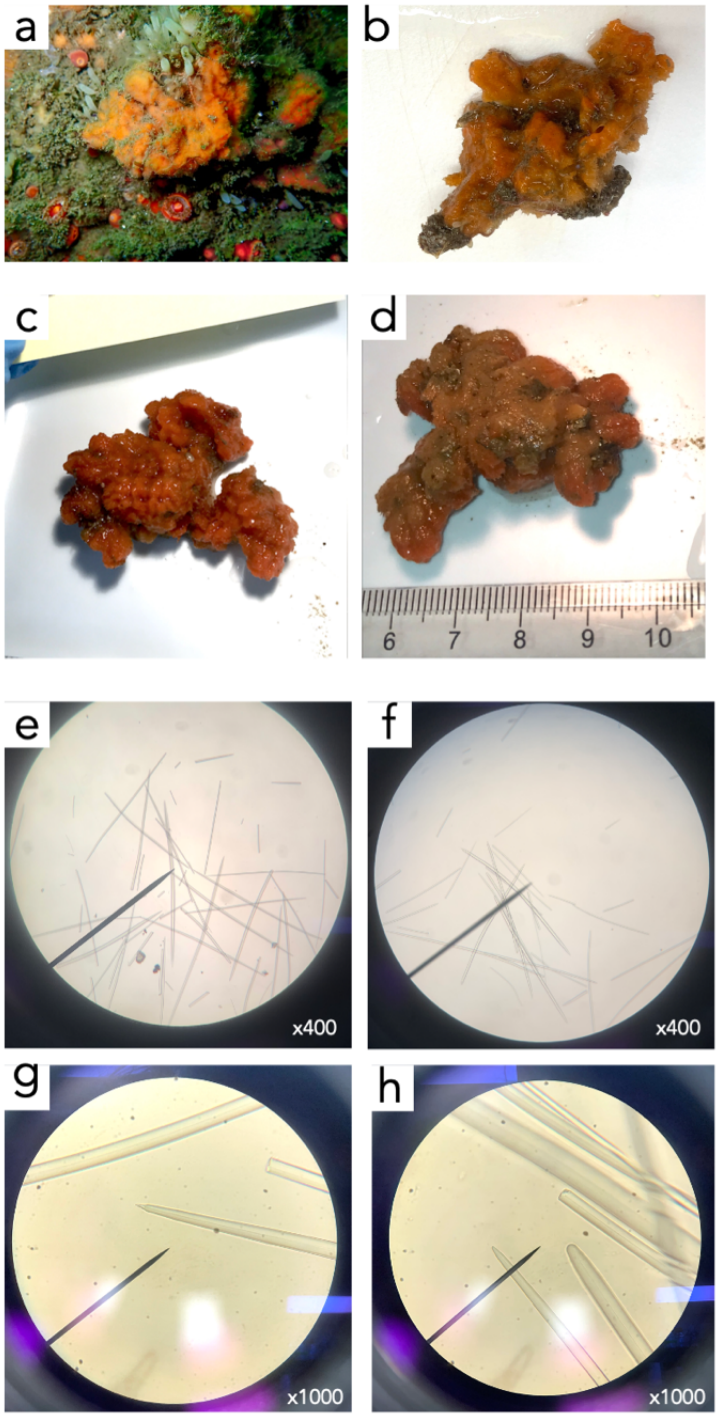


Figure S3. A. Underwater *in situ* photo of the undescribed Bubarida at God's Rock (N32° 51.344', W117° 16.621') in San Diego, California, United States and B. the same undescribed Bubarida specimen in the laboratory post-collection. C. Top view of another undescribed Bubarida specimen and D. the underside of that same undescribed Bubarida specimen with metric ruler for scale. E-F. Light microscopy images of undescribed Bubarida spicules at x400 magnification and G-H. at x1000 magnification.

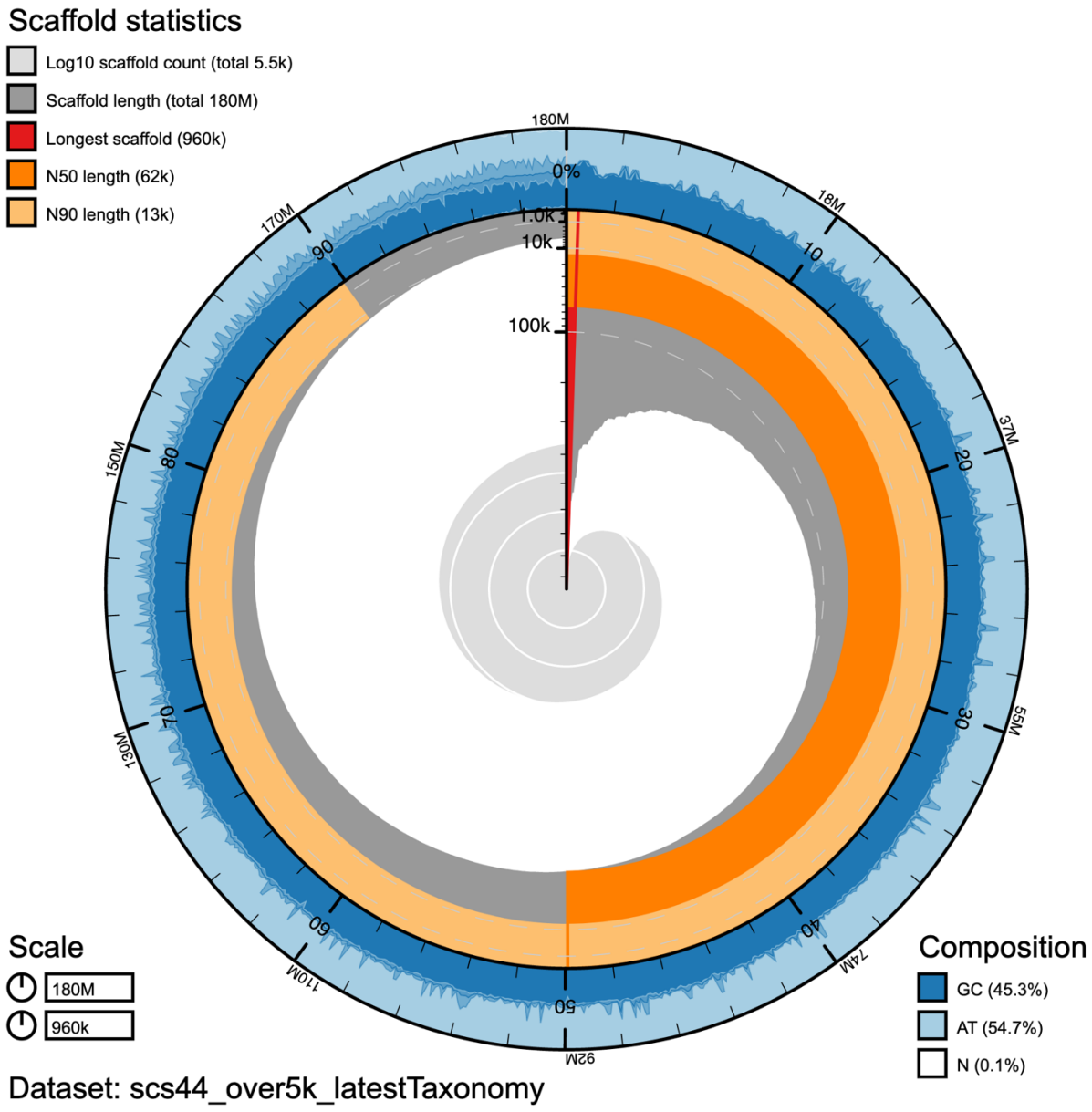


Figure S4. BlobTools snail plot of contigs 5kb and longer from the hybrid Illumina short-read and Oxford Nanopore long-read undescribed Bubarida metagenome assembly. The contig lengths are shown in dark grey, with the radius scaled to the length of the longest contig (960kb), which is shown in red. The N50 (62kb) is shown in orange and the N90 (13kb) is shown in pale orange. The entire 180 Mb assembly is represented in the outer ring, with GC/AT percentage represented in dark and light blue.

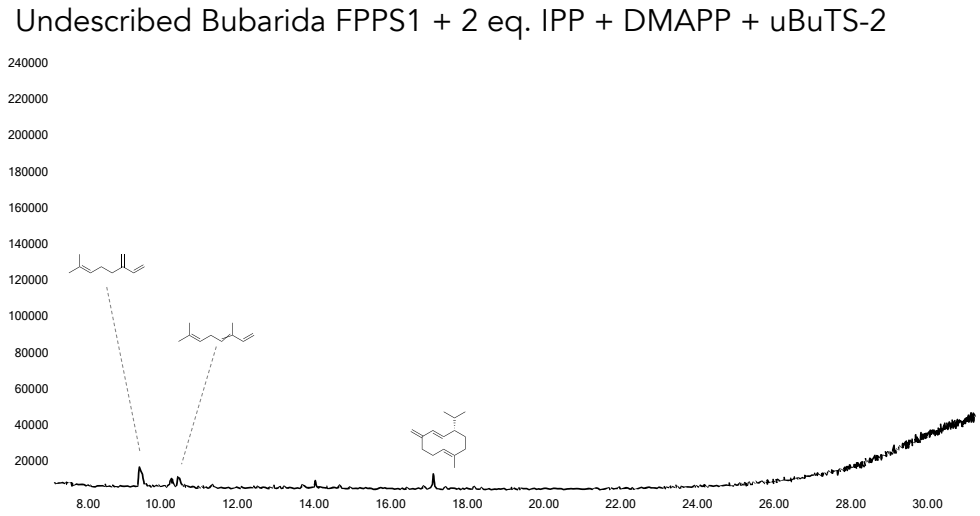


Figure S5. GCMS TIC trace of paired FPP synthase 1/uBuTS-2 assay. Structure annotations are based on NIST spectral library comparison and comparison with (-)-germacrene D **8** standard.

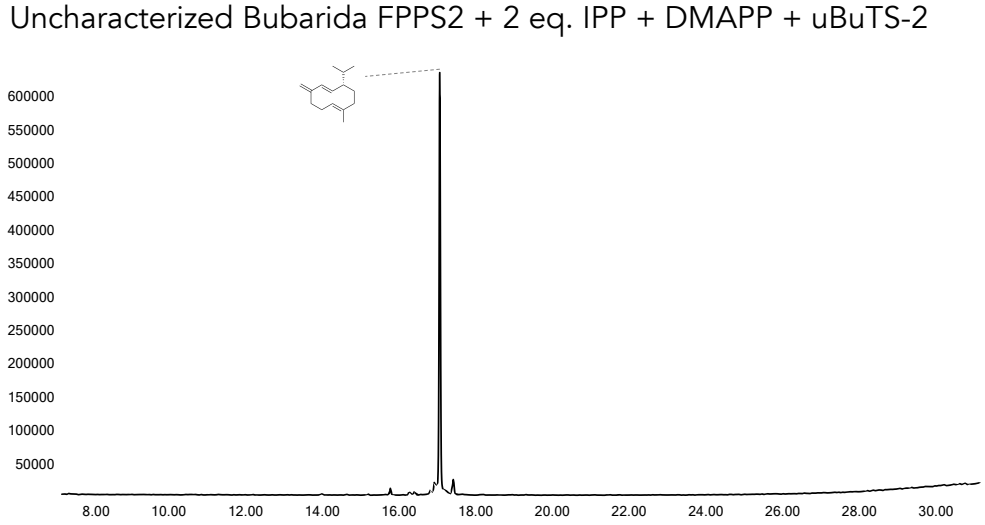


Figure S6. GCMS TIC trace of paired FPP synthase 2/ uBuTS-2 assay. Structure annotations are based on NIST spectral library comparison and comparison with (-)-germacrene D **8** standard.

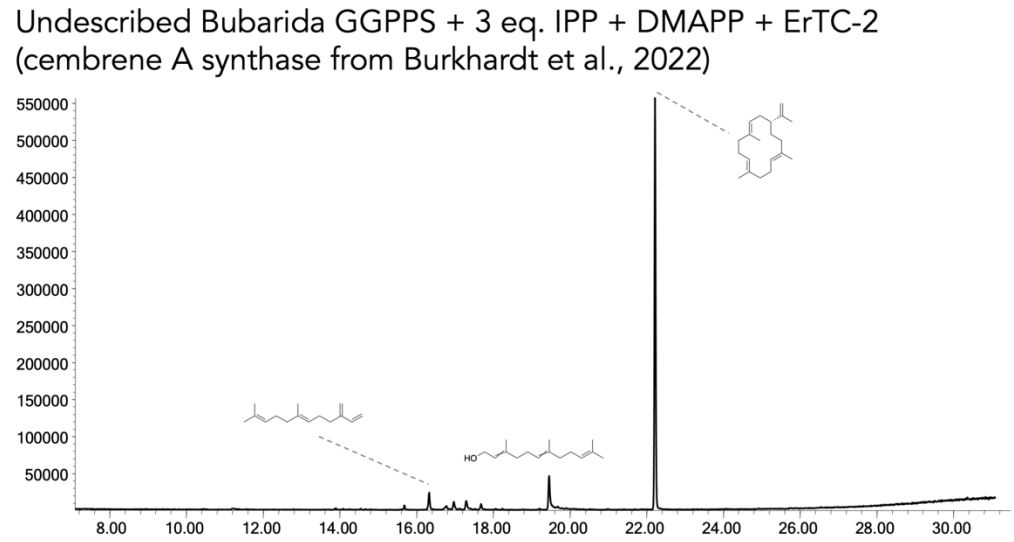


Figure S7. GCMS TIC trace of paired GGPP synthase/ErTC-2 assay(13). Structure annotations are based on NIST spectral library comparison and comparison with cembrene A standard.

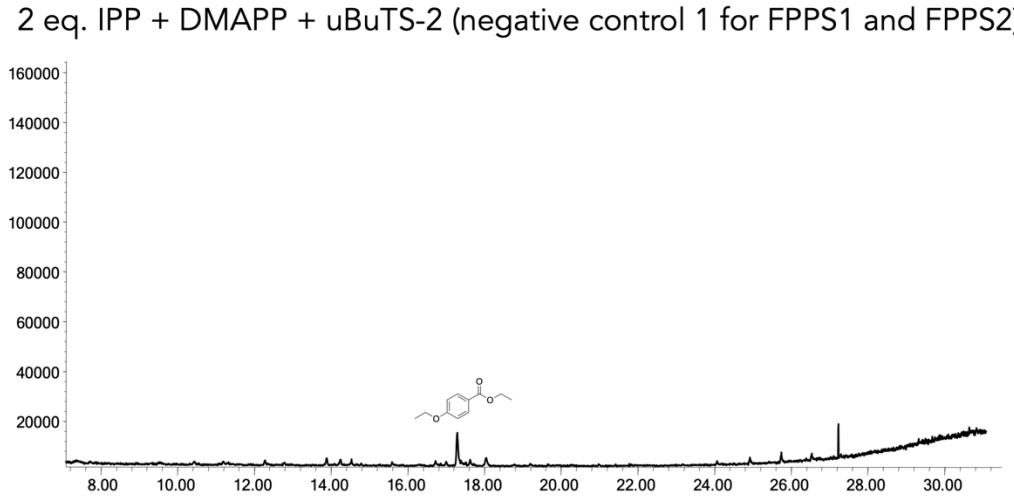


Figure S8. GCMS TIC trace of negative control 1 for FPPS1 and FPPS2. Structure annotations are based on NIST spectral library comparison.

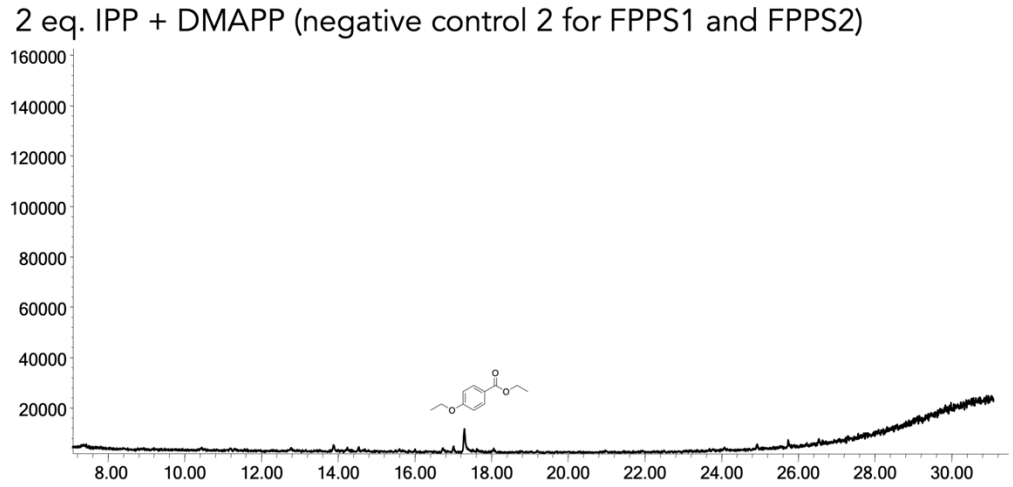


Figure S9. GCMS TIC trace of negative control 2 for FPPS1 and FPPS2. Structure annotations are based on NIST spectral library comparison.

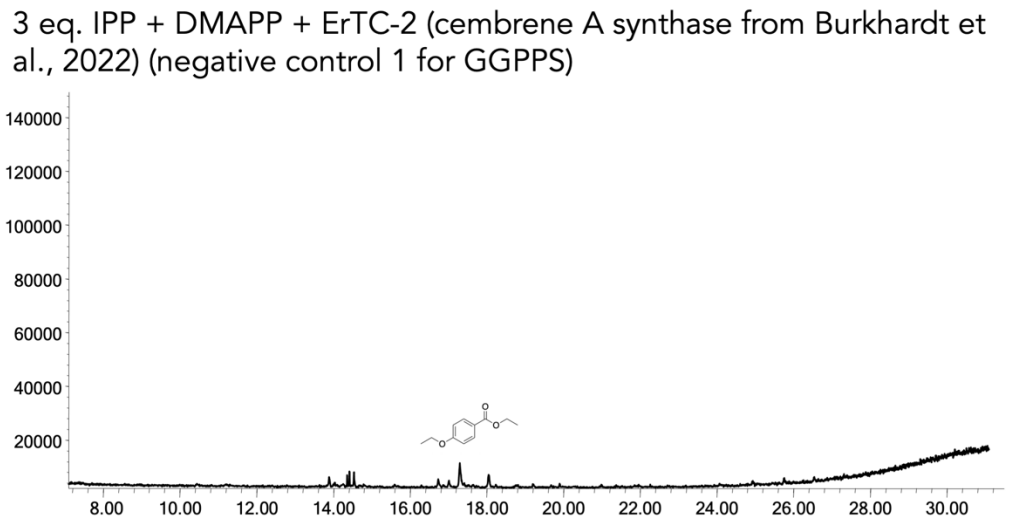


Figure S10. GCMS TIC trace of negative control 1 for GGPPS(13). Structure annotations are based on NIST spectral library comparison.

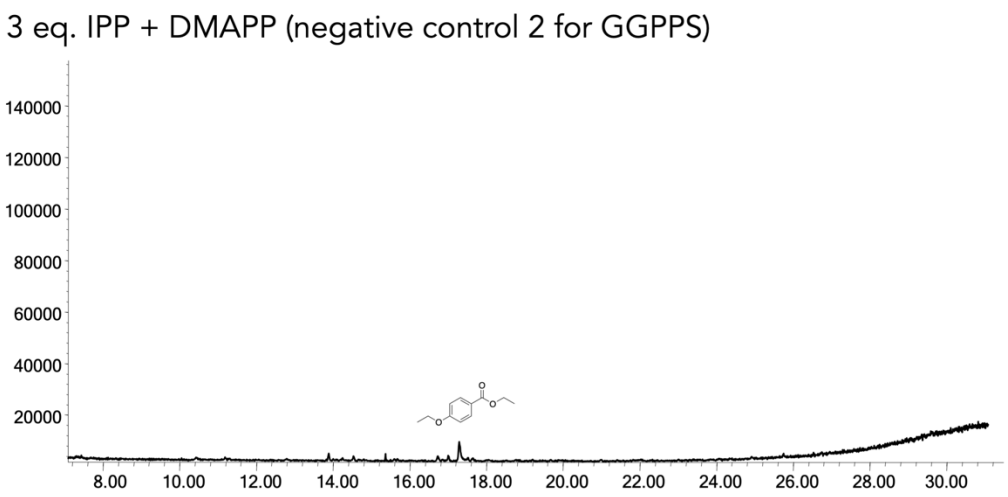


Figure S11. GCMS TIC trace of negative control 2 for GGPPS. Structure annotations are based on NIST spectral library comparison.

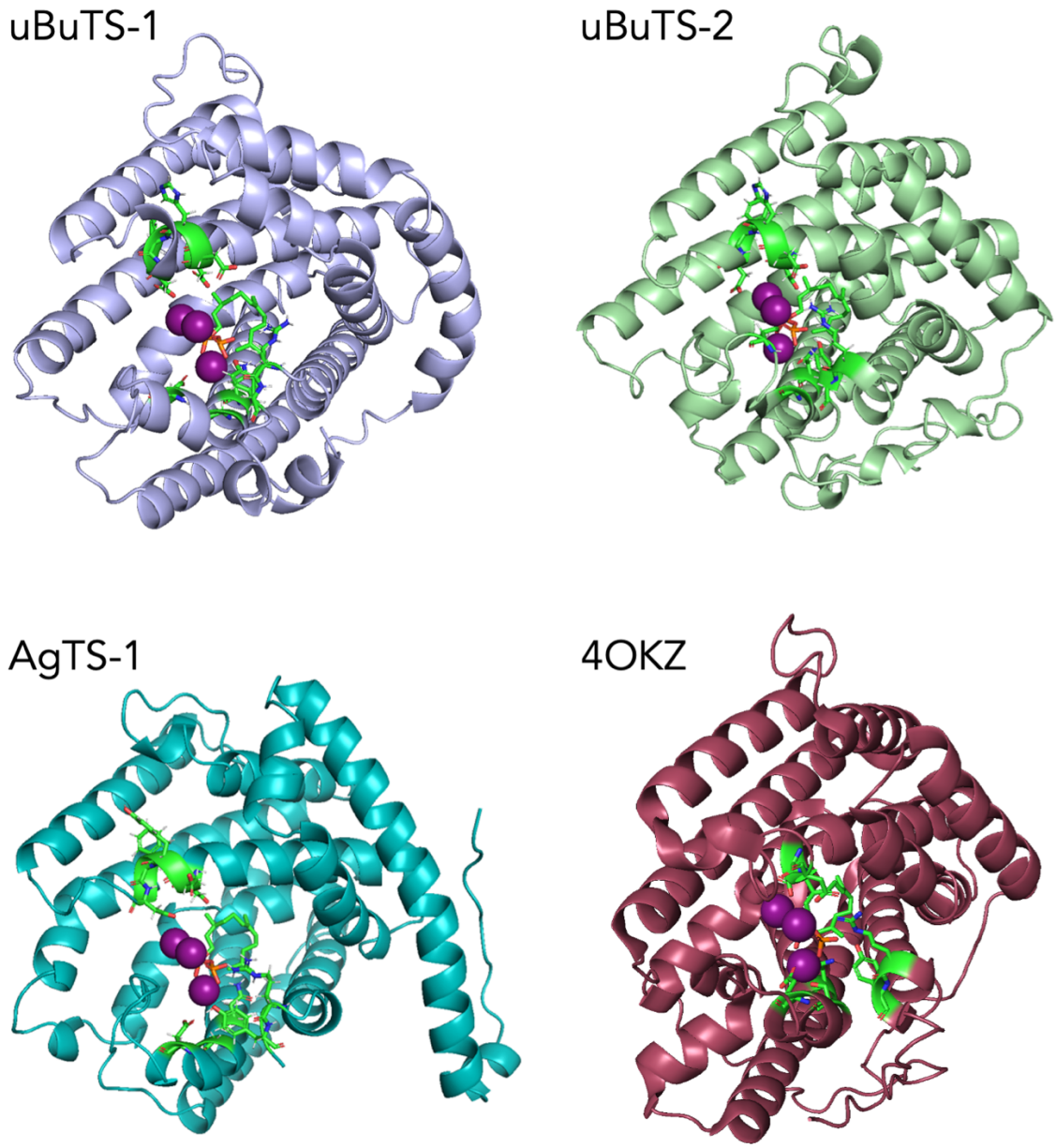


Figure S12. Homology models of uBuTS-1, uBuTS-2, and AgTS-1 and crystal structure of selinadiene synthase (PDB entry 4OKZ) (28). Amino acid sequences of sponge TSs were submitted to Robetta (https://robetta.bakerlab.org/) for structure prediction using *Eleutherobia rubra* cembrane A synthase (PDB entry 7S5L) as a template. Sponge TS homology models were visualized using PyMOL (https://pymol.org/). FPP and Mg^{2+} locations are based on selinadiene synthase (PDB entry 4OKZ). Typical conserved residues for type I TSs are shown as sticks and highlighted in green.

>TRINITY_DN66204_c0_g1_i4 len=1370 path=[3:0-267 6:268-316 7:317-371 9:372-397 11:398-485 12:486-575 13:576-584 15:585-596 16:597-605 18:606-606 19:607-620 21:621-621 22:622-633 23:634-668 24:669-669 26:670-675 27:676-676 29:677-678 30:679-748 33:749-749 36:750-765 39:766-766 42:767-824 45:825-901 48:902-975 51:976-1000 54:1001-1037 58:1038-1039 59:1040-1045 61:1046-1049 62:1050-1073 64:1074-1077 65:1078-1093 68:1094-1369]
GCAGGTAGGAAGTTAGTTAAAGGCTATTAAGGATGAAATGCTACACAGCCAGCTCAGTAGTCTAACTTCCTCTCTAAATGTATTT
GCACTTTGACCTTAGGGCAGTATATAAAGGTGTTCTGCAGGCTCTGAATAATTGCCTGTTTAAGACCTGCATGCAGCTGCAAAT
GTAAATTGTATTGTTCAGTCATC**ATGTCACTTAGTGTTTTTGCCAAGATTCCTGTCCCTACATTGGAGCTTGAGCATCCAATGCA**
TTATGTAGACTCTACTGTTAGCGTGGCTAAAGTTCGAATATCATCAGTGAGTGCAGGACTAAACAATAGATTGTATCGCCGTG
TATTGGATTTAAAACTCAACCCGGAATACATCAACCACATGGATATTAATGACTTTGTTGCCAAAGCCTGCAATCTCGGTCTT
CCAACCCCTGACCAGGAACGGCTGTGGATTTTCACTTCAGTTAATGCACTTTTTATATTCGAGTTCGACGATCACTTTGACAA
GCCACTACTCATCGCACCAGAGTATGTTGCAAGGCGATCCAAGCAGATGAGGGCTGTGTTGAGAAGCCTAAGCAGCCACAA
GCTTAGTGGCCTGCAGGGAAGCCTGGAGGATTGGCCTGCTGAAGTTCCATGCAAGGAGGCTTACCTTTGGCTTTTAAGGGAA
GCTAAAGATCTGAAGGAGGGCGCAGCAGAGCTGCTTCATGATATTTTTGTTGACTATTGTTTCGGTGTTGAAGAGGAGGTTAT
CGAGTGGAAATCAGATATGTATCGGGGTGATCTTACTGCCTGGAGTCTTGACCGCTACAAGGAGGTTCGCAAACGTGCAGCA
GGTGTAGTTTTTGCAATCGTAGTACCGCTCTTCGTCACCCGTAAGTGGATACAGAAGGAGCATGTCAACACATGCACTGACC
TTCTGTATGAGGCTGCACTCCTGGTTGGACTCTCGAACGACATACTAGGAATCCCAAGGGACCTTAGAGACAGCGGAACTAT
GACAGTGCTCAAGATAGCAAGCACCAGTGAGGTTGTCCAGCATCATAATATGATGGTTGAAGTACTCCACAAGAAAGTTCTA
GCGCTTGAAGGCAACACAATGCACTTCATGGAAGAGGTGGAGACTAGTGTAGCGGGGGGTTTTCCTGTGGCAGTGCCATTCA
AGACGTTATACTTTGTGAATTGTATTTTTTCCTCTATAGCTGTATAATCTGTGTGATTCCCTCATGATAATTCAGAACAATCAAGTT
TTAGTAATATCATCTTTTATAAATATTGTCATGCTCATACTTAATTGATGAATTAACAAATTTGTAAGTTTTTGCAATTGTCTCCTGC
CACTGACTGAATAAAATTTGCAAAATTAAAAAAAAAAA

>TRINITY_DN66204_c0_g1_i4.p1 type:complete len:328 gc:universal TRINITY_DN66204_c0_g1_i4:193-1176(+)
MSLSVFAKIPVPTLELEHPMHYVDSTVSVAKVRISSVSAGLNNRLYRRVLDLKLNPEYINHMDINDFVAKACNLGLPTPDQERLWIFTS
VNALFIFEFDDHFDKPLLIAPEYVARRSKQMRAVLRSLSSHKLSGLQGSLEDWPAEVPCKEAYLWLLREAKDLKEGAAELLHDIFVDY
CFGVEEEVIEWKSDMYRGDLTAWSLDRYKEVRKRAAGVVFAIVVPLFVTRKWIQKEHVNTCTDLLYEAALLVGLSNDILGIPRDLRDS
GTMTVLKIASTSEVVQHHNMMVEVLHKKVLALEGNTMHFMEEVETSVAGVFLWQCHSRRYTL*

Figure S13. Polyadenylated transcript with uBuTS-10 sequence bolded. Translated uBuTS-10 amino acid sequence shown below.

```
CLUSTAL O(1.2.4) multiple sequence alignment


Ga0070431_10856351      ------------------------------------------------------------    0
Ga0072505_11956581      MSRFVEVPVKALALQHPFKYAHSALSVAKLREASVSSLLKTRLRNHLSALNLNPKYMEAL    60
StTS-1                  MSRFVEVPVKALALQHPFKYAHSALSVAKLREASVSSLLKTRLRNHLSALNLNPKYMEAL    60


Ga0070431_10856351      --------------------------------CFFVFDDHFDEQVGTPEIVAKLSLEMRD    28
Ga0072505_11956581      NIPDFIAKIYGIECPSTEDKERQWVFSATFTNCFFVFDDHFDEQVGTPEIVAKLSLEMRD    120
StTS-1                  NIPDFIAKIYGIECPSTEDKERQWVFSATFTNCFFVFDDHFDEQVGTPEIVAKLSLEMRD    120
                                                ****************************


Ga0070431_10856351      ILRALSRHTLSGLQGVLDDWPTDVPCKKAYLWLLQEAEGLRKGAAELVHATFNDYCLGVE    88
Ga0072505_11956581      ILRALSRHTLSGLQGVLDDWPTDVPCKKAYLWLLQEAEGLRKGAAELVHATFNDYCLGVE    180
StTS-1                  ILRALSRHTLSGLQGVLDDWPTDVPCKKAYLWLLQEAEGLRKGAAELVHATFNDYCLGVE    180
                        ************************************************************


Ga0070431_10856351      SEIVEWAPDVHRGDMSAWNLDRCTEVRKRSAAGNMALAPLYVINKWMTREHYRACNDLLY    148
Ga0072505_11956581      SEIVEWAPDVHRGDMSAWNLDRCTEVRKRSAAGNMALAPLYVINKWMTREHYRACNDLLY    240
StTS-1                  SEIVEWAPDVHRGDMSAWNLDRCTEVRKRSAAGNMALAPLYVINKWMTREHYRACNDLLY    240
                        ************************************************************


Ga0070431_10856351      DVSLIIALPNDVIQSVKRNNHSISMETTKILSADEIVQCHNKKVECLRKDILELDGDTRR    208
Ga0072505_11956581      DVSLIIALPNDVIQSVKRNNHSISME----------------------------------    266
StTS-1                  DVSLIIALPNDVIQSVKRNNHSISMETTKILSADEIVQCHNKKVECLRKDILELDGDTRR    300
                        *************************


Ga0070431_10856351      FMEEIEISAVGVFLWQCNCKRYVD*       232
Ga0072505_11956581      ------------------------       266
StTS-1                  FMEEIEISAVGVFLWQCNCKRYVD-       324
```

Figure S14. Clustal Omega (1.2.4) multiple sequence alignment of synthesized gene StTS-1 and the *Stylissa massa* TS fragments of which StTS-1 is comprised: Ga0070431_10856351(IMG study ID Gs0114528, project ID Gp0112735) and Ga0072505_11956581 (IMG study ID Gs0114528, project ID Gp0114746). Predicted key active site motifs are highlighted in yellow.

