Supplementary Information for

A widespread metabolic gene cluster family in metazoans

Authors
Natalie E. Grayson[1,8], Paul D. Scesa[2,8], Malia L. Moore[1,3], Jean-Baptiste Ledoux[4], Jessica Gomez-Garrido[5,6], Tyler Alioto[5,6], Todd P. Michael[1,3], Immo Burkhardt[1*], Eric W. Schmidt[2*], Bradley S. Moore[1,7*]

Affiliations
[1]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA
[2]Department of Medicinal Chemistry, University of Utah, Salt Lake City, UT, USA
[3]The Plant Molecular and Cellular Laboratory, Salk Institute for Biological Studies, La Jolla, CA, USA
[4]CIIMAR, Centro Interdisciplinar de Investigação Marinha e Ambiental, Universidade do Porto, Porto, Portugal
[5]Centro Nacional de Análisis Genómico, Barcelona, Spain
[6]Universitat de Barcelona, Barcelona, Spain
[7]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, USA
[8]These authors contributed equally

*Correspondence: iburkhardt@ucsd.edu (IB), ews1@utah.edu (EWS), bsmoore@ucsd.edu (BSM)

Contents

Methods

Sample collection. Divers collected *Renilla koellikeri* and *Stylatula elongata* (both in the Pennatuloidea superfamily) from La Jolla, CA*, Briareum asbestinum* (family Briareidae) and *Erythropodium caribeorum* (family Erythropodiidae) coral from Broward County, Florida, and *Corallium rubrum* (family Coralliidae) coral from Cap Castel, Illes Medes, Catalunya, Spain. We also obtained a *Dichotella gemmacea* (family Ellisellidae) specimen from Two Little Fishies Inc. All samples were washed 3 times for 5 min each with Ca^{2+}- and Mg^{2+}-free artificial seawater and then flash frozen with liquid nitrogen as live organisms and kept at -80 °C for less than 2 weeks until DNA extraction. At the time of flash freezing, aliquots of each organism were preserved in RNA*later*. In the case of *S. elongata*, taxonomy was uncertain, so an aliquot of tissue from the base of the polyp leaf was collected, dissolved in bleach, and analyzed under a microscope. Given the absence of slender rods with homology to those characterized by *Stylatula,*[24] with help from Catherine S. McFadden (Harvey Mudd College) we were able to discern with certainty that the organism was a *S. elongata* (**SI Figure 17**).[1]

High molecular weight DNA extraction and quantification. High molecular weight (HMW) DNA was isolated from tissue from *R. koellikeri*, *S. elongata*, *B. asbestinum*, and *D. gemmacea* as previously described.[2] In brief, tissue was homogenized by mortar and pestle while frozen with liquid nitrogen. The Qiagen genomic-tip kit 500/G 9 (Cat# 10262) was used, following the protocol for the "Preparation of Tissue Samples" with the recommendations of tissue amounts for mouse tail. 1 µl he extracted DNA was of the sample on a BioDrop µLITE (BioDrop, Holliston, MA, USA) to ensure the purity of DNA, confirming that the OD 260/280 was 1.8 and the OD 260/230 was between 2.0–2.2. DNA concentrations were assessed using the dsDNA HS assay on a Qubit fluorometer v3.0 (Thermo Fisher Scientific, Singapore). DNA size distribution was assessed with a TapeStation (Agilent 4150, Santa Clara, CA, USA). Based on the degree of purity and size distribution of the DNA, samples were either purified and size selected further using the Oxford Nanopore Technology (ONT) short fragment eliminator kit (Cat # EXP-SFE001) and Beckman-Coulter Ampure RNA clean XP beads (Cat # A63987) or directly used to build sequencing libraries. Depending on post quality control concentration and size distribution, samples were either used to prepare libraries for Pacific Biosciences (PacBio) or ONT sequencing. For all four organisms an aliquot of HMW DNA was also used to obtain Illumina short read data.

***E. caribaeorum* DNA extraction and sequencing.** HMW DNA was extracted from frozen, RNA*later* preserved tissue using chloroform-phenol extraction as previously described.[3] Crude DNA precipitate from chloroform-phenol extraction was dissolved in nuclease free water (200 ml) and loaded onto an equilibrated Qiagen Genomic Tip (20/G, Cat# 10223) column. The column was washed and eluted, and the DNA was precipitated from the eluent according to the Genomic Tip Kit protocol. DNA quality and quantity was assessed by Nanodrop and gel electrophoresis. Illumina reads had been previously obtained and are publicly available for this sample (**SI Table 4**)[3].

RNA extraction. RNA was extracted from the tissues of *R. koellikeri*, *S. elongata*, and *D. gemmacea* preserved in RNA-later previously described.[2] In brief, the RNeasy Plant Mini Kit was used (Qiagen, Cat # 74904) including the optional step of an on column DNAse digestion. Quality and quantity were both assessed via nanodrop and qubit methods as described in DNA extraction. For *E. caribaeorum* and *B. asbestinum,* RNA was extracted using methods previously described.[3]

Sequencing methods. The *R. koellikeri* and *S. elongata* genomes were sequenced with PacBio. PacBio Sequel IIe running SMRT Link v11.1 with SMARTbell prep kit 3.0 (Cat # 102-182-700) produced 28.7 Gb with a mean length of 11,734 bp and 85.4% HiFi reads (Q>20) for *R. koellikeri* and 26.2 Gb with a mean length of 12,845 bp and 77.9% HiFi reads for *S. elongata*. Due to lower yields and quality of DNA, *D. gemmacea* and *B. asbestinum* genomes were sequenced with ONT. Following library prep with ligation chemistry (Cat # SQK-LSK114), an ONT Promethion running MinKNOW v23.04.6 produced 77.8 Gb for *D. gemmacea* and 59.4 Gb for *B. asbestinum*, quantified by NanoPlot v1.40.0. *E. caribaeorum* HMW DNA (3 mg) was used to prepare an ONT library using the ligation sequencing kit (SQK-LSK114) according to the manufacturer's protocol with size selection using long fragment buffer from the kit. The library was loaded onto a Promethion flow cell (FLO-PRO114M), and data collected using MinKNOW (v. 24.02.10) with real time, high accuracy base calling. The run yielded 17.46 Gb called bases with an N50 of 3.09 kb.

We generated short reads for genome size estimates and assembly polishing using an Illumina NovaSeq6000 with an S4 300 cycle run configuration, following Illumina DNA library prep (Cat # 20060059), which produced 150 million reads for *R. koellikeri*, 206 million reads for *S. elongata*, 146 million reads for *D. gemmacea*, and 114 million reads for *B. asbestinum*. For *R. koellikeri*, *S. elongata*, and *D. gemmacea*, the same Illumina run yielded RNAseq for transcriptome assembly and genome annotations. cDNA libraries were prepared from total RNA with Illumina TruSeq Stranded Total RNA kit (Cat # 20020599) with rRNA depletion steps, barcoded with IDT TruSeq RNA indexes (Cat # UDI-20022371). This yielded 25.8 million reads for *R. koellikeri* peduncle tissue, 54.8 million reads for *R. koellikeri* body tissue, 68.8 million reads for *S. elongata*, and 47.7 million reads for *D. gemmacea*. Given insufficient depth for terpene synthase transcript mapping across several samples, an additional sequencing run with an S2 300 cycle configuration yielded 114 million reads for *R. koellikeri peduncle tissue*, 64 million reads for *R. koellikeri body tissue,* 226 million reads for *S. elongata*, and 618 million reads for *D. gemmacea*. In the case of *E. caribeorum,* genomic short read sequencing data from the same individual were generated in a previous study and already publicly available (**SI Table 4**).[3]

Genome size estimation. Genome size was estimated by K-mer frequency distribution analysis (genome Size=K-mer_num/Peak_depth). First, the short reads were filtered using fastp (v 0.19.4)[4] with default parameters and quality checked with FastQC (v 0.11.9). Subsequently, the K-mers (k=17 bp) were counted using Jellyfish (v 2.2.10)[5] with the parameter "-C -m 51 -s 10000000000 -t 50". The output histogram file was then analyzed

with the following equation: Sum of the K-mer number*K-merCount – (minus) the first peak/ (divided) by the K-mer peak. This equation assumes that all K-mers that occured exactly once in the read set are erroneous and should be ignored, computes the depth of non-erroneous K-mers, and estimates diploid and haploid genome sizes using the two different peak coverages. GenomeScope was also used to estimate the genome size with default parameters[6]. The frequency distribution of 17-mers is based on the genome characteristics and in the light of the pattern of Poisson distribution (**SI Figure 3, SI Table 6**).

Genome assembly. Nanoplot (v 1.41.0)[7] was used to determine the quality of the sequencing data obtained for all five organisms.
ONT: While we used the entirety of ONT generated reads for *B. asbestinum* and *E. caribeorum*, in the case of *D. gemmacea,* seqtk (v 1.3) was used to remove reads that were shorter than 10 kb. We used flye (v 2.9.1)[8] with default parameters to assemble the genomes of *B. asbestinum, E. caribeorum,* and *D. gemmacea*. We then used Minimap2 (v 2.24)[9] to align the long reads to the assembly, and with these aligned reads, we generated a consensus assembly running racon (v 1.5.0)[10] with long reads three times. We next trimmed and removed adapter sequences from Illumina short reads. FastQC was used to assess quality and fastp was used for cleaning. Then used Minimap2 to map the short reads against the consensus assembly. Lastly, we used pilon (v 1.24)[11] to polish the ONT assemblies with our mapped short read data.
PacBio: *R. koellikeri* and *S. elongata* were assembled with hifiasm (v 0.19.5)[12] with default parameters; however, in the case of *R. koellikeri*, hifiasm incorrectly calculated the homozygous coverage peak. The hom-cov parameter was adjusted to 165, where the homozygous peak appeared in the K-mer based plot **(SI Figure 3).** All assembled genomes were checked for duplication (D%) rate and completeness (C%) with Benchmarking Universal Single-Copy Orthologs (BUSCO v 5.4.7)[13]s using a metazoan database as reference. Assembly contiguity was determined by QUAST (v 5.2.0)[14].

Transcriptome assembly. The trimmed and adapter removed (FastQC and FastP with default parameters) cDNA sequences from *R. koellikeri*, *S. elongata*, *B. asbestinum*, and *D. gemmacea* were assembled with trinity (v 2.8.5)[15] using the stranded parameter (--SS_lib_type RF). The *E. caribaeorum* transcriptome was assembled as described previously using Trimmomatic[16] for adapter removal and SPADES[17] for assembly.[3] The processed cDNA sequences were also mapped to the assembled genomes using HiSat2[18] (**SI Table 6**). A poor alignment score for *B. asbestinum* (40.36%) hints at potential contamination or poor extraction of RNA from the sample; however, the alignment was sufficient for our genome mining pipeline.

Syntenic analysis. We found the terpene cyclases in our genomic and transcriptomic assemblies using traditional blast searches with the *R. muelleri* cembrene B cyclase previously characterized by Burkhardt et al.[19] We then used the Conserved Domain Database (NIH web-based search tool)[20] to annotate other functional proteins in genomic neighborhood of the terpene cyclases. We determined intron bounds of all genes with their

hits in the respective transcriptomes. In the case that the gene of interest was not transcriptionally active and could not be found by blasting into the transcriptome, we used fgenesh, an HMM-based gene structure prediction tool in the softberry suite[21], with *Nematostella vectensis* as a protein model. We further validated the predicted genes with hmmscan[22] against the profile HMM database to confirm the presence of the correct conserved motifs (see **SI Figure 18**). We subsequently annotated the contigs in benchling and exported them as a GenBank file. To simplify the syntenic analysis, we removed all introns from gene coding regions and shortened intergenic regions. The non-modified BGCs are depicted in **SI Figure 4**. We then used Clinker (v 0.26)[23] to calculate the sequence similarity and visualize the genetic synteny (**Figure 3b** of synteny and **SI Figure 5** of percent identity matrices).

Phylogenetic analysis. For phylogenetic analysis of the cytochrome P450s, we followed the method outlined by Pankov, K. V. et al.[24] for genome wide annotation of cytochrome P450 genes. In short, we used TransDecoder (v 5.7.1) to identify candidate coding regions within transcript sequences of *R. koellikeri, S. elongata, B. asbestinum*, *D. gemmacea, Heliopora coerulea, Corallium rubrum,* and *E. caribaeorum*. The Basic Local Alignment Search Tool (BLAST)[25] was used to identify local alignments between the coding regions of the seven octocoral and a query that consisted of all annotated CYPs from *Hydra vulgaris, Acropora digitifera, Aurelia aurita,* and *Nematostella vectensis* in a blastp search (protein query against a proteome). Given that not all CYPs are transcriptionally active and cDNA sequencing only offers a snapshot of an organism's biosynthetic potential, we used tblastn (protein query against a genome nucleotide) searches to identify CYP alignments between the original cnidarian CYP list amended with the CYPs discovered in the seven octocoral transcriptomes and the respective genomes of those octocoral (**SI Table 8** of list of CYPs included and sources). To pull out the scoring pairs from the genomic blast searches, we used bedtools (v 2.26.0)[26]. Further, we used transeq (v 5.7.1) to translate the nucleotide hits in the genomes. To ensure that all protein hits were in fact cytochrome P450s, we used hmmscan to ensure they had the correct motifs. Sequences with alignment scores and coverage of 80% or lower and percent identity of 20% or lower were discarded. We then used cd-hit (v 4.8.1)[27] with a 99% cut off to remove duplicates.

To pull out candidate cbCYPa, -b, and -c sequences from the poor contiguity genome of the *Pteroeides caledonicum,* we used BLAST with the homologous *R. muelleri* sequences. Finally, we used Kalign (v.2.04)[28] to generate alignments of the CYP amino acid sequences found in the octocoral genomes and transcriptomes and the *Hydra vulgaris, Acropora digitifera, Aurelia aurita,* and *Nematostella vectensis* CYPs annotated by Pankov, K. V. et al.[24] and the putative briarane biosynthesis CYPs in *Pteroeides caledonicum*. The phylogenetic analysis was performed using IQ-TREE multicore (v 2.0.3) substitution model Q.pfam+F+R10 (determined by ModelFinder as included in IQ-TREE)[29] and 1,000 bootstrap replicates. The resulting tree was visualized with iTOl (v 6.9.1)[30] (**Figure 3C, SI Figure 6**).

Chemical isolation and elucidation of the *R. koellikeri* briarane diterpene. *R. koellikeri* (approximately 1 kg wet weight) was extracted with methanol:dichloromethane 2:1 (600 ml) overnight three times with stirring. The solvent was evaporated, and the residue was

partitioned between water and hexane (100 ml each) and then water and ethyl acetate (100 ml each). The hexanes, ethyl acetate, and aqueous fractions were all analyzed via LCMS for the presence of chlorinated compounds. While the organic fractions both had chlorinated compounds, the ethyl acetate fraction was much less complex and was used for the following stages of isolation and purification. After evaporating the solvent, 1.30 g of crude extract remained. The residue was then resuspended in 100 ml of 40% methanol and loaded onto a 20 g C_{18} (reversed phase silica) column and fractionated with a gradient of water:methanol from 40%-100% stepwise with 20 ml per fraction. Fractions were analyzed via liquid chromatography–MS (LC–MS), and those that contained chlorinated compounds with the fragmentation patterns of terpenes in the retention time window matching the expected polarity of a functionalized briarane were pooled. Further purifications were carried out with an Agilent Technologies 1260 Infinity series HPLC equipped with a degasser, quaternary pump, autosamplers, diode array, and fraction collector. A Kinetex 5 µm C18 100 Å, 250 x 10.0 mm column was used with an isocratic hold of 47% acetonitrile/water at a flow rate of 2 ml/minute. Purified compounds were then characterized via NMR (Bruker Avance III spectrometer- 600 MHz using a 1.7-mm inverse detection triple resonance (H-C/N/D) cryoprobe and JEOL ECZ spectrometer- 500 MHz using a 3mm probe) (**NMR supplementary note, SI Table 2, SI Figure 1**).

NMR analysis of R. koellikeri compounds. We identified 11-hydroxyptilosarcenone (**6**) as the major terpenoid present in *R. koellikeri* tissue (**NMR supplementary note, SI Table 2, SI Figure 1**). The identity of this compound, which had been previously isolated from the sea pen *Ptilosarcus gurneyi,* yet not previously from *Renilla* species, was supported by HRMS and NMR data. We collected full 1D and 2D NMR data sets on the mixture, including 2D COSY, NOESY, HSQC and HMBC spectra, in solvents CD_3OD and $CDCl_3$ to determine the main compound as **6**, representing 75% of the mixture. The second compound was determined to be ptilosarcenone. We found inconsistencies in the reported positional assignments of ^{1}H and ^{13}C NMR chemical shifts. The structure and chemical shifts of **6** were assigned using the 2D data as follows: COSY correlations between H-13 and H-14 along with an HMBC correlation from H-14 to C-12 indicated the presence of an enone fragment. HMBC correlations from H_3-15 to C-1, C-2, C-10, and C-14 along with HMBC correlations from H_3-20 to C-10, C-11, and C-12 confirmed the presence of the cyclohexanone fragment and the positions of the angular methyl and tertiary alcohol. An HMBC correlation from H-2 to an acetate carbonyl (δ_C 169.3) and a COSY correlation from H-2 to H-3 and from H-3 to H-4 established the position of the allylic acetate, which in turn was connected to the 1,1-disubstituted olefin based on an HMBC correlation from H-16_a and H-16_b to C-4. HMBC correlations from H-16_a and H-16_b to the conspicuous C-6 (δ_H 5.05 and δ_C 60.9 in the HSQC) suggested the likely position of a chloromethine. In turn, H-6 showed a COSY correlation to H-7 of the butrylactone oxymethine (δ_H 4.98 and δ_C 78.8 in the HSQC). A COSY correlation between H_3-18 and H-17 along with HMBC correlations from H_3-18 to C-8, C-17 and C-19 allowed full assignment of the methylated butryl lactone. A COSY correlation between the H-9 acetoxymethine (δ_H 5.81 and δ_C 69.4 in the HSQC) and H-10 suggested the closure of the decane ring. The decane ring closure and lactone position were confirmed by a wealth of HMBC correlations from H-9. This included an

HMBC correlation from H-9 to the acetate carbonyl (δ_C 170.0) as well as to C-7, C-8, C-10, C-11 and C-17 (full NMR data and 2D correlations **NMR supplementary note, SI Figure 1**). This led us to corroborate the structure of **6** as previously described, with the newly assigned positions of NMR chemical shifts presented in **SI Table 2.**

Crude chemical extraction and molecular networking. *D. gemmacea*, *C. rubrum*, *E. caribaeorum*, *S. elongata, B. asbestinum*, and *R. koellikeri* were used as samples for this work. Approximately, 1g of each frozen coral sample was crushed and extracted overnight in 2 mL of a 2:1 methanol:dichloro-methane mixture at room temperature while stirring. All extracts were first passed through a cotton filled Pasteur pipette to remove large remaining fragments and then evaporated at 40 °C using a rotary evaporator. The remaining residue was dissolved in methanol (HPLC grade, Fisher Scientific) and prepared for LC-MS analysis, by filtering through a Mini Spin Filter Column (American Chromatography Supplies). The purified 11-hydroxyptilosarcenone (**6**) from *R. koellikeri,* in addition to the tissue from 6 crude extract samples were analyzed by LCMS method A. The obtained LC-MS-MS data were converted to .mzML format using msConvert (ProteoWizard)[31] and analyzed via MzMine (v 3.9.0)[32]. The feature based molecular network parameters used and raw data are available through the MASSIVE dataset (ID: MSV000094792). The processed spectra were uploaded to GNPS[33] and visualized with cytoscape[34].

LCMS method A: LC-MS-MS analysis was conducted on a 6530 Accurate-Mass Q-TOF mass spectrometer (Agilent Technologies) coupled to the 1290 Infinity leveled-up 1260 Infinity LC system (Agilent Technologies) with a Kinetex® LC column (150 x 4.6 mm, 5 µm C18 100 Å, Phenomenex). Dual ESI was used as ion source in positive ionization mode with a carrier gas temperature of 300°C and the ionization voltage Vcap set to 3000 V. The applied collision energy voltages can be described by the following formula 2.6· m/z + 14.75. The LC gradient consisted of a 2-minute isocratic period at 20 % acetonitrile in water at 1.0 mL/min flow rate followed by a linear increase to 95% acetonitrile after 18 minutes total then an additional linear increase from 95% to 100% acetonitrile after 2 minutes total, then a 3-minute isocratic period at 100% acetonitrile (25 minutes total).

Plasmids, strains, and media. Candidate terpene cyclase genes were codon-optimized for yeast and purchased from Twist Biosciences in the pESC-leu2d vector backbone (Addgene #20120) along with a copy of the geranylgeranyl pyrophosphate synthase gene XdCrtE.[35–37] Candidate cytochrome P450 genes codon optimized for yeast were purchase from Twist Biosciences in the pESC-URA vector backbone (Agilent Technologies) along with a copy of the *H. coelurea* CPR gene. Candidate SDH genes were purchased from Twist Biosciences in the pET-28b+ vector with N-terminal His-tags at the NdeI/XhoI site. Haploid *S. cerevisiae* YPH499 and YPH500 were purchased from Agilent Technologies (included with pESC-URA plasmid as a kit). Plasmids were transformed into yeast cells using a lithium acetate method and into *E. coli* BL21(DE3) using chemical transformation.[38] Synthetic complete (SC) media[39] was prepared using yeast nitrogen base (Gibco) and synthetic complete medium supplement (Aldrich). The appropriate nutrient, either uracil, leucine or both, was omitted from the recipe for auxotrophic selection. SC liquid

production medium was made in a similar manner but containing 20 ml/L glycerol (2% final concentration), 20 ml/L ethanol (2% final concentration), and 40 g/L galactose (4% final concentration) with the appropriate nutrient dropped out. Yeast peptone (YP) production media was prepared using 15 g/L peptone, 15 g/L yeast extract, 2 g/L glucose (0.2% final concentration), 20 ml/L glycerol (2% final concentration), 20 ml/L ethanol (2% final concentration), and 40 g/L galactose (4% final concentration). Luria-Bertani (LB) medium was purchased from Fisher. Terrific Broth (TB) medium was purchased from Sigma and supplemented with 8 ml/L of glycerol (0.8% final). Kanamycin was added to LB and TB media at 50 mg/L final concentration.

Biochemical characterization of terpene cyclase genes. Terpene cyclase plasmids were transformed into *S. cerevisiae* strain YPH499 and then grown on SC (-leu) dropout agar plates for 3 days. Colonies (3 to 5 pooled) were seeded into 10 mL of SC (-leu) liquid medium and grown overnight at 30 °C. Overnight seed cultures were centrifuged then resuspended in YP liquid production medium and used to inoculate 50 ml total of YP liquid production medium (OD_{600} = 1.50 ± 0.06) in 250 mL baffled flasks with a vented metal cap. Cultures were grown at 22 °C for five days, then whole culture broths were extracted with 1:1 diethyl ether/hexane (50 ml) by shaking followed by centrifugation. The separated organic layer was collected and dried by shaking with a sodium sulfate and silica gel mixture. The volume was then reduced to about 3 ml by evaporating under a stream of air in a test tube and filtered with a PTFE syringe filter. A portion (2 ml) was evaporated, dissolved in hexane (0.2 ml), and analyzed by GCMS.

GCMS was performed on an Agilent 7890B GC equipped with a 5977B MSD and 7693 autosampler and an Agilent DP5 MS +DG column. The oven program consisted of a 1-minute period at 75 °C, followed by a linear increase to 250 °C after 11 minutes, then an increase to 300 °C after 1 minute, and finally a 1-minute hold at 300 °C (14 minutes total) with a helium carrier gas flow rate of 1 ml/min. A 6-minute solvent delay and 1:50 split were used.

Biochemical screening of cytochrome P450 genes. Cytochrome P450 plasmids were transformed into *S. cerevisiae* strain YPH500 and then grown on SC (-ura) dropout agar plates for 2 days. Colonies were seeded into 3 ml of SC (-ura) liquid medium and grown overnight at 30 °C. Overnight cultures (1 ml portions each) containing P450 plasmids were added to a 24-well plate and centrifuged, washed with SC (-leu, -ura) medium, centrifuged again and resuspended with 4 ml SC (-leu, -ura) medium. Yeast cells harboring the DgemTC plasmid were harvested from agar plates grown as described above by scraping up several colonies with an inoculating loop, then added to each well containing P450 harboring cell suspensions. Cell mixtures were incubated overnight at 30 °C, then 0.2 ml of each cell suspension was plated onto SC (-leu, -ura) dropout agar in 6-well dishes and grown for 2 days, after which colonies formed. Colonies were seeded into 5 mL of SC (-leu, -ura) liquid medium and grown overnight at 30 °C. Overnight seed cultures were used to inoculate SC (-leu, -ura) liquid production medium (25 ml) with HP20 resin added (1 g) in 125 ml flasks to an OD_{600} of about 1.0, then grown at 20 °C for three days. The whole

culture broths were filtered through cheese cloth loaded into a syringe barrel. The separated resin was eluted with acetonitrile (15 ml). Aliquots (1 ml) were centrifuged and then analyzed by LCMS method B.

LCMS method B: UPLC analysis was performed and high-resolution mass spectra (HR-MS) collected using a Waters Acquity UPLC linked to a Waters Xevo G2-XS Q-tof. A Waters BEH C_{18} column (2.1 x 50 mm, 1.7 µm) was used. The LC gradient consisted of a 1-minute isocratic period at 50 % acetonitrile in water at 0.6 mL/min followed by a linear increase to 75% acetonitrile after 5 minutes total followed by an increase to 100% acetonitrile after 6 minutes total then a 2-minute isocratic period at 100% acetonitrile (8 minutes total).

Scale up fermentation to produce 10, 11 and S1. Dgem cbTC, Ecar cbCYPb, and Selo cbCYPc showed relatively high levels of production and were used as examples for scale up fermentation. Yeast cultures harboring either TC plasmid alone or a hybrid strain harboring a TC and a CYP plasmid were prepared by inoculating colonies into SC dropout liquid media (-leu for TC strain and -leu, -ura for TC/CYP strains, 75 ml each) followed by growth overnight at 30 °C. The overnight cultures were split into 30 ml aliquots then seeded into a 2.8 L baffled Fernbach flask containing YP production media without added galactose (800 ml) then grown out overnight at 20 °C to an OD_{600} of about 1.0. Cultures were induced by addition of 200 mL of 20% galactose to a final concentration of 40 g/L, and then approximately 20 mL of HP20 was added as a slurry in ethanol. The culture was grown at 20 °C for 3 days, then harvested in the manner described above by collecting the resin and eluting with acetone (100 ml each). A second and third elution were performed using 50 ml each of acetone/hexane 1:1 and 50 ml each diethyl ether, respectively. The first and second elutions were pooled and evaporated, then combined with the third (ether) elution and 25 ml of water added. A separatory funnel was used to separate the top organic layer which was then evaporated on a rotary evaporator. The syrupy residue was extracted with hexane and the hexane layer collected with a glass pipette. Aliquots of these crude extracts were evaporated and checked by TLC and C (see below). Extracts in hexane were fractionated by flash chromatography using a Teledyne Isco CombiFlash system on a column of silica gel (24 g, 230-400 mesh grade 60) using hexane/isopropanol (0 to 20% isopropanol gradient over 12 minutes at 35 ml/min flow rate) with UV monitoring at 215 nm. Fractions were checked by TLC and evaporated to produce the following (in duplicate): Dgem cbTC strain, 4.2 ± 2.3 mg cembrene B (**10**). Dgem cbTC/Ecar cbCYPb strain 2.7 ± 0.1 mg cembrene B (**10**) and 11.1 ± 3.5 mg 19-hydroxy cembrene B (**11**). Dgem cbTC/Selo cbCYPc strain, 2.8 ± 0.1 mg cembrene B (**10**) and 22.1 ± 4.2 mg 7-hydroxy cembrene B (**S1**).

LCMS method C: UPLC analysis was performed and high-resolution mass spectra (HR-MS) collected using a Waters Acquity UPLC linked to a Waters Xevo G2-XS Q-tof. A Waters HSS T3 C_{18} column (2.1 x 100 mm, 1.8 µm) was used. The LC gradient consisted of a 1-minute isocratic period at 60 % acetonitrile in water at 0.4 mL/min flow rate followed by a linear increase to 100% acetonitrile after 8 minutes total then a 2.5-minute isocratic period at 100% acetonitrile (10.5 minutes total).

UV spectra were recorded using a Thermo Scientific NanoDrop in cuvette mode using quartz cuvettes. IR spectra were recorded on a Nicolet iS5 FT-IR spectrometer operating in transmittance mode using polyethylene cards (Thermo Scientific). NMR data for **1**, **2**, **3, 4, 5,** and **6** were collected and processed using the following instruments, software and parameters: ^{1}H NMR and ^{13}C NMR spectra were recorded on a Varian iNOVA 500 (^{1}H 500 MHz) NMR spectrometer equipped with a 3 mm Nalorac MDBG probe or Varian iNOVA 600 (^{1}H 600 MHz) NMR spectrometer equipped with a 5 mm Nalorac HCN probe operated using OpenVnmrJ. Data were processed and analyzed using MestreNova (v 14.3.1). Chemical shifts were referenced to the solvent residual proton for ^{1}H NMR (δ 7.26 for $CDCl_3$) and the ^{13}C signal for ^{13}C NMR (δ 77.2 for $CDCl_3$).

cembrene B (10): Clear oil; ^{1}H NMR (500 MHz, C_6D_6) δ 5.15 (m, 2H), 5.03 (t, *J* = 6.7 Hz, 1H), 2.89 (d, *J* = 6.9 Hz, 2H), 2.28 (t, *J* = 7.6 Hz, 2H), 2.16 (m, 2H), 2.12 (m, 2H), 2.10 – 2.00 (m, 6H), 1.69 (s, 3H), 1.66 (s, 3H), 1.60 (s, 3H), 1.55 (s, 3H), 1.52 (s, 3H). ^{13}C NMR (125 MHz, C_6D_6) δ 134.8, 133.7, 133.0, 131.7, 126.8, 125.4, 124.7, 124.7, 40.4, 39.3, 37.9, 32.0, 31.0, 25.4, 24.3, 20.9, 20.7, 16.0, 15.4, 15.1.; EIMS *m/z* 272.1.

19-hydroxy cembrene B (11): Clear oil; UV (MeCN) λ_{max} (log ε) 216 (3.3); IR (polyethylene film) ν_{max} 3337, 2962, 2922, 1436, 1383, 1001 cm^{-1}; NMR see **SI Tables 10 and 11**; HRESIMS *m/z* 289.2540 (calcd for $C_{20}H_{33}O^{+}$, 289.2531).

7-hydroxy cembrene B (S1): Clear oil; $[\alpha]_D^{20}$ -8.73° (*c* 0.275, MeOH); UV (MeOH) λ_{max} (log ε) 215 (3.2); IR (polyethylene film) ν_{max} 3334, 2916, 2852, 1462, 1383, 994 cm^{-1}; NMR see **SI Tables 10 and 11**; HRESIMS *m/z* 271.2426 (calcd for $C_{20}H_{31}^{+}$, 271.2426).

Cembrene 19-hydroxylase substrate selectivity assay. The cembrene A synthase Basb caTC was identified from a *B. asbestinum* sequencing dataset along with the cembrene B synthase Basb cbTC. The cembrene C synthase Stro ccTC was identified from a publicly available *Sarcophyton trocheliophorum* transcriptomic dataset. The cembrene 19-hydroxylase Ecar cbCYPb was co-expressed with Basb cbTC, Basb caTC and Stro ccTC in a manner as described above using SC liquid production media and HP20 resin. The crude extracts were analyzed by LCMS as described above using method A. Only co-expression of Ecar cbCYPb with Basb cbTC (cembrene B) showed production of oxidized products, while co-expression with Basb caTC and Stro ccTC (cembrene A and cembrene C, respectively) did not show any detectable oxidized products.

Biotransformation screen of cembrene alcohols 11 and S1 using yeast harboring CYP genes. YPH 500 yeast cells harboring either cbCYPb, cbCYPc or empty pESC-URA were seeded into SC (-Ura) liquid medium containing dextrose and grown overnight at 30 °C. An aliquot (0.5 ml) of each culture was seeded into SC (-Ura) liquid medium containing dextrose (10 ml). Cells were grown 4 hours at 30 °C then centrifuged and resuspended in SC (-Ura) liquid medium containing galactose. To each culture was added bovine serum albumin (BSA) at 0.7 mg/ml followed by either **11** (to cbCYPc and empty vector) or **S1** (to cbCYPb and empty vector) as a solution in ethanol to a final concentration of 100 µg/ml.

Cultures were grown at 20 °C and 180 RPM in shake flasks for 3 days, then extracted with diethyl ether. The ether was evaporated, and the residue dissolved in methanol and analyzed by LCMS using method B.

Scaled-up biotransformation to produce diol 12. YPH 500 yeast cells harboring either cbCYPb or cbCYPc were seeded into SC (-Ura) liquid medium containing dextrose (10 ml) and grown overnight at 30 °C. The seed cultures were added to SC (-Ura) liquid medium containing galactose (250 ml) followed by albumin (200 µg/ml) then **11** (22 mg in ethanol) to cbCYPc and **S1** (44 mg in ethanol) to cbCYPb. The cultures were shaken at 20 °C and 180 RPM for 4 days, then each was extracted with ethyl acetate (100 ml). The organic layer was evaporated and the residue purified by combiflash using a C_{18} column and water:acetonitrile. The fractions were evaporated to produce **12** (11.3 mg from **11**/cbCYPc and 1.6 mg from **S1**/cbCYPb).

7,19-dihydroxy cembrene B (12): Colorless crystals; $[\alpha]_D^{20}$ +3.54° (*c* 0.226, MeOH); UV (MeOH) λ_{max} (log ε) 215 (3.1); IR (polyethylene film) ν_{max} 3357, 2919, 2852, 1711, 1462, 1383, 994 cm^{-1}; NMR see **SI Tables 10 and 11**; HRESIMS *m/z* 287.2381 (calcd for $C_{20}H_{31}O^+$, 287.2370).

Biochemical characterization of SDH genes in vitro. Plasmids bearing SDH genes were transformed into *E. coli* BL21(DE3) strain and grown on LB kanamycin plates overnight at 37 °C. Colonies were inoculated into LB liquid medium (20 mL) with 50 mg/L kanamycin and 0.2% dextrose and grown overnight at 30 °C. Cultures were used to inoculate TB liquid medium (1 L) with 50 mg/L kanamycin and grown at 28 °C for 5 hours resulting in varying OD_{600} values between 0.5 and 1.5. Cultures were cooled to 16 °C and induced with 0.5 mM IPTG, then an additional 50 mg/L kanamycin was added, and the cells were grown overnight at 16 °C. Some strains grew noticeably slower, including those harboring *E. caribaeorum* cbSDH, *R. koellikeri* cbSDH and *S. elongata* cbSDH, correlating with higher relative SDH activity. These strains were grown for an additional 24 hours to improve protein yields. Cells were then harvested by centrifugation and stored at -80 °C until use. Cell pellets were suspended in 25 ml of lysis buffer (HEPES 25 mM pH 8.2, NaCl 100 mM, glycerol 5 % v/v, TCEP 100 µM, lysozyme 0.6 mg/ml and DNase 0.05 mg/ml) and incubated at 4 °C for 1 hour. The cells were disrupted using sonication then centrifuged at 52,000 x *g* for 20 minutes at 4 °C. The supernatants were transferred to 50 ml conical tubes and about 0.25 ml of Ni-NTA resin added, then the mixture was incubated on ice for 1 hour with periodic mixing. The suspensions were centrifuged at 5000 x *g* for 5 minutes and the supernatant pipetted off. The pellets of Ni-NTA resin were suspended in 2 ml wash buffer (HEPES 25 mM pH 8.2, NaCl 100 mM, glycerol 5 % v/v, TCEP 100 µM, and imidazole 25 Mm) and transferred to 2 ml tubes then centrifuged at 13,000 x *g*. This wash step was repeated once, then the Ni-NTA resin was eluted by adding 2 ml of elution buffer (HEPES 25 mM pH 8.2, NaCl 100 mM, glycerol 5 % v/v, TCEP 100 µM, and imidazole 250 Mm). After centrifugation, the supernatant was pipetted into a separate tube and the elution repeated. Using a 2 ml centrifugal concentrator with 3 kDa MW cutoff membrane, the eluents were pooled for each sample then concentrated and buffer exchanged twice in

assay buffer (HEPES 25 mM pH 8.2, NaCl 100 mM, glycerol 5 % v/v, TCEP 100 µM). Assays were performed by dissolving NADP (200 µM) and BSA (1 mg/ml) in assay buffer, then adding 100 µg of cembrene alcohol (**11**, **12** or **S1**) as a 10 µL aliquot in ethanol followed by addition of the SDH enzyme to a final volume of 1 ml. The reaction was shaken at 200 RPM and 20 °C for 16 hours then extracted with diethyl ether (3 x 0.5 ml) with vortexing followed by centrifugation at 13,000 x *g* for 3 minutes. The pooled organic layers were evaporated under a stream of air and the organic residue dissolved in 2 ml of methanol. Enzyme reactions were analyzed by LCMS method C.