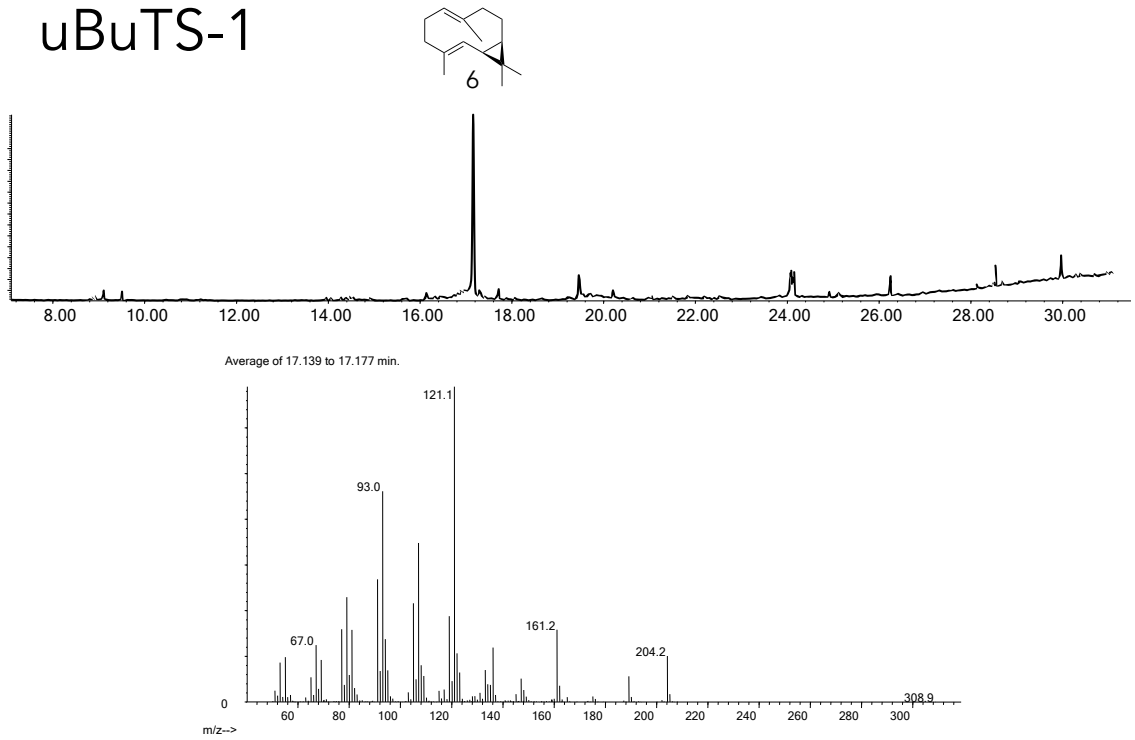


Figure S15. – GCMS TIC trace of the hexanes extract of *in vitro* assay with purified uBuTS-1 and FPP. The mass spectrum of the major peak is shown below. The structure of the main product, (+)-bicyclogermacrene **6**, is based on full structural characterization by NMR and OR (Supplementary Note).

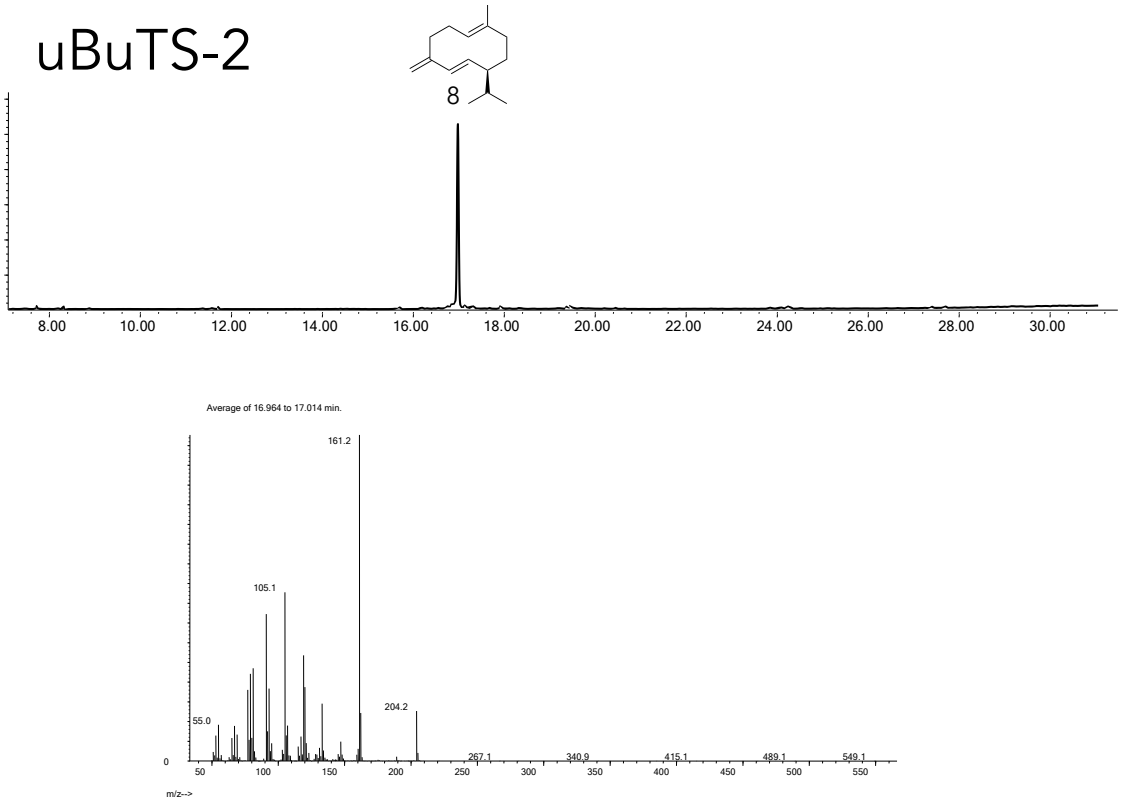


Figure S16. GCMS TIC trace of the hexanes extract of *in vitro* assay with purified uBuTS-2 and FPP. The mass spectrum of the major peak is shown below. The structure of the main product, (-)-germacrene D **8**, is based on full structural characterization by NMR and OR (Supplementary Note).

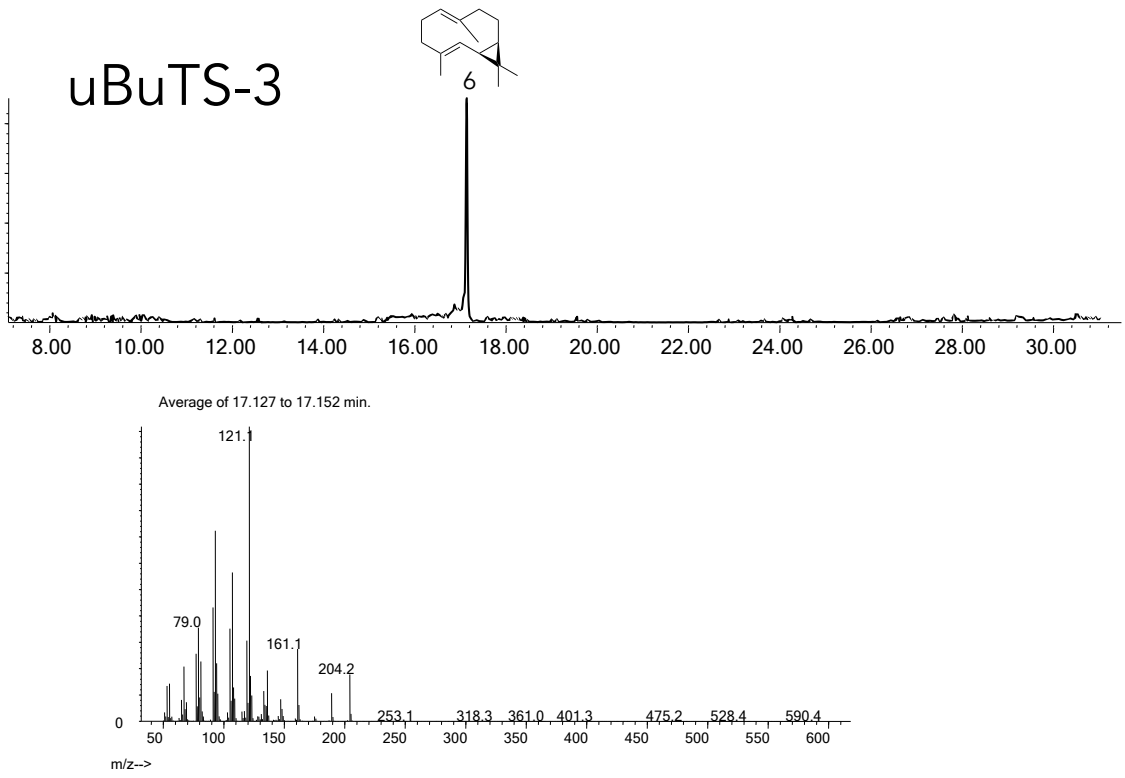


Figure S17. GCMS TIC trace of the hexanes extract of an *in vitro* assay with purified uBuTS-3 and FPP. The mass spectrum of the major peak is shown below. The structure of the main product, (+)-bicyclogermacrene **6**, is based on GCMS with purified standard.

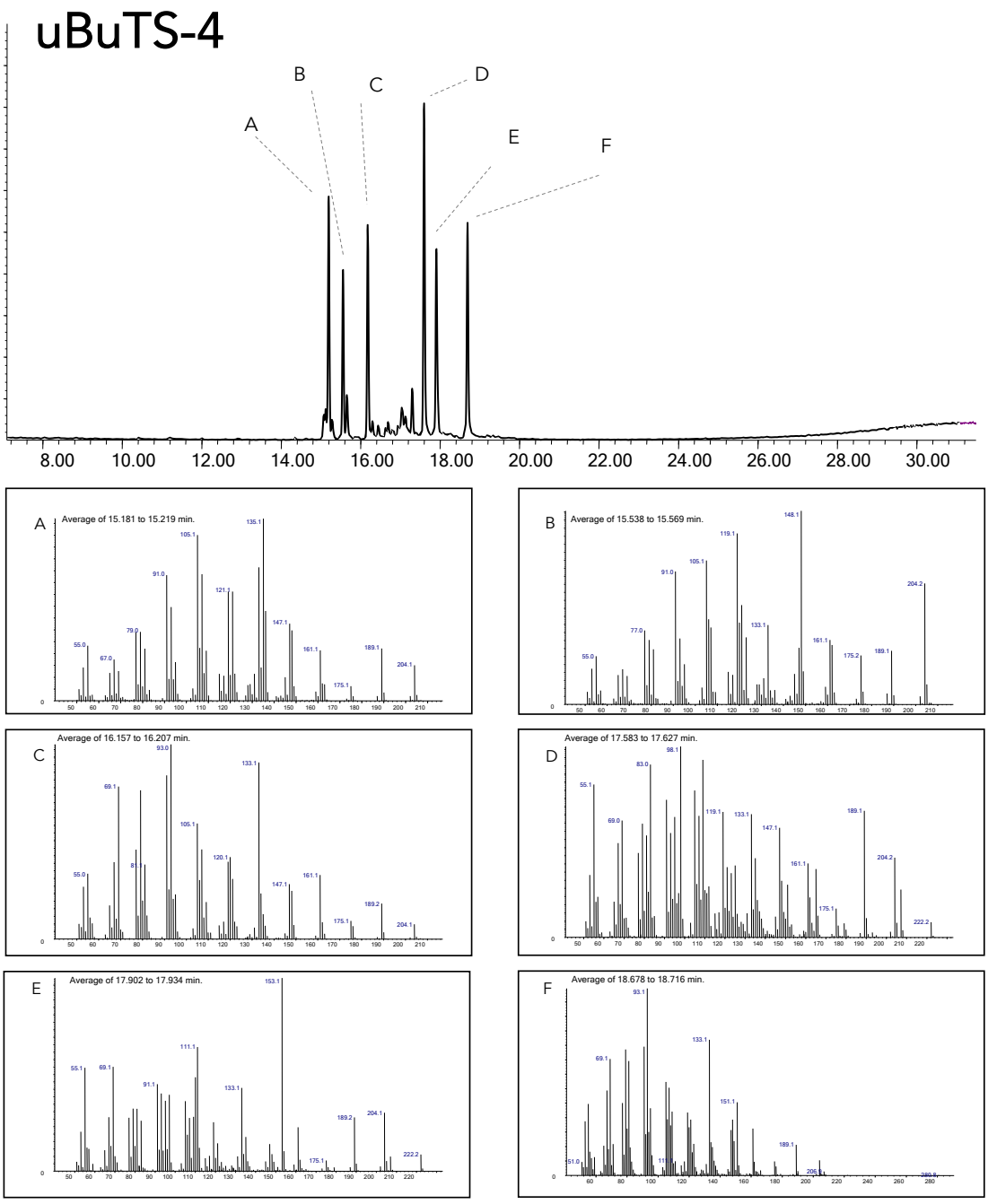


Figure S18. GCMS TIC trace of the hexanes extract of an *in vitro* assay with purified uBuTS-4 and FPP. Structure annotations are based on NIST spectral library comparison. Mass spectra for major peaks are shown in alphabetized boxes below the GCMS trace.

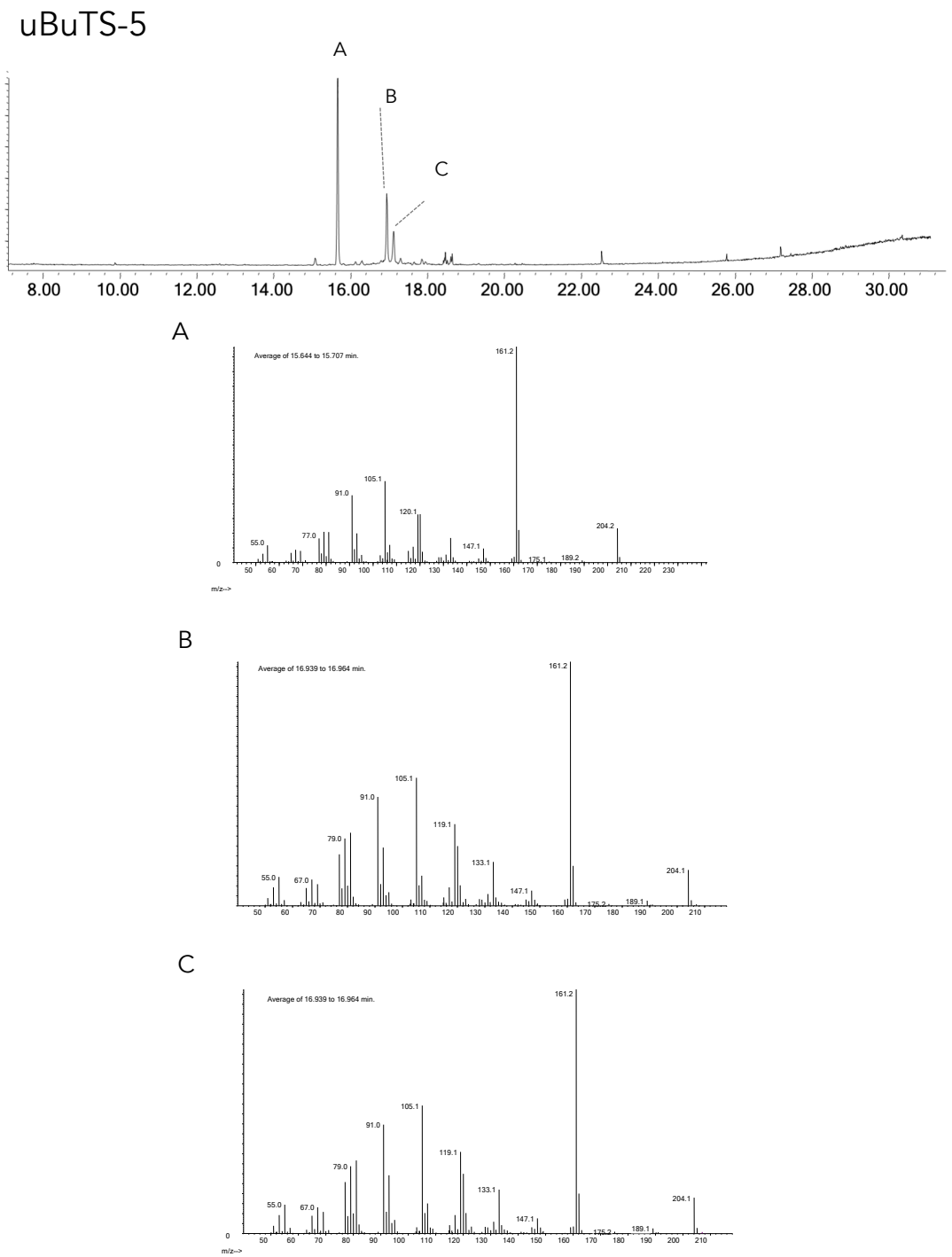


Figure S19. GCMS TIC trace of the hexanes extract of an *in vitro* assay with purified uBuTS-5 and FPP. Structure annotations are based on NIST spectral library comparison and comparison with (+)-bicyclogermacrene **6** and (-)-germacrene D **8** standards. Mass spectra for major peaks are shown in alphabetized boxes below the GCMS trace.

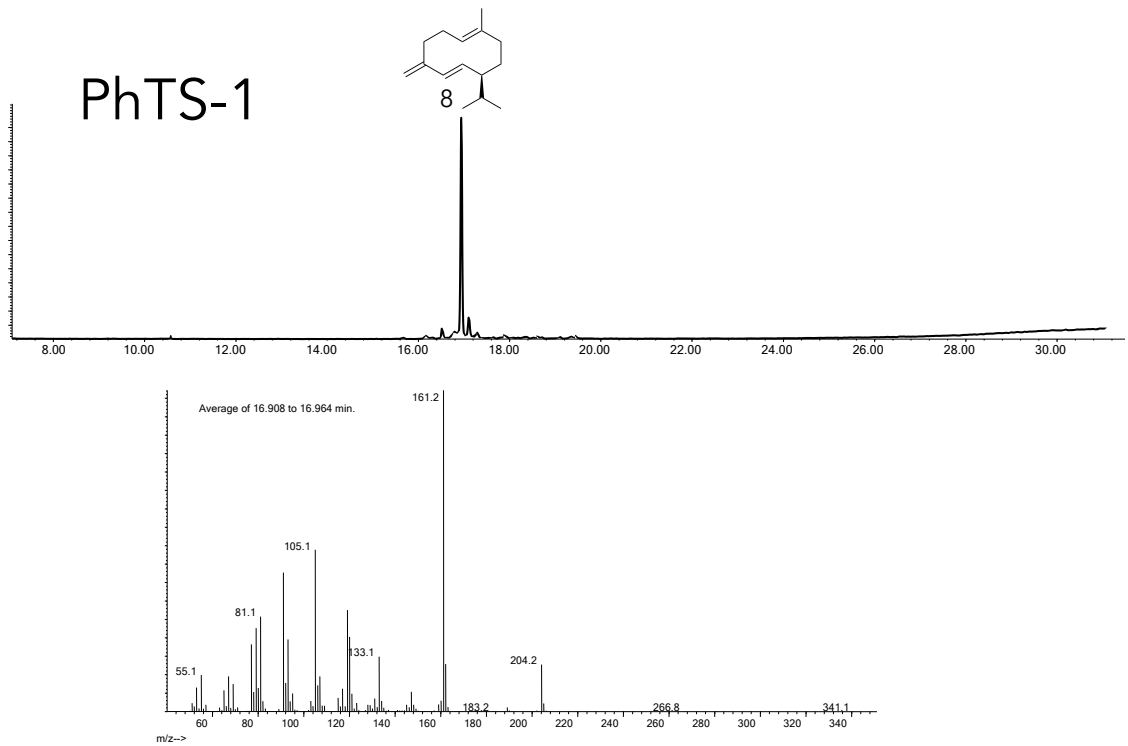


Figure S20. GCMS TIC trace of the hexanes extract of an *in vitro* assay with purified **PhTS**-1 and FPP. The mass spectrum of the major peak is shown below. The structure of the main product, (-)-germacrene D **8**, is based on GCMS with purified standard.

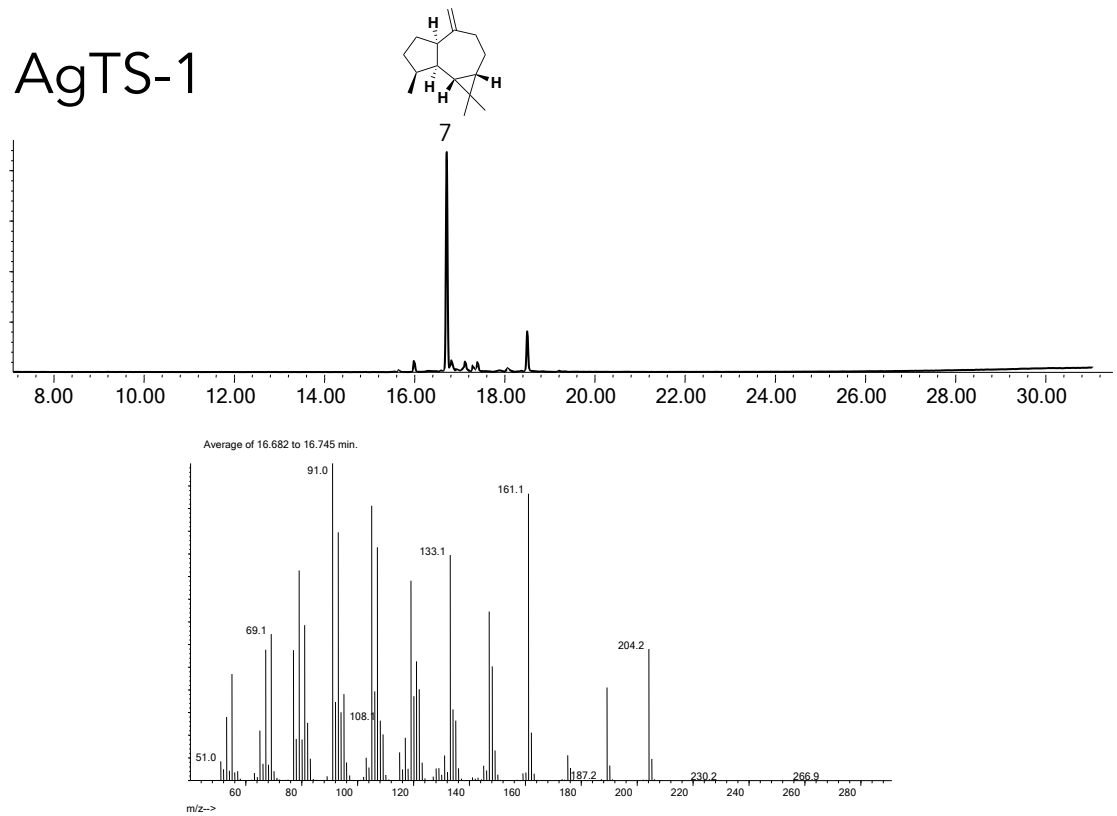


Figure S21. GCMS TIC trace of the hexanes extract of an *in vitro* assay with purified AgTS-1 and FPP. The mass spectrum of the major peak is shown below. The structure of the main product, (+)-allo aromadendrene **7**, is based on full structural characterization by NMR and OR (Supplementary Note).

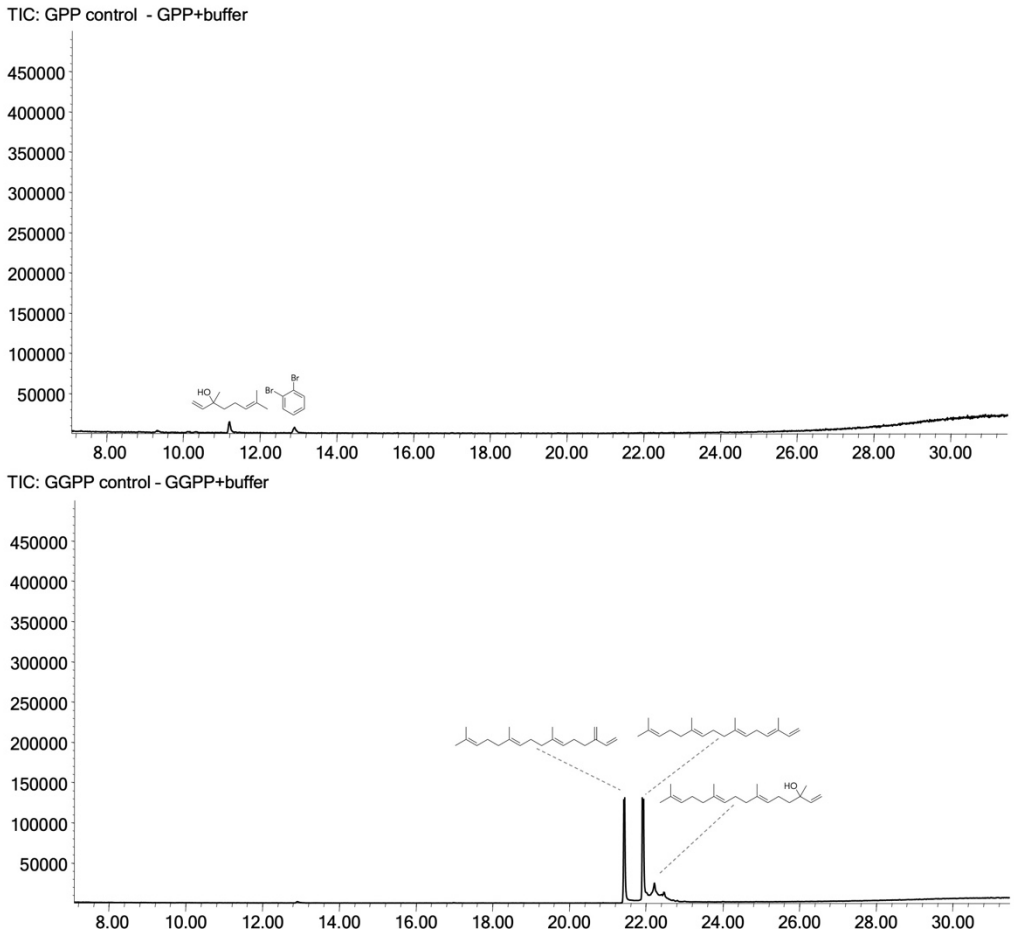


Figure S22. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (lower) controls. Annotations based on NIST spectral library comparison.