Semisynthesis of S2. To a solution of **11** (25 mg, 0.086 mmol, 1 eq) in dichloromethane (2 ml) was added Dess-Martin periodinane (183.2 mg, 0.432 mmol, 5 eq) and the suspension stirred at room temperature overnight. The reaction was quenched by addition of 200 µl of isopropanol followed by addition of saturated sodium chloride solution. The dichloromethane layer was separated and dried on a rotary evaporator then purified by silica gel chromatography to produce **S2** (20 mg, 0.070 mmol, 81% yield).

Aldehyde S2: Clear oil; UV (MeCN) λ_{max} (log ε) 215 (3.8), 252 (3.7); IR (polyethylene film) ν_{max} 2964, 2918, 2852, 1668, 1261, 1020 cm^{-1}; NMR see **SI Tables 10 and 11**; HRESIMS *m/z* 269.2264 (calcd for $C_{20}H_{29}^{+}$, 269.2264).

Preparative scale in vitro reaction with SDH and diol 12. *E. coli* strain BL21(DE2) harboring *S. elongata* cbSDH on pET-28b(+) was grown in 6 separate 1 L cultures using TB medium, as described above. A cell pellet (100 ml in volume) was suspended in HEPES lysis buffer (50 mM HEPES, 150 mM NaCl, 10 mM imidazole, pH 8.0) and treated with lysozyme (0.6 mg/ml) and DNase (0.1 mg/ml). No detergents or reducing agents were added, as these were found to inhibit the reaction in previous experiments. The suspension was sonicated then centrifuged at 52,000 x *g* for 20 minutes. The clarified lysate was incubated over 20 ml of Ni-NTA in a beaker with stirring then collected in a glass column. The resin bed was washed with HEPES lysis buffer (50 ml), HEPES wash buffer (50 ml [50 mM HEPES, 150 mM NaCl, 50 mM imidazole, pH 8.0]) then eluted with HEPES elution buffer (5x15 ml [50 mM HEPES, 150 mM NaCl, 250 mM imidazole, pH 8.0]). Fractions 2, 3, 4, and 5 were diluted up to 200 ml with HEPES reaction buffer (50 mM HEPES, 150 mM NaCl, pH 8.0) in an Erlenmeyer flask with gentle magnetic stirring at room temperature, then was added BSA (0.1 mg/ml), NADP (2 mM), and **6** (15 mg in 1000 µl DMSO). The reaction mixture was stirred for two days. The aqueous reaction was extracted with 2x100 ml of dichloromethane, using centrifugation to break the emulsion, and the organic extract dried over sodium sulfate and evaporated. The residue was filtered through a short bed of silica gel using hexane:ethyl acetate 1:1 and then purified by preparative HPLC using a phenylhexyl column and water:acetonitrile as mobile phase to provide lactone **13** (1.5 mg).

Cembrene B lactone 13: Clear oil; $[\alpha]_D^{20}$ +42.4° (*c* 0.250, MeOH); UV (MeOH) λ_{max} (log ε) 230 (4.1), 260 (shoulder); IR (polyethylene film) ν_{max} 2919, 2850, 1757, 1710, 1461, 1385,

1289, 1089, 1029 cm^{-1}; NMR see **SI Tables 10 and 11**; HRESIMS *m/z* 301.2183 (calcd for $C_{20}H_{29}O_2^+$, 301.2163).

Supplemental

Table S1 CMNPD results for Scleralcyonacea coral.

Current Family Name	Terpenoids reported	Briaranes reported	Producer organism	Example briarane
Sarcodictyonidae	yes	no	N/A	N/A
Spongiodermidae	no	no	N/A	N/A
Ideogorgiidae	no	no	N/A	N/A
Parisididae	no	no	N/A	N/A
Parasphaerascleridae	yes	no	N/A	N/A
Chelidonisididae	no	no	N/A	N/A
Chrysogorgiidae	no	no	N/A	N/A
Primnoidae	yes	no	N/A	N/A
Keratoisididae	yes	no	N/A	N/A
Isidoidae	no	no	N/A	N/A
Mopseidae	no	no	N/A	N/A
Pleurogorgiidae	no	no	N/A	N/A
Ifalukellidae	yes	no	N/A	N/A
Helioporidae	yes	no	N/A	N/A
Cornulariidae	no	no	N/A	N/A
Pennatuloidea	yes	yes	*Stylatula elongata*	stylatulide[40]
Briareidae	yes	yes	*Briareum asbestinum*	briareolate ester J [41]
Erythropodiidae	yes	yes	*Erythropodium caribeorum*	eythrolide D[42]
Coralliidae	yes	yes	*Minabea sp.*	minabein 1[43]
Ellisellidae	yes	yes	*Dichotella gemmacea*	gemmacolide R[44]

Table S2 ^{1}H and ^{13}C-NMR data for 11-hydroxyptilosarcenone (**6**) in $CDCl_3$.

Position	Reported[45,46]		This study	
	δ_C	δ_H	δ_C [e, f]	δ_H
1	43.1[a]	-	29.6	-
2	76.2	5.89 (d, 9.0)	76.2	5.93 (d, 7.2)
3	130.3 [b]	5.54 (dd, 11.7, 9.0)	127.4	5.57 (dd, 11.9, 10.2)
4	127.3 [b]	5.98 (d, 11.7)	130.5	6.00 (d, 11.9)
5	136.5	-	[g]	-
6	60.9	4.96 (br d, 3.2) [d]	60.9	5.05 (br d)

7	69.2 [c]	5.02 (br s) [d]	78.8	4.98 (d, 3.7)
8	81.6	-	81.9	-
9	78.9 [c]	5.77 (d, 5.6)	69.4	5.81 (d, 6.1)
10	30.9[a]	2.92 (d, 5.6)	43.2	2.95 (d, 6.1)
11	76.1	-	76.2	-
12	199.2	-	199.5	-
13	121.8	6.08 (d, 10.5)	121.7	6.13 (d, 10.8)
14	156.7	6.70 (d, 10.5)	156.7	6.73 (d, 10.8)
15	13.7	1.23 (s)	13.8	1.22 (s)
16	117.3	6.00 (br s)	117.5	6.03 (br s)
	-	5.88 (br s)	-	5.91 (br s)
17	45.8	2.35 (q, 7.1)	45.8	2.39 (q, 7.4)
18	6.7	1.21 (d, 7.1)	6.8	1.31 (d, 7.6)
19	174.8	-	175.0	-
20	25.2	1.46 (s)	25.2	1.49 (s)
CH_3COO-	170.0, 169.3	-	169.5/169.1	-
CH_3COO-	21.9, 21.1	2.20, 2.12	21.8, 21.1	2.24, 2.16
-OH	-	4.88 (s)	-	[g]

[a] Interchanged. [b] Interchanged. [c] Interchanged. [d] Interchanged. [e] Multiplicity determined by HSQC. [f] Detected by HSQC or HMBC. [g] Not detected.

Table S3 Samples used for molecular networking analyses.

Organism	Sample Type	Exact mass of network briarane	Corresponding literature exact mass
Stylatula elongata	crude extract	NA	NA
Briareum asbestinum	encrusting morphology crude extract	NA	NA
	branching morphology crude extract	629.2659	NA
Erythropodium caribeorum	fraction	554.1547 (erythrolide D)	554.1555[47]
Renilla koellikeri	crude extract	496.1449 (11-hydroxyptilosarcenone)	496.1500[45]
	pure compound	496.1449 (11-hydroxyptilosarcenone)	496.1500[45]
Dichotella gemmacea	crude extract	696.2207 (gemmacolide R)	696.2629[44]
Corallium rubrum	crude extract	NA	NA

Table S4 Basic assembly statistics for octocoral genomes.

Organism	Source	Accession	Genome Size (Mb)	Busco C (%) metazoan DB	contig N50 (kb)
R. koellikeri	This paper	SAMN40621396	144	86.9	4,895.0
B. asbestinum	This paper	SAMN40621398	1,155	87.3	331.8
D. gemmacea	This paper	SAMN40621399	401	87.5	195.9
S. elongata	This paper	SAMN40621397	269	88.0	5,139.0
C. rubrum	doi: 10.1101/2024.07.13.603384	PRJEB54017	475	88.5	1,637.0
E. caribaeorum	This paper	SAMN41659149	300	86.4	233.0
R. muelleri	doi: 10.1093/gigascience/giz026	http://rmue.reefgenomics.org	172	84.5	70.5
P. caledonicum	NA	SAMN14247561	292	72.4	19.2
H. coerulea	doi: 10.1038/s41597-023-02291-z	SAMN33360050	429	88.9	1,428.0
D. gigantea	DOI: 10.1093/gbe/evz043	SAMN10516674	286	88.5	1,446.0
Xenia sp.	DOI: 10.1038/s41586-020-2385-7	SAMN23741995	227	92.0	1,100.0
P. clavata	DOI: 10.1534/g3.120.401371	SAMEA6085997	607	72.4	19.2
Trachythela sp.	DOI: 10.1093/gbe/evaa265	SAMN16072316	578	88.5	3,564.0
R. reniformis	NA	SAMEA103980763	132	39.8	2.0
P. stokvisi	NA	SAMN16093713	689	81.6	1.4
P. marumi	NA	SAMN16114184	583	67.9	1.3

Table S5 Sequencing Accession numbers.

	Raw illumina RNA reads	Raw ONT reads	Raw PacBio reads	Raw illumina gDNA reads

	Source	Accession	Source	Accession	Source	Accession	Source	Accession
R.koellikeri	This paper	SRR28507355	NA	NA	This paper	SRR28507351	NA	NA
B.asbestinum	This paper	SRR28507352	This paper	SRR28507348	NA	NA	This paper	SRR28507346
D.gemmacea	This paper	SRR28507353	This paper	SRR28507349	NA	NA	This paper	SRR28507347
E.caribaeorum	doi: 10.1038/s41589-023-01352-z	SRR15783032	This paper	SRR29682103	NA	NA	doi: 10.1038/s41589-023-01352-z	ERR4660993
S.elongata	This paper	SRR28507354	NA	NA	This paper	SRR28507350	NA	NA

Table S6 Long and short read sequencing statistics across all in house genomes.

	B. asbestinum	***D. gemmacea***	***R. koellikeri***	***S. elongata***	***E. caribeorum***
Short Read Genomic Sequencing					
Total Sequences (M)	114	147	150	207	56
Total Bases (Gb)	17	22	23	31	5
Sequences flagged as poor quality	0	0	0	0	0
Sequence length	151	151	151	151	100
%GC	38	39	39	40	42
Kmer Based Genome size (Mb)	1,155	401	144	269	-
Genome Coverage	14	41	185	141	18
Long Read Genomic Sequencing					
Raw bases from WGS-Nanopore (Gb)	59	78	NA	NA	17
Raw bases from WGS- Pacbio (Gb)	NA	NA	29	26	NA
Mean Read length (kb)	2	4	12	13	2
%Reads Q>20	NA	NA	85	78	NA
Transcriptomic Short Read Sequencing					
Total Sequences (M)	184	618	178	226	NA
Coverage (x)	47	250	268	165	NA
HiSat Alignment	40.36%	92.43%	90.51%	60.42%	NA

Table S7 Genbank Accessions for annotated BGC contigs.

Organism	Accession
B. asbestinum	PP584625.1, PP584626
D. gemmacea	PP584627
S. elongata	PP584628, PP584629
R. koellikeri	PP584630
C. rubrum	PP584631
E. caribaeorum	PP584632, PP584633

Table S8 All CYPs used in phylogenetic analysis.

Organism	# Annotated CYPs	Reason Included	CYP Sequence Source
Briareum asbestinum	65	Briarane Producer	This paper
Corallium rubrum	40	Briarane Producer	This paper
Dichotella gemmacae	42	Briarane Producer	This paper
Erythropodium caribaeorum	41	Briarane Producer	This paper
Renilla koellikeri	31	Briarane Producer	This paper
Stylatula elongata	40	Briarane Producer	This paper
Heliopora coerulea	40	Other octocoral	This paper
Xenia sp.	19	Other octocoral	This paper
Hydra vulgaris	24	Model Cnidarian	Pankov, K. V. et al.[24]
Acropora digitifera	24	Model Cnidarian	Pankov, K. V. et al.[24]
Aurelia aurita	37	Model Cnidarian	Pankov, K. V. et al.[24]
Nematostella vectensis	70	Model Cnidarian	Pankov, K. V. et al.[24]

Table S9 Recombinant enzymes biochemically characterized in this study.

Entry	Enzyme	Organism	Classification	Function
1	Basb_cbTC	*Briareum asbestinum*	terpene cyclase	cembrene B synthase
2	Rkoe_cbTC	*Renilla koellikeri*	terpene cyclase	cembrene B synthase
3	Selo_cbTC	*Stylatula elongata*	terpene cyclase	cembrene B synthase
4	Dgem_cbTC	*Dichotella gemmacea*	terpene cyclase	cembrene B synthase
5	Crub_cbTC	*Corallium rubrum*	terpene cyclase	cembrene B synthase
6	Dgem_cbCYPa	*Dichotella gemmacea*	clan 46 cytochrome P450	unknown
7	Dgem_cbCYPb	*Dichotella gemmacea*	octocoral specific cytochrome P450	cembrene B 19-hydroxylase
8	Dgem_cbCYPc	*Dichotella gemmacea*	octocoral specific cytochrome P450	cembrene B 7-hydroxylase
9	Selo_cbCYPc	*Stylatula elongata*	octocoral specific cytochrome P450	cembrene B 7-hydroxylase
10	Rkoe_cbCYPa	*Renilla koellikeri*	clan 46 cytochrome P450	unknown
11	Rkoe_cbCYPb	*Renilla koellikeri*	octocoral specific cytochrome P450	cembrene B 19-hydroxylase
12	Rkoe_cbCYPc	*Renilla koellikeri*	octocoral specific cytochrome P450	cembrene B 7-hydroxylase
13	Crub_cbCYPa	*Corallium rubrum*	clan 46 cytochrome P450	unknown
14	Crub_cbCYPb	*Corallium rubrum*	octocoral specific cytochrome P450	cembrene B 19-hydroxylase
15	Ecar_cbCYPa	*Erythropodium caribaeorum*	clan 46 cytochrome P450	unknown
16	Ecar_cbCYPb	*Erythropodium caribaeorum*	octocoral specific cytochrome P450	cembrene B 19-hydroxylase
17	Basb_cbCYPa	*Briareum asbestinum*	clan 46 cytochrome P450	unknown
18	Basb_cbCYPb	*Briareum asbestinum*	octocoral specific cytochrome P450	cembrene B 19-hydroxylase
19	Basb_cbSDH1	*Briareum asbestinum*	short chain dehydrogenase	unknown
20	Basb_cbSDH2	*Briareum asbestinum*	short chain dehydrogenase	cembrene B alcohol dehydrogenase
21	Dgem_cbSDH1	*Dichotella gemmacea*	short chain dehydrogenase	cembrene B alcohol dehydrogenase
22	Ecar_cbSDH1	*Erythropodium caribaeorum*	short chain dehydrogenase	cembrene B alcohol dehydrogenase
23	Rkoe_cbSDH1	*Renilla koellikeri*	short chain dehydrogenase	cembrene B alcohol dehydrogenase
24	Rkoe_cbSDH2	*Renilla koellikeri*	short chain dehydrogenase	unknown
25	Selo_cbSDH1	*Stylatula elongata*	short chain dehydrogenase	cembrene B alcohol dehydrogenase
26	Basb_caTC	*Briareum asbestinum*	Terpene synthase	Cembrene A synthase
27	Stro_ccTC	*Sarcophyton trocheliophorum*	Terpene synthase	Cembrene C synthase

Table S10. ^{1}H-NMR data for **11** – **13**, **S1** and **S2** measured in $CDCl_3$ at 500 MHz.

Position	**11**	**12**	**13**	**S1**	**S2**
2	2.05 (m)	2.05 (m) [a]	1.98 (m) 2.15 (m)	2.04 (m)	2.02 (m)
3	2.08 (m)	2.07 (m) [a]	2.04 (m) 2.17 (m)	2.08 (m)	2.06 (t, 6.5)
4	5.03 (t, 7.4)	5.07 (t, 5.1)	5.08 (t, 6.4)	4.98 (m) [a]	5.05 (t, 7.4)
6	2.00 (t, 7.9)	2.37 (t, 13.2) 2.23 (m) [a]	2.21 (dd, 14.6, 7.0) 2.38 (dd, 14.6, 4.4)	2.26 (dd, 10.6, 1.0) 2.34 (dd, 10.8, 10.7)	2.10 (t, 8.8)
7	2.28 (t, 7.9)	4.86 (dd, 11.0, 4.4)	4.75 (m)	4.87 (dd, 10.9, 3.8)	2.73 (t, 7.9)
9	2.84 (d, 7.2)	2.98 (dd, 16.1, 5.1) 2.75 (dd, 16.1, 6.6)	3.06 (d, 6.8)	2.73 (dd, 16.8, 8.1) 2.87 (dd, 16.8, 8.1)	3.07 (d, 7.0)
10	5.05 (t, 7.4)	4.97 (m) [a]	5.00 (t, 6.4)	4.98 (m) [a]	5.08 (t, 7.4)
12	2.16 (m)	2.09 (m) [a] 2.15 (m) [a]	2.14 (m)	2.08 (m)	2.18 (m)
13	2.20 (m)	2.25 (m) [a] 2.17 (m) [a]	2.27 (m) 2.10 (m)	2.13 (m) 2.21 (m)	2.19 (m)
14	4.94 (t, 6.6)	5.00 (m) [a]	4.86 (t, 6.4)	4.98 (m)	4.91 (t, 6.6)
15	1.56 (s)	1.53 (s)	1.53 (s)	1.52 (s)	1.54 (s)
16	1.60 (s)	1.64 (s)	1.72 (s)	1.62 (s)	1.60 (s)
18	1.76 (s)	1.79 (s)	1.80 (s)	1.70 (s)	1.74 (s)
19	4.09 (s)	4.16 (d, 11.7) 4.01 (d, 11.7)	-	1.66 (s)	10.08 (s)
20	1.60 (s)	1.68 (s)	1.61 (s)	1.66 (s)	1.62 (s)

[a] Overlapping signal.

Table S11. ^{13}C-NMR data for **11** – **13**, **S1** and **S2** measured in $CDCl_3$ at 125 MHz.

Position	**11**	**12**	**13**	**S1**	**S2**
1	133.2 (C)	133.0 (C)	133.9 (C)	133.2 (C)	133.2 (C)
2	40.0 (CH_2)	39.0 (CH_2)	39.5 (CH_2)	39.2 (CH_2)	39.7 (CH_2)
3	23.8 (CH_2)	23.9 (CH_2)	23.9 (CH_2)	24.3 (CH_2)	23.7 (CH_2)
4	125.1 (CH)	127.8 (CH)	127.5 (CH)	127.1 (CH)	125.5 (CH)
5	134.4 (C)	130.2 (C)	131.3 (C)	130.3 (C)	133.4 (C)
6	38.3 (CH_2)	44.8 (CH_2)	42.2 (CH_2)	44.8 (CH_2)	39.0 (CH_2)
7	31.4 (CH_2)	69.1 (CH)	83.9 (CH)	69.5 (CH)	29.7 (CH_2)
8	136.7 (C)	137.9 (C)	163.1 (C)	130.3 (C)	161.9 (C)
9	31.3 (CH_2)	25.9 (CH_2)	25.7 (CH_2)	25.4 (CH_2)	33.6 (CH_2)
10	123.7 (CH)	123.3 (CH)	119.1 (CH)	124.6 (CH)	120.6 (CH)
11	134.5 (C)	133.4 (C)	137.5 (C)	132.3 (C)	136.9 (C)
12	39.1 (CH_2)	38.1 (CH_2)	38.7 (CH_2)	38.1 (CH_2)	39.1 (CH_2)
13	25.0 (CH_2)	24.8 (CH_2)	24.8 (CH_2)	25.0 (CH_2)	24.8 (CH_2)
14	126.7 (CH)	127.3 (CH)	125.6 (CH)	127.3 (CH)	126.1 (CH)
15	15.0 (CH_3)	14.9 (CH_3)	15.1 (CH_3)	15.1 (CH_3)	14.9 (CH_3)
16	15.7(CH_3)	16.7 (CH_3)	16.5 (CH_3)	16.9 (CH_3)	15.8 (CH_3)
17	129.3 (C)	133.8 (C)	123.1 (C)	133.2 (C)	133.0 (C)
18	16.9(CH_3)	18.0 (CH_3)	8.6 (CH_3)	20.3 (CH_3)	10.8 (CH_3)
19	64.2 (CH_2)	62.9 (CH_2)	174.7 (C)	21.6 (CH_3)	191.1 (CH)
20	16.0 (CH_3)	17.0 (CH_3)	15.6(CH_3)	17.2 (CH_3)	15.6 (CH_3)

[c] Overlapping signal.

Table S12 X-ray crystallographic data for diol **12.**

Formula	**C20H32O2**
*Dcalc.***(**g cm-3**)**	1.068
μ (mm-1**)**	0.513
Formula Weight	304.45
Colour	colourless
Shape	needle
Size **(**mm3**)**	0.15×0.08×0.03
T (K)	100.00(10)
Crystal System	orthorhombic
Flack Parameter	0.00(9)
Hooft Parameter	0.01(6)
Space Group	*P212121*
a (Å)	7.36900(10)
b (Å)	14.6718(2)
c (Å)	17.5054(3)
α (°)	90
β (°)	90
γ (°)	90
V **(**Å3**)**	1892.62(5)
Z	4
Z'	1
Wavelength **(**Å**)**	1.54184
Radiation type	Cu Kα
Θmin (°)	3.931
Θmax (°)	74.501
Measured Refl.	27879
Independent Refl.	3857
Reflections with I > 2(I)	3705
Rint	0.0467
Parameters	203
Restraints	0
Largest Peak	0.149
Deepest Hole	-0.135
GooF	1.078
wR2 (all data)	0.0893
wR2	0.0885
R1 (all data)	0.0350
R1	0.0336

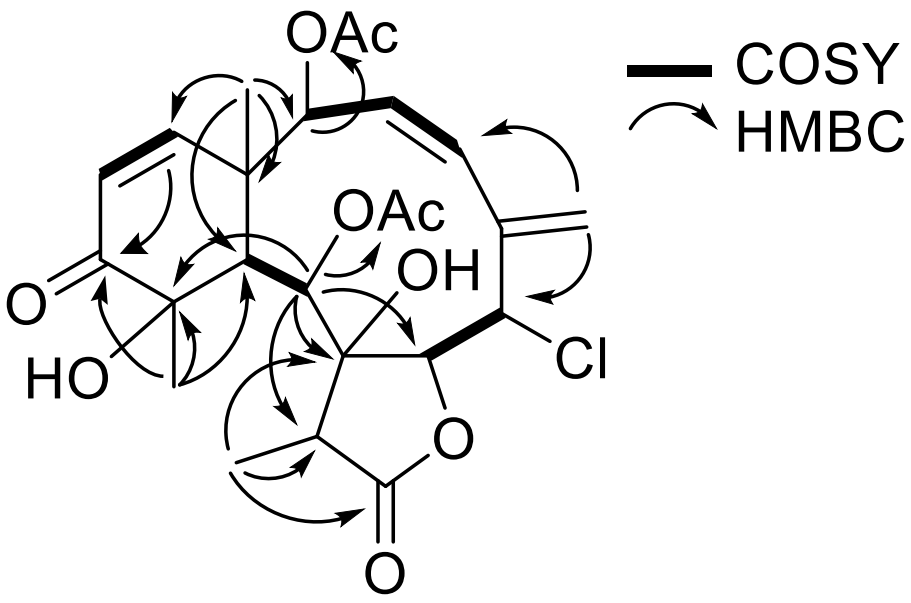


Figure S1 2D NMR correlations of 11-hydroxyptilosarcenone (6).

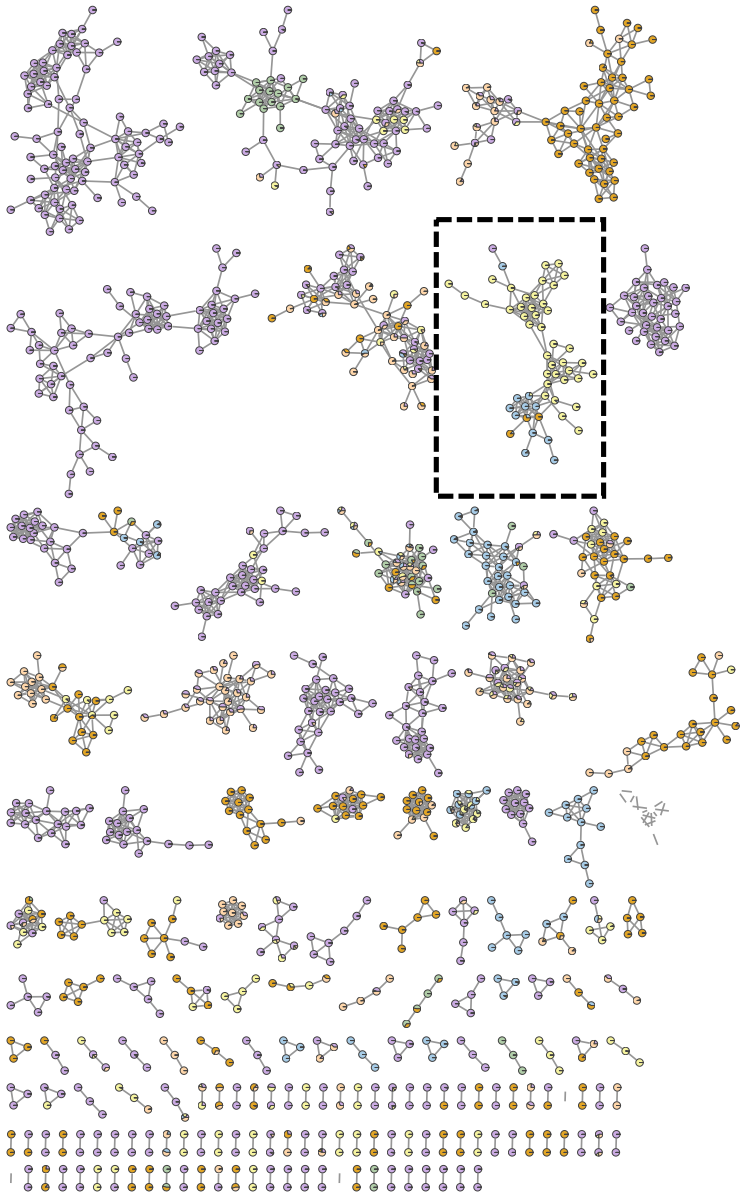


Figure S2 Full molecular network. Box indicates briarane diterpene network shown in **Figure 2c**.

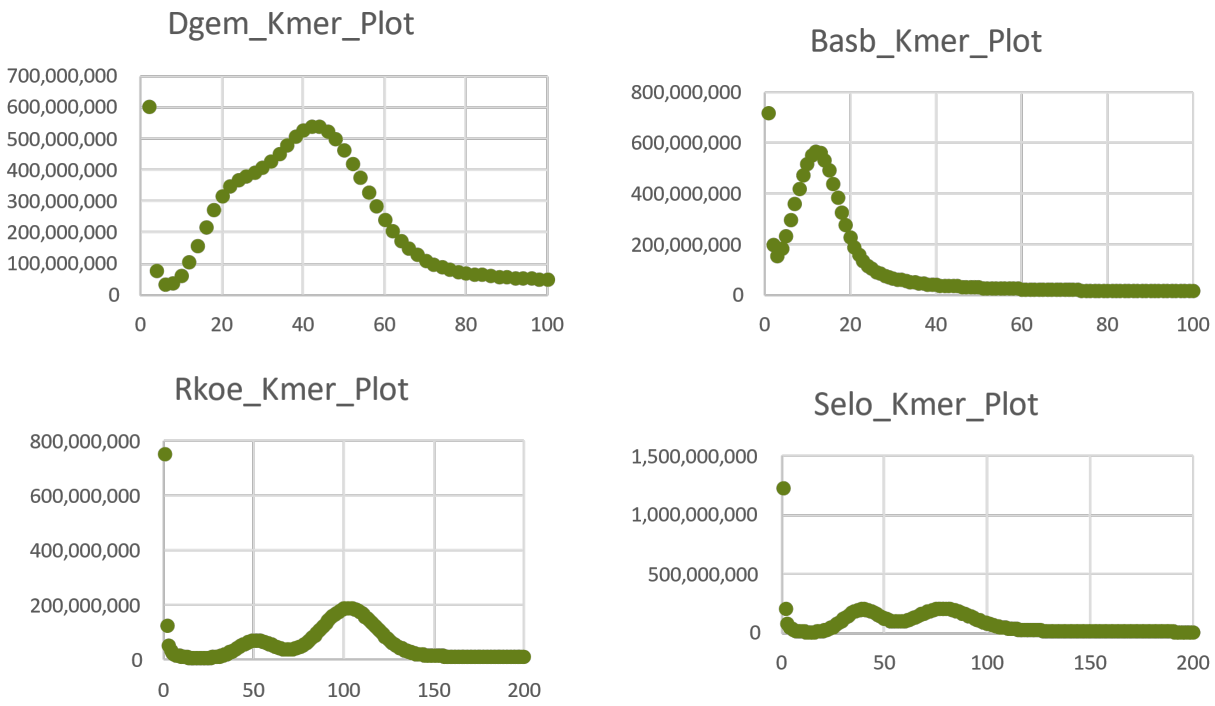


Figure S3 Kmer plots for genome size predictions.

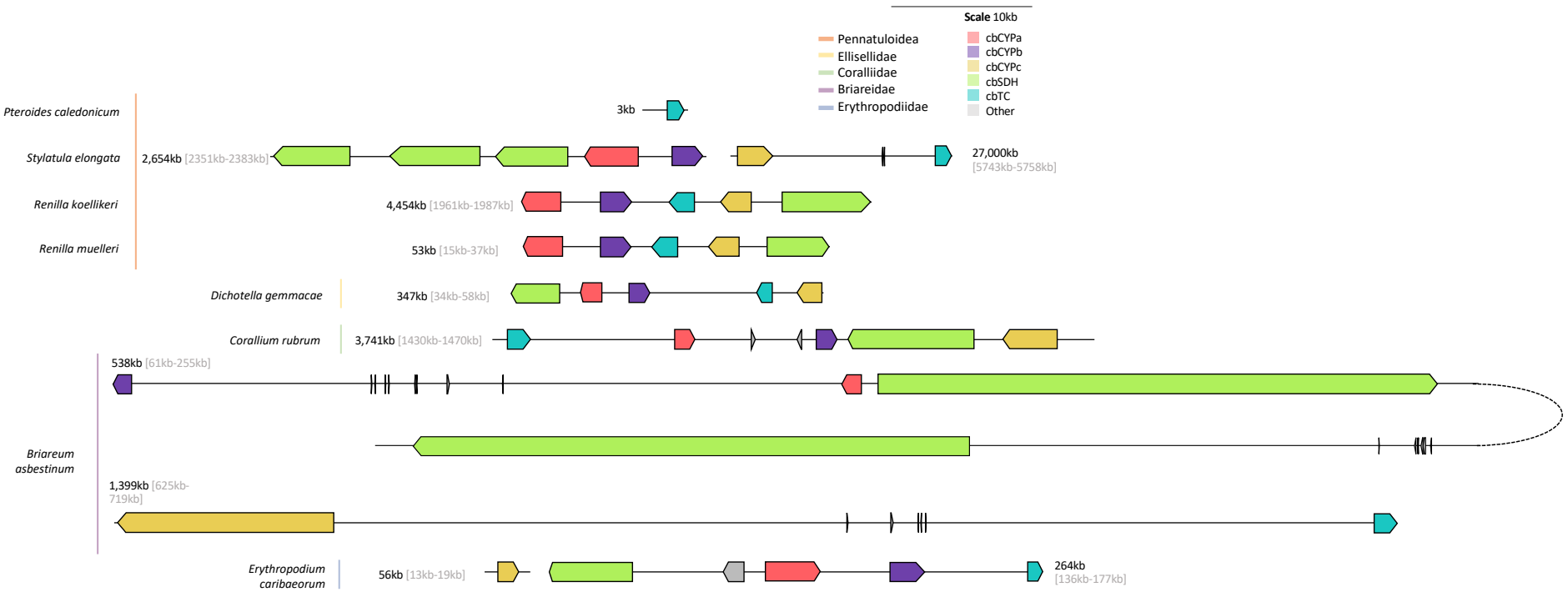


Figure S4 Full to scale synteny.

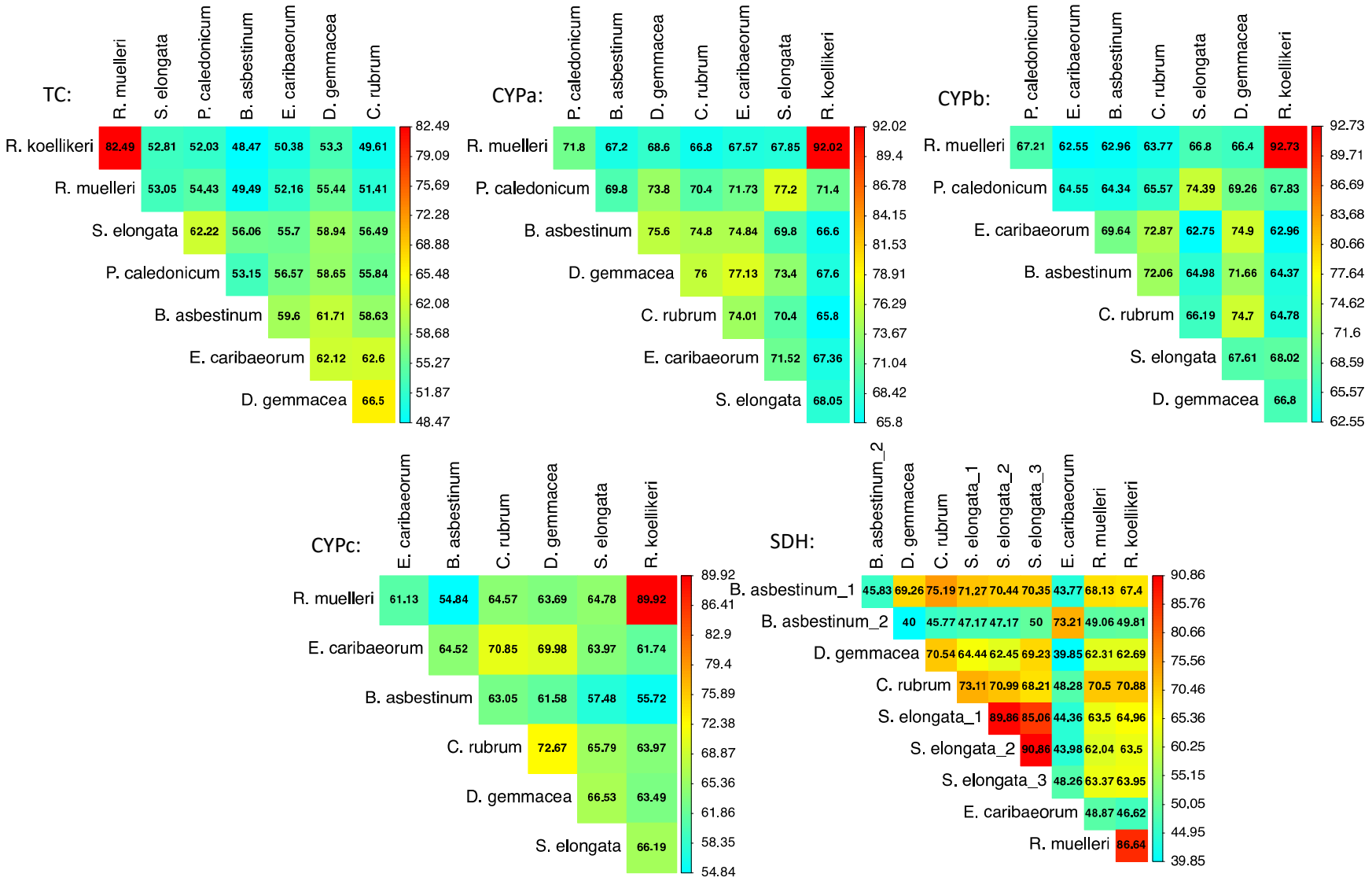


Figure S5 %ID matrices for all clinker genes.