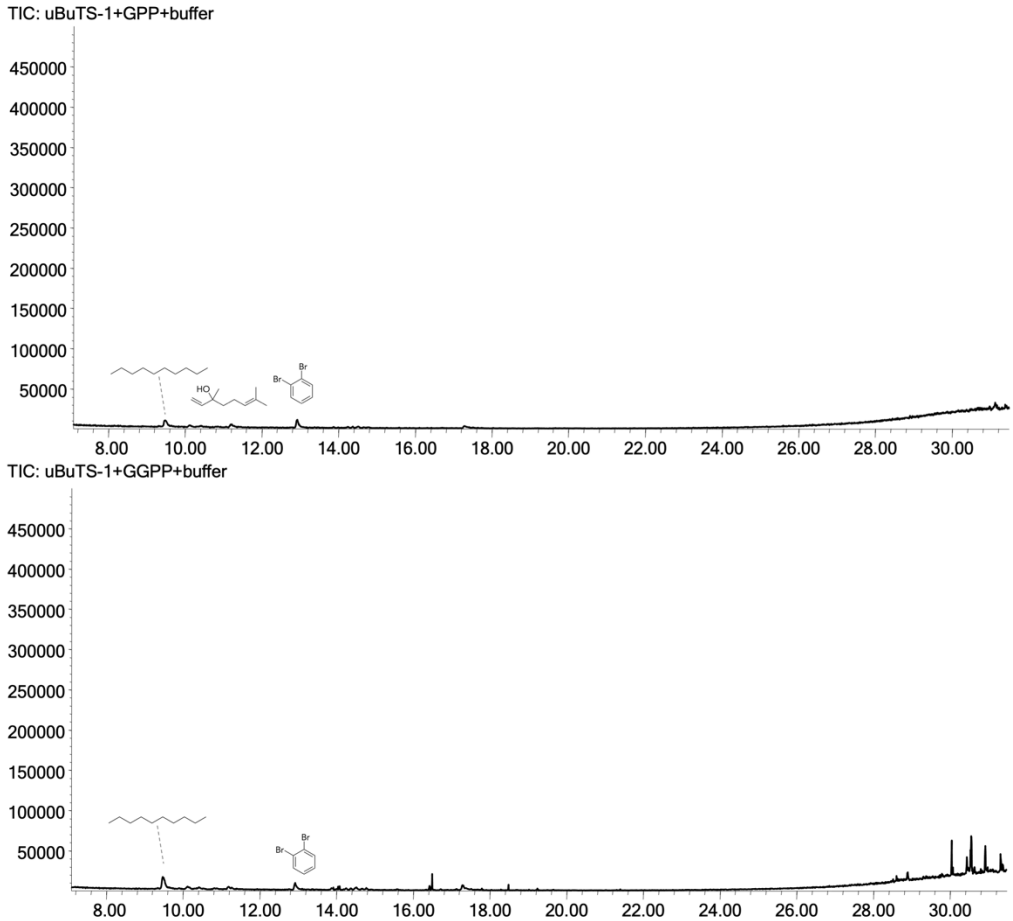


Figure S23. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (lower) assays with uBuTS-1. Annotations based on NIST spectral library comparison.

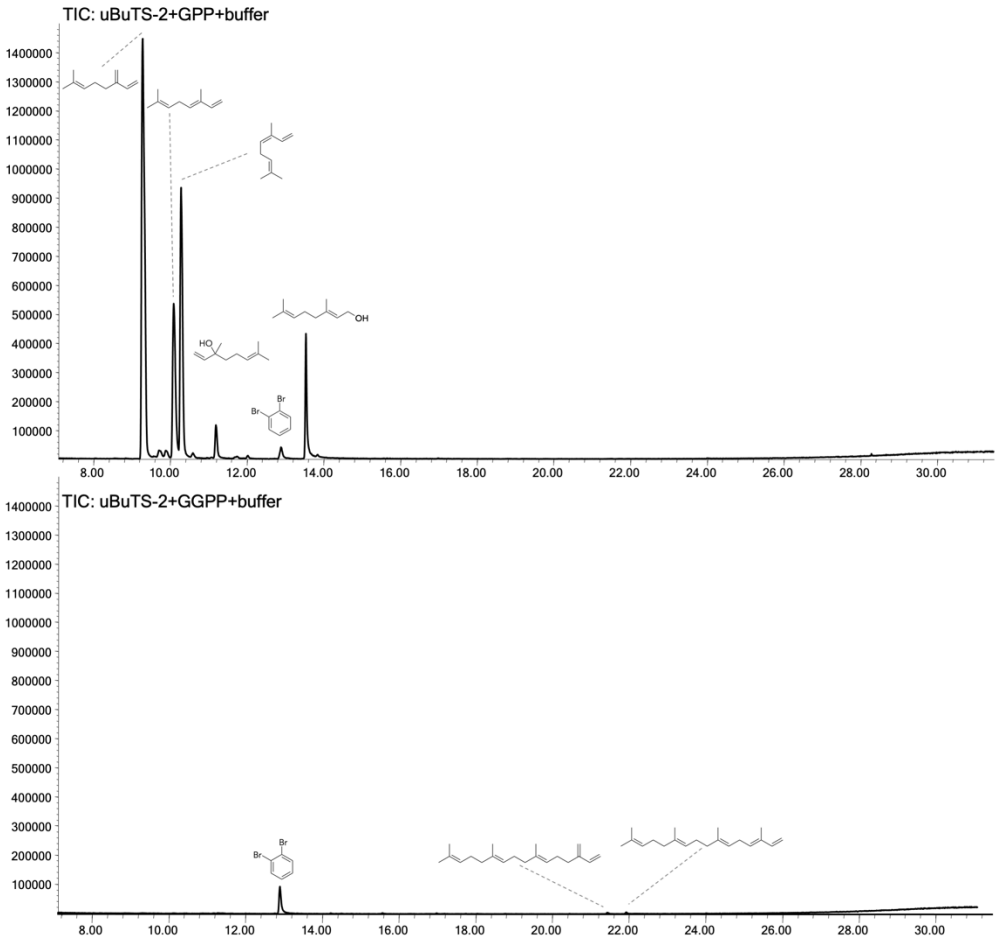


Figure S24. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (lower) assays with uBuTS-2. Annotations based on NIST spectral library comparison.

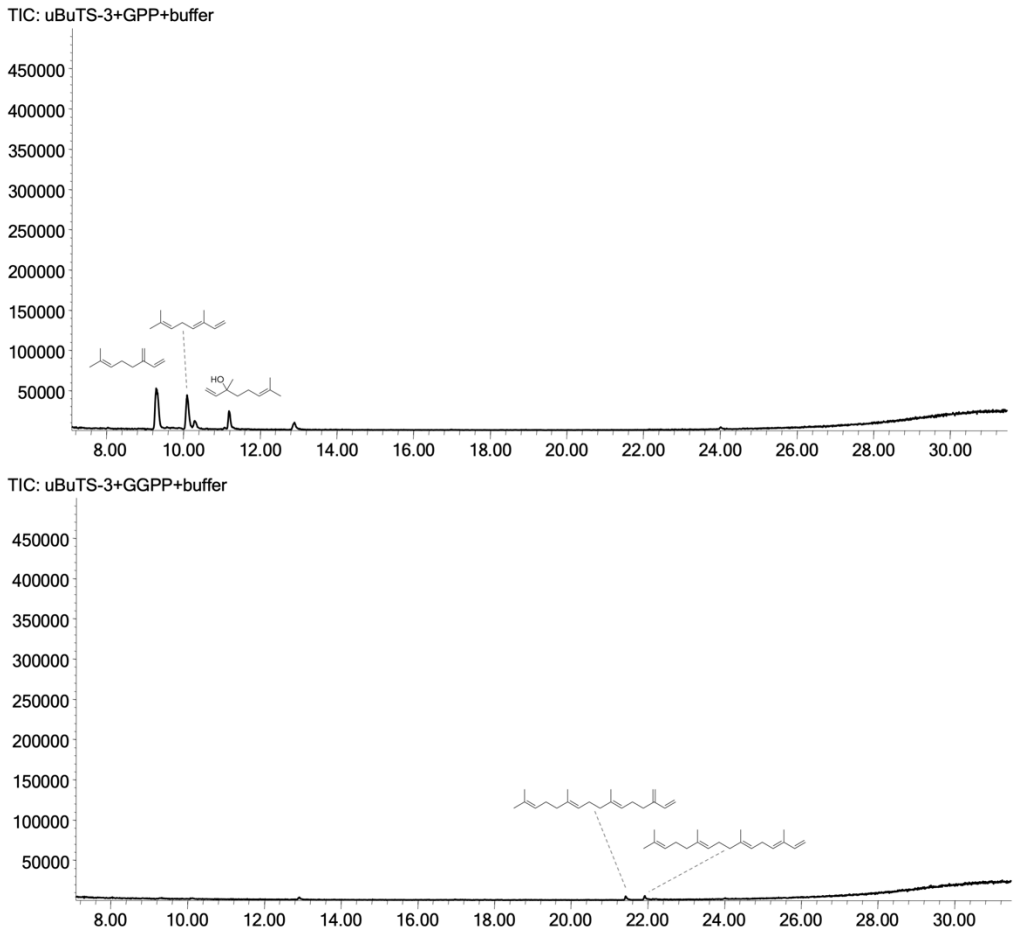


Figure S25. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (lower) assays with uBuTS-3. Annotations based on NIST spectral library comparison.

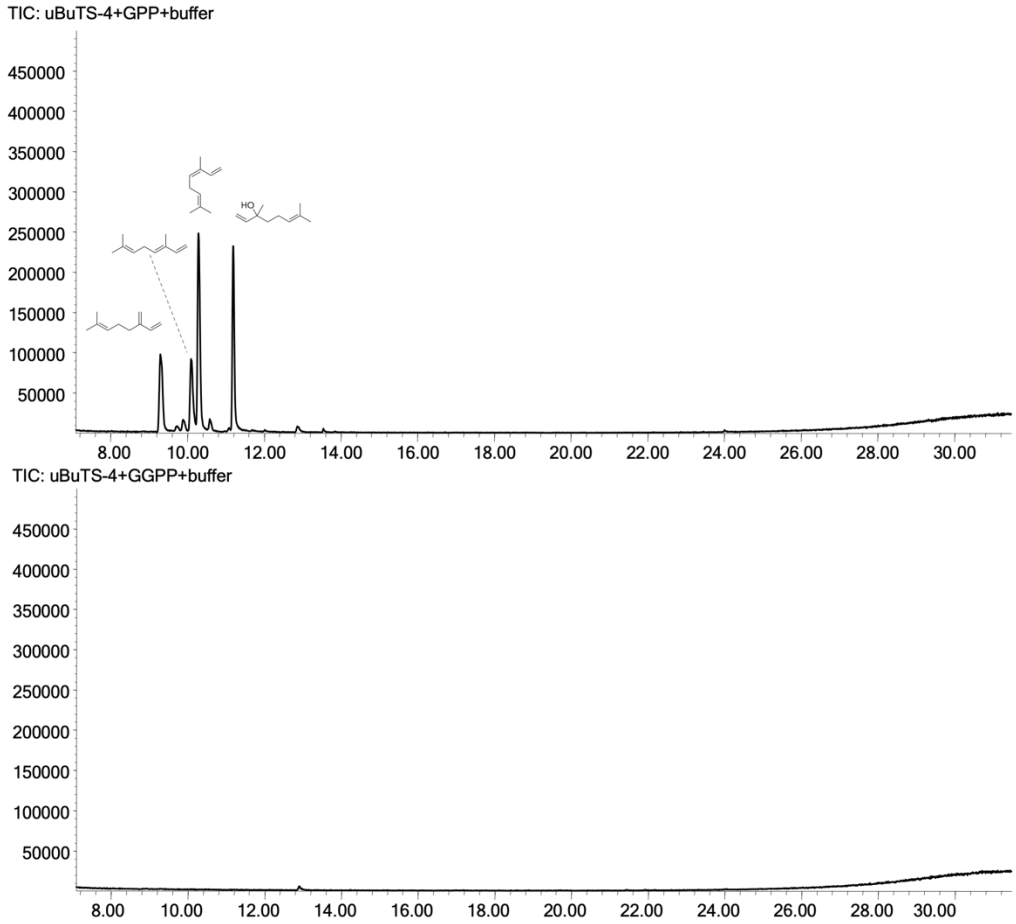


Figure S26. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (lower) assays with uBuTS-4. Annotations based on NIST spectral library comparison.

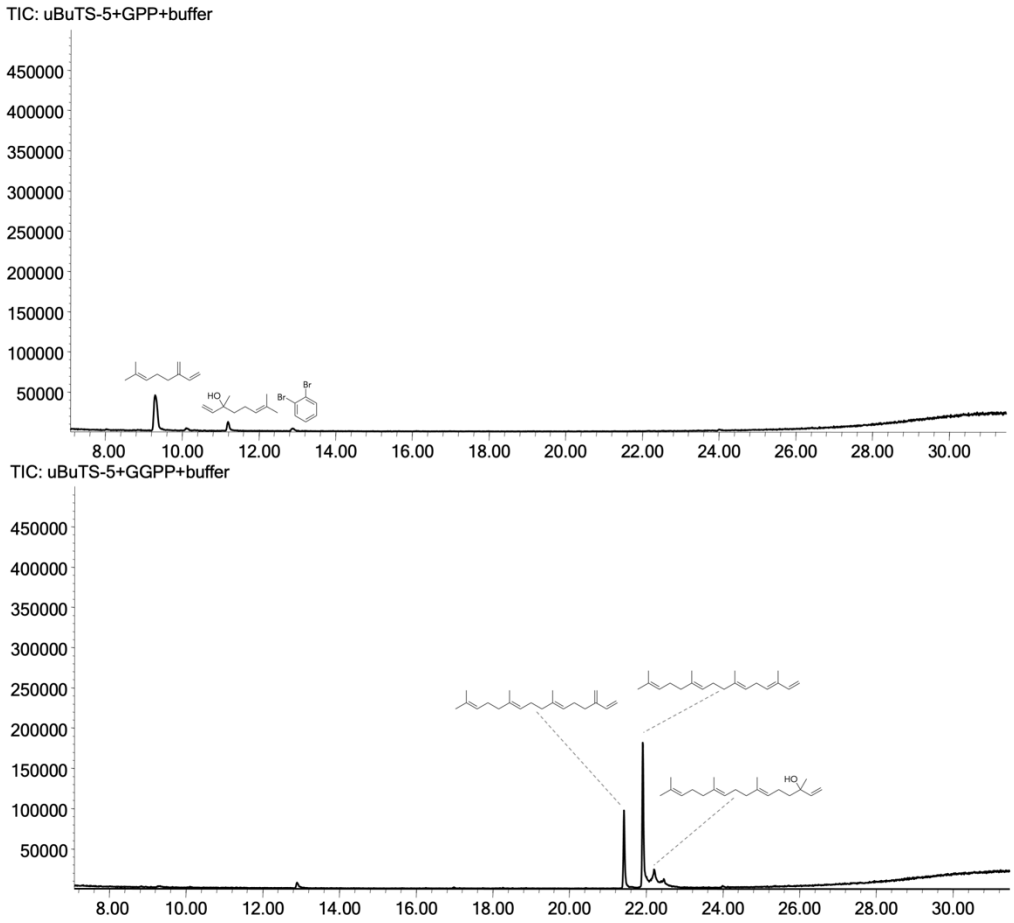


Figure S27. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (lower) assays with uBuTS-5. Annotations based on NIST spectral library comparison.

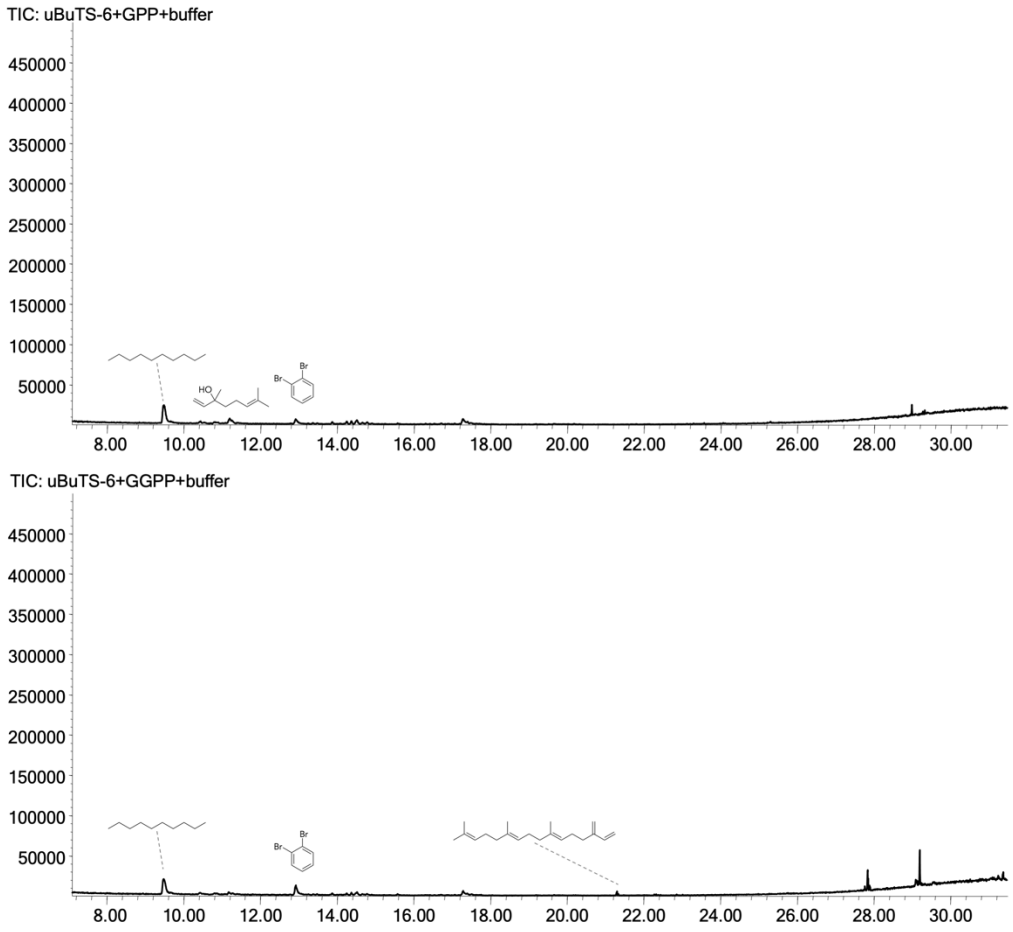


Figure S28. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (lower) assays with uBu-6. Annotations based on NIST spectral library comparison.

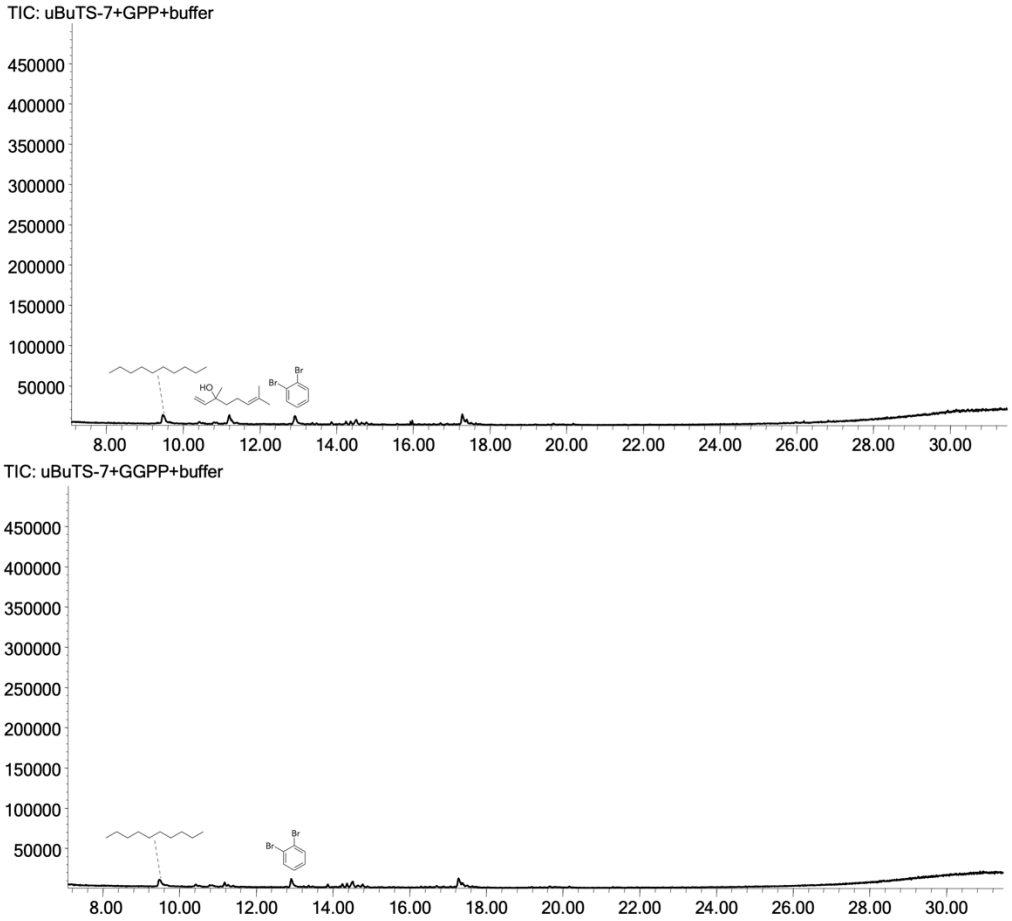


Figure S29. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (lower) assays with uBuTS-7. Annotations based on NIST spectral library comparison.

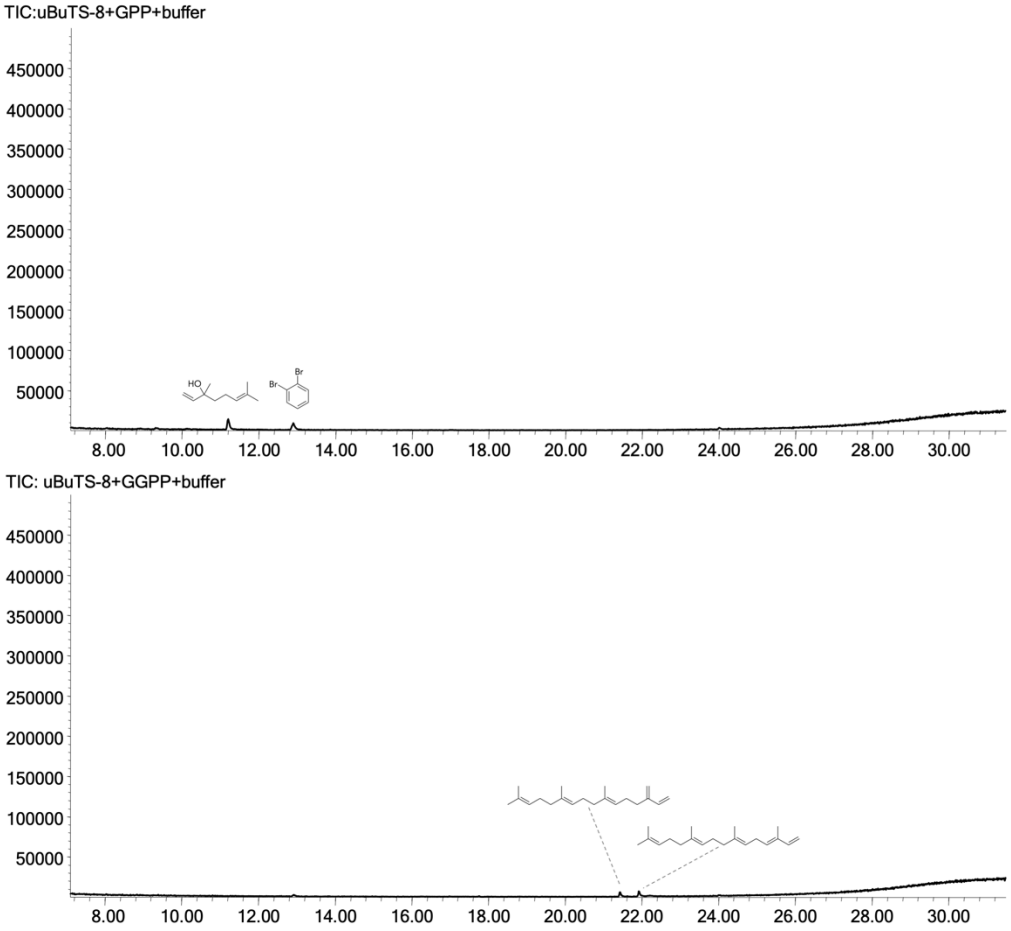


Figure S30. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (lower) assays with uBuTS-8. Annotations based on NIST spectral library comparison.

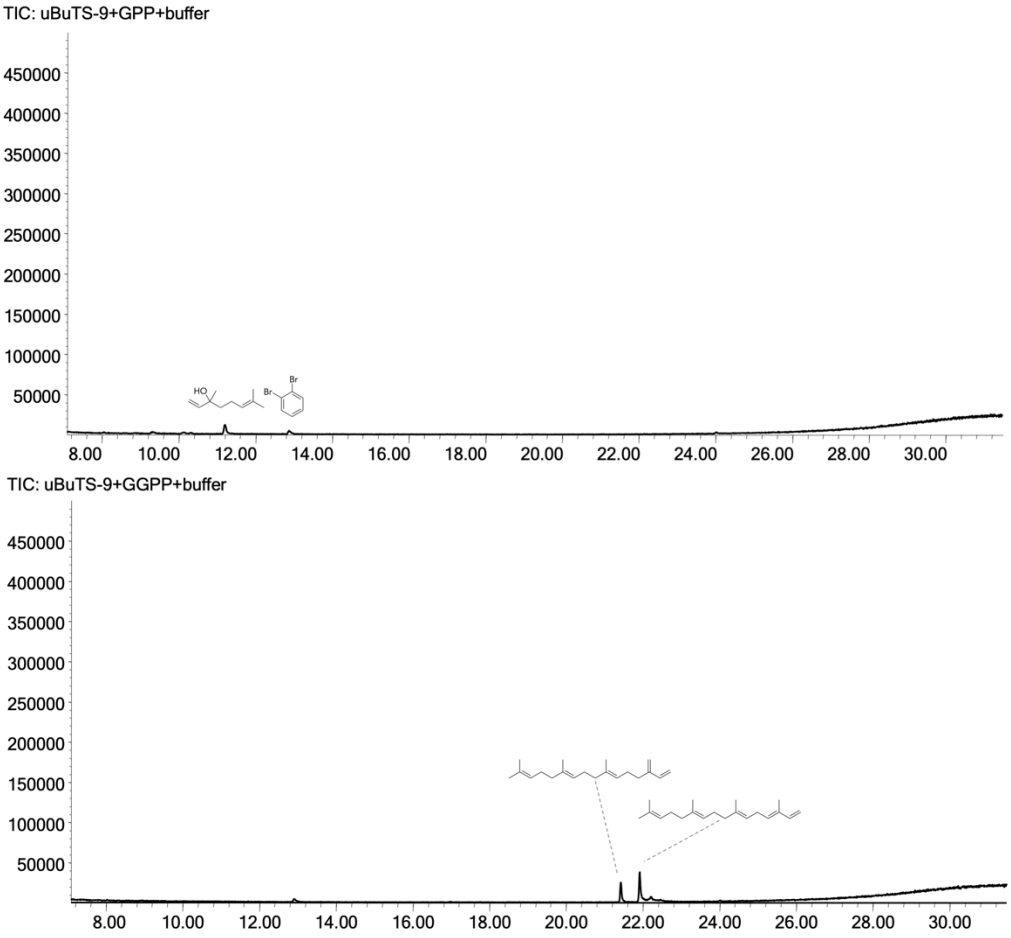


Figure S31. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (lower) assays with uBuTS-9. Annotations based on NIST spectral library comparison.