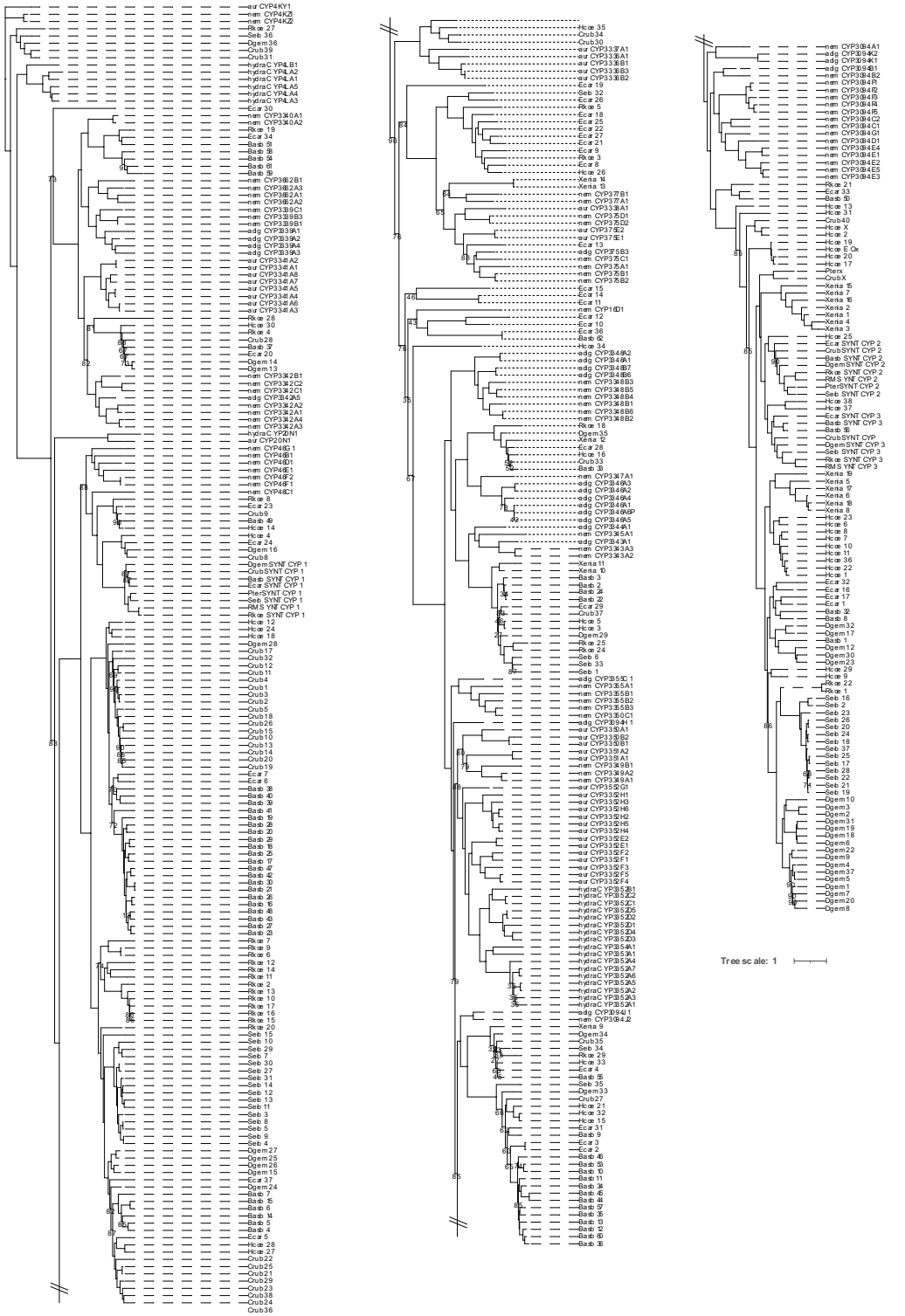


Figure S6 Detailed CYP phylogeny. Bootstraps lower than 90 are labeled.

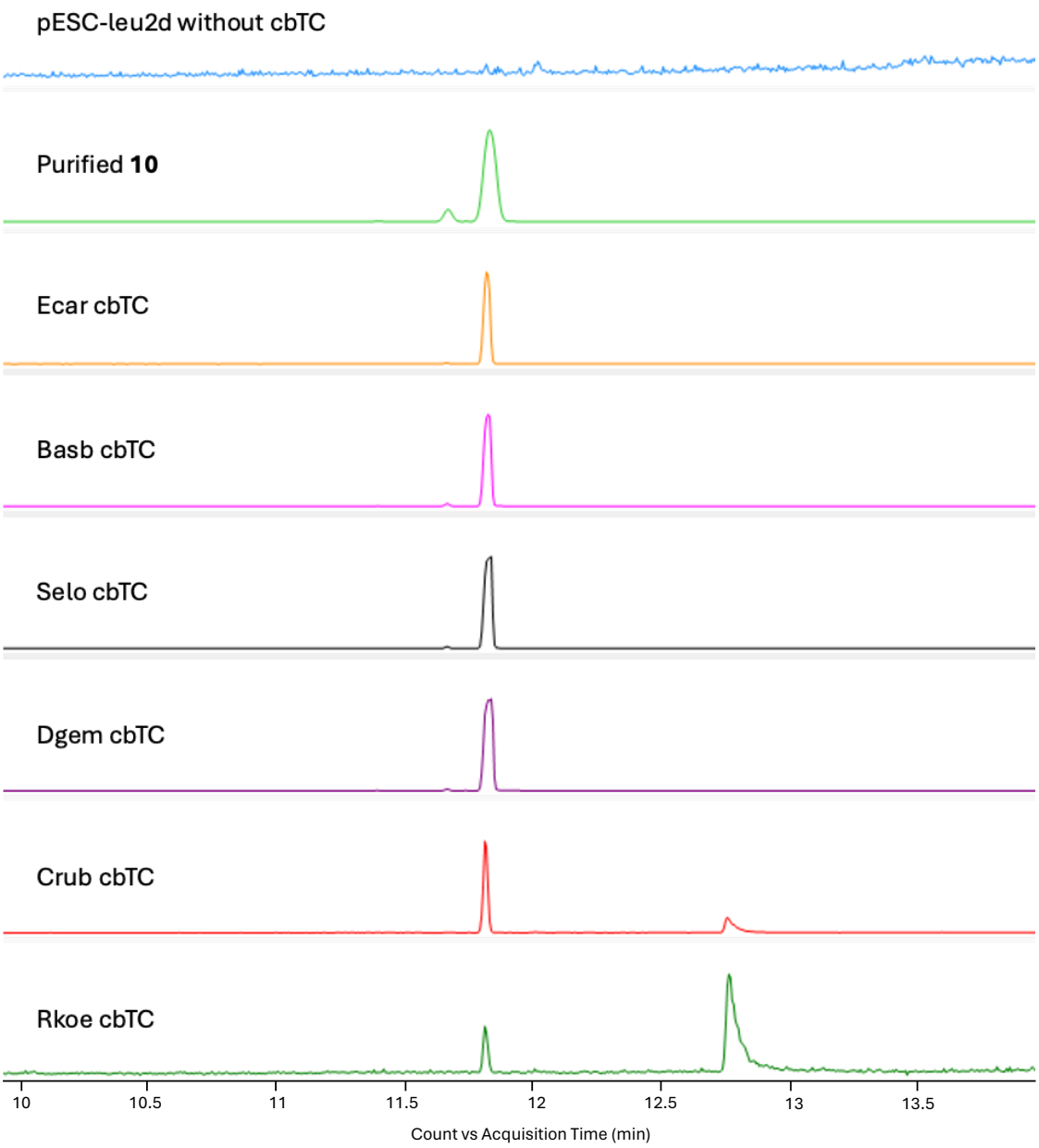


Figure S7 GCMS EIC traces at 272.2 m/z for crude extracts of yeast harboring a cbTC or pESC-leu2d without cbTC. A trace of purified 10 is shown for comparison. The peak for 10 is present at 11.8 min. The peak at 12.8 min in Crub and Rkoe samples is GGOH.

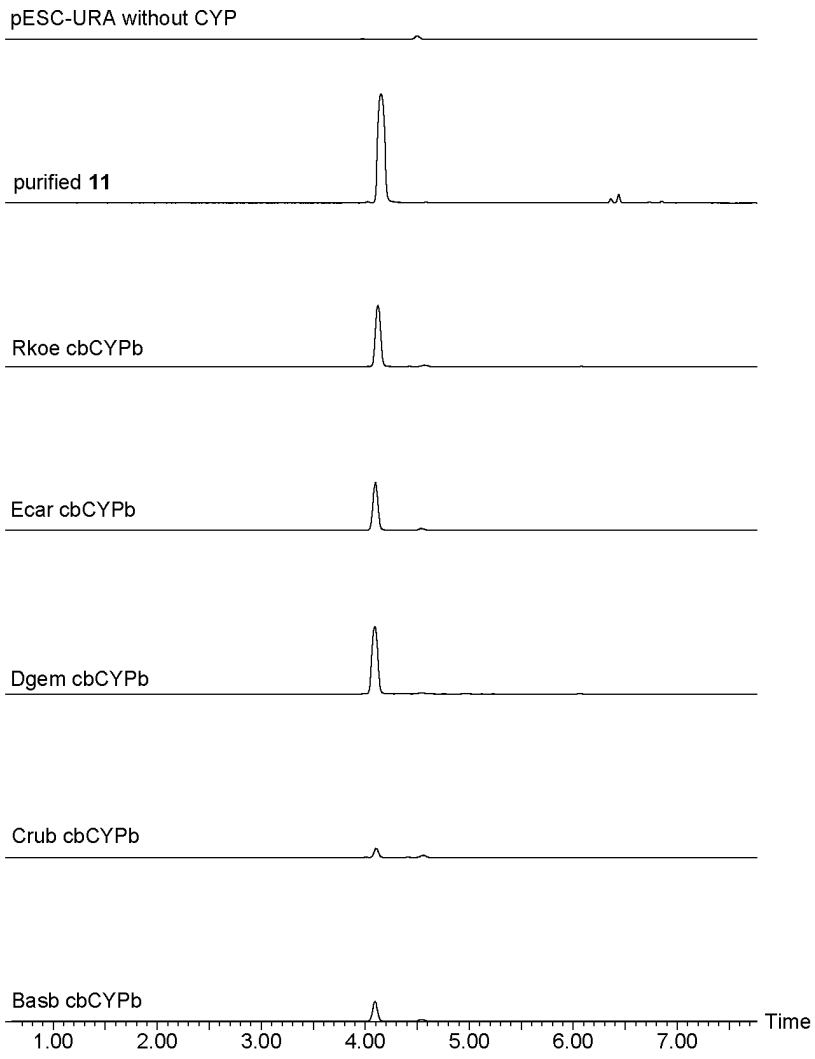
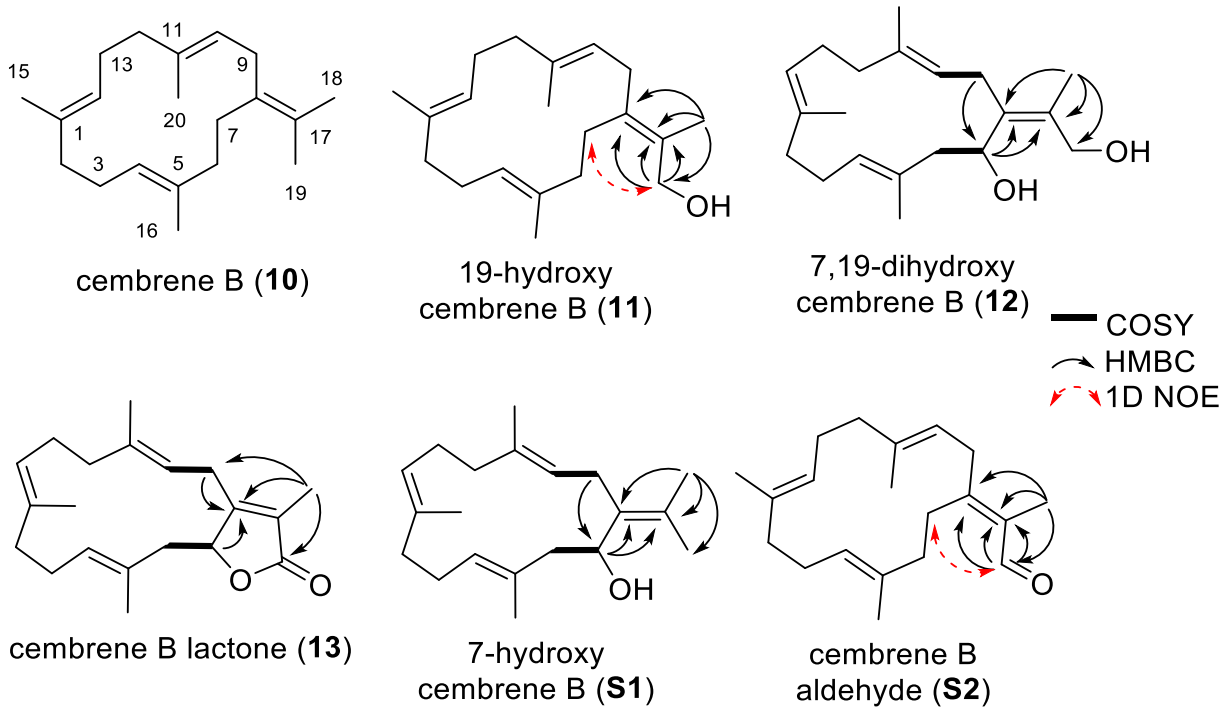


Figure S8 LCMS extracted ion chromatogram (EIC) traces for cbCYPb in vivo assays. Using co-expression with a cbTC at m/z 271.2 in positive mode electrospray ionization (ESI).



cembrene B (**10**)

19-hydroxy cembrene B (**11**)

7,19-dihydroxy cembrene B (**12**)

cembrene B lactone (**13**)

7-hydroxy cembrene B (**S1**)

cembrene B aldehyde (**S2**)

Figure S9 2D NMR correlations for biosynthetic products.

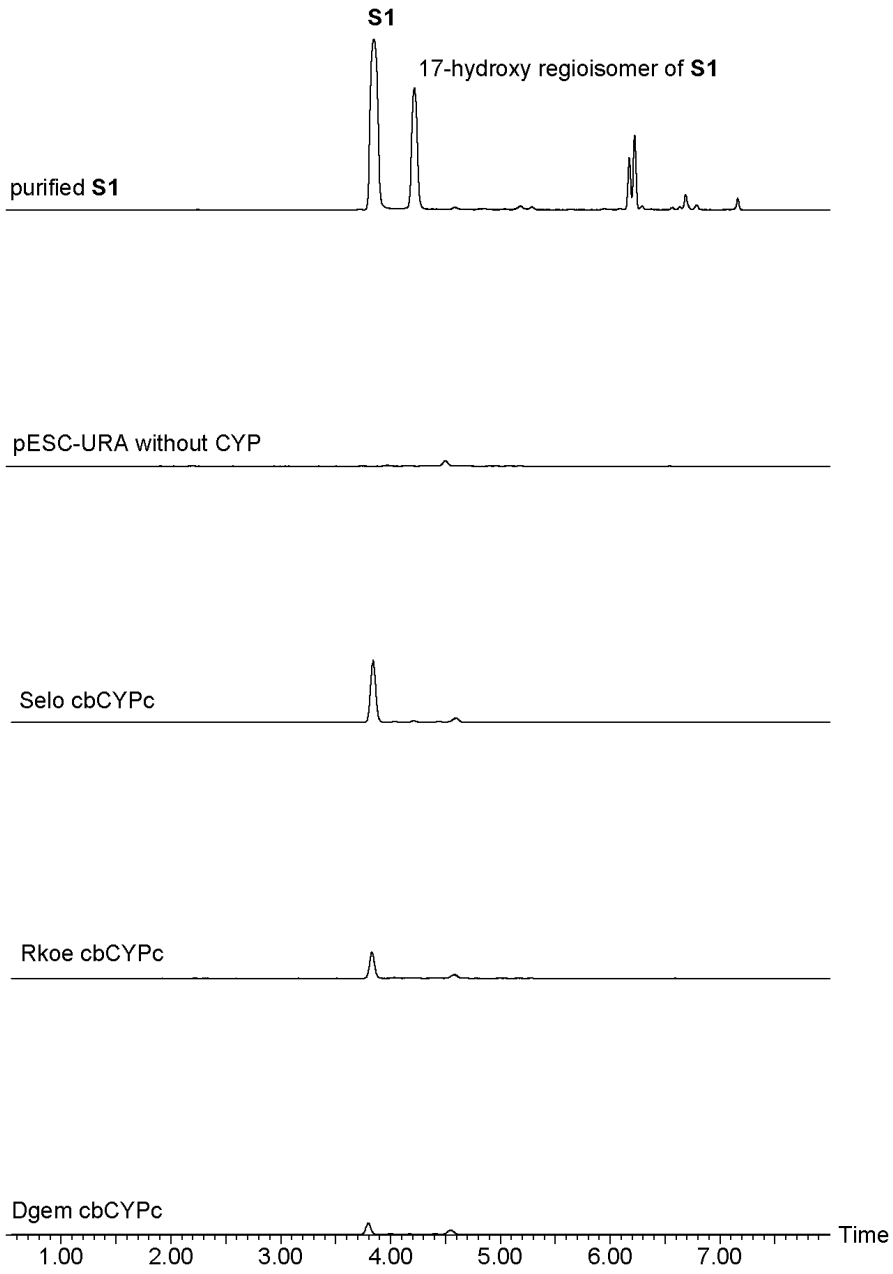


Figure S10 LCMS EIC traces for cbCYPc in vivo assays. Using co-expression with a cbTC at m/z 271.2 in positive mode ESI. The peak after S1 in the trace is the 17-hydroxy regioisomer of S1.