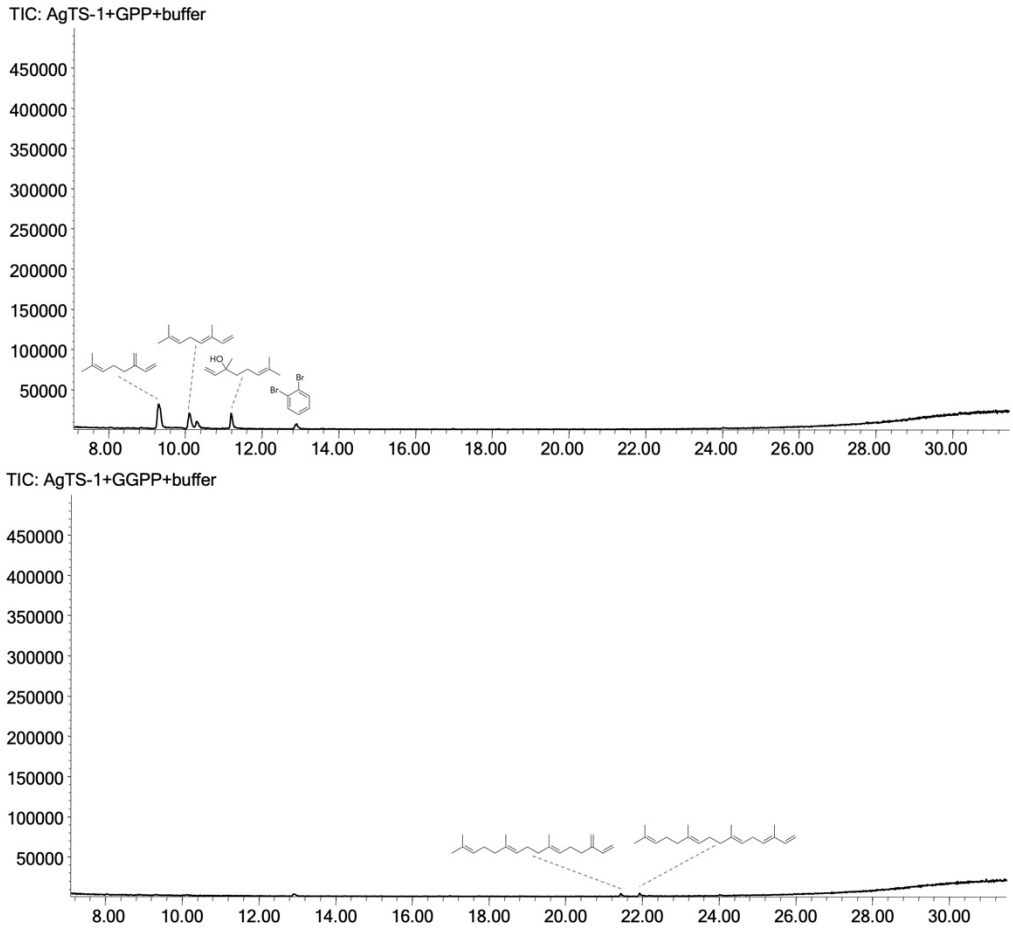


Figure S32. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (lower) assays with AgTS-1. Annotations based on NIST spectral library comparison.

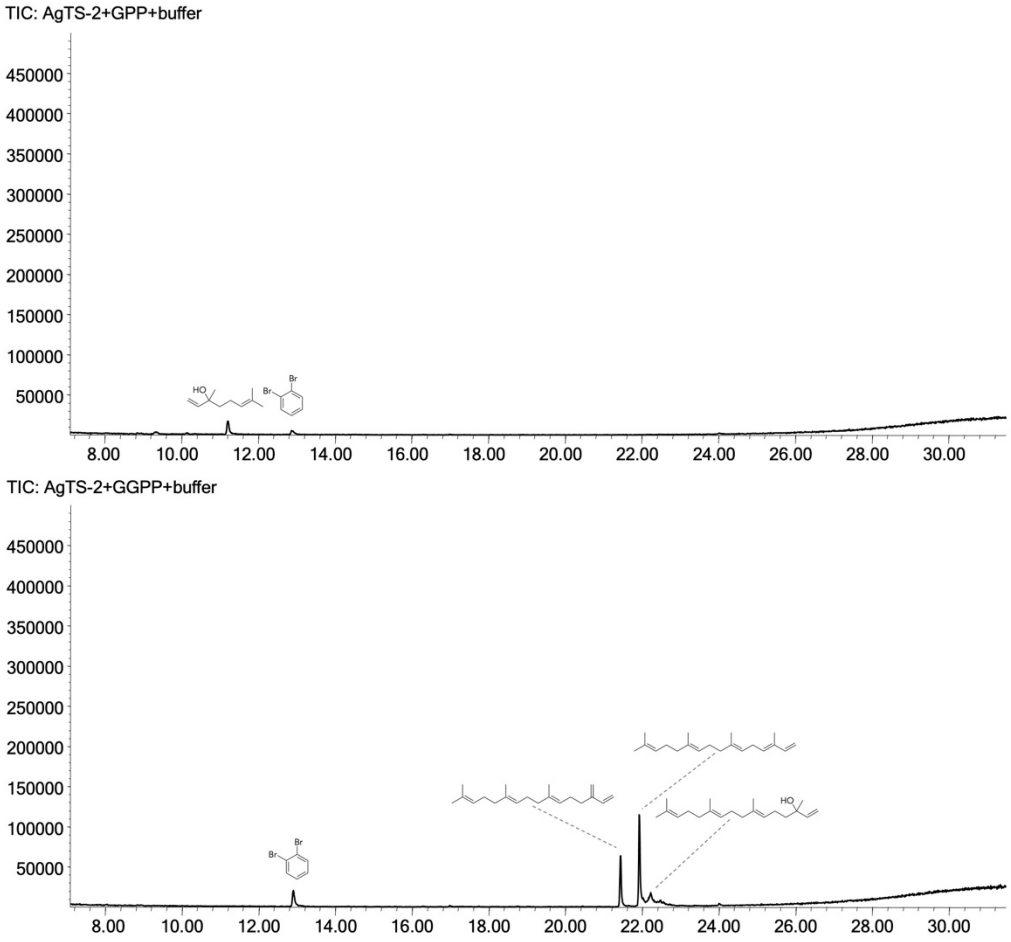


Figure S33. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (lower) assays with AgTS-2. Annotations based on NIST spectral library comparison.

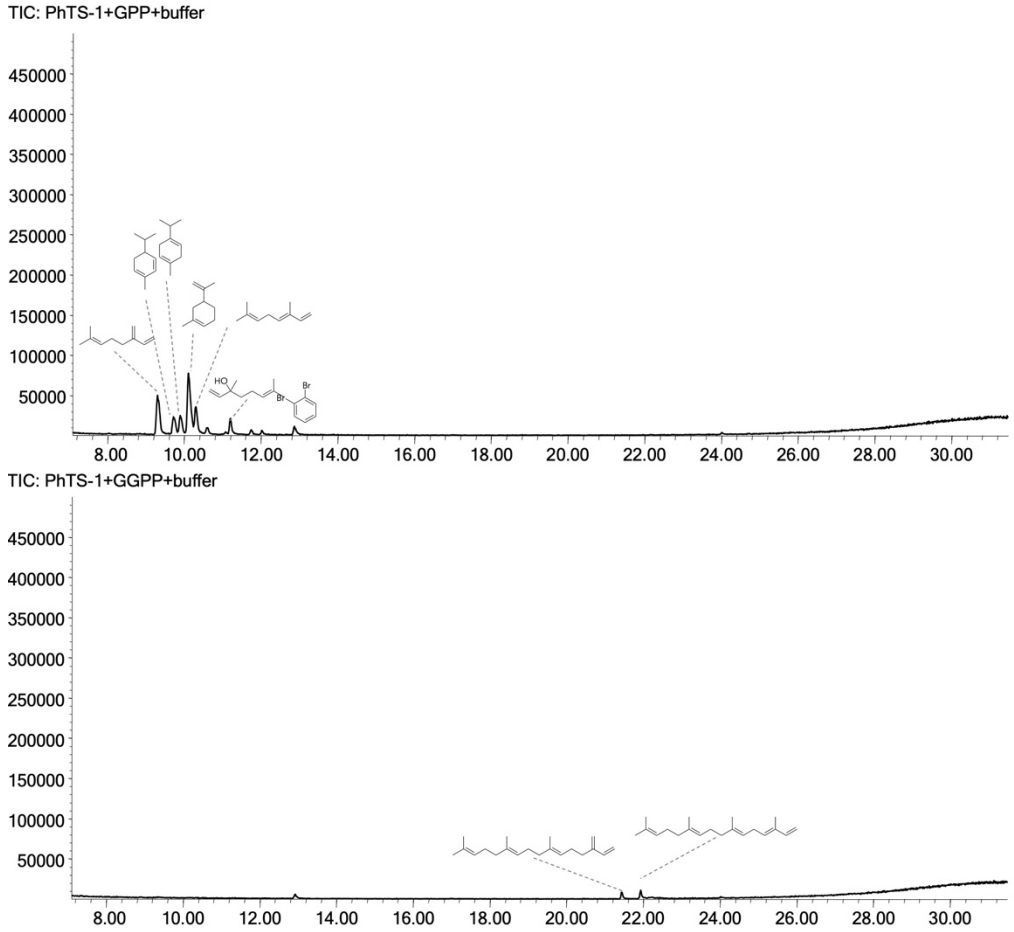


Figure S34. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (lower) assays with PhTS-1. Annotations based on NIST spectral library comparison.

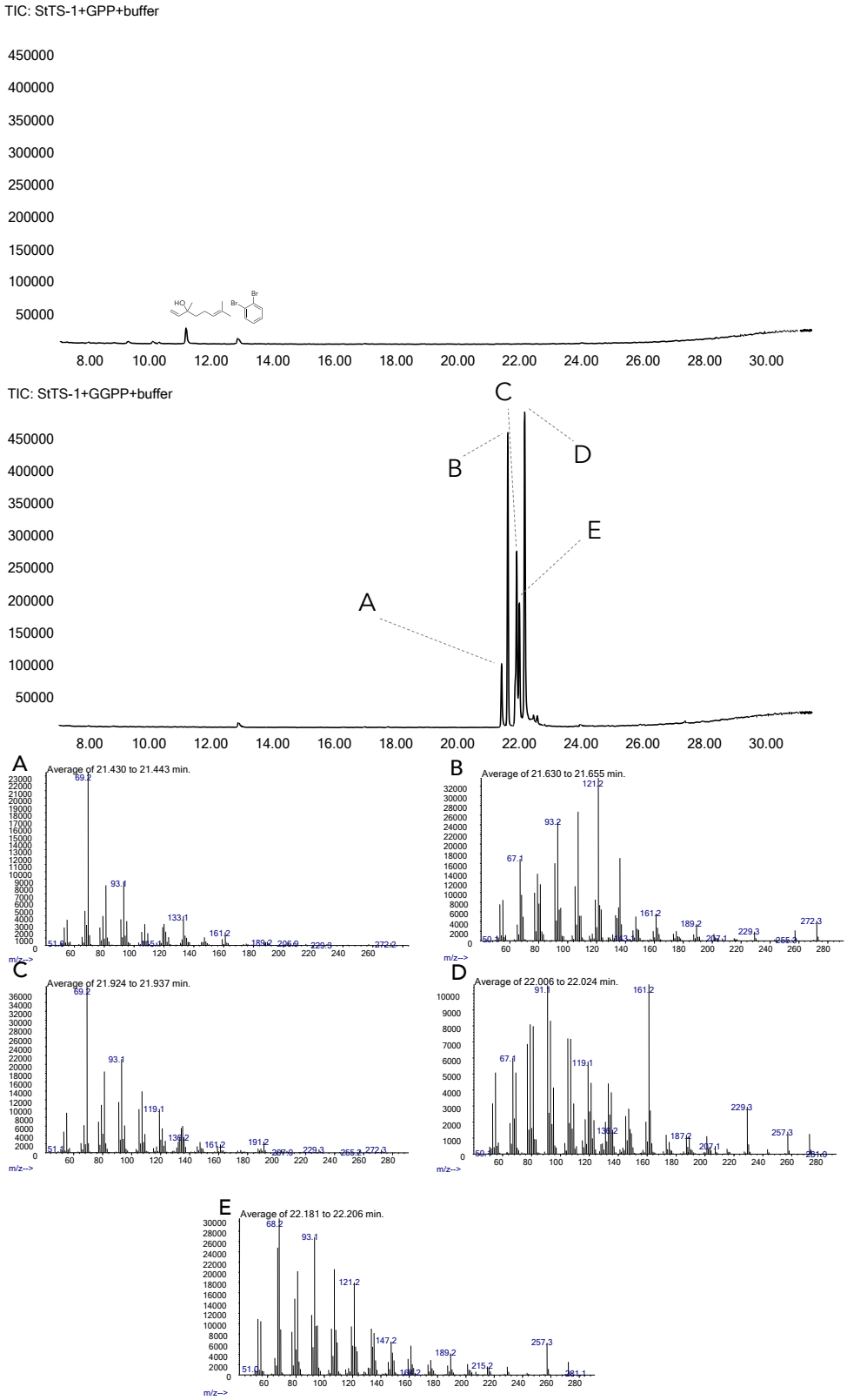


Figure S35. GCMS TIC traces of hexanes extracts of GPP (upper) and GGPP (middle) assays with StTS-1. MS of major GGPP assay peaks are shown in the lower panel. Annotations based on NIST spectral library comparison.

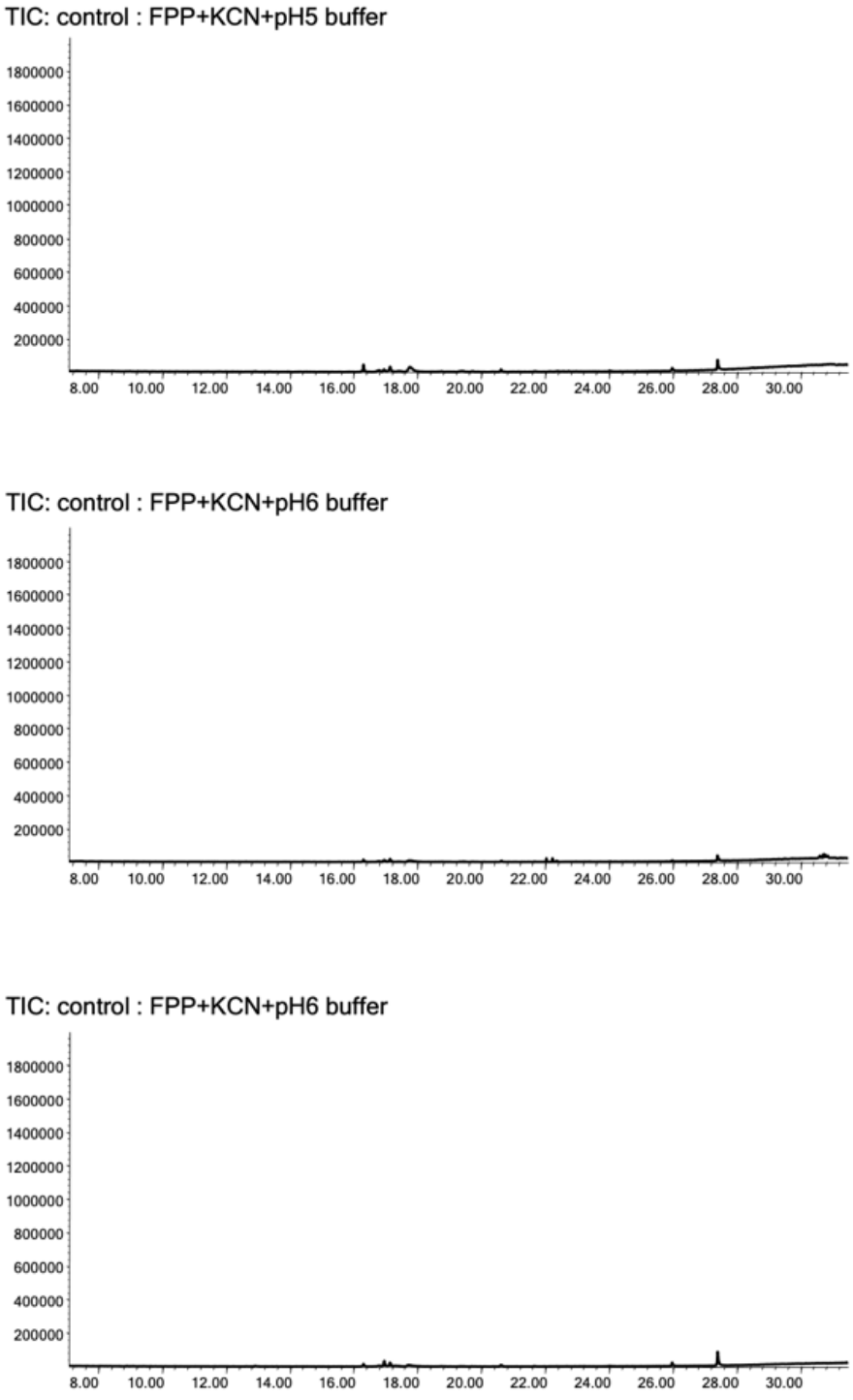


Figure S36. GCMS TIC traces of hexanes extracts of no enzyme FPP and KCN controls at pH5, 6, and 7.

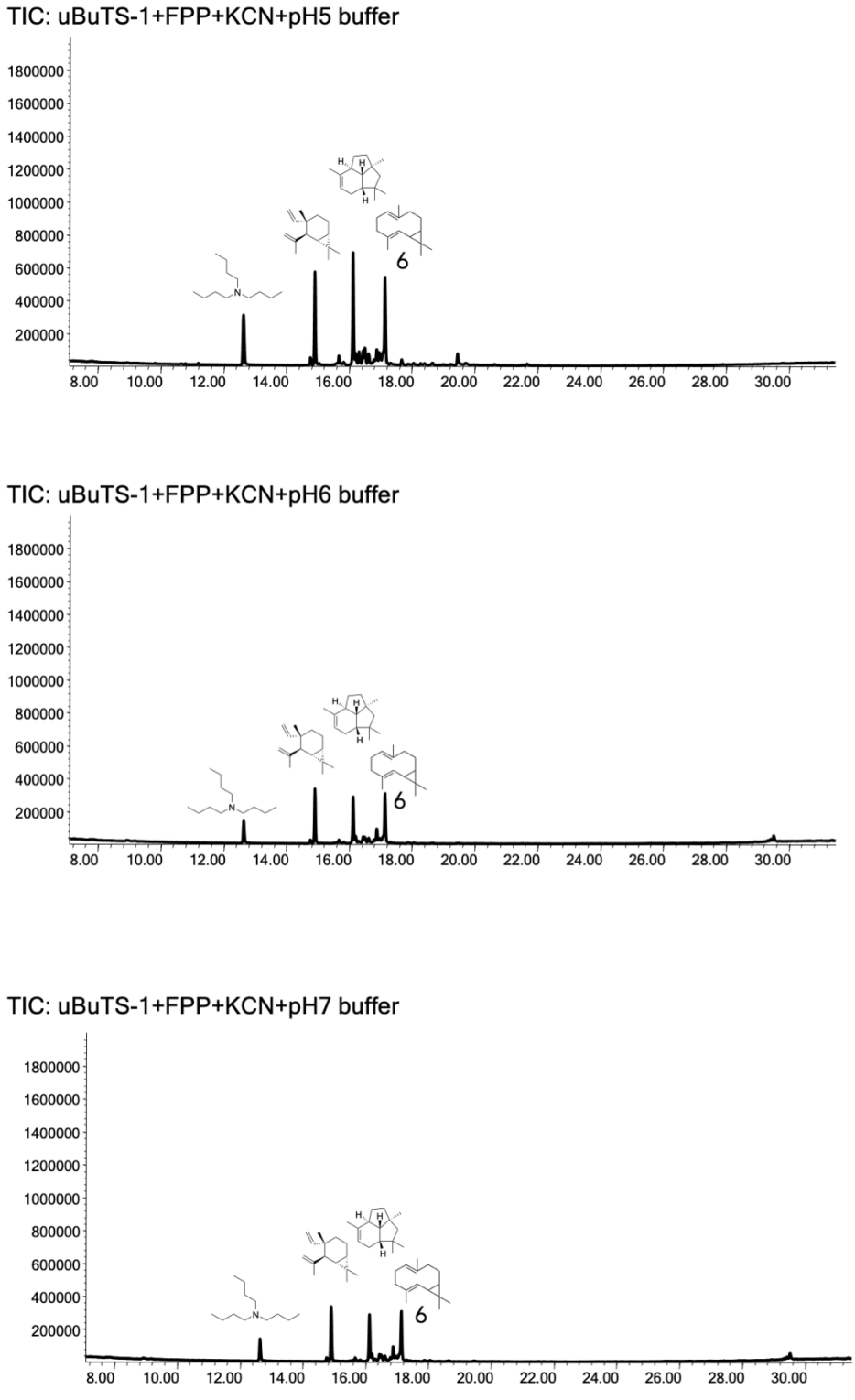


Figure S37. GCMS TIC traces of hexanes extracts of KCN and FPP assays with uBuTS-1 at pH5, 6, and 7. Annotations based on NIST spectral library comparison and comparison with bicyclogermacrene **6** standard.

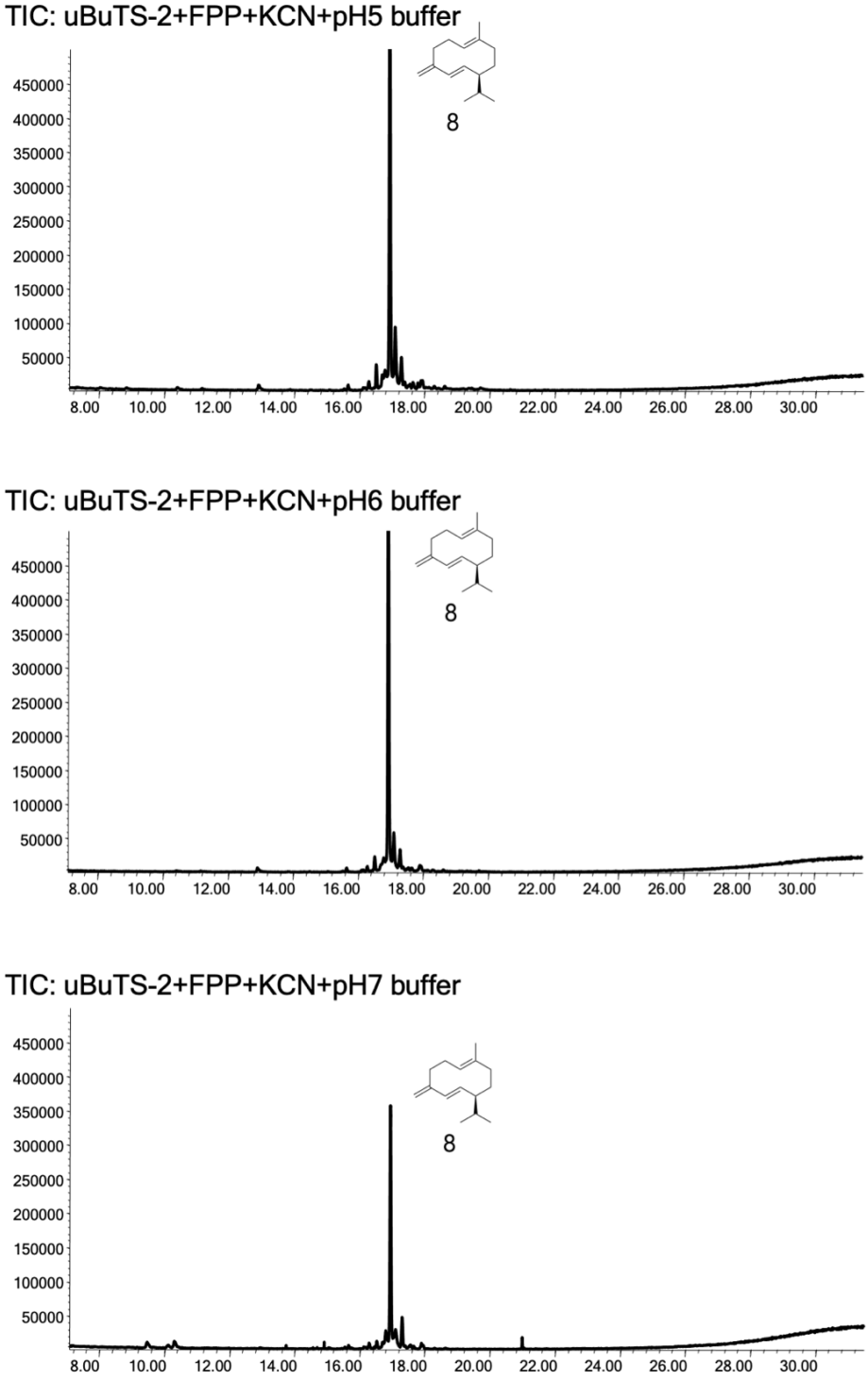


Figure S38. GCMS TIC traces of hexanes extracts of KCN and FPP assays with uBuTS-2 at pH5, 6, and 7. Annotations based on NIST spectral library comparison and comparison with germacrene D **8** standard.

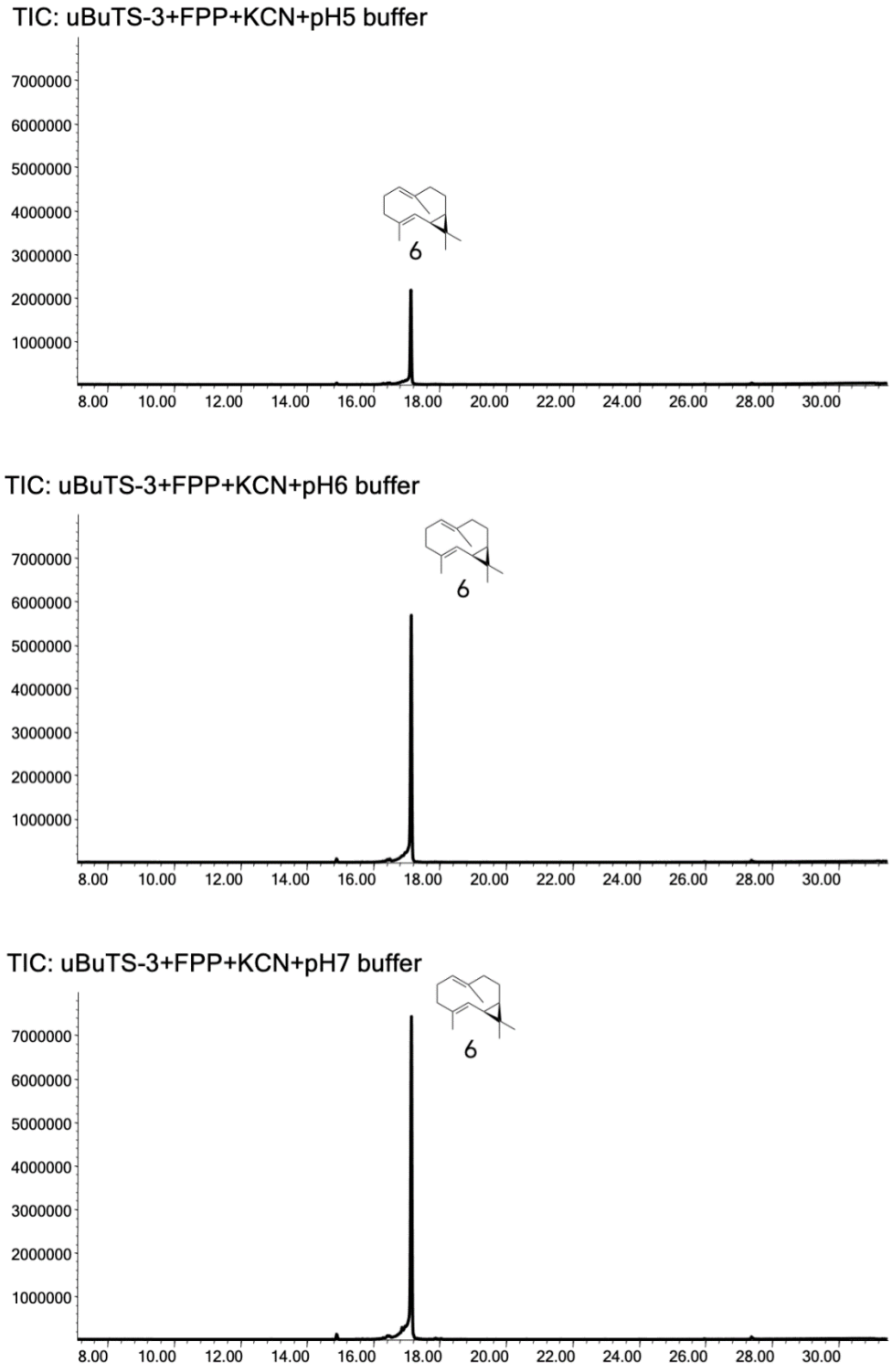


Figure S39. GCMS TIC traces of hexanes extracts of KCN and FPP assays with uBuTS-3 at pH5, 6, and 7. Annotations based on NIST spectral library comparison and comparison with bicyclogermacrene **6** standard.

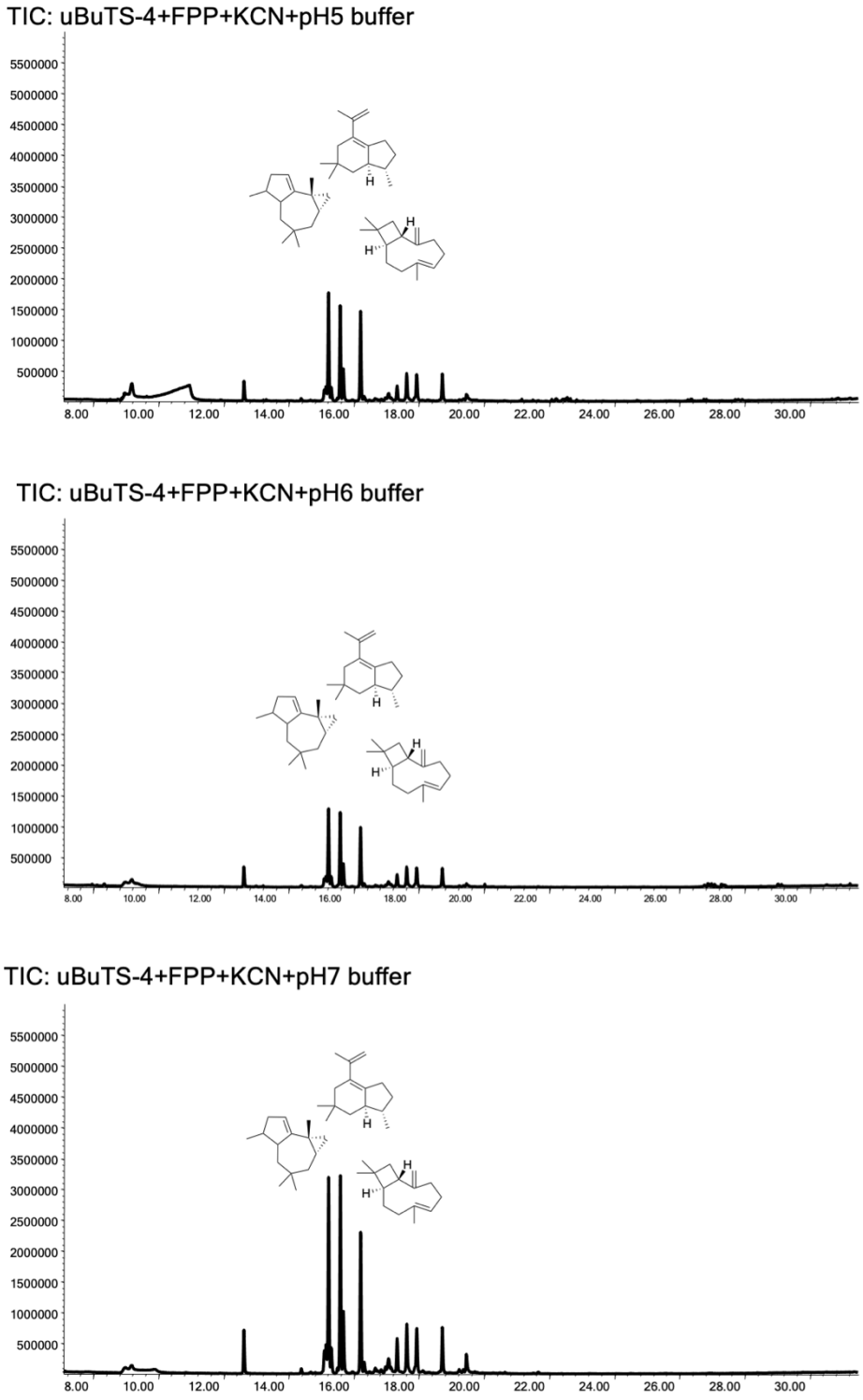


Figure S40. GCMS TIC traces of hexanes extracts of KCN and FPP assays with uBuTS-4 at pH5, 6, and 7. Annotations based on NIST spectral library comparison.

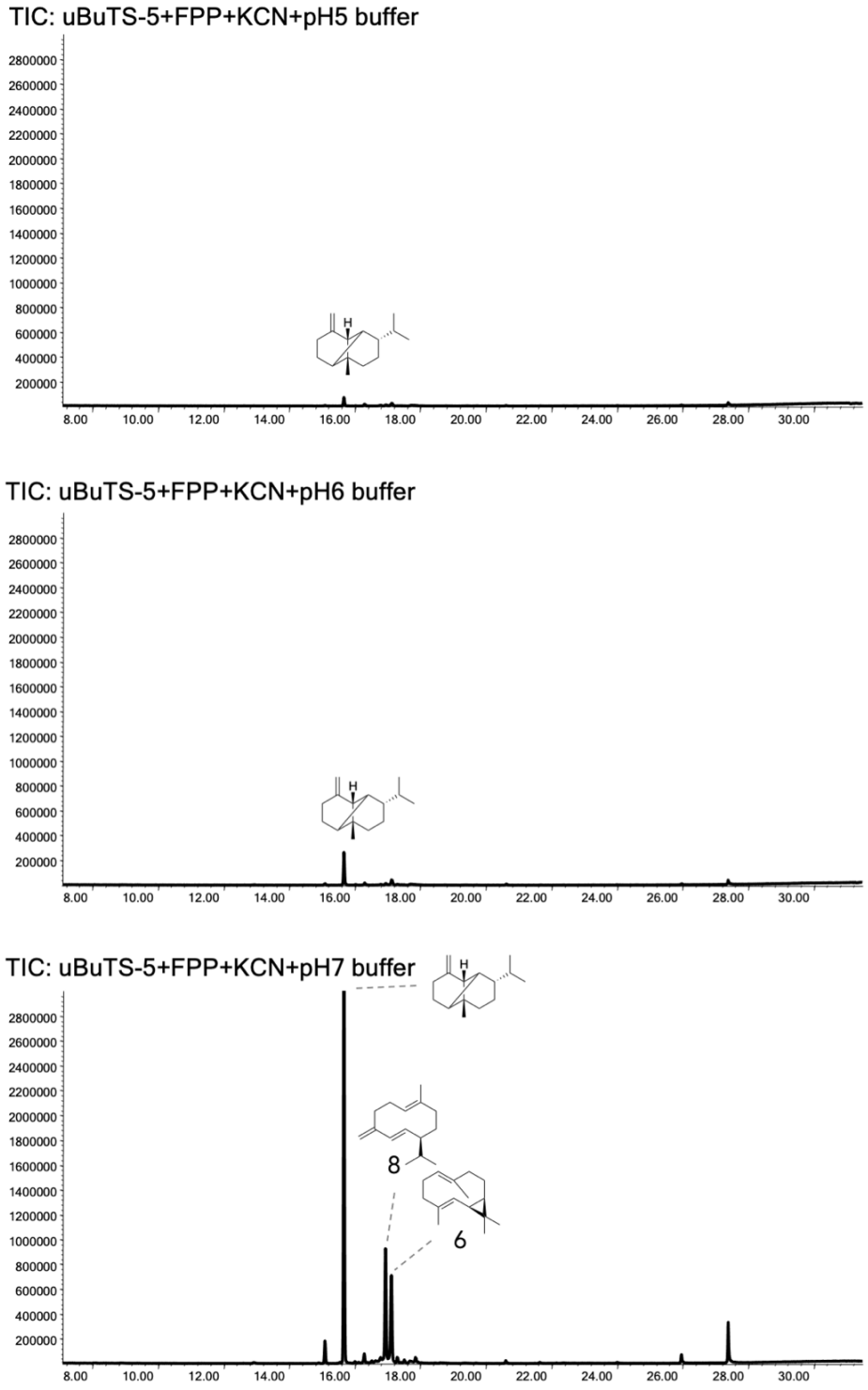


Figure S41. GCMS TIC traces of hexanes extracts of KCN and FPP assays with uBuTS-5 at pH5, 6, and 7. Annotations based on NIST spectral library comparison and comparison with bicyclogermacrene **6** and germacrene D **8** standards.

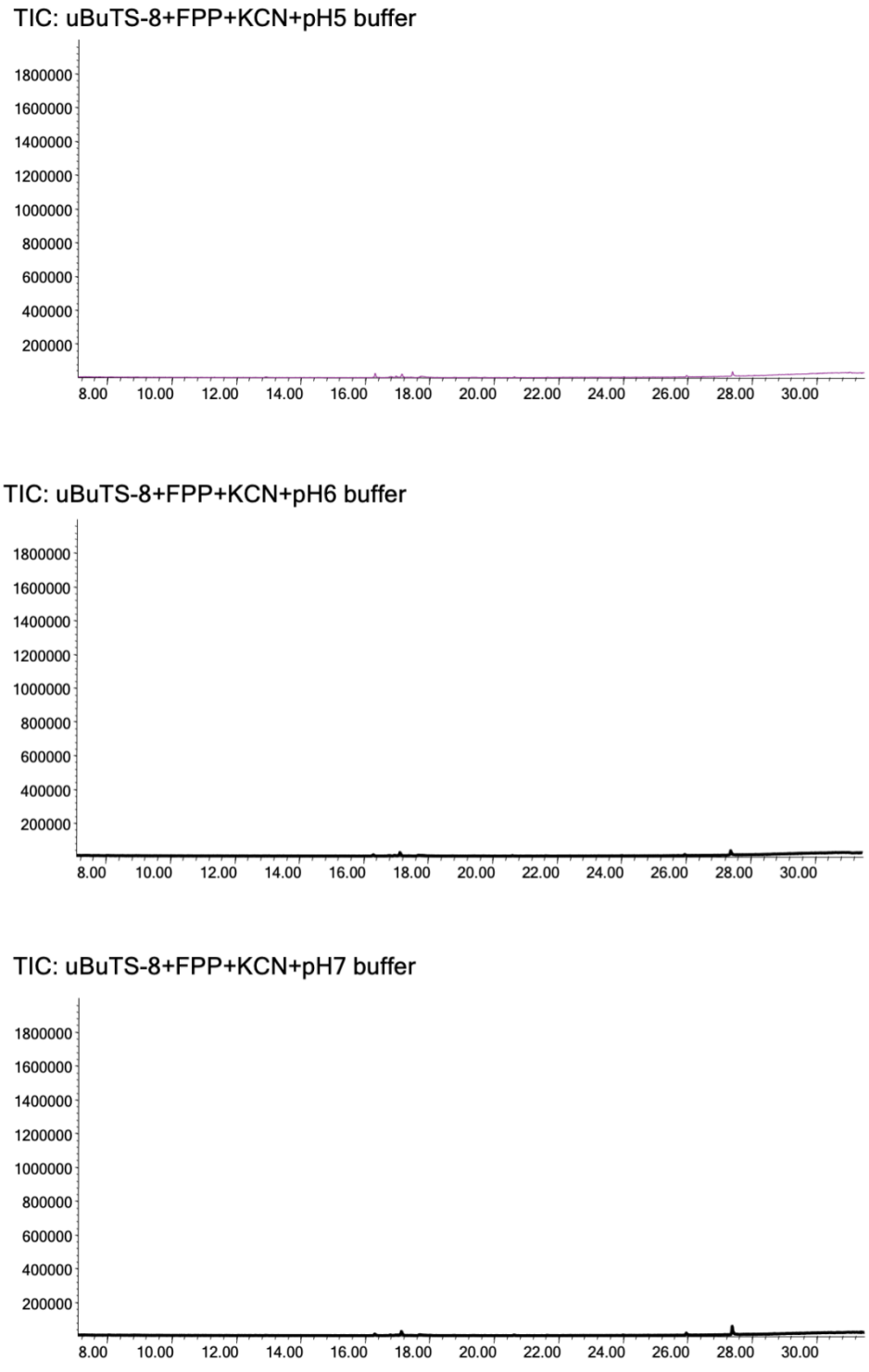


Figure S42. GCMS TIC traces of hexanes extracts of KCN and FPP assays with uBuTS-8 at pH5, 6, and 7.

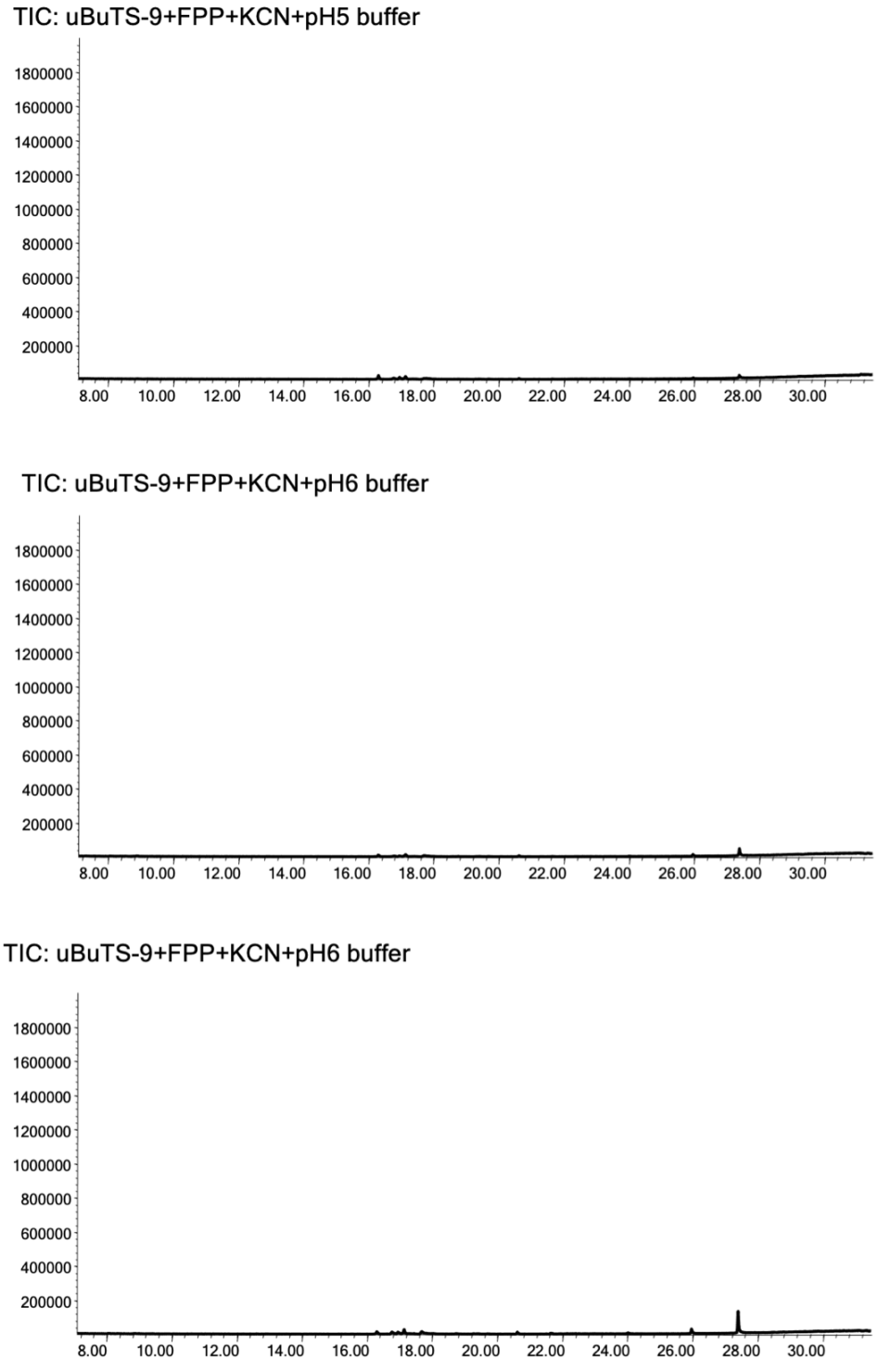


Figure S43. GCMS TIC traces of hexanes extracts of KCN and FPP assays with uBuTS-9 at pH5, 6, and 7.

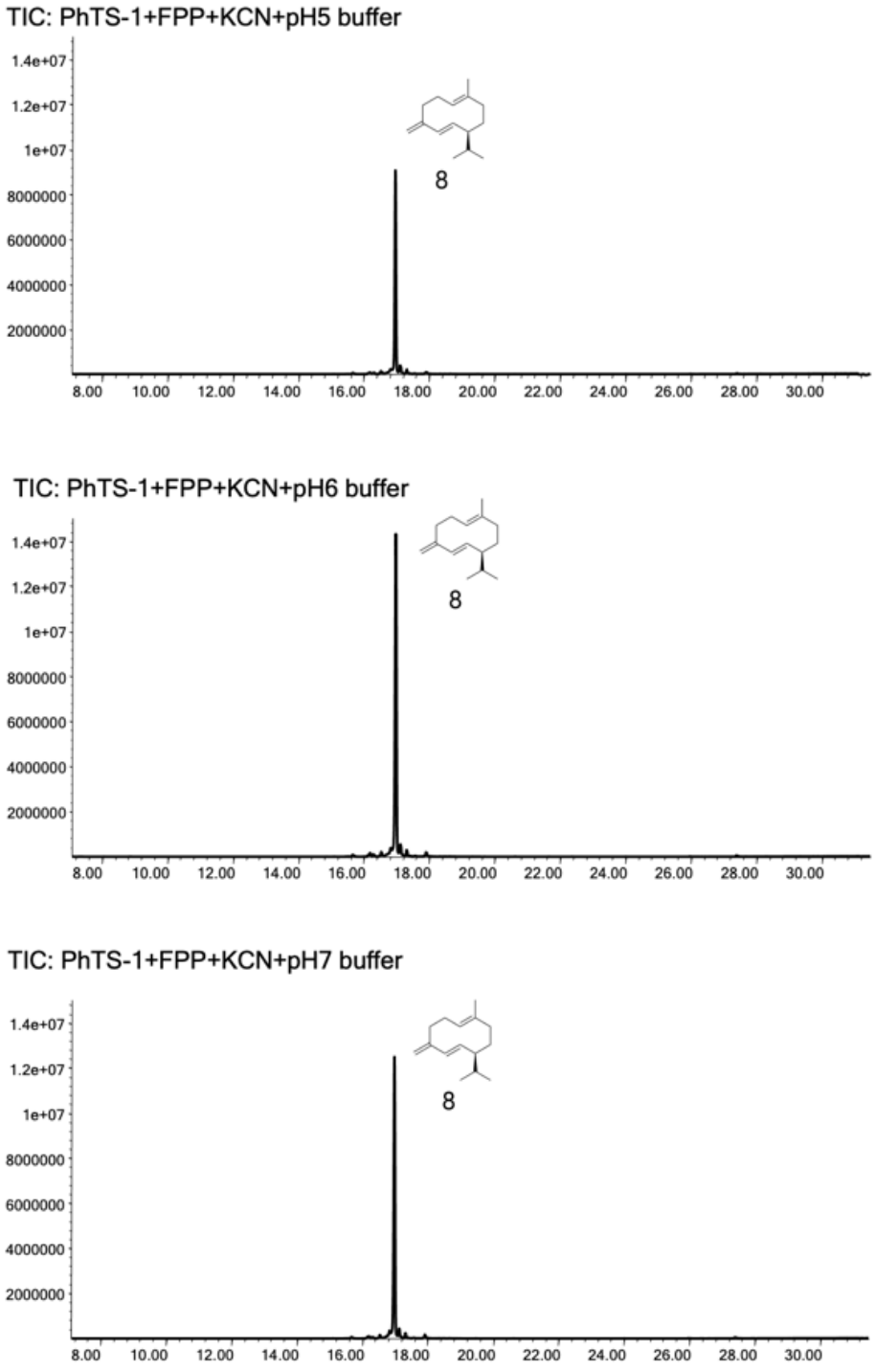


Figure S44. GCMS TIC traces of hexanes extracts of KCN and FPP assays with PhTS-1 at pH5, 6, and 7. Annotations based on NIST spectral library comparison and comparison with germacrene D **8** standard.

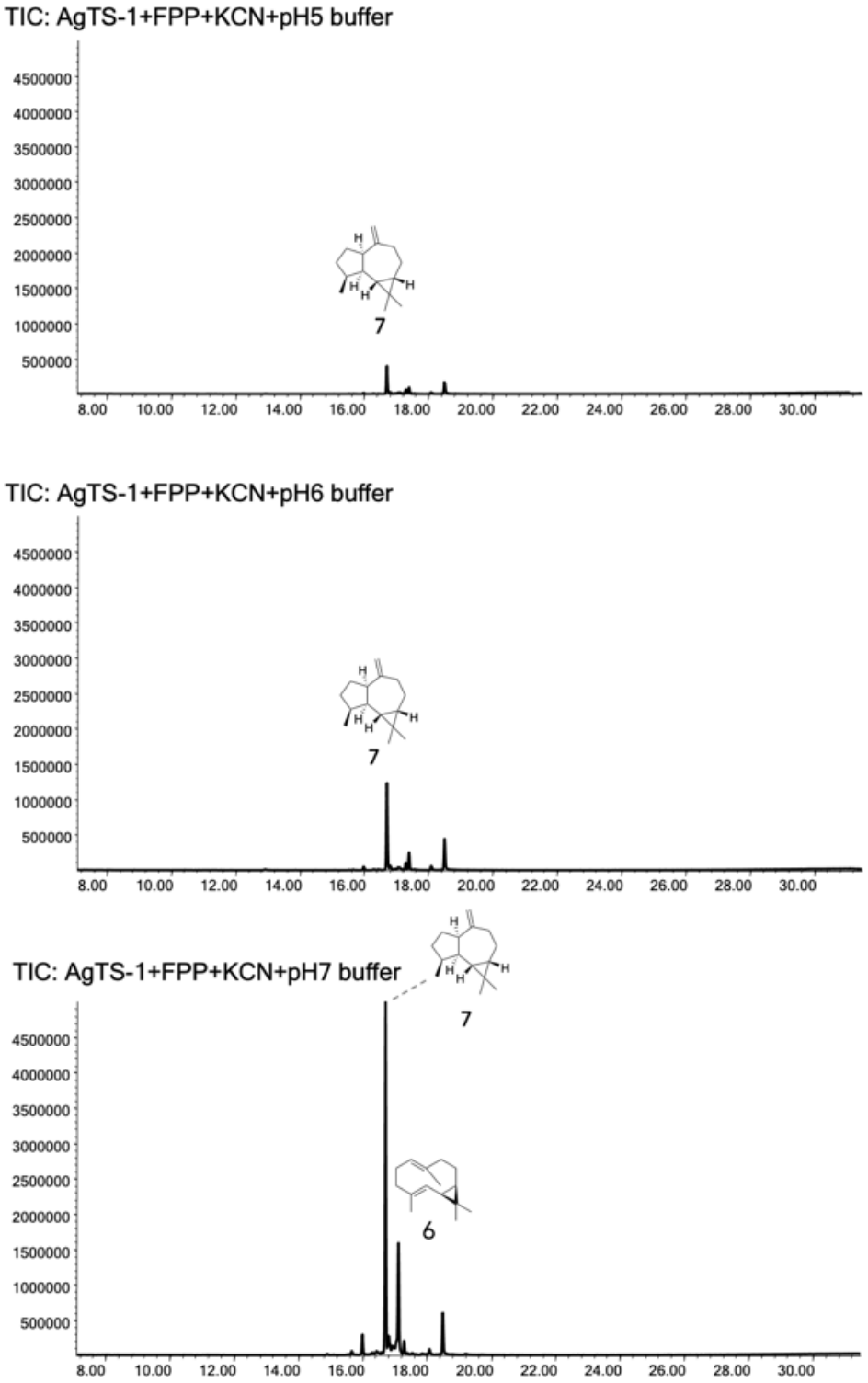


Figure S45. GCMS TIC traces of hexanes extracts of KCN and FPP assays with AgTS-1 at pH5, 6, and 7. Annotations based on NIST spectral library comparison and comparison with alloaromadendrene **7** standard.

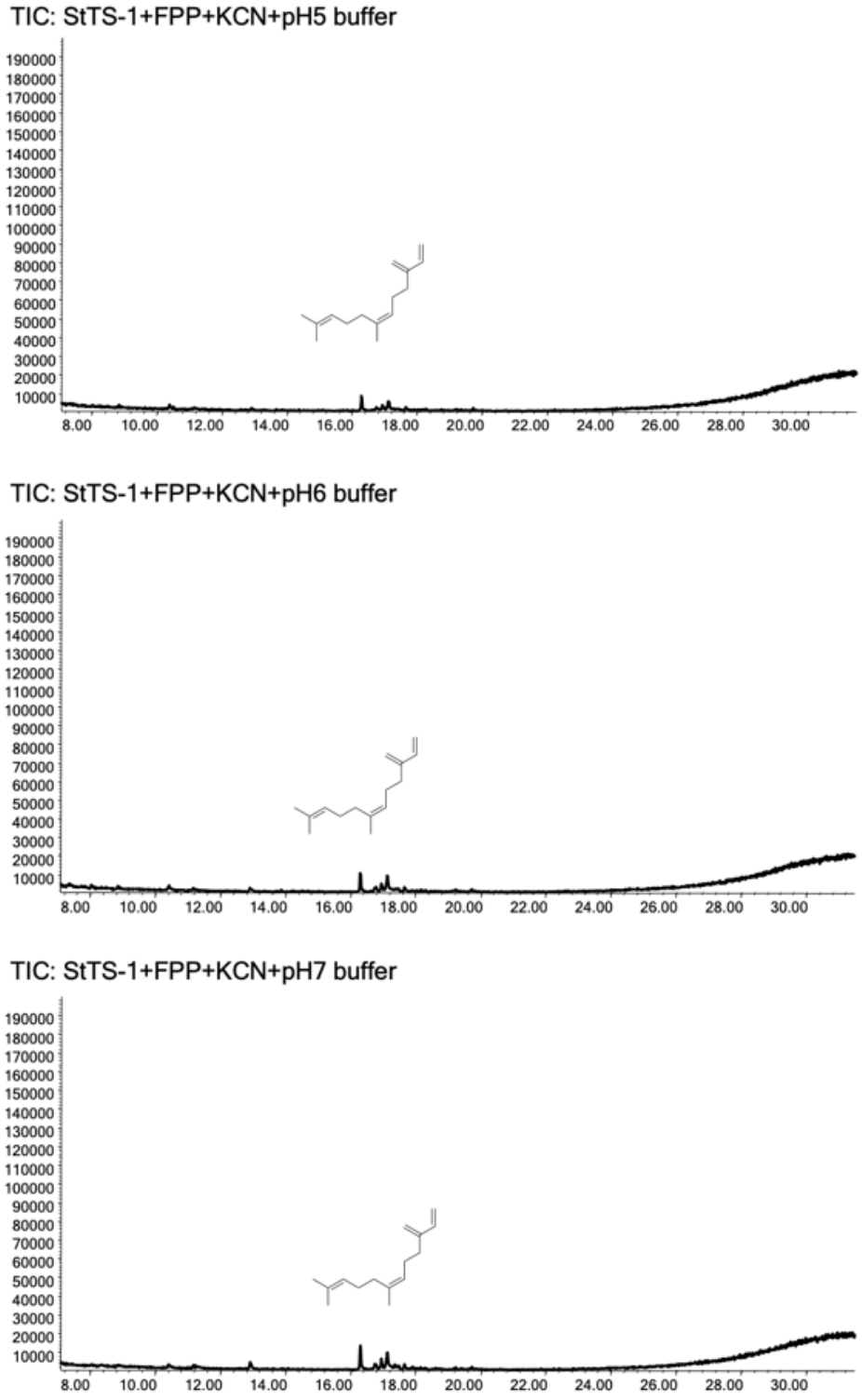


Figure S46. GCMS TIC traces of hexanes extracts of KCN and FPP assays with StTS-1 at pH5, 6, and 7. Annotations based on NIST spectral library comparison.

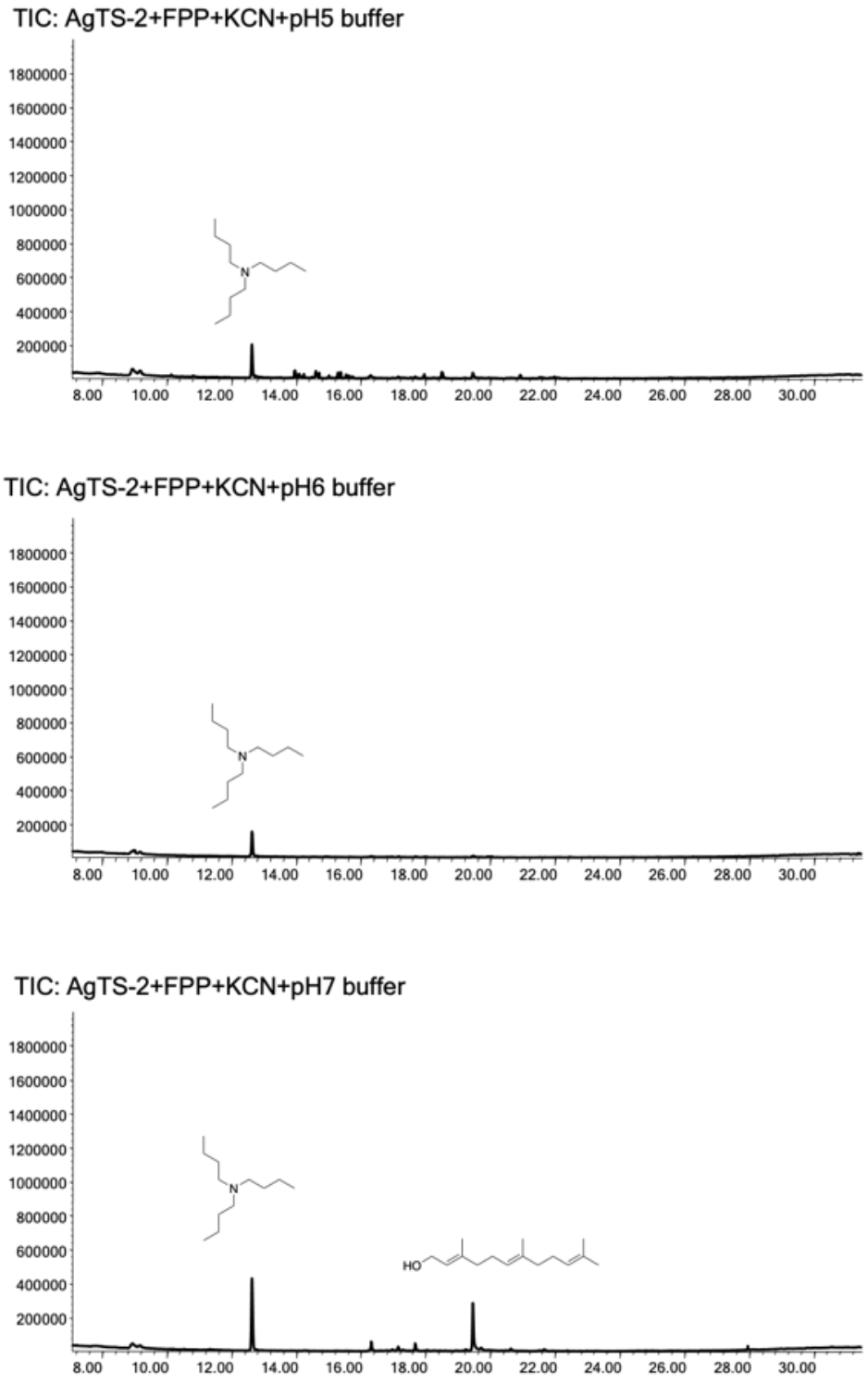


Figure S47. GCMS TIC traces of hexanes extracts of KCN and FPP assays with AgTS-2 at pH5, 6, and 7. Annotations based on NIST spectral library comparison.