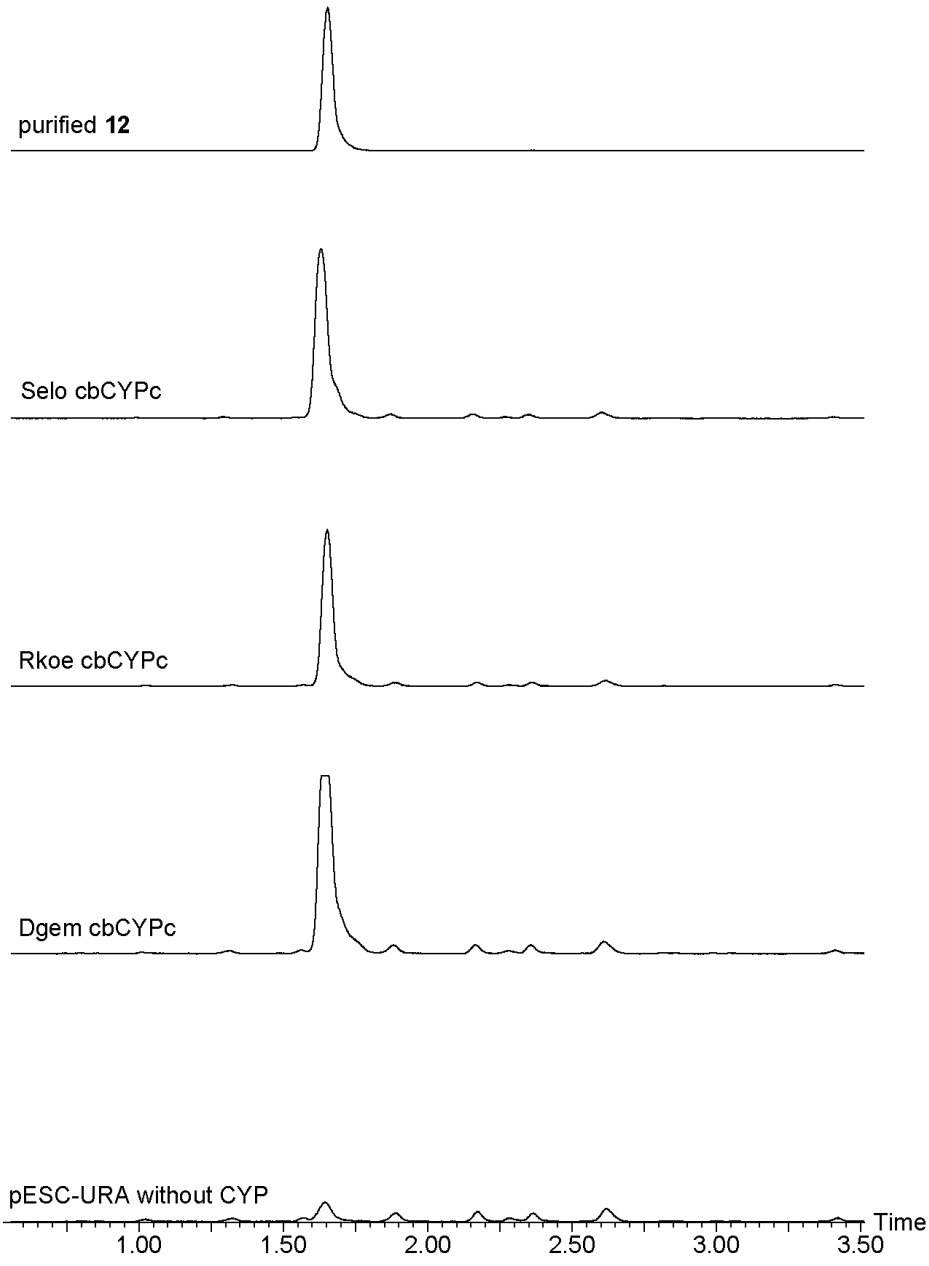


Figure S11 LCMS EIC traces for cbCYPc enzyme in vivo biotransformation assays. Using added 19-hydroxy cembrene B (11) monitored at m/z 287.2 in positive mode ESI.

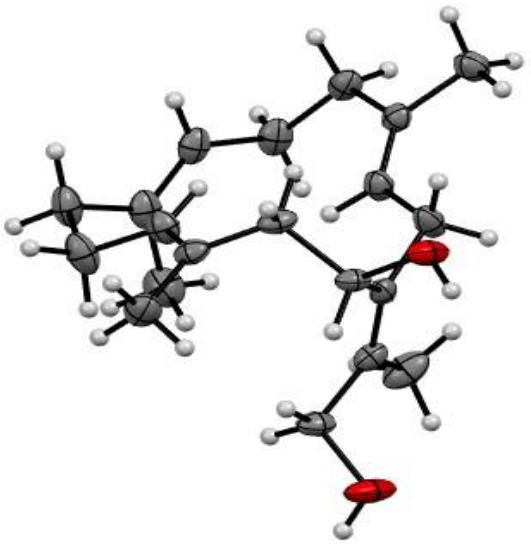


Figure S12 X-ray crystal structure of diol **12**.

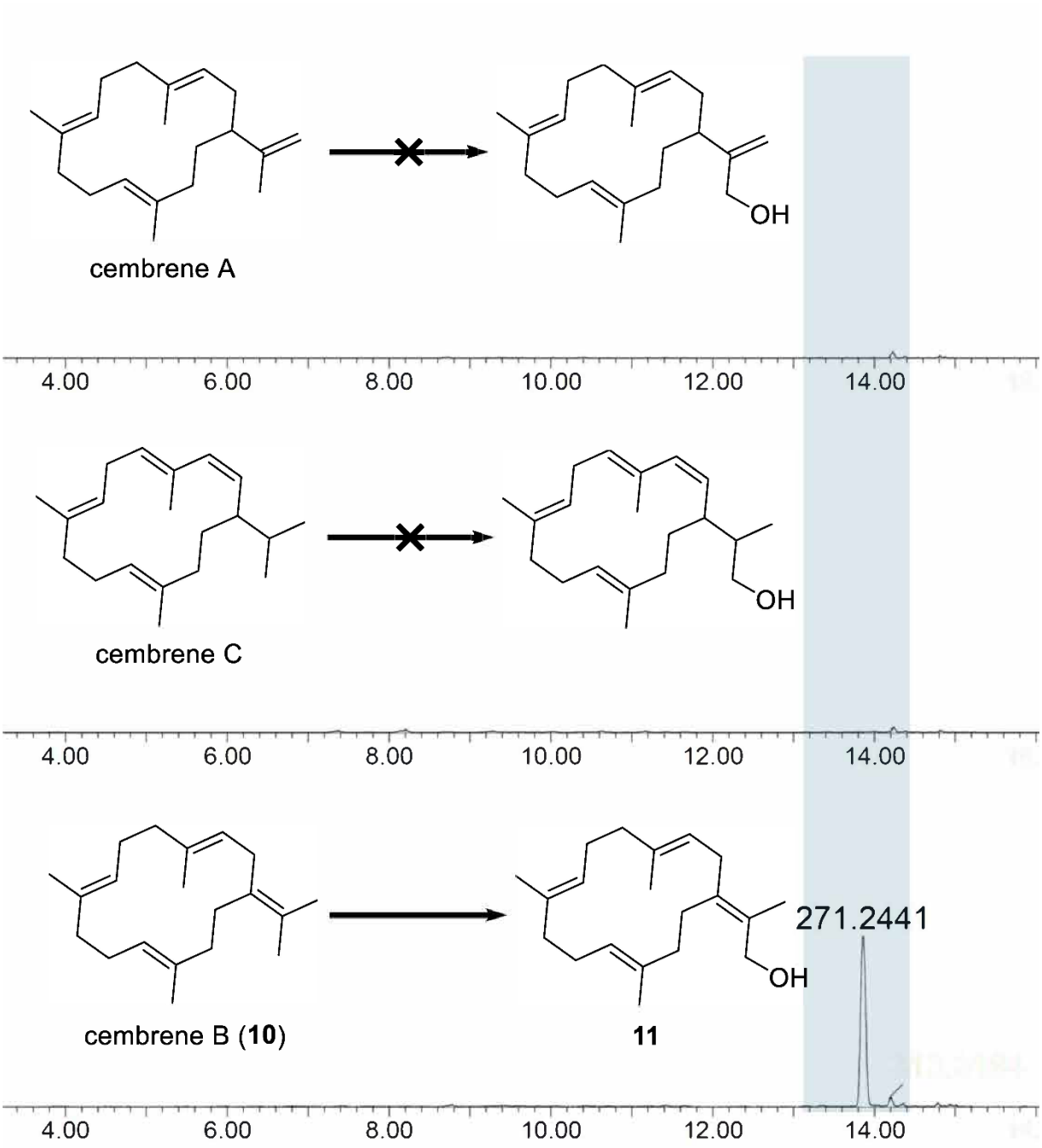


Figure S13 LCMS for co-expression of cembrene A, B, and C synthases with cembrene 19-hydroxylase Ecar cbCYPb. EIC trace monitored at m/z 271.2 Co-expressed TC genes are as follows: Top) Basb caTC Middle) Stro ccTC Bottom) Basb cbTC. Oxidized products are only produced in the presence of cembrene B.

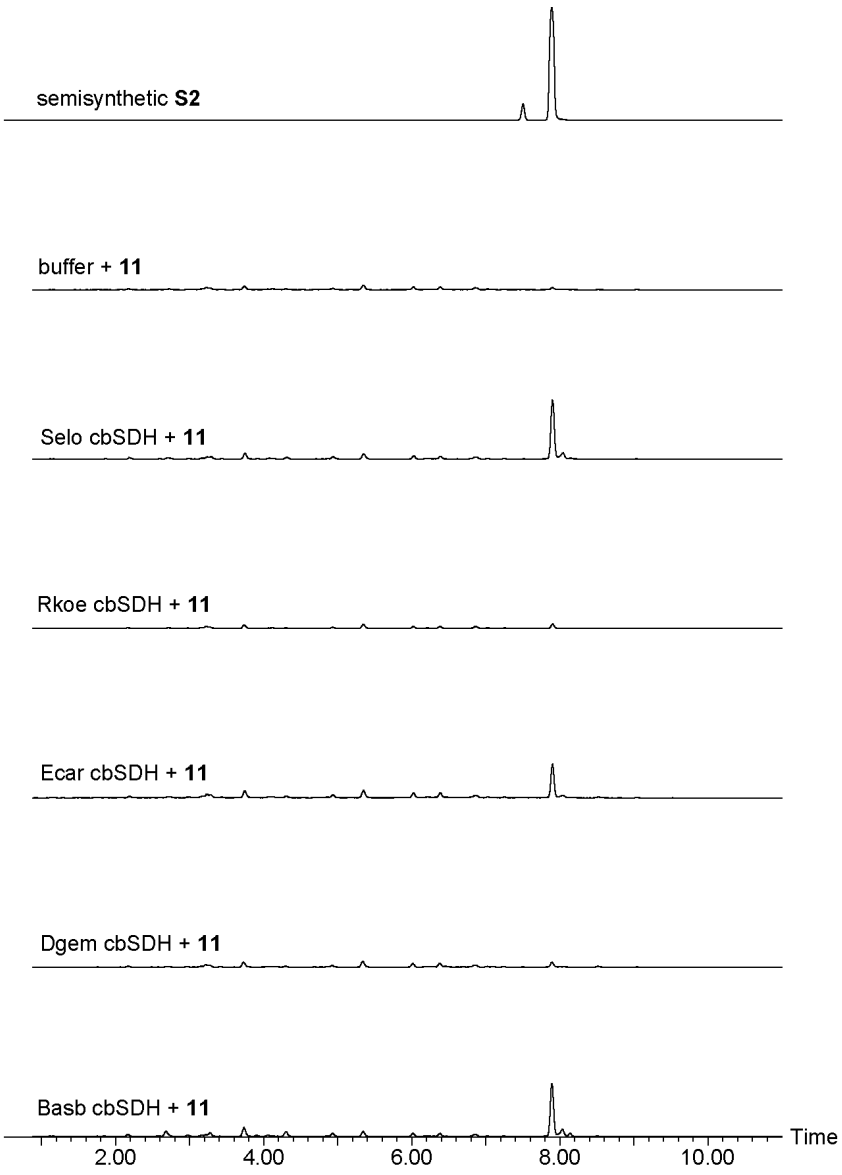


Figure S14 LCMS EIC traces for cbSDH in vitro assays using added 19-dihydroxy cembrene B (**11**). Monitored at m/z 287.2 in positive mode ESI.

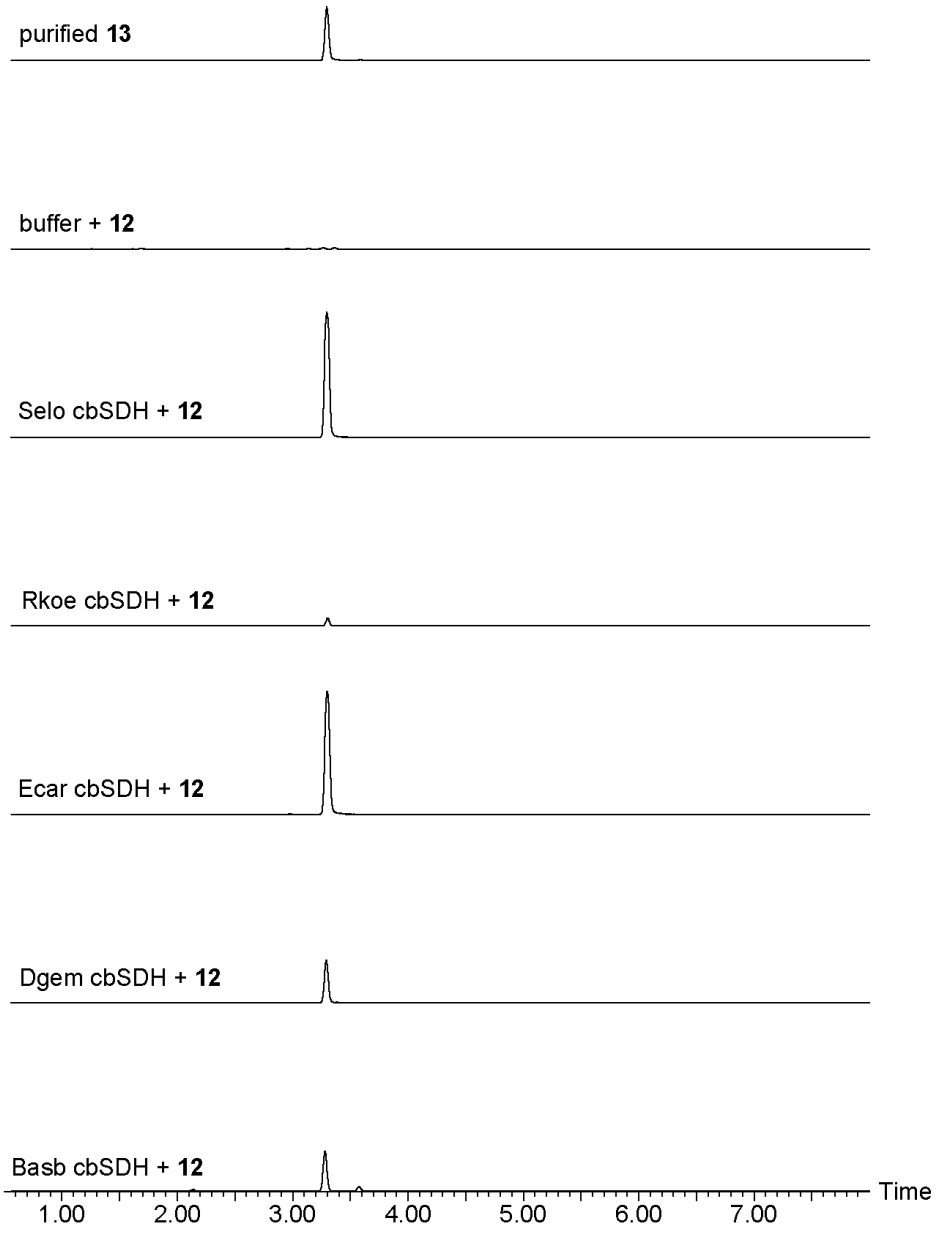
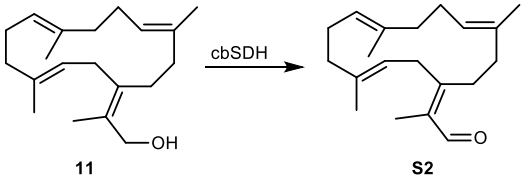


Figure S15 LCMS EIC traces for cbSDH in vitro assays using added 7,19-dihydroxy cembrene B (**12**). Monitored at *m/z* 323.2 ($[M+Na]^+$)in positive mode ESI.

A



B

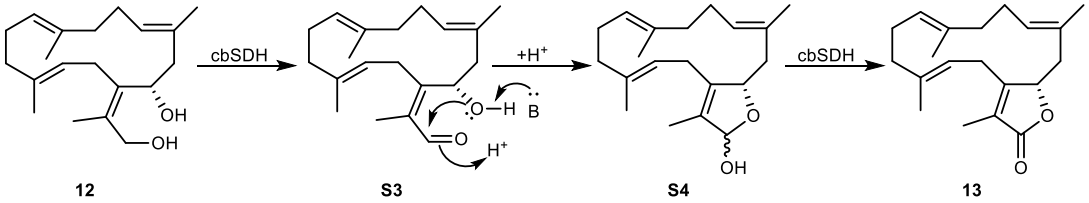


Figure S16 Proposed lactonization mechanism. A: formation of shunt product S2 by C19 alcohol oxidation by cbSDH. B: stepwise formation of the lactone by double oxidation at position 19 carried out by cbSDH.

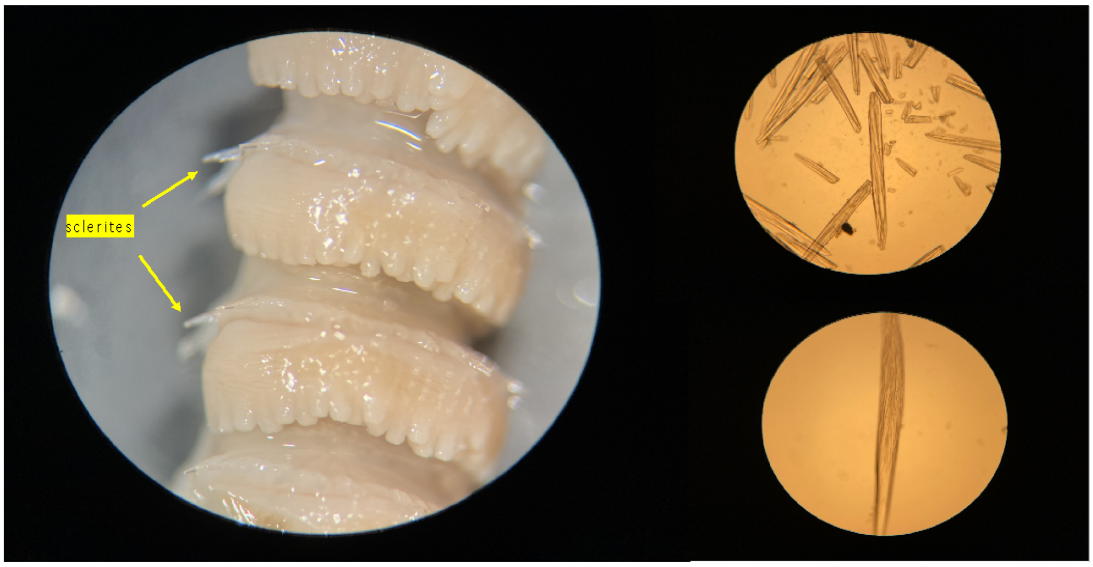


Figure S17 Stylatula elongata species identification. To distinguish between the two similar southern California sea pens, *S. elongata* and *Virgularia*, the presence or absence of sclerites at the base of the polyp leaves was used. The sclerites, or carbonate structures of about 0.5-1.5mm long, depicted below under 10x magnification establishes it as *S. elongata*.

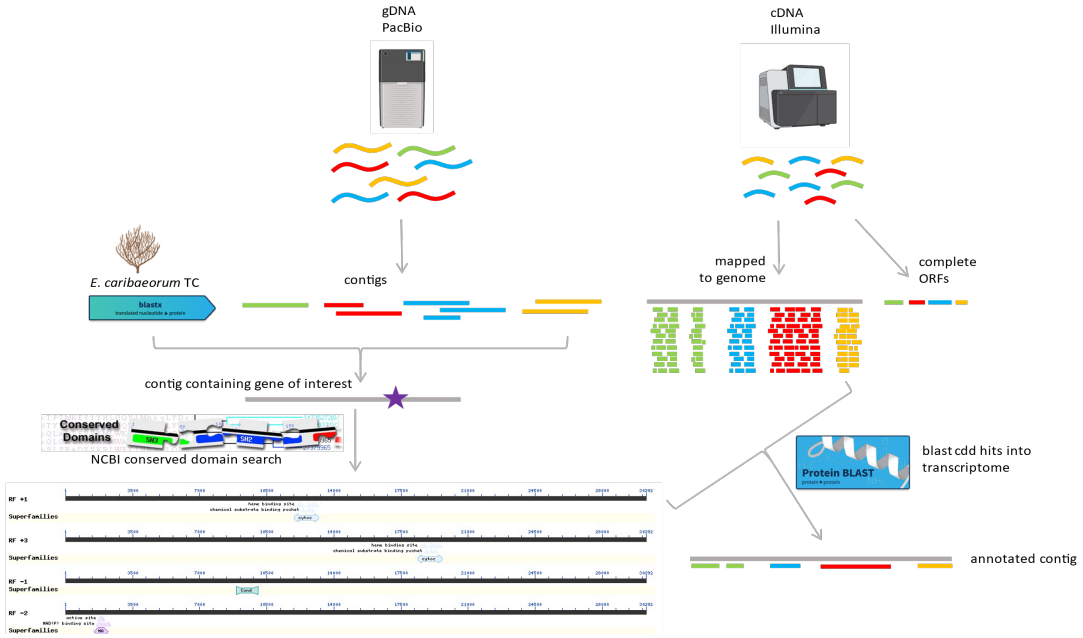


Figure S18 Pipeline of BGC annotation.

List of sequences tested in this study.

Cembrene synthase (TC) protein sequences tested in this study
(*=stop codon)

>Rkoe cbTC
MLRKEVKIPEQWCIPCEVEAVTPPKLRKDIIDWCVCTGITADRRTAEKAVDRLNPYMYMKILFPKLKPDEPLIASQYELNA
SVIVTGMDDVLETYTLSAIEELDHTFKETENWLLSLPSKDYPDLDAILDHIPTTKPTYSRSVVAMFTDYFNKYCAAIRAGYK
PPDDRLRRFRARLAAAVTAYLEMAKRKRRRDGDLDEEEFLWQRSADNLSFPVLMLAEVFTGVLANGEIPNTTLHYYHFY
SNLFAIVLNDLNSYHRDRDNDDNSLVKLWLKLGVAEKFEDAAAKTVDFLNSIIVRMYVAMTENSESHPKSESFRRFREA
VGYTFNGWILVHTTAVERYKSSPFQTVLSPVARGAEEAWLKSETAFGKRCVTIFEELISRRVEEMNKLYGLDST*
>Basb cbTC
MSCKNLVRFPVNWCVPLSECHKSEEARKKHVVDWCVNTGVATKLKAEKMVSQLNPFLYMATLFPKVNPDDAAAAELY
LINASLIVGGYIVDDVLETYSVSGLEELDRFFRFTEGWVSKLLPDDYPKLDEILCSLSEIKVPYSKAVICMFIDNFNKYCYEI
FHKYTSSVQTVGNFRKRLSSSVTAYFEMALAKGRKSVKVDEKEFLWQRSGDNLGFPVLMLTEVISGLLQKESTKIPQVITI
HYFHMYSNLFAHVLNDLNSYYRDVATNHNNLVKLWLQNGIATDFDDAAIKIVEFLNSVIINIHKYCQKLKREFSNSTETD
NFLESIKYTFTGWVHVHTTAVERYRLSPHQITLAKVESGEVNQWLQDESEFGKRCVQLFDKLLKERQAEYTVIYGME*
>Crub cbTC
MSTDREIYVPDSWCVPLYNKPSIKLRERIADWMTKTGLASKTLAQKAVEKLNPFQYMNILFPKVDLSDSSAARLYEINAAL
IVGGYTADDVIETYSLEALKELDYFFRSTEAWVAKLEPSDYPTVKEIAESLSEFEMEMSYSKAVICMFIDSFNQYCYEIWRA
HSVSFSKMSQFRRRLSASVTAYLEIAMETRRGDANIDEKEFLWQRSADNLGFPVLMLTEVLSGLLNNECEIPLVTLHYFH
MYSNLFAHVLNDMNSYHRDIHTDHNSIVKLWKKKGIAIDFDDAATKVKVVQFLNSVTKHMQKSFKKVQQDYPNCLALR
KFLENISYTFTGWVLVHSTAVERYKLSPYQANLVKVEPNGEEEWLEDDSEFGNRCVQMFDQLLCERENEMVFLYGLE*
>Selo cbTC
MSSEWTVSIPVSWYVPVTNTTSLNLFRKEIVSWCTRTGIAGRSVAEKAVSRLNPAHYMMTLFPKVDPLQPTLSKLYEMNA
AIIVGGYIMDDVLETYTLEALVELDAFFRSTEAWLSQLQPSEYPKIEEIVEQLQEITVPFSRAVVCMFTDYFNQCCLALSKI
DEISPSNVRHYRKRLSASVTAYLEMAMEKRQDTAVKLDEKEFLWQRSADNLGFPVLMFIEVFSNLLQEAELAKIPTTIFHY
FHMYSNLFAHVLNDLNSFHRDIDTDHNNLIKLWIQNGLASDFEDAAKKVVQFLSSAIIRMYKAFAECSKQFPNSPALER
FLESIGYSFNGWIVVHTTTVERYRLSPYQISLTIVQQDKQDEWLKNESNFGKRCVQLFEELMKEREEKMIEIYGLSN*
>Dgem cbTC
MSKNFNIGAVYVPVDWCSPLTHKSSDTTELKKGIVEWCVSTGIASKDVAEKAVVQLNPFQYMNMLFPKVDPSDSNSAE
LYKINASLIVAGYIMDDVLETYSLEALEELDKFFRATEEWVSKLNPVDYPNLGRISETLSEIEMTYSRAVICMFIDNYNKYSF
EIWRNYKVSLSKVSQFRRRLSASVTAYLEMAMRKRRVSENIDEKEFLWRRSADNLGFPVLMLSEVLSNVLNKEDENIPST
TFHYFHMYSNLFAHVLNDLNSYHRDIETNHNSLVKLWMKDGVAVDFNDAATKIVQFLNSVIIRMHKAFQETERRFPECP
PLQAFLRNTAYTFAGWILVHTTAVERYKLSPYQTTLAKVGASRFGEWLESESDFGNRCVQLFDELMQEREEEMVKIYGM
DGLN*

CYP450 protein sequences tested in this study

>Dgem cbCYPa
MELCDVLLYFATFFLILVSILSTYIAYAFGYRRLKYNHIPGPPLLHFIKGHAGKILKVIDDGYPLDEYLLDCHKIYGNTILVWL
FHEPLISVTSPELVKQTLVVENFPKSSKLFACGVRLFGERLAGHGLITNFDHENWVKRREFMNPAFHRRYLRNMVPVFN
SCSKQLERKLESIADGKTKVKLLEELGPLTLDVICQAGYGANFDIVNNPNSQFPTAFNLTMQGMEMFVFNPFHRIDFRTF
GFQRKVIDACRLLRKTGKDLIEQRRRDIQIGKDVPNDILSFILKSADKDMFLTDEDLIDEFVTFFYAGHETTARTLAFCLME
LAFHPEVKAKLVEEITSVLGKNDDVSFDDLAKLEYTWMCIRETLRLYPIAAGIRRFIDKPTVLDGISIPGDTDLYVSSYVSAR
HPNHFDDAEEFKPERWSSEEVVTSSRLTAFPFSVGPRQCIGQTFAQIESKILLSRLLRKFVIDLVPGHTKKITQRITIQPRD
GIECVLTERM*
>Dgem cbCYPb
MFFEVAASLLVVWLSWYVMTTYQSRRNMPPGPYPLPIIGNTHQLGPDPPFTMDKLWEKYGDVYQVRFPVGTFVIVNSF
EAAREALVTKKDDFAGRPIELMYPVDIITEAKSVASVDFGVKLMFRKKIMKSALRVFGDGLTQVQDRVNSAVKELLEEIEK
TNGKPFPVKHFVSAALTTQLWGWISSQRYPFGHRILNALIEFQEKMIYLMRQGSYFQLLPFMKFFPTKFMRTLKEVLQM
REDIFGAELEEHLRTYNKETCRDIIDSLISSYEIEKSKTTTKDIGSIDDIKFLLLDAVVAGADTSITAVSWFILYVVLNKDLQEK
LHAELDDVVGRDRLPCWSDIKNLSYLQAVVCEVMRHTGFLPWMPHKTIRDTTIQGYHVPKDTPVFINFYRIHRDPEEYH
EPLKFNPDRFLNSDGTFVGWTAVSAFLPFGVGRRACLGQDLGKMQVFSVISCLLHQFVFQLADNESKPSLEDGMLGS
LRYPVDYEIIAKKRM*

>Dgem cbCYPc
MIVEIITSLIAVWFAYFFIKTYRERKCKPPGPFPLPFIGNLHQLGSDPPFTMDNLWLKYGDVYQVTFPVGTFVIVNSGEAAR
EALVTRKDDFAGRPTALMYPIDVVMEGKNIGNEDYGTQLMFHRKVLKSALHVFGEGVQQVERRVQRAVTELLEETEKM
NGEAFYLQDQVAAAIASQLWEWLSSKKCSFDDKSLKEIVQFNENITSLARQGNYYQMLPLLKYLPTKFMKNIQDTVKLR
DKIFSRELEEHRRTYTEGVIRDVTDALICSYEMEKSKASRKDIGQIEDLKFLMIDVILAGADTSRSVLAWCILFLSLRKELQT
KLQKEIDEVVGRDGLPRWKDVKNLPKLQSFVCEVLRYSSIFPWMPHKTIRNTTINGYLIPKYTPVFLNFYRIHRDPKEWN
DPESFEPDRFLDSAGNFIGWSTVKAFLPFGIGRRSCVGEDLGKMQVFGILSSLIHQFSLELPDNHPRSPLLETVSGFIHQ
PKNHKIIANKRE*
>Ecar cbCYPa
MLILFLVAGLMLFVYIGYLLIWRRLKYRHIPGPPLAHVIKGHAGDMLKKIADGFPLDEYLLDCHKSYGNVVLIWLYHEPLL
SVTSPAMVKETLVLNRFPKSSKLFQCGVRLFGARLAGHGLITNFDHENWMKRRELMNPAFHRRYLRNLVPIFNSCCKR
LERKLESSANGVTKVRFLDELGLVTLDVMCLAGYGKDFDIVNNPNSQFPKAFNLTMEGMEMFAFNPFHQLDVRTFKFQ
REVIEACRFLRKTGNEMITQRRQDLLNGVDVPNDILTFILKSADKATDLSDEDLIDEFVTFFYAGHETTARTLAFCLMELAF
YPEVKAKLLEEIANVLGERDDDVTFENLPKLEYTWFCIRETLRLYPIAAGIRRYIDKPSILNGIKIPGKTDLYVSSYVSARHPD
HFDHPEEFIPERWASHDIVTSSGLTSFPFSVGYRTCIGQTFAQIESKILLARILRKFSIQLVPGHTKKITQRITIQPREGIECSL
KRRIHHHHHH*
>Ecar cbCYPb
MRKNAHSWQTMFLEVCLSLTVVWFLWYMMTTYHSRRLTPPGPFPLPLIGNTHLIGSNPPFTMDKLWETYGDVYQVKFP
VGTFVIVNSFEAAREALITRKDDFAGRPIDLMYPVDIITEGKSVASADFGLKLMFRKKITKSAIRVFGDNVHVVQGRMTNA
VKELLSEIEALDENPFSVKSYISSAITTQLWEWISSRQYRHGNKTLRSLVEFQEKMIYLMRQGSYFQMLPFMKFLPTKFMR
TLQEVLEMRETIFGAELEEHRRTYTTGVTRDIIDSMIASYEIEKAKTTDKDVGTVEDIKFLLLDAVVAGADTSITVVSWFILHM
ILDKDLQTRLQNELDDVVGRDRLPCWGDVKNLHYLQATVCEVMRHTGFLPWMPHRTIRDTSIKGYHVPKHTSVFINFY
RVHRDPQEFCEPMLFDPGRFVNSDGTFKGWTAVSAFLPFGVGRRACLGQDLGKMQVLSVLSCLLYQFTFHEVESESR
PSLEEGDTGSLYHPTDYKVIAKRRDHHHHHH*
>Selo cbCYPc
MLLEVVIAIATTWLVYWLGKTYQARKWGPPGPFPLPGIGNLHQLGASPPFTLDHLWQTYGDVYQITFPVGTFVVVNSGR
AAREALVTRKDDFAGRPTELMYPIDVILEGKNLGNGDYGPQLVFKRKILKSALHVFGEGVERAAVRVNRAVIDLLKHTEN
KRKEAFYIQDYISSAIASQLWEWLTSKKTSFDDASLHEVIAFNENVTSLARQGNYYQMLPLLTYLPTAFVKQLQNTLRMR
DDIYACELDEHRLTYVDGCIRDVTDALISSYHVEKAKCSDREVGPIEDLKFLMIDVVLAGADTSRTVVAWCVFFLVVQKE
LQNELQQEIDRAADEDHYLLWEQVKNLPKLQSFVCEVLRYTAIFPWMPHKTIRDTTLNGYHIAKNTPIFINFYRIHRDINE
WDEPEVFKPHRFLDSDGNFVGWKAKKGFLPFGIGRRFCLGQELGKMQVFGMVANLLHQFTLELPAGDPIPSILDSKKS
FVHQPPSHRVIATKRR*
>Crub cbCYPb
MFLEVAASLLVVWFVWYMISTIQLRRNMPPGPFPLPLIGNTHQIGSDPPFTMDNLWEKFGDVYQVKFPIGTFVIINSFEA
AREALVTKKDDFAGRPIDLMYPVDIITEAKSVASCDYGLKLMFRRKIIKSALHVFGDNLQRVEERVNSAVMELLTEIQEMN
GDAFPIKYFISAAVTSQLWGWISSERYRFGDKRLHSLIEFQENMIFLLRQGSYFQMLPLMKYLPTKFMKTLQEVLVMREEI
FGAELEKHCRTYNEDTSRDIIDSLISSYNKEKSKTSARDVGTVEDIMFLLLDAVVAGADTSITVISWFILHMILKKDLQTKLQ
KELDEAVGRNRLPRWDDIKSLPRLQATVCEVMRHTGFLPWMPHRAIRDTTIQGYHVPKNTSVFINFYRIHRDAKEFDEP
SLFNPDRFLNSDGTFVGWSAISAFLPFGVGRRACLGADLGKMQVFSVISSLLHQFTFDVNDQEPTPNLVDSISGSLRHP
KDFKVIAKNRL*
> Crub cbCYPa
MEFSLVLSYSITCFLIIALSLLTYVGYTFGWRRMQYRHIPSPPLSHFFKGHAGEMFKYMKNGYPLDEYLLDCHKKYGNIILI
WIYHEPLLSVTSPALVKETLVLRNFPKSSKLFQCGVKLFGARLAGHGLITNFDHENWLKRREFMNPAFHRRYLRNMVPT
FNRCCERLERKLQRIANGVTKVRLLDELGSLTLDVICEAGYGKSFDIVNNPKSQFPTAFNRTMEGMELSVFHPLHRIDFR
TFPFQRKVIEACHFLRKTGKDLITERRRDIQNGYDVPNDILTFILKSADKGLLLSDDDLIDEFVTFFYAGHETTARTLAFCL
MELAFHPECKVKLLEEISRNLGKDGDVSFEDLPKLQYTGMCILESLRLYPIAAGIRRVIKTPTVLRGLHIPENSDLYVSSYV
SARLPEHFADADQFKPERWSSEEVVLASRLTAFPFSTGPRTCIGQTFANMETKILLSRLLRKFSIELVPGHTKKITQRITIQP
RDGIECTLTNKV*
>Rkoe cbCYPc
MFLEVVVALMAVWLIHFAVKTWRSRLGDGPPGPIPLPLIGNLHQLGSNPPYTIDHLWRRYGDVYQITFPVGTFVVINSG
KTAREALVERKDDFAGRPTEIMYPIDVILEGKNIGNGDYRPLLLFQRKVLRSALHVWGEGAKRAEARVNRAVVTLLRETE
AKNGAPFHIQDYVVSAIASQLWEWLTSKRTDVRQESLKEVIRFNESITSLARQGNYYQMLPLLTYLPTEFTKKIRETVELRD
KIYARELDEHLASYVEGRIRDVTDALIAAYKAETAKGLDKDMGTIEDLKFLMIDVMLAGADTSRTVVSWSILFLVSDTELQG