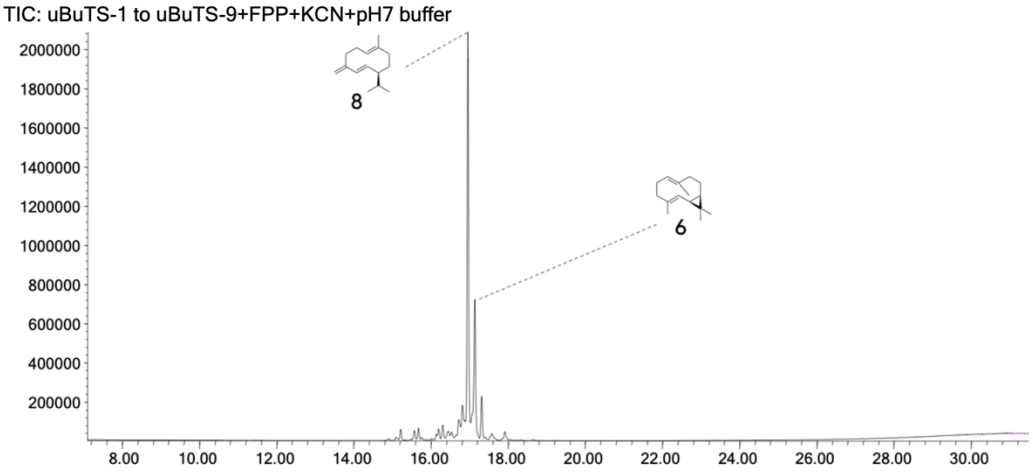


Figure S48. GCMS TIC traces of hexanes extracts of KCN and FPP assays with uBuTS-1 to uBuTS-9 at pH5, 6, and 7. Annotations based on comparison with bicyclogermacrene **6** and germacrene D **8** standards.

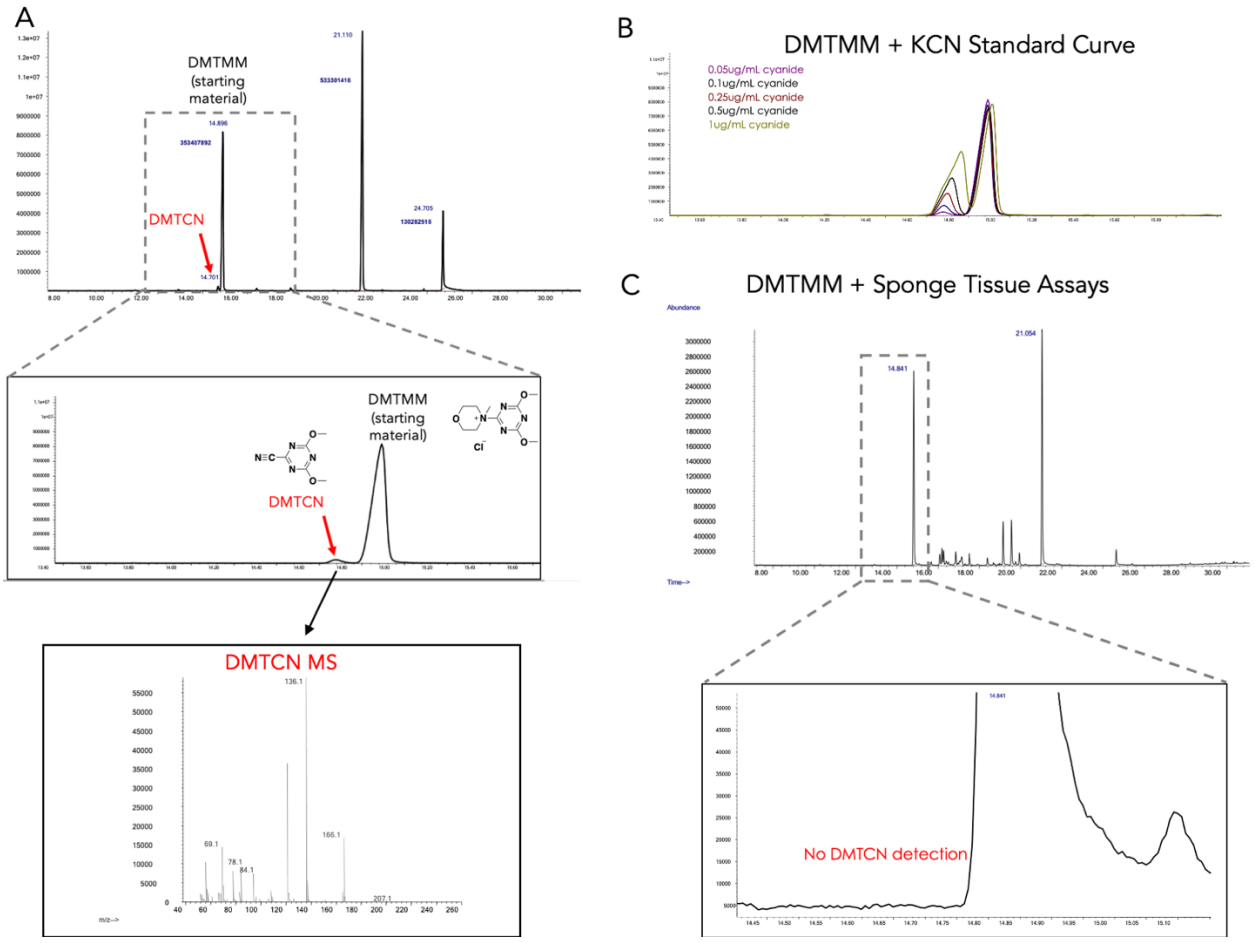


Figure S49. – A. GCMS TIC trace of ethyl acetate extract of a 900µL reaction of 80mM 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMTMM) and 0.05µg/mL potassium cyanide as described in Yamaguchi and Miyaguchi, 2020. The product, 2-cyano-4,6-dimethoxy-1,3,5- triazine (DMTCN) is shown in red. B. GCMS TIC trace of the ethyl acetate extracts of reactions with DMTMM and several concentrations of potassium cyanide, from 0.05µg/mL to 1µg/mL. C. GCMS TIC trace of the ethyl acetate extract of a 900uL reaction of 80mM DMTMM with approximately 0.2g homogenized undescribed Bubarida sponge tissue. No DMTCN is detectable.

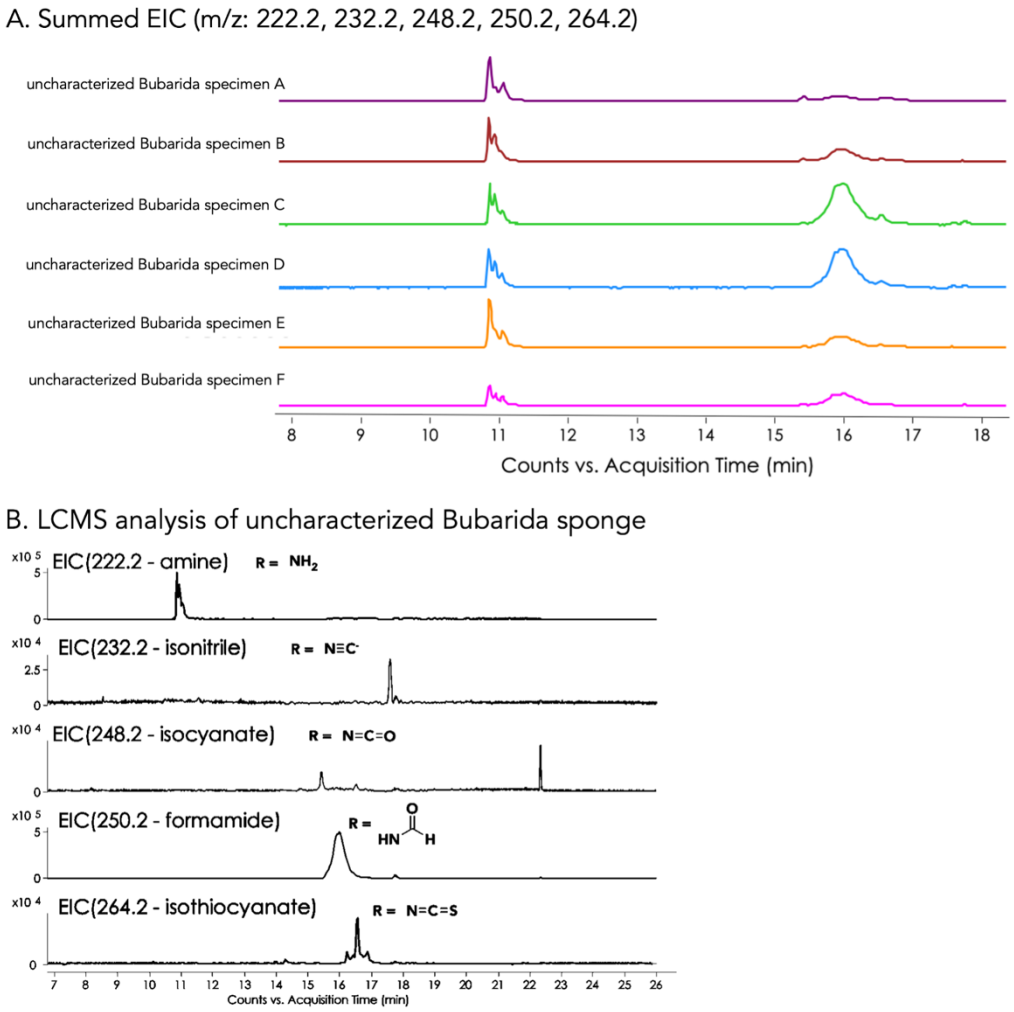


Figure S50. A. Summed LCMS EIC of m/z values representing nitrogenous sesquiterpenes (amine, isonitrile, isocyanate, formamide, and isothiocyanate) for multiple undescribed Bubarida specimen crude extracts. B. EICs of crude undescribed Bubarida extract for m/z values representing the sesquiterpene amine, isonitrile, isocyanate, formamide, and isothiocyanate.

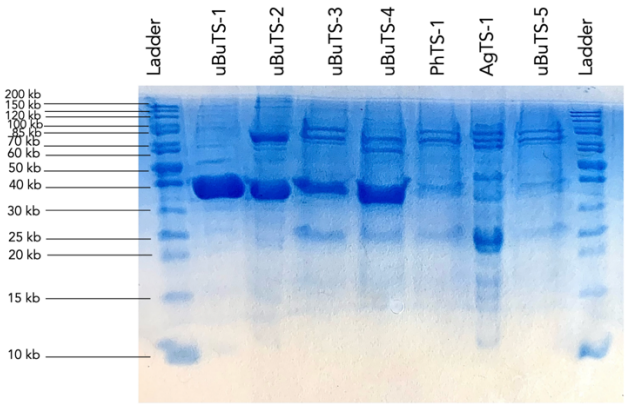


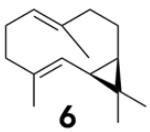
Figure S51. SDS-PAGE gel of active sesquiterpene synthases uBuTS1-5, PhTS-1, and AgTS-1. The ladder is EZ-Run™ Rec Protein Ladder (Fisher BioReagents™, cat. no. BP3602500).

Supplementary Note – Analytical data for isolated terpenes.

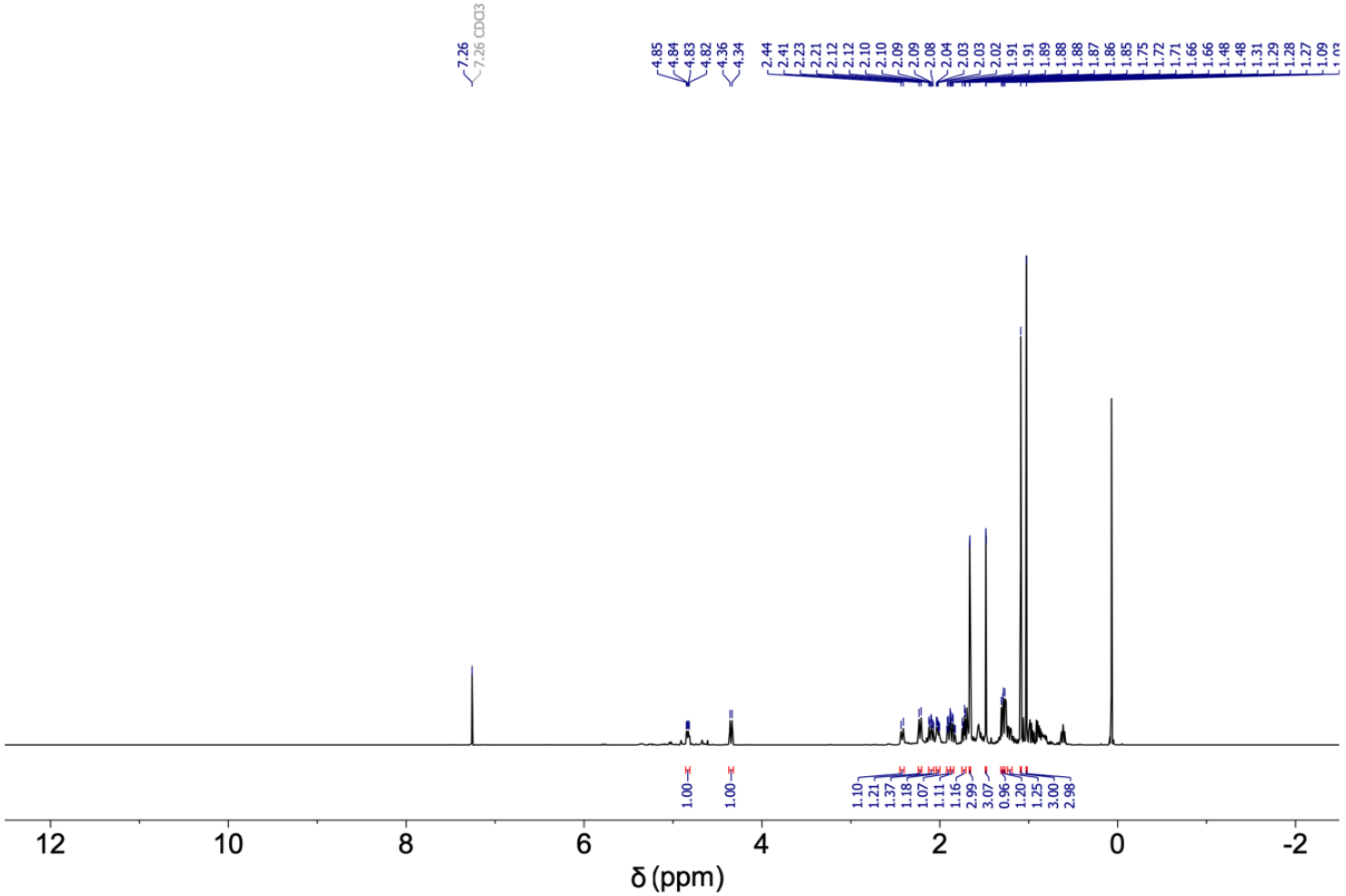
(+)-Bicyclogermacrene **6**: **^{1}H-NMR** (500 MHz, Chloroform-*d*): δ = 4.83 (dd, *J* = 11.2, 5.3 Hz, 1H), 4.35 (d, *J* = 11.5 Hz, 1H), 2.42 (d, *J* = 13.2 Hz, 1H), 2.22 (d, *J* = 12.0 Hz, 1H), 2.10 (ddd, *J* = 12.5, 11.1, 3.8 Hz, 1H), 2.04 – 2.00 (m, 1H), 1.92 – 1.88 (m, 1H), 1.86 (dd, *J* = 10.5, 3.4 Hz, 1H), 1.73 (dd, *J* = 13.0, 4.1 Hz, 1H), 1.66 (d, *J* = 1.4 Hz, 3H), 1.48 (d, *J* = 1.6 Hz, 3H), 1.30 (d, *J* = 8.8 Hz, 1H), 1.28 (d, *J* = 8.7 Hz, 1H), 1.22 (ddd, *J* = 12.5, 4.8, 1.4 Hz, 1H), 1.09 (s, 3H), 1.03 (s, 3H). **^{13}C-NMR** (125 MHz, Chloroform-*d*): δ = 141.00, 128.12, 126.65, 124.97, 41.31, 37.39, 30.21, 29.37, 27.09, 26.16, 21.02, 20.01, 16.71, 15.58 (29). **Optical rotation**: $[\alpha]_D^{26.0}$ = +16(*c* 0.82, $CHCl_3$) (30, 31). **EI-MS** (70eV)**: *m/z*** (%) = 204 (15), 189 (8), 162 (5), 161 (23), 148 (4), 147 (6), 136 (17), 135 (5), 134 (5), 133 (9), 123 (9), 122 (15), 121 (100), 120 (6), 119 (26), 117 (3), 109 (8), 108 (11), 107 (50), 106 (6), 105 (29), 95 (9), 94 (20), 93 (65), 92 (9), 91 (37), 82 (3), 81 (20), 80 (8), 79 (31), 78 (5), 77 (22), 69 (6), 68 (3), 67 (17), 65 (7), 55 (12).
See NMR spectra below.

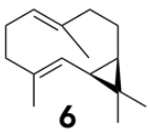
(+)-Alloaromadendrene **7**: **^{1}H-NMR** (500 MHz, Chloroform-*d*): δ = 4.72 (d, *J* = 13.0 Hz, 2H), 2.67 (q, *J* = 8.1 Hz, 1H), 2.34 – 2.27 (m, 2H), 2.07 – 2.04 (m, 1H), 1.88 – 1.82 (m, 3H), 1.72 (d, *J* = 4.2 Hz, 2H), 1.31 (d, *J* = 1.8 Hz, 1H), 1.23 – 1.21 (m, 1H), 1.00 (s, 3H), 0.95 (s, 3H), 0.94 (d, *J* = 6.9 Hz, 3H), 0.55 (ddd, *J* = 10.9, 9.2, 6.2 Hz, 1H), 0.24 (dd, *J* = 10.9, 9.3 Hz, 1H). **^{13}C-NMR** (125 MHz, Chloroform-*d*): δ = 152.79, 109.79, 50.91, 42.29, 37.92, 35.84, 31.33, 29.85, 28.76, 28.35, 24.91, 23.63, 22.26, 16.55, 16.01(32). **Optical rotation:** $[\alpha]_D^{26.0}$ = +8.58 (*c* 0.167, $CHCl_3$) (33). **EI-MS** (70eV)**: *m/z*** (%) = 205 (7), 204 (45), 190 (5), 189 (32), 176 (4), 175 (9), 162 (16), 161 (96), 149 (10), 148 (37), 147 (56), 146 (3), 145 (5), 136 (4), 135 (20), 134 (23), 133 (73), 132 (3), 131 (8), 129 (4), 128 (4), 123 (6), 122 (30), 121 (38), 120 (27), 119 (65), 118 (4), 117 (14), 116 (4), 115 (9), 109 (15), 108 (19), 107 (75), 106 (28), 105 (90), 104 (4), 103 (7), 96 (6), 95 (27), 94 (21), 93 (77), 92 (24), 91 (100), 83 (7), 82 (17), 81 (47), 80 (12), 79 (64), 78 (13), 77 (40), 69 (42), 68 (5), 67 (39), 66 (5), 65 (15), 55 (30), 53 (18)
See NMR spectra below.

(-)-Germacrene D **8**: **^{1}H-NMR** (500 MHz, Chloroform-*d*): δ = 5.78 (d, *J* = 15.8 Hz, 1H), 5.25 (dd, *J* = 15.8, 10.0 Hz, 1H), 5.16 – 5.08 (m, 1H), 4.81 – 4.72 (m, 2H), 2.44 – 2.34 (m, 2H), 2.24 (dd, *J* = 12.0, 3.2 Hz, 1H), 2.12 – 2.07 (m, 1H), 1.99 (ddd, *J* = 15.3, 7.1, 3.4 Hz, 2H), 1.51 (s, 3H), 1.50 – 1.39 (m, 4H), 0.86 (d, *J* = 6.7 Hz, 3H), 0.81 (d, *J* = 6.8 Hz, 3H). **^{13}C-NMR** (125 MHz, Chloroform-*d*): δ = 149.02, 135.63, 134.16, 133.70, 129.79, 109.19, 53.04, 40.84, 34.63, 32.88, 29.38, 26.64, 20.91, 19.47, 16.06 (34). **Optical rotation:** $[\alpha]_D^{26.0}$ = -159 (c 0.51, $CHCl_3$) (35, 36). **EI-MS** (70eV)**: *m/z*** (%) = 204 (15), 162 (15), 161 (100), 160 (4), 147 (6), 134 (3), 133 (18), 131 (4), 121 (5), 120 (23), 119 (33), 117 (8), 115 (4), 107 (11), 106 (8), 105 (52), 103 (3), 95 (5), 93 (22), 92 (9), 91 (45), 81 (29), 80 (7), 79 (27), 78 (7), 77 (22), 69 (8), 67 (11), 65 (7), 55 (11), 53 (8)
See NMR spectra below.



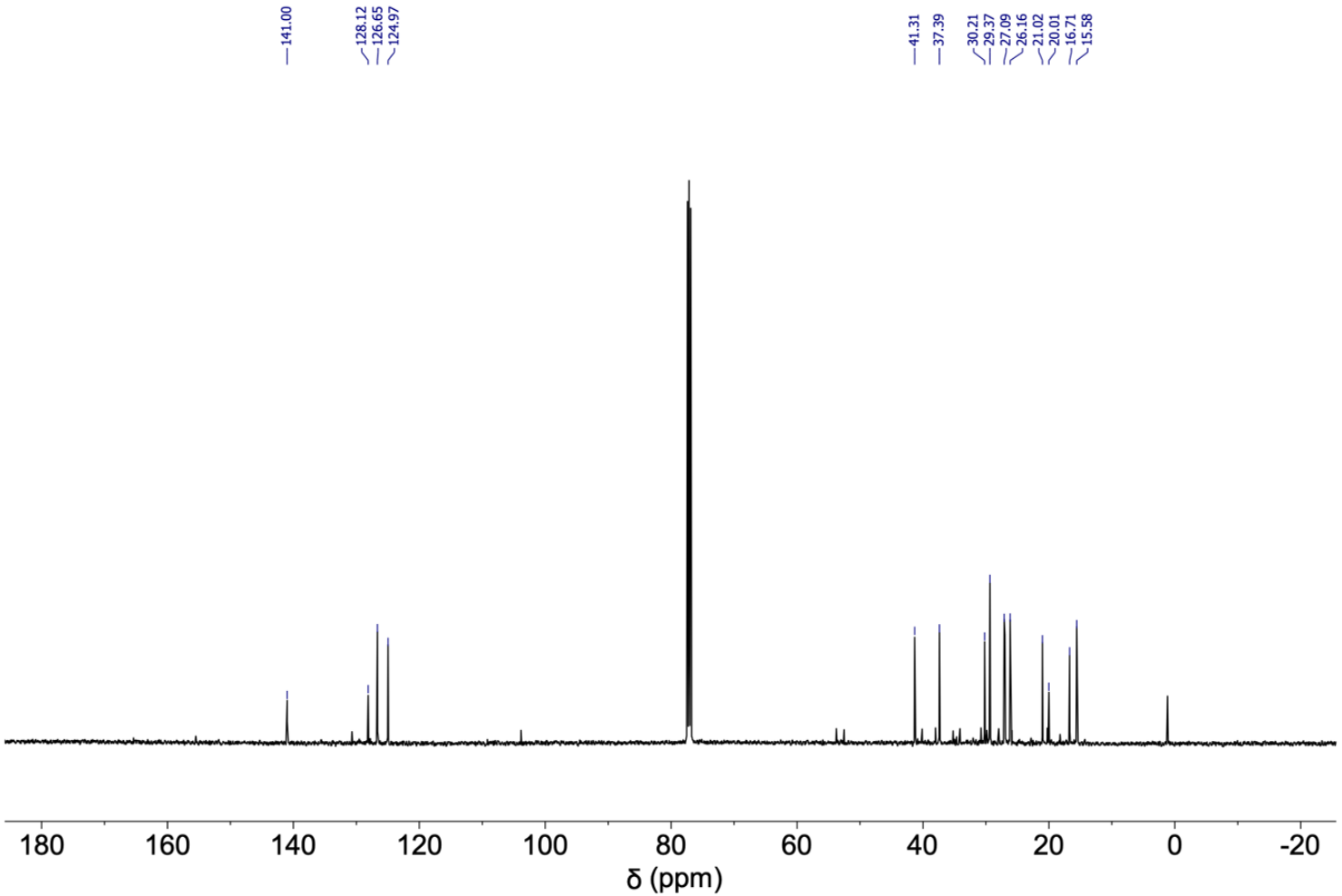
(+)-Bicyclogermacrene ^{1}H

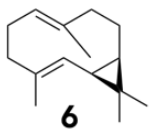




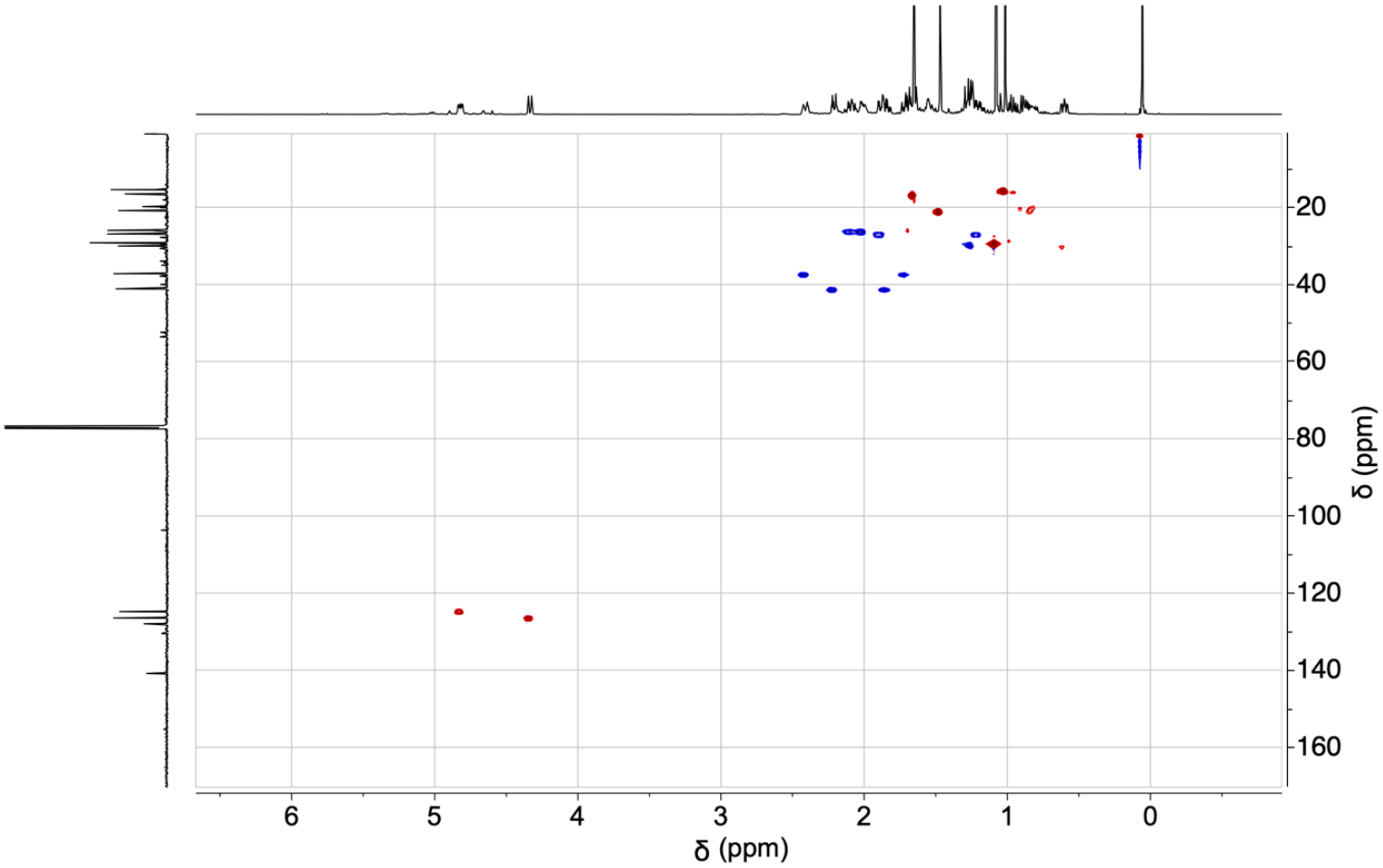
6

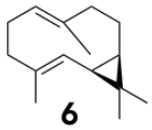
(+)-Bicyclogermacrene ^{13}C



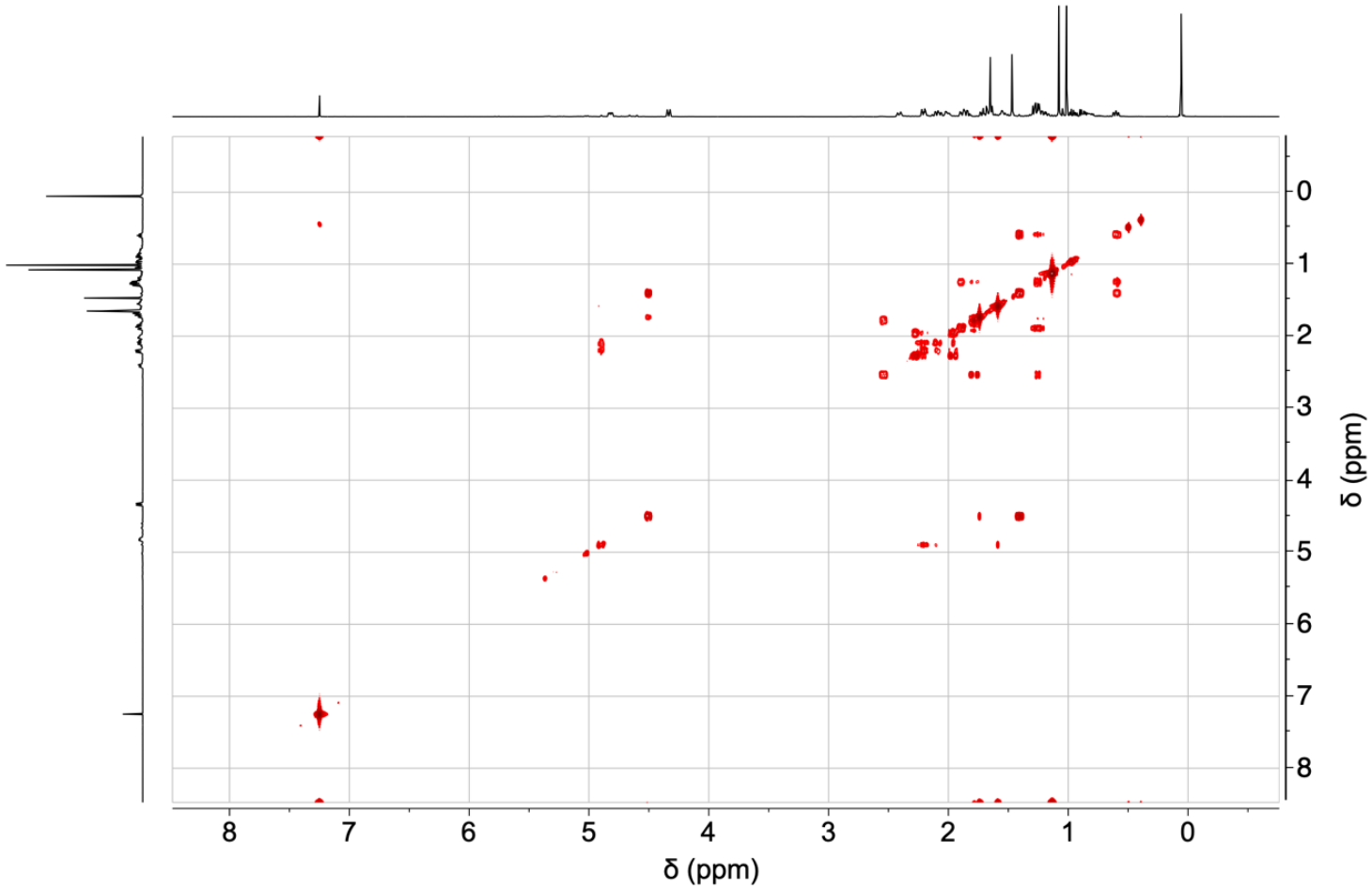


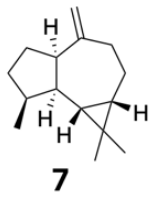
(+)-Bicyclogermacrene HSQC





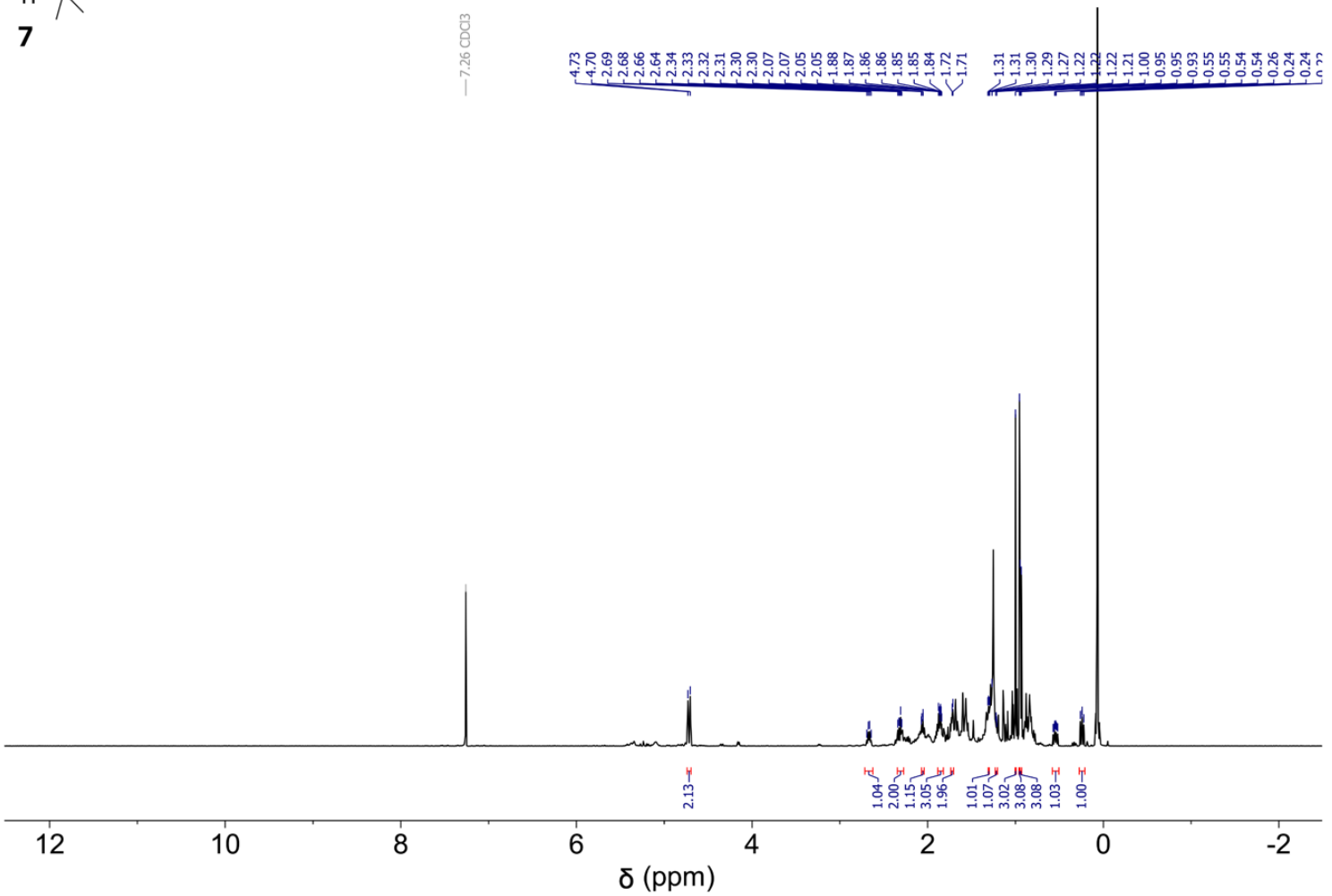
(+)-Bicyclogermacrene COSY

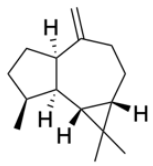




7

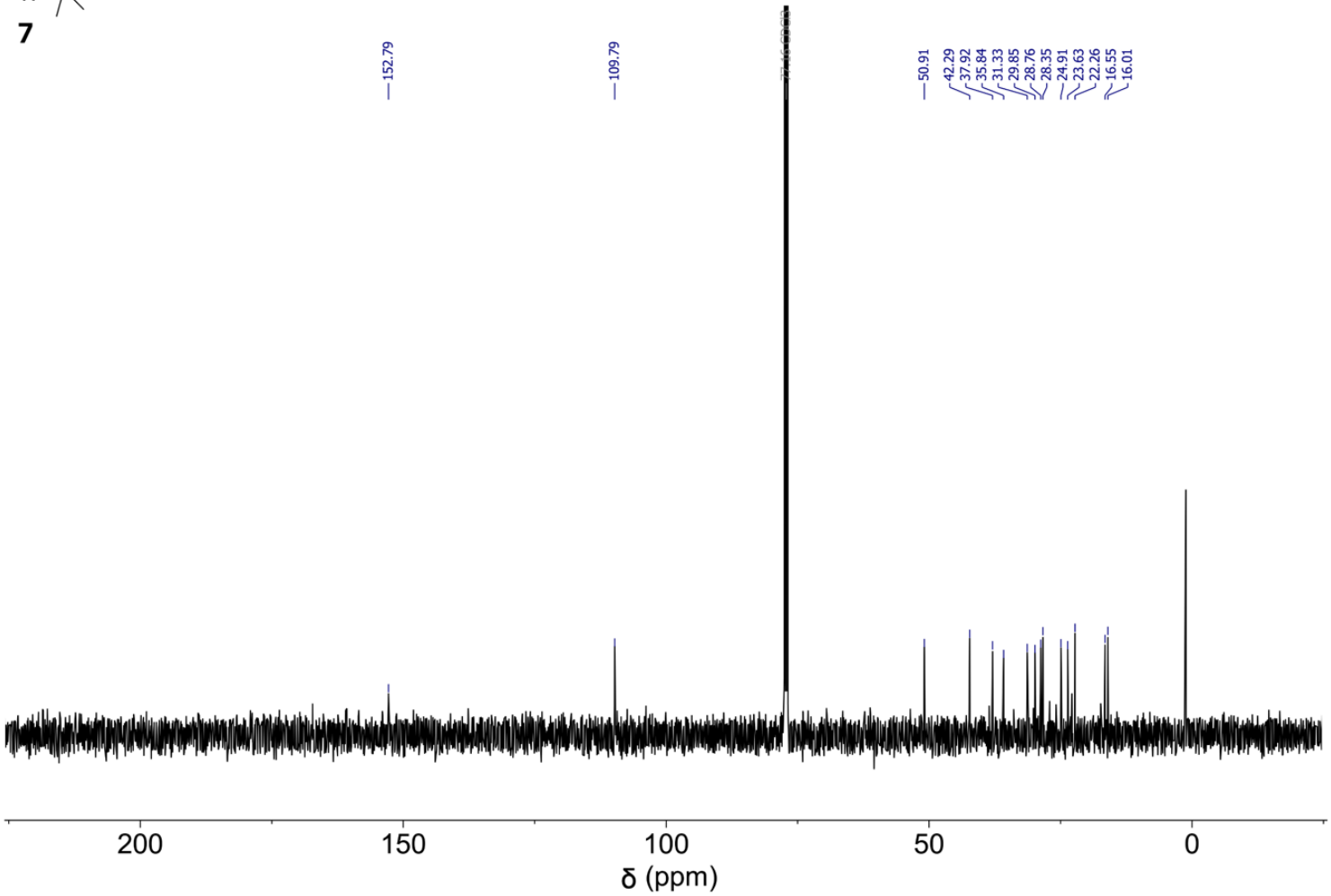
(+)-Alloaromadendrene ^{1}H

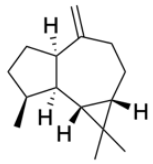




7

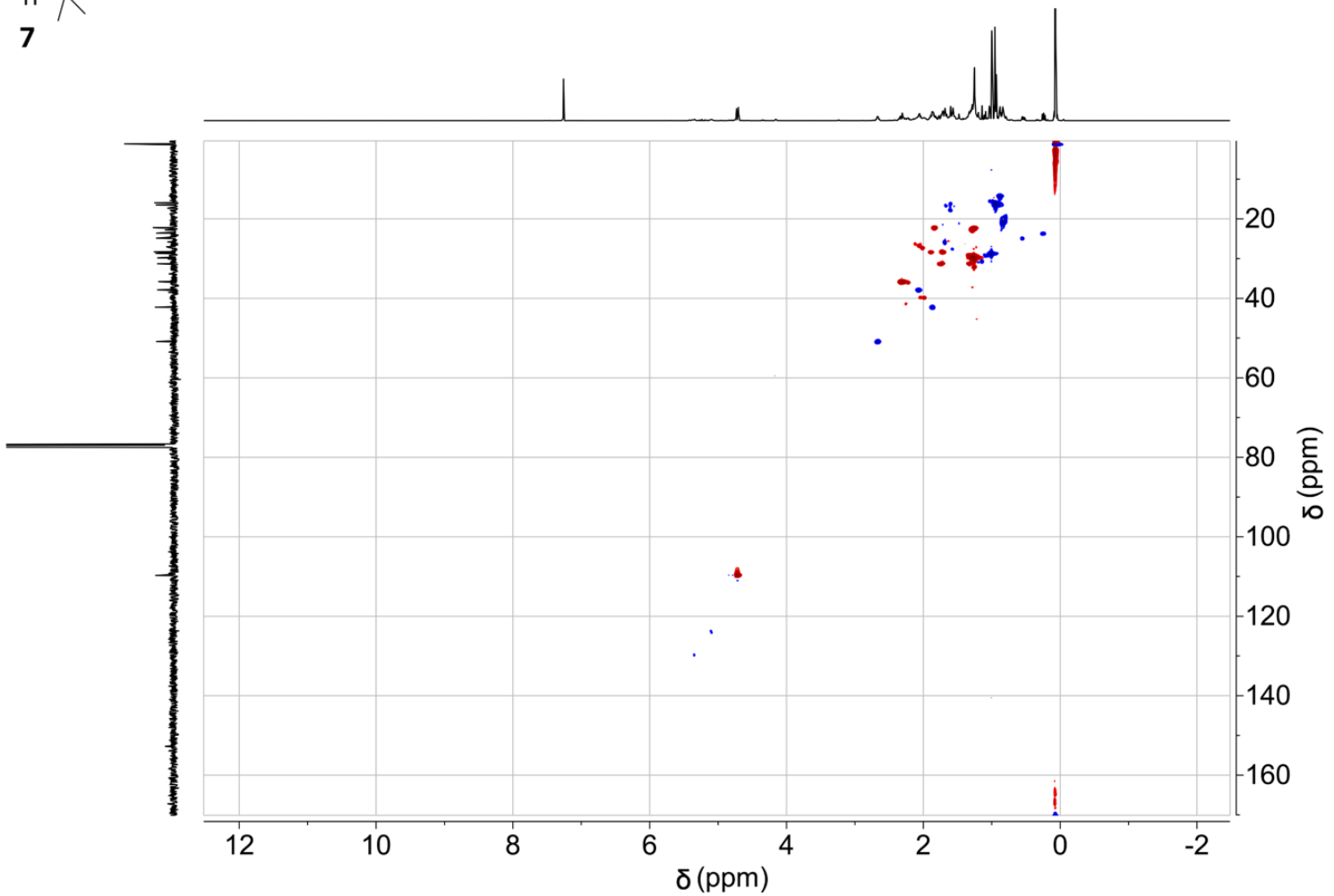
(+)-Alloaromadendrene ^{13}C

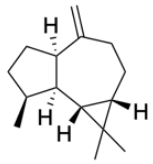




7

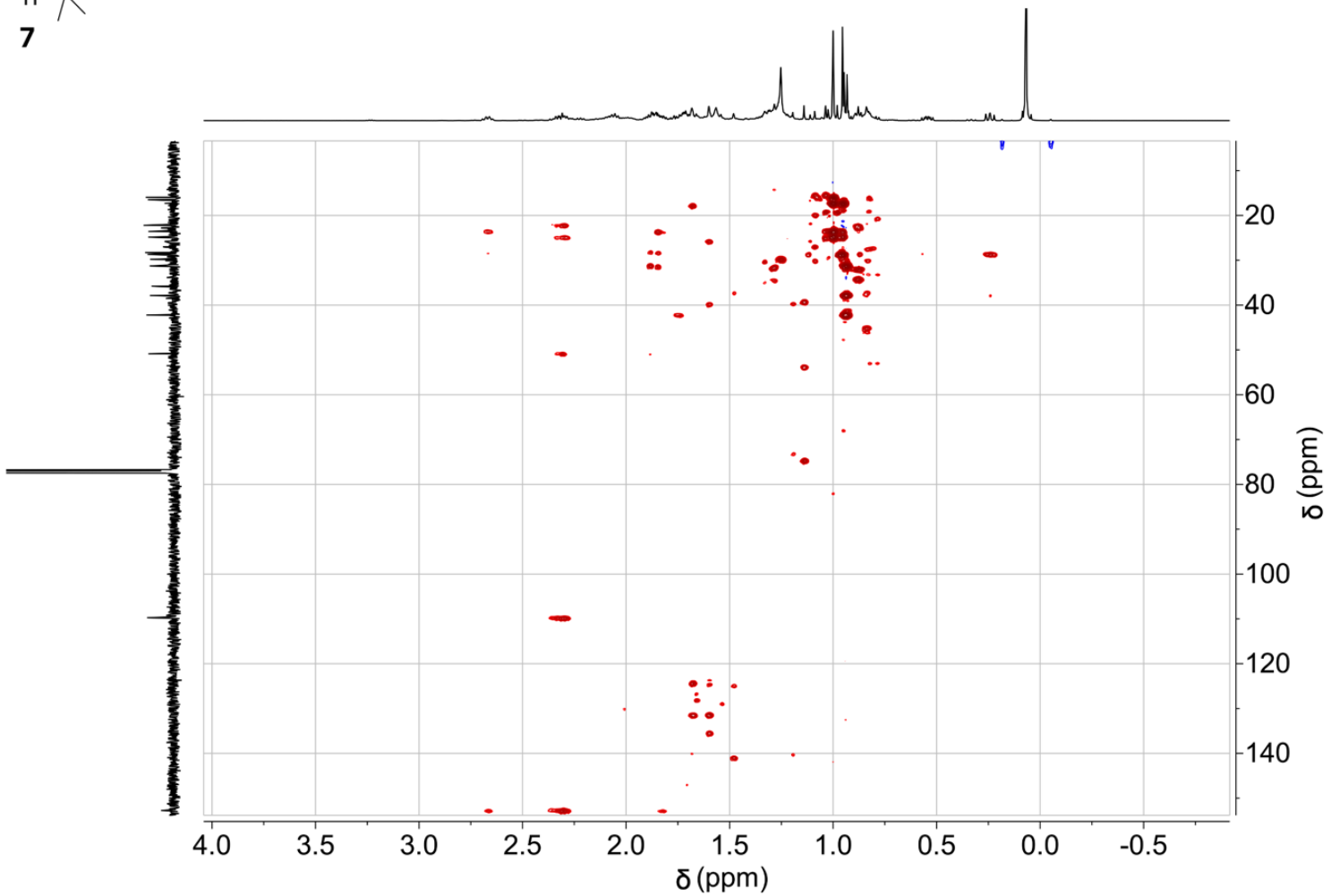
(+)-Alloaromadendrene HSQC

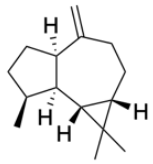




7

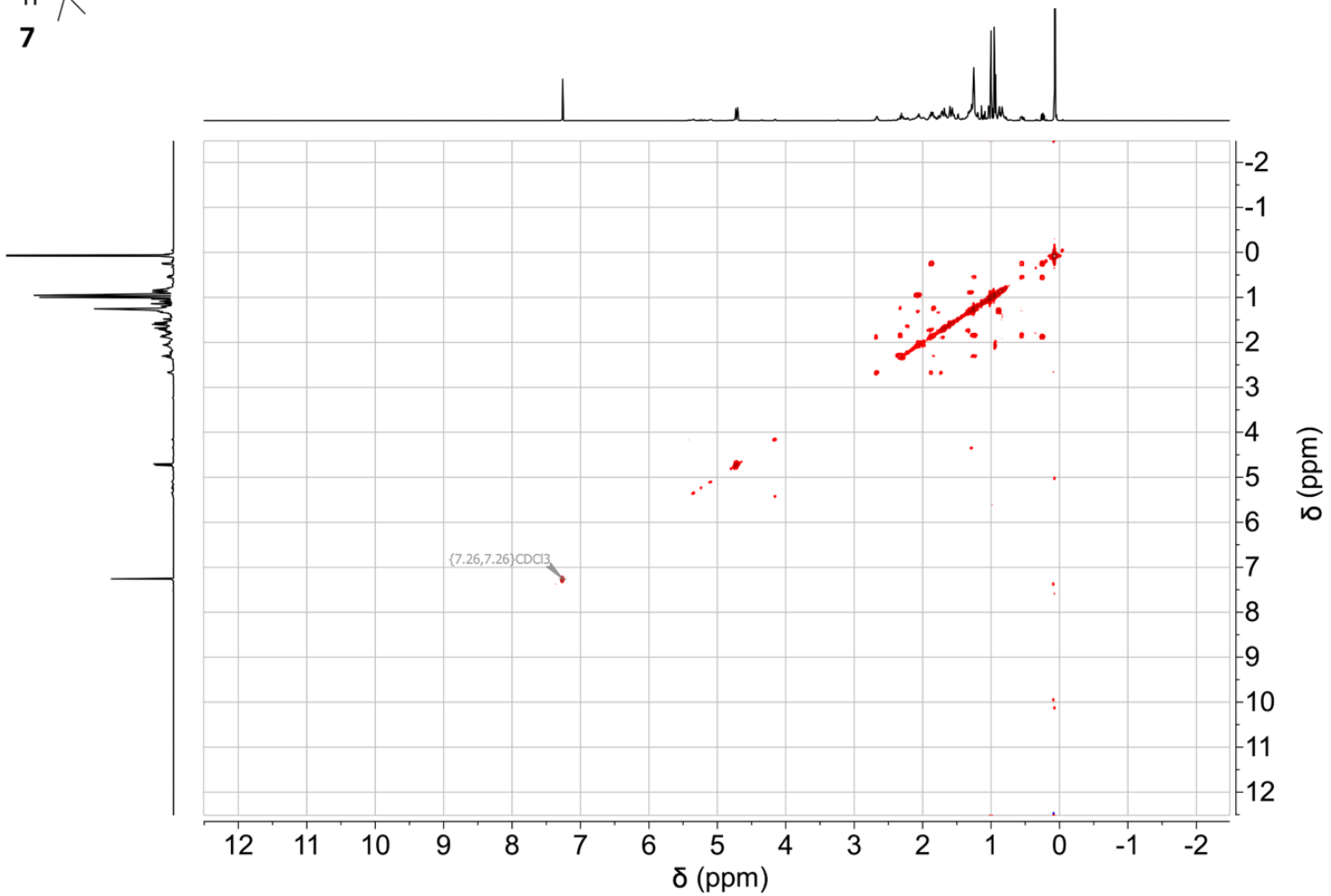
(+)-Alloaromadendrene HMBC

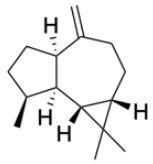




7

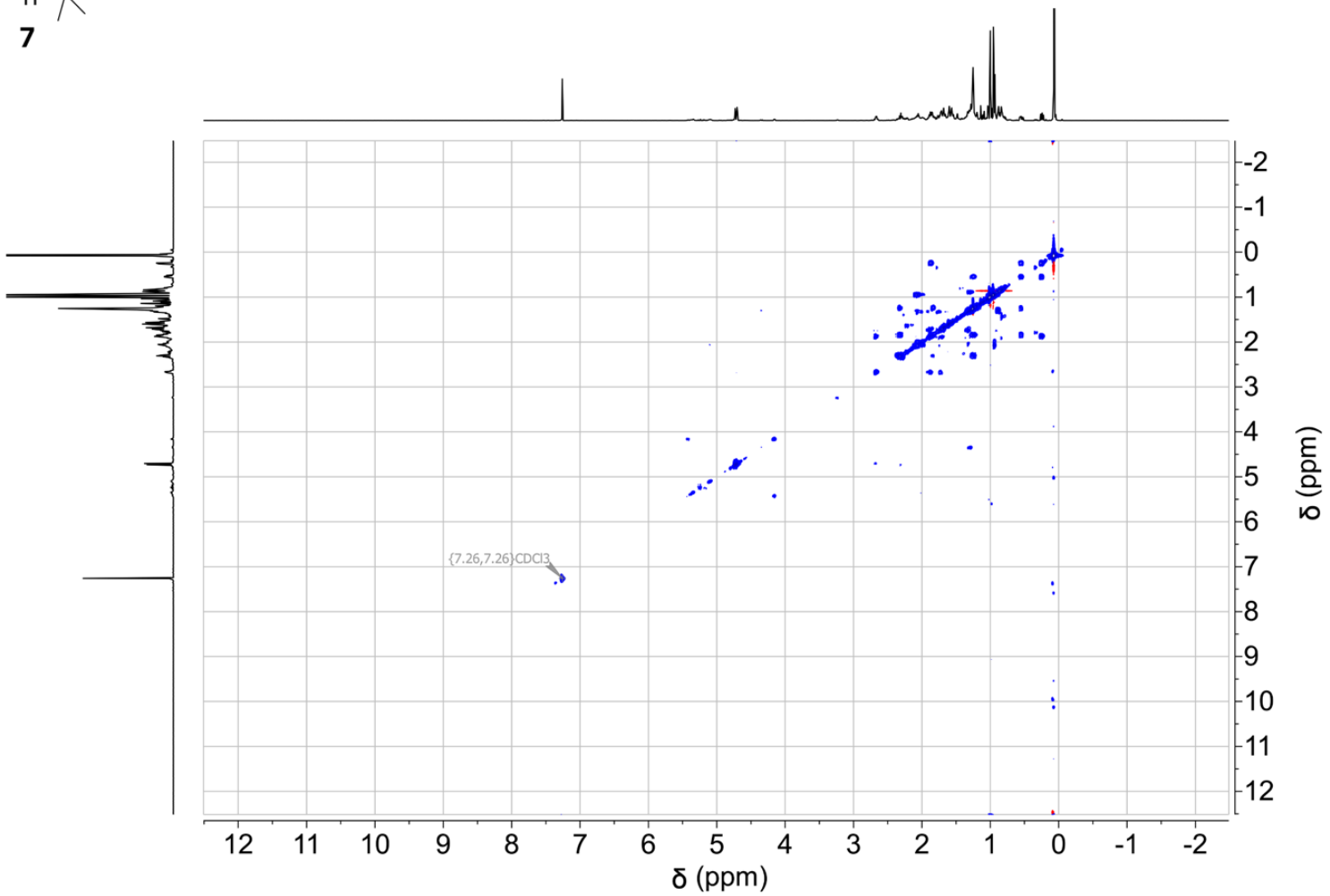
(+)-Alloaromadendrene COSY

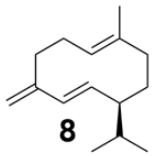




7

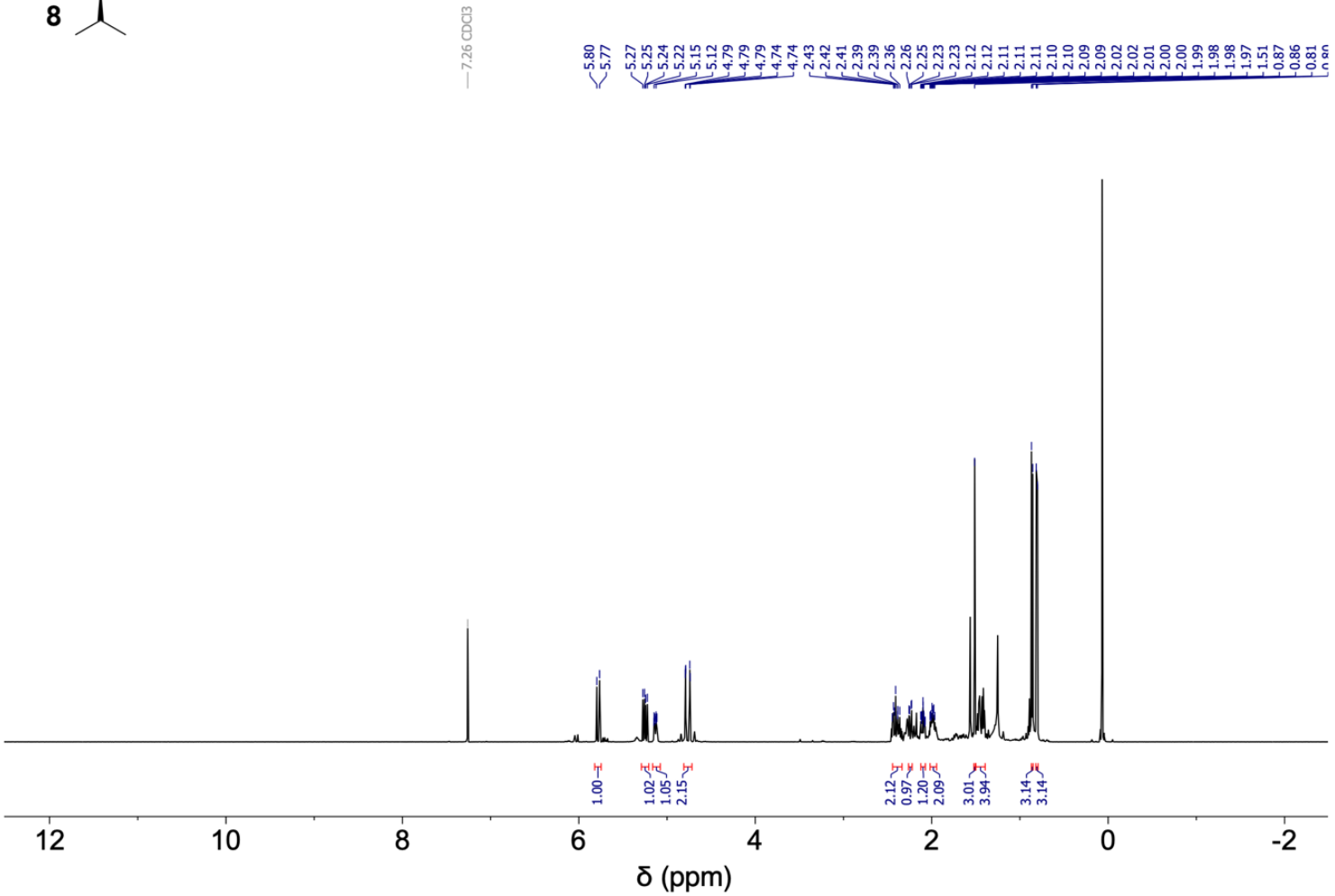
(+)-Alloaromadendrene NOESY

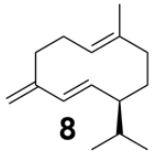




8

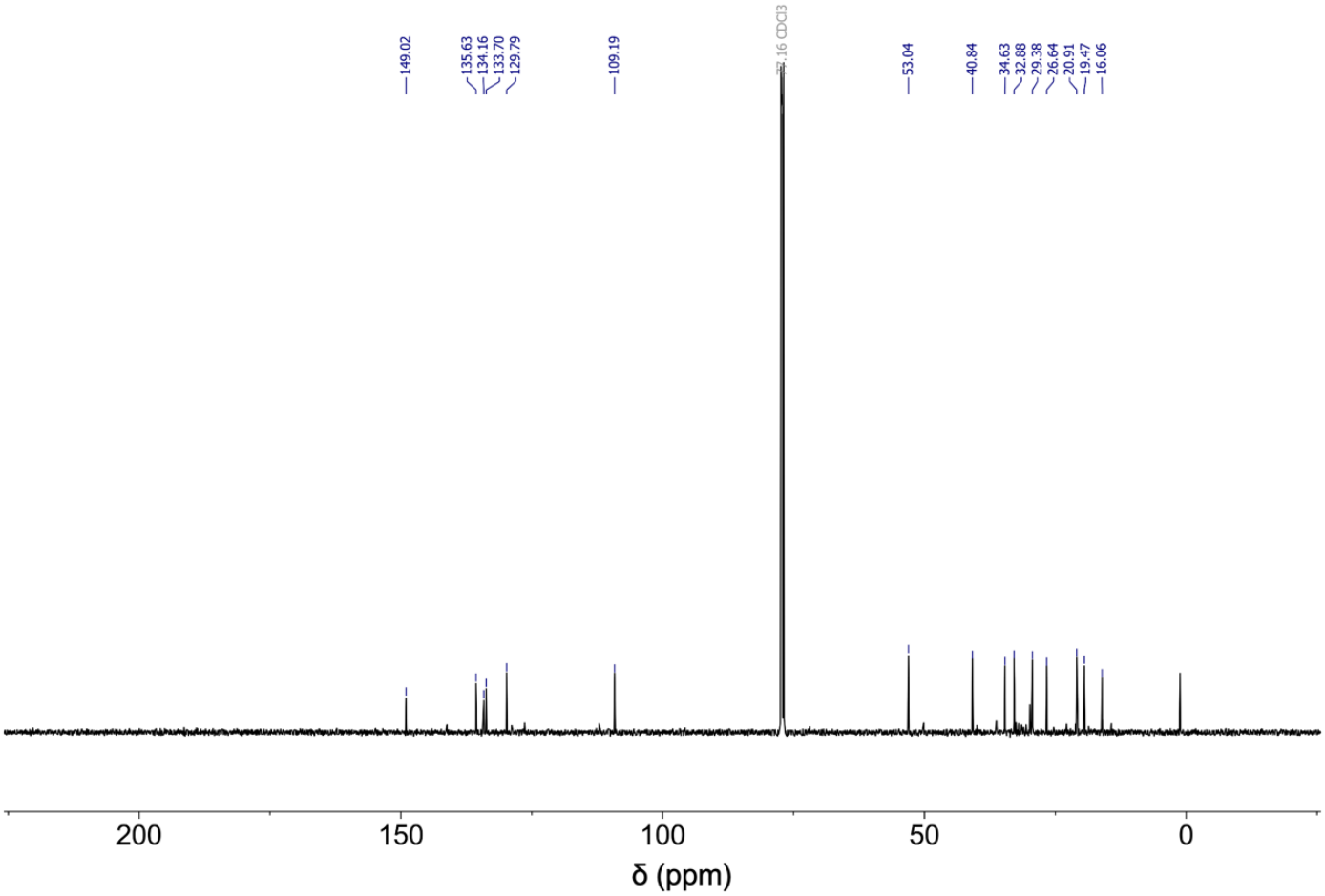
(-)-Germacrene D ^{1}H

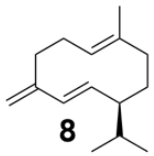




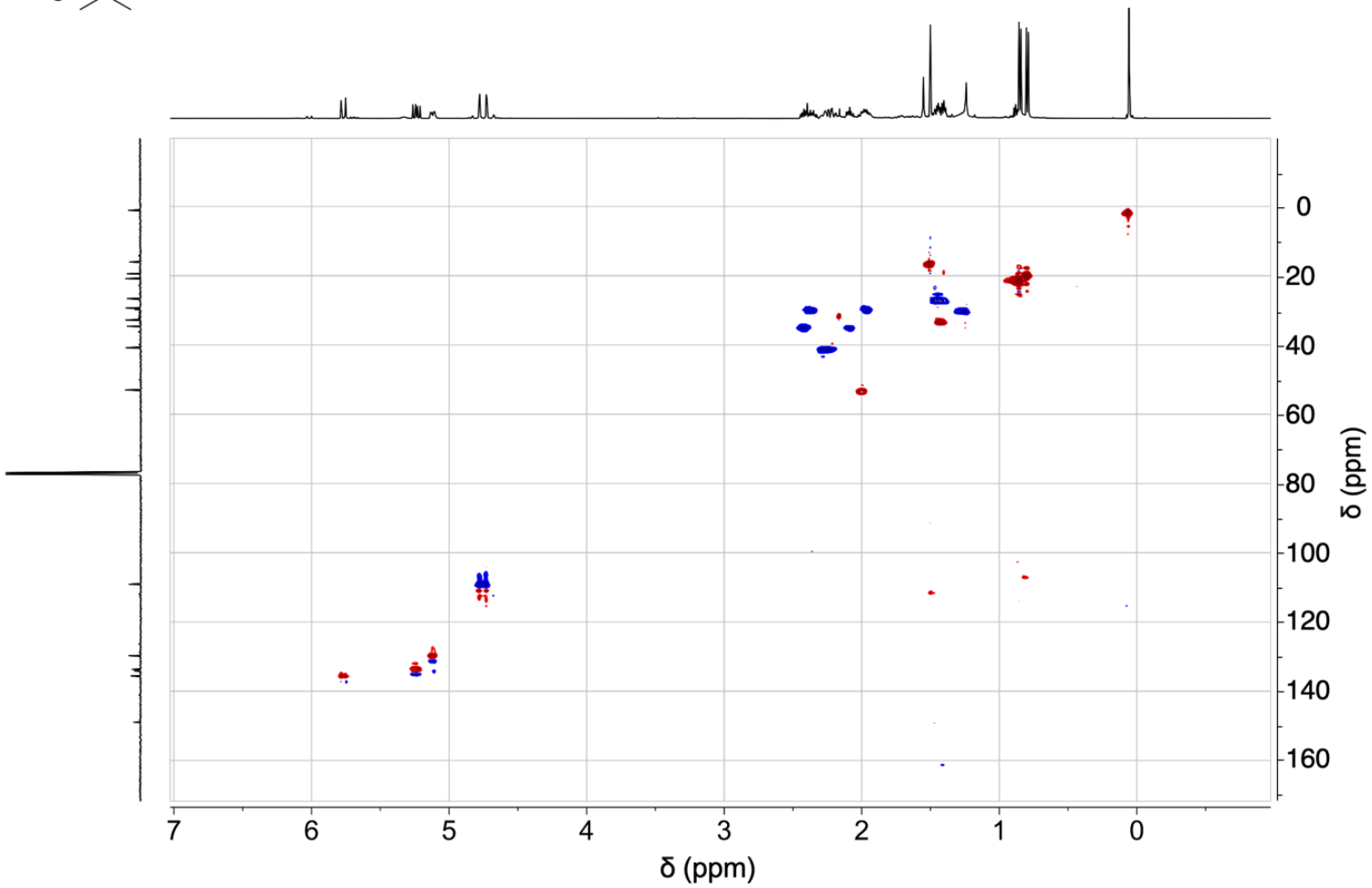
8

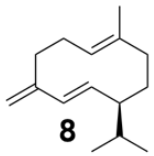
(-)-Germacrene D ^{13}C



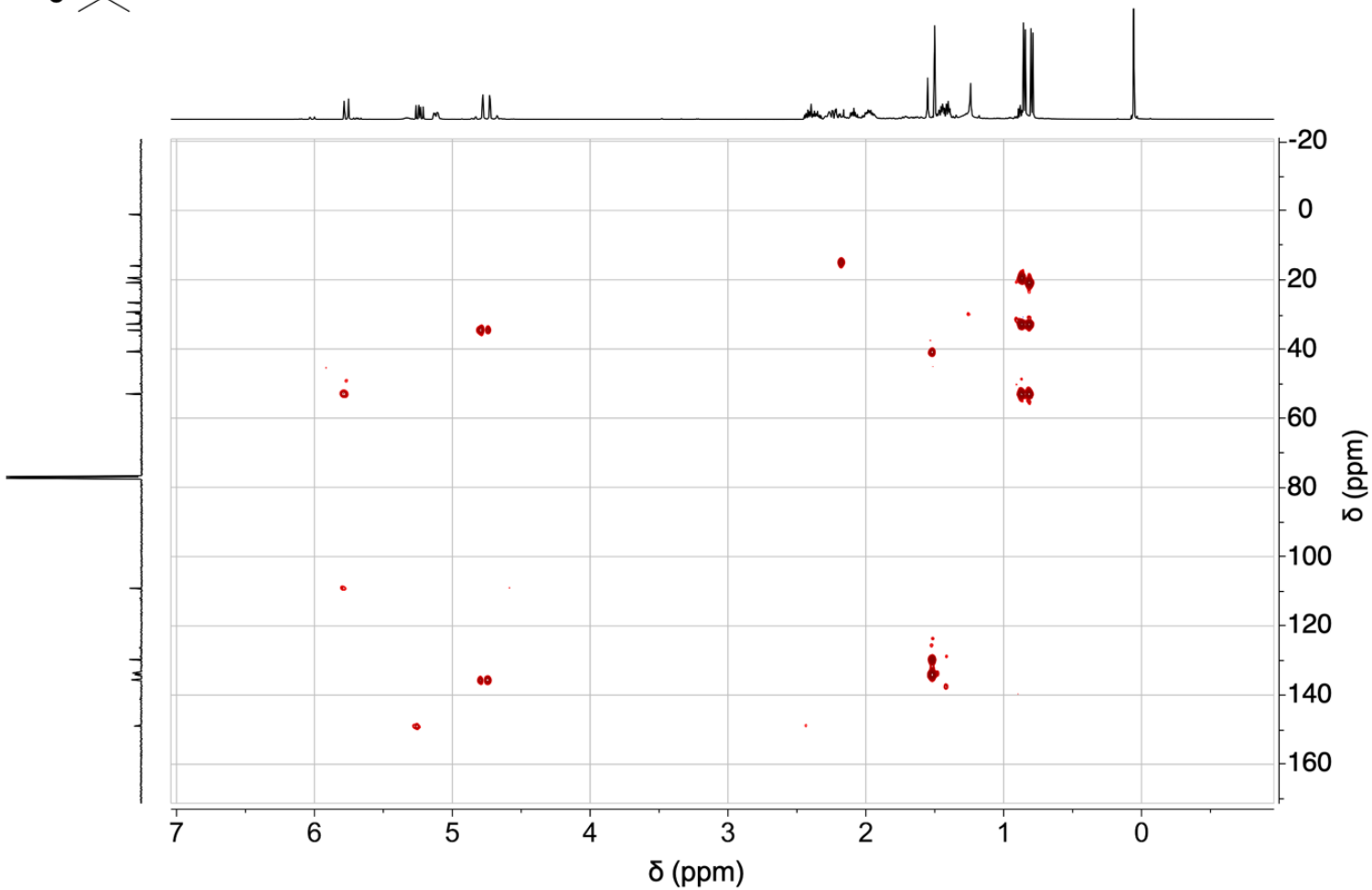


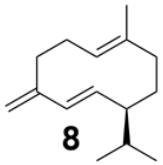
(-)-Germacrene D HSQC





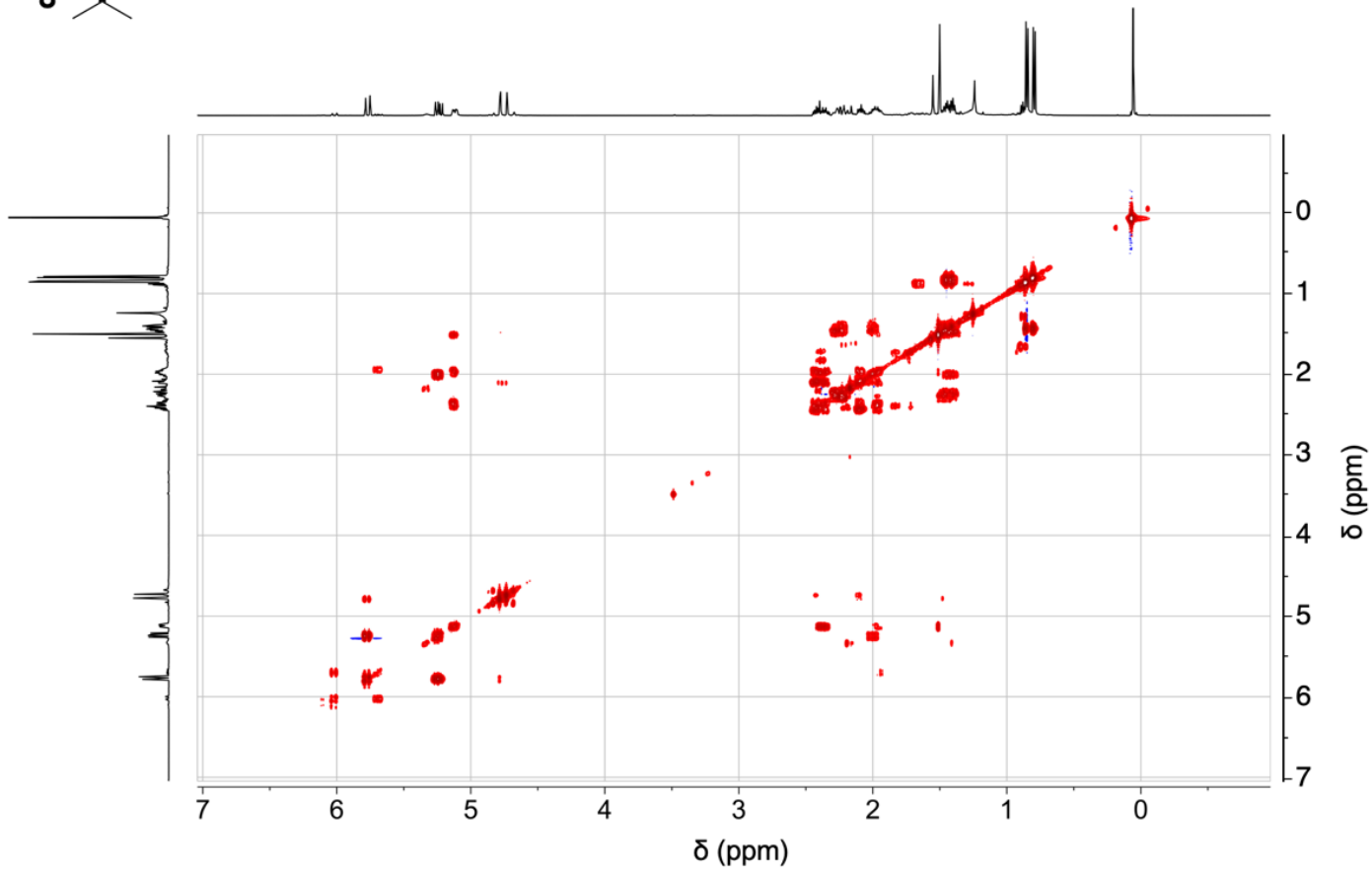
(-)-Germacrene D HMBC

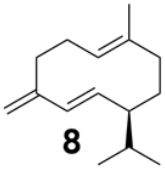




8

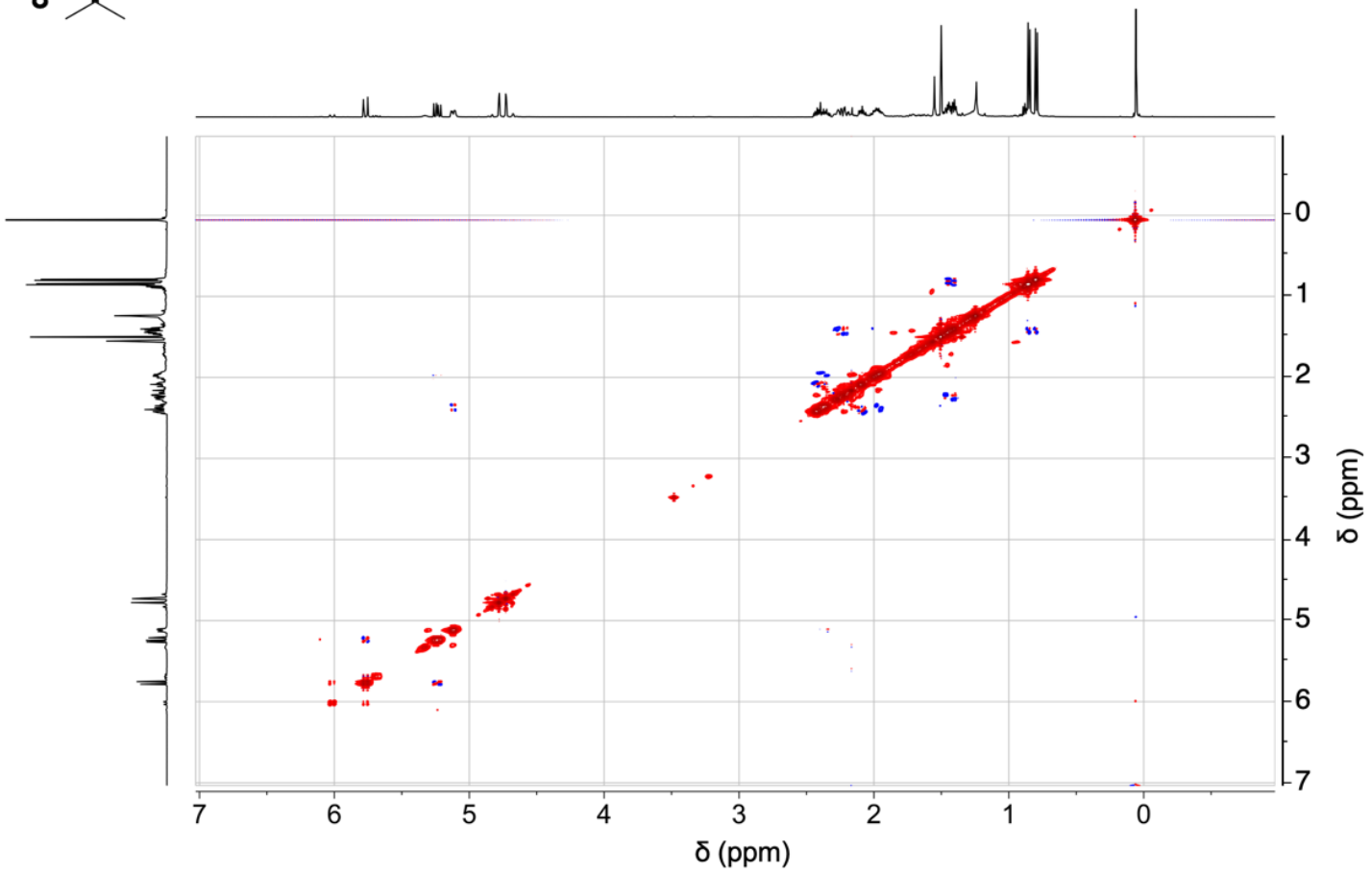
(-)-Germacrene D COSY





8

(-)-Germacrene D NOESY



Supplementary References

1. S. Podell, *et al.*, Pangenomic comparison of globally distributed *Poribacteria* associated with sponge hosts and marine particles. *ISME J.* **13**, 468–481 (2019).
2. D. Erpenbeck, *et al.*, Molecular biodiversity of Red Sea demosponges. *Mar. Pollut. Bull.* **105**, 507–514 (2016).
3. T. L. Turner, S. Lonhart, The Sponges of the Carmel Pinnacles Marine Protected Area. *bioRxiv* (2022) https:/doi.org/https://doi.org/10.1101/2022.11.02.514922.
4. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
5. A. V. Zimin, *et al.*, The MaSuRCA genome assembler. *Bioinformatics* **29**, 2669–2677 (2013).
6. B. J. Walker, *et al.*, Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* **9,** 1–14 (2014).
7. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
8. R. Challis, E. Richards, J. Rajan, G. Cochrane, M. Blaxter, BlobToolKit - interactive quality assessment of genome assemblies. *G3 Genes, Genomes, Genet.* **10**, 1361–1374 (2020).
9. L. Schmittmann, S. Franzenburg, L. Pita, Individuality in the Immune Repertoire and Induced Response of the Sponge *Halichondria panicea*. *Front. Immunol.* **12**, 1–13 (2021).
10. A. D. Finoshin, *et al.*, Iron metabolic pathways in the processes of sponge plasticity. *PLoS One* **15**, 1–25 (2020).
11. B. Buchfink, K. Reuter, H. G. Drost, Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 366–368 (2021).
12. S. R. Eddy, Accelerated profile HMM searches. *PLoS Comput. Biol.* **7** (2011).
13. I. Burkhardt, T. De Rond, P. Y. Chen, B. S. Moore, Ancient plant-like terpene biosynthesis in corals. *Nat. Chem. Biol.* **18**, 664–669 (2022).
14. I. M. A. Chen, *et al.*, The IMG/M data management and analysis system v.6.0: New tools and advanced capabilities. *Nucleic Acids Res.* **49**, D751–D763 (2021).
15. T. Ryu, *et al.*, Hologenome analysis of two marine sponges with different microbiomes. *BMC Genomics* **17**, 1–11 (2016).
16. M. Srivastava, *et al.*, The *Amphimedon queenslandica* genome and the evolution of animal complexity. *Nature* **466**, 720–726 (2010).
17. N. J. Kenny, *et al.*, Tracing animal genomic evolution with the chromosomal-level assembly of the freshwater sponge *Ephydatia muelleri*. *Nat. Commun.* **11**, 1–11 (2020).
18. B. Plese, *et al.*, Mitochondrial evolution in the Demospongiae (Porifera): Phylogeny, divergence time, and genome biology. *Mol. Phylogenet. Evol.* **155,** 1–14 (2021).
19. T. Lassmann, O. Frings, E. L. L. Sonnhammer, Kalign2: High-performance multiple alignment of protein and nucleotide sequences allowing external features. *Nucleic Acids Res.* **37**, 858–865 (2009).
20. B. Q. Minh, *et al.*, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
21. J. Piel, *et al.*, Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei. Proc. Natl. Acad. Sci.* **101**, 16222–16227 (2004).
22. K. M. Fisch, *et al.*, Polyketide assembly lines of uncultivated sponge symbionts from structure-based gene targeting. *Nat. Chem. Biol.* **5**, 494–501 (2009).
23. T. Mori, *et al.*, Single-bacterial genomics validates rich and varied specialized metabolism of uncultivated *Entotheonella* sponge symbionts. *Proc. Natl. Acad. Sci.* **115**, 1718–1723 (2018).
24. V. Agarwal, *et al.*, Metagenomic discovery of polybrominated diphenyl ether biosynthesis by marine sponges. *Nat. Chem. Biol.* **13**, 537–543 (2017).
25. Y. Nakashima, Y. Egami, M. Kimura, T. Wakimoto, I. Abe, Metagenomic analysis of the sponge Discodermia reveals the production of the cyanobacterial natural product kasumigamide by "Entotheonella." *PLoS One* **11**, 1–15 (2016).
26. M. F. Freeman, A. L. Vagstad, J. Piel, Polytheonamide biosynthesis showcasing the metabolic potential of sponge-associated uncultivated "Entotheonella" bacteria. *Curr. Opin. Chem. Biol.* **31**, 8–14 (2016).
27. R. Ueoka, *et al.*, Metabolic and evolutionary origin of actin-binding polyketides from diverse

organisms. *Nat. Chem. Biol.* **11**, 705–712 (2015).
28. P. Baer, *et al.*, Induced-fit mechanism in class i terpene cyclases. *Angew. Chemie - Int. Ed.* **53**, 7652–7656 (2014).
29. J. E. McMurry, G. K. Bosch, Synthesis of Macrocyclic Terpenoid Hydrocarbons by Intramolecular Carbonyl Coupling: Bicyclogermacrene, Lepidozene, and Casbene. *J. Org. Chem.* **52**, 4885–4893 (1987).
30. T. J. A. Bruce, *et al.*, Response of economically important aphids to components of *Hemizygia petiolata* essential oil. *Pest Manag. Sci.* **61**, 1115–1121 (2005).
31. Z. Y. Huang, Q. Y. Wu, C. X. Li, H. L. Yu, J. H. Xu, Facile Production of (+)-Aristolochene and (+)-Bicyclogermacrene in Escherichia coli Using Newly Discovered Sesquiterpene Synthases from Penicillium expansum. *J. Agric. Food Chem.* **70**, 5860–5868 (2022).
32. R. Faure, Two-Dimensional Nuclear Magnetic Resonance of Sesquiterpenes. 4*-Application to Complete Assignment of ’H and "C NMR Spectra of Some Aromadendrane Derivatives. *Magn. Reson. Chem.* **29**, 969–971 (1991).
33. M. Yasumoto, K. Mada, T. Ooi, T. Kusumi, New terpenoid components from the volatile oils of the soft corals Clavularia viridis and Sarcophyton acutangulum. *J. Nat. Prod.* **63**, 1534–1536 (2000).
34. T. Røstelien, A. K. Borg-Karlson, J. Fäldt, U. Jacobsson, H. Mustaparta, The plant sesquiterpene germacrene D specifically activates a major type of antennal receptor neuron of the tobacco budworm moth *Heliothis virescens*. *Chem. Senses* **25**, 141–148 (2000).
35. H. Itokawa, H. Matsumoto, S. Mihashi, Constituents of Torilis japonica D.C. I. Isolation and Optical Purity of Germacra-4 (15), 5 (E), 10 (14)-trien-1β-ol. *Chem. Pharm. Bull.* **31**, 1743–1745 (1983).
36. M. Niwa, M. Iguchi, S. Yamamura, Co-occurrence of (-) and (+)-Germacrene-D in *Solidago altissima* L. : Determination of the Optical Rotation of optically Pure Germacrene-D. *Chem. Pharm. Bull.* **28**, 997–999 (1980).