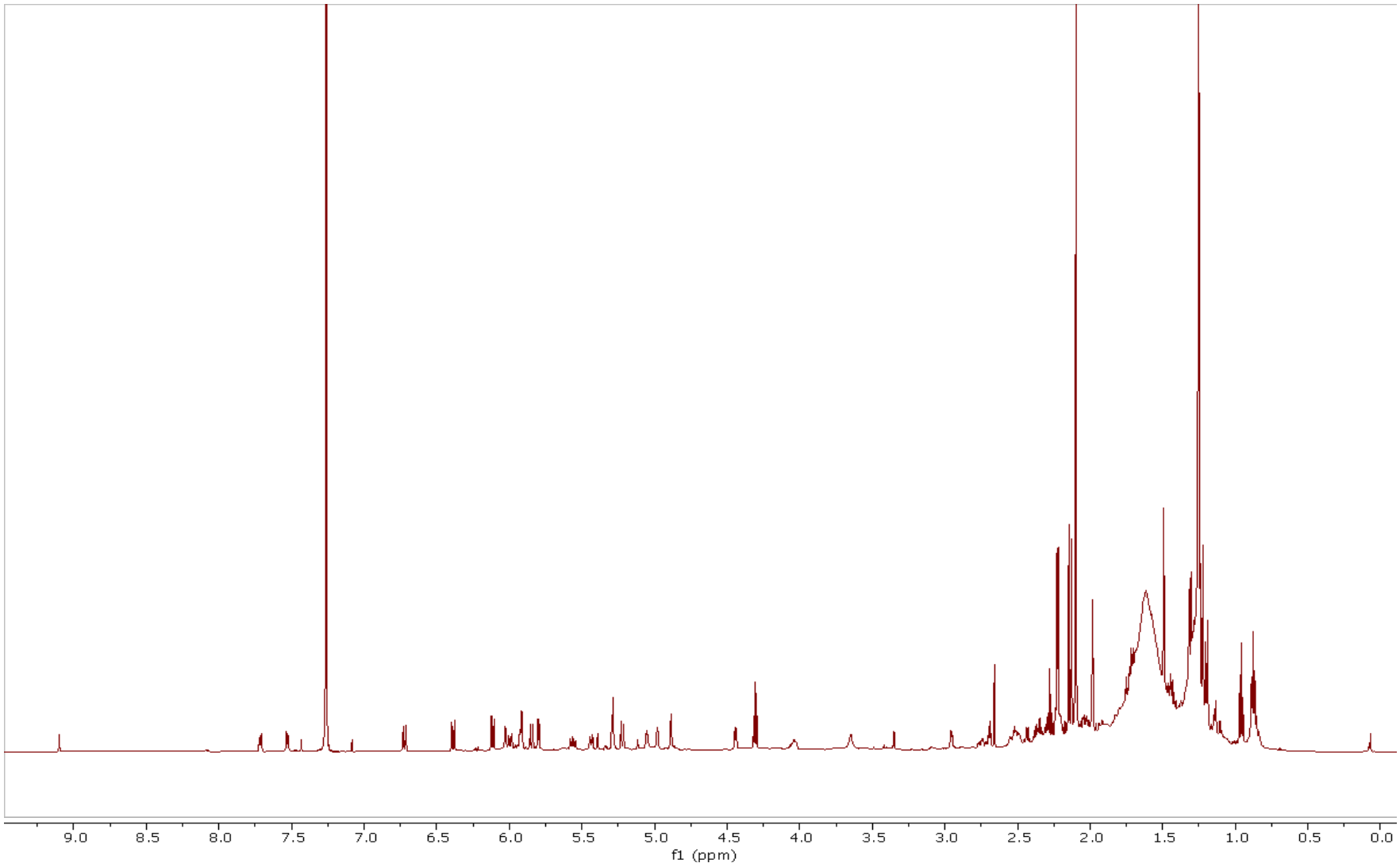
KLQQEIDRVFGGEERLLLAEDSKCLDRLQSFVCEVLRFVSLFPWMPHRTIRDTEVGGYKIAKDTPVFVNYYRIHRDPEEF
EAPDLFLAERFLDTGGHFVGWTAVSGFIPFGAGRRSCSGQDLGRMQVFAVLANLIRQFRLESPPGEPAPSIDDTVSGFI
NQPKSHRVVAIKRR*
>Rkoe cbCYPb
KTMILEVFLALSVVWLTWYYITVFYSRRGNAPPGPIPLPFIGNALQLGQHPPFTMDALWRAYGDVYQIQFPVGTFVIVNG
YQATKEALVTKRDDFAGRPIDIMYPVDVITQAKSVASADYGVKFMFRKKLLKSAFRVFGESLERVEERVNSAVVELIGVVE
KSGGEPLYIKHHISSAITAQLWDWISSQRCRFGDETLNSLIEFQDKMIFLLMQGSLHQMLPFMKYLPTKFNRYLADVLR
MREEIFGTELHQHLQTYNNDEPRDIIDNLIASYEAEVDKGTTRDIGSIEDIKYLLLDIVVAGADTSITSVSWFVLYMVLYRDV
QKTLQAELDEVVGVDRLPRWSDIKNLHRLQSAVCEVMRITGFLPWMPHRAIRDTTVQGYRIPKGTSVFINYYRTHRDPK
EFAAPEEFRPNRFLDADGKFVGWTGVAAFMPFGLGRRACIGQDLGKMQTFSVISSLLHQFSLALPELAPTPSLDEGKLG
SLRHPVDYHVIATRRH*
>Rkoe cbCYPa
MEYTTILWLLLALFCCLVLVVMGYVCYAFIYRRYLYRHIPGPPITHFIKGNAVEMLKMIANGYPLDEYLLNCHKAYGPVVCI
WLFHEPLLSVTTPELVRETLAVRNFPKSSKLFQCGVRLFGARLAGHGLITNFDHENWLVRRETLNPAFHRRYLKNIMPIF
NECCAAFENKLARYADGRTRVKLLDEICPLTLDVLSRAAYGTDFQIVNNPDSPFPTAFNLTMSGMEMFVFNPLHWLDF
RTFPFQRKVVEACKFLRKTGNDMLVQRRRDIEAGKEVPSDLLTFIVKSGEKLSDILTDVDLIDELVTFFYAGHETTARTLAF
TFMELGFHPEVMEKLVSEITSVLPPTGDVTMEHLNKLQYTWLCLRETLRLYPIAAGLRRYINTPCELGGYQIPDKTDLYVS
SYVSARMPGNFEDAKEYRPERWLSEEVVVNSRLTQFPFSVGYRTCIGQTFAMMETKTLLARILRKYVIDLVPGHSKKITQ
RITIQPRDMVECFVSKREWPQTTREPNEKSP*
>Basb cbCYPa
MEFSNALYFSVILLVIVVTTSLTYLGYLFGWRKLKYRHIPGPPLAHFIKGHAGEIFKVIDDGYPLDEYLLTCHKTYGNVVLIW
LFHEPFLSVTSPELVKETLGARRFPKSSKLFQCGVKLFGARLAGHGLITNFDHENWLKRREFMNPAFHRRYLKNMIPIFN
GCCESLERKLASSVNCTTKVRLLDELGYLTLDVICQAGYGKDFDIVNNPSSKFPTAFNLTMQGMEMFAFNPLHWIDFRT
FPFQRKVVDACHFLRQTGKELITERRELLKGGADIPNDILTFILKSADKNMILTDEDLIDEFVTFFYAGHETTARTLGFCLM
ELAFHPEVKEKLLEEILSVLPETGDISFDDLPKLEYTSMCIRETLRLYPIAAGIRRFIDKPTVLAGIFIPENTDLYVSSYASARH
PNHFEDAEEFKPERWKSQDVLLSSRLTAFPFSVGYRTCIGETFANMEAKILLSRLLRKFEFHLVPGHTKKISQSITIRPRDG
IECRISKRA*
>Basb cbCYPb
MFLEVVTALVVTWLLWYLIHTYQCRRNMPPGPFPLPLLGNTHQIGTDPPFSMDHLWKKYGDVYQVTFPIGTFVIINGFD
AAREALVTRKDDFAGRPIDLMYPVDIITQGKSVASADYGLKFIFRKKIVKSALHIFGNNLSQVEERMRSAVKDLLTEIDETQ
GRPFSLKYYISAALTSQLWGWISSQRYRHGNETLQALIQFQEKMIFLMRQGSYFQMLPFMKFFPTKFTKTLKDVLKMREE
IFGQELNEHLRTYNNEETRDIIDSLIFSYETEKKKTTVKDVGTVDDIKYLLLDAVVAGADTSITSVSWFMLYLVSKNIDQSKI
HDELDKVVGPDRPPRWEDVKKLHYLQASVCEVMRITGFLPWMPHRTIRDTTVQGYHVPKNTSVFVNFYRIHRDPKVYS
EPSVFNPDRFLNVDGAFNGWTTLPAFVPFGVGRRACLGQDLGKLQVLSVVSSLLHQFTFDADNQHPLPSLDDGTLGS
LRHPIDFKVIAKRRQ*

SDH protein sequences identified in this study

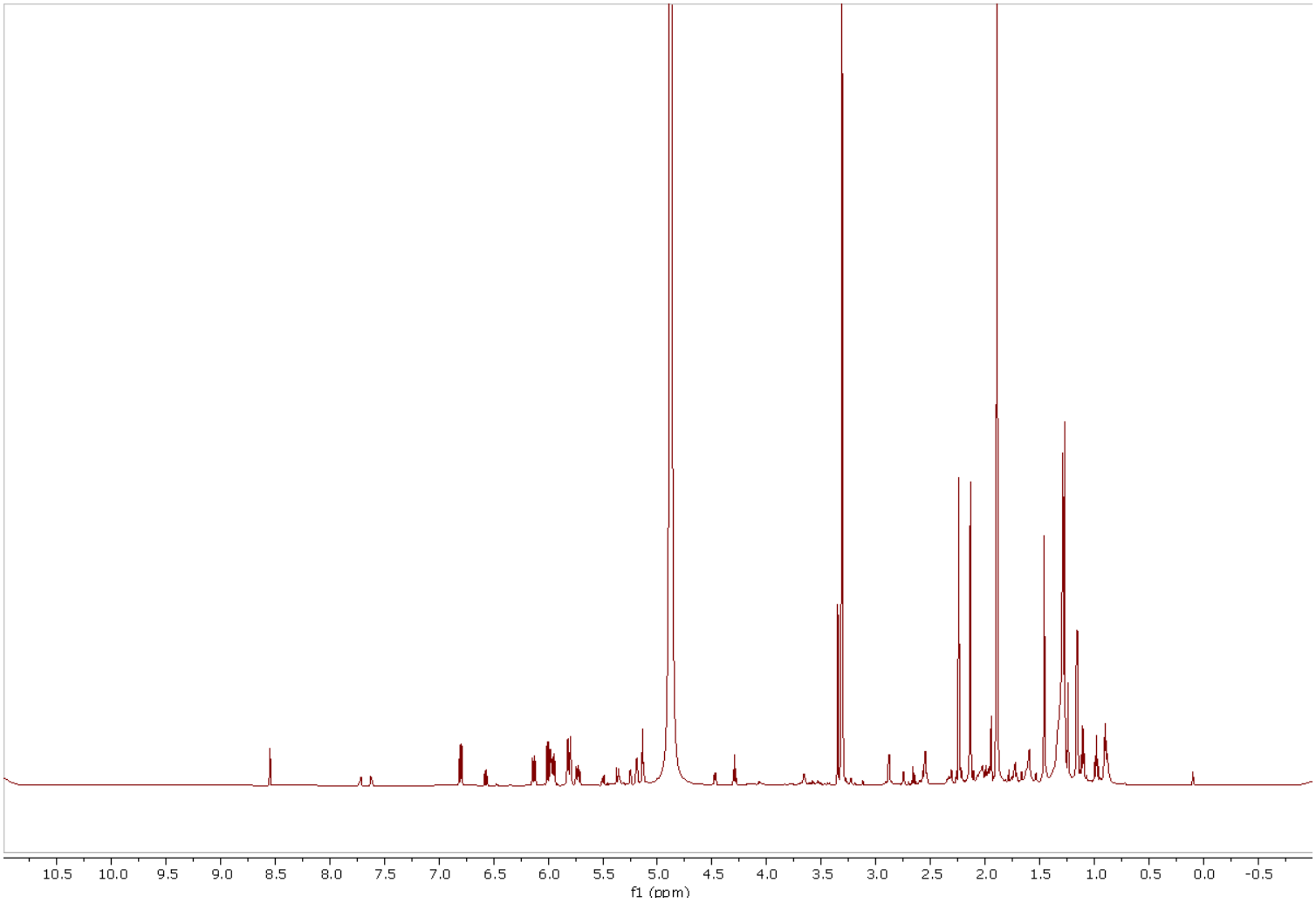
>Basb cbSDH2
MVQYLGVVTGGAQGIGLAIVEEILQSGGKAYFFDVKEEQGKTTEAELQSKYGKNRVKFLKCDVANADEFKGAIDKAKKE
MGGLNVFCNNAGFARPFYDISRWEQEIDVNLKASIRGSLVALDLMSAEKRGLGGVIVNTSSFVALNPVPLFPTYCAAKA
GVIAFMRSMAPLARAANVRINCVCPDATDTPQYQKTLPELAATFSQEKYKSFVESMASGGGKQAISPNAQKPKDVAVG
VMKIIEDETKNGEALVVSKEHGTYYHIFK*
>Selo cbSDH1
MACTEYLAIVTGGAQGIGRAIVSALLEAGGKVHFLDVKVDQGNQTQKEFDKQFGIGRAKFVVCDVSDRTQLQKAFQDA
KLDLGGLNVVCNNAGIFSPVFDRDRWTKVVAVNLTATIHGTLLGLDMMSTNKGGKGGVIINTSSVEGLRPFPIGPVYSA
TKYGVIGFSRSLAPVSIRDNVRINCICPGRIDTELMYSGVQDCLQVYSHDEIQSLLTVGRGSNWADAKLEPKENAPALAS
SVVGEEVLKLVKDETKNGEALVISKVLGTYFQKFE*
>Dgem cbSDH1
MTTTYLAIVTGGAQGIGRAITSALLEAGGKVHFLDVNIEEGKKAQNDFDKKYGAGQSKFVFCDVSDKDQLEEAFREAKS
ELEGLNVVCNNAGITCPVFDRSRWPRVVEVNLSATIHGTLLGLDLMGTSKGGTGGVVINTSAIAGLRPVPVAPVYSATK
HGVIGFSRSLAPVAVAENVRINCICPDGTNTTLVANSVAELLPVFSIDEIQSLLAEHVYGMDISAGQGPRVSADQSDTLLP
SAIGEQVLELVKDESKNGEALVISKVLGTYFQKFN*
>Basb cbSDH1

MADPYLAIVTGGAQGIGYAITVALLETGAKIHFLDVNSEEGKAAQKTFDKKYGAGKAKFVVCDVSDRAQLGKAFKDAKS
QLGGLNVVCNNAGITCPVFDRTRWPKVVSVNLSAAIHGTFLGLDLMGTSKGGNGGVVINTAAIAGLRPVPIAPVYSATK
HGIIGFSRSLAPIAISDNVRINCICPDGTDTTLVSNSVAELLPVFSVAEIQSLLAKHVYGIDTNNKQSTQSSNDMSSSLLLPI
SVAEEVLKLIQDETKTGQAIVVSKVLGTYSHDFK*
>Rk cbSDH1
MATYLAIVTGGAQGIGHAITSAILQSGGKVHFLDIEEELGIKAQAEFDEKYGVGKSKFVMCDVSNREELKEVFIAAKREMG
GLNVVCNNAGIHAPVFDRSHWPKVVDVNLTAVIQGTLLGLDLMAVSKGGQGGVVINTSAMAGLRPVPLAPVYSATKH
GIIGLSRSLAPMSIADKVRINCVCPDGTNTTLVANSMAELLPVFSVEEIQKILVKHIYQLDLNEETMQAGPDLKGDPADAL
LRPSEVGEEVFKLIQDDSKNGEALVVSKQLGTYYWQFN*
>Ecar cbSDH
MVEYLAVVTGGAQGIGLAIVMELLEAGGKVYFIDMDEEQGKKAEKDLQHRYGQDRVKFIKCDVSNEKEFEASIRKAREE
MRGLNVMCNNAGFARPFYDMSRWELEIDVNLKASIRGSLIALELMNPENNGSGGVIVNTSSYTGLNPAPLFPTYCAAK
AGVIAFARGFAPIAVTSNVRINCVLPDATMTPMYQKTLPELASTLSQEKFQSFMKTMTSSQGSKQAISPNTQEPRDVALG
VIKILQDETKNGEALVVSKEHGTYYHQFN*
>Rk cbSDH2
MAEKRVELERLASSLVAIKWTIVGIALLMLLKNLLSTKRQRTRRKLPGKVVVITGANSGIGFATAMEFARRRARVVLACRR
DGQRTAARIRRAADNNDVVWRQLDLSCCDSVRRFARDVLTSERRLHVLVNNAGVMTPPDRHTRDGRDTQLAVNYLG
HFLLTNLLLNLLKKSDGKILNLTSHTYMRATHVDLTDPRPCDPWVAHHRSKLAVVMSTRELSRRLRGTNVTANVIDPGTT
NTSLRRHTWLGFPPLRYFLVPVTWALWQTPEHVGQSIANLLASRALDEHSGQYFADGKIQHIAEHALNENAAKQLWE
NSVKLVGLQENAVL*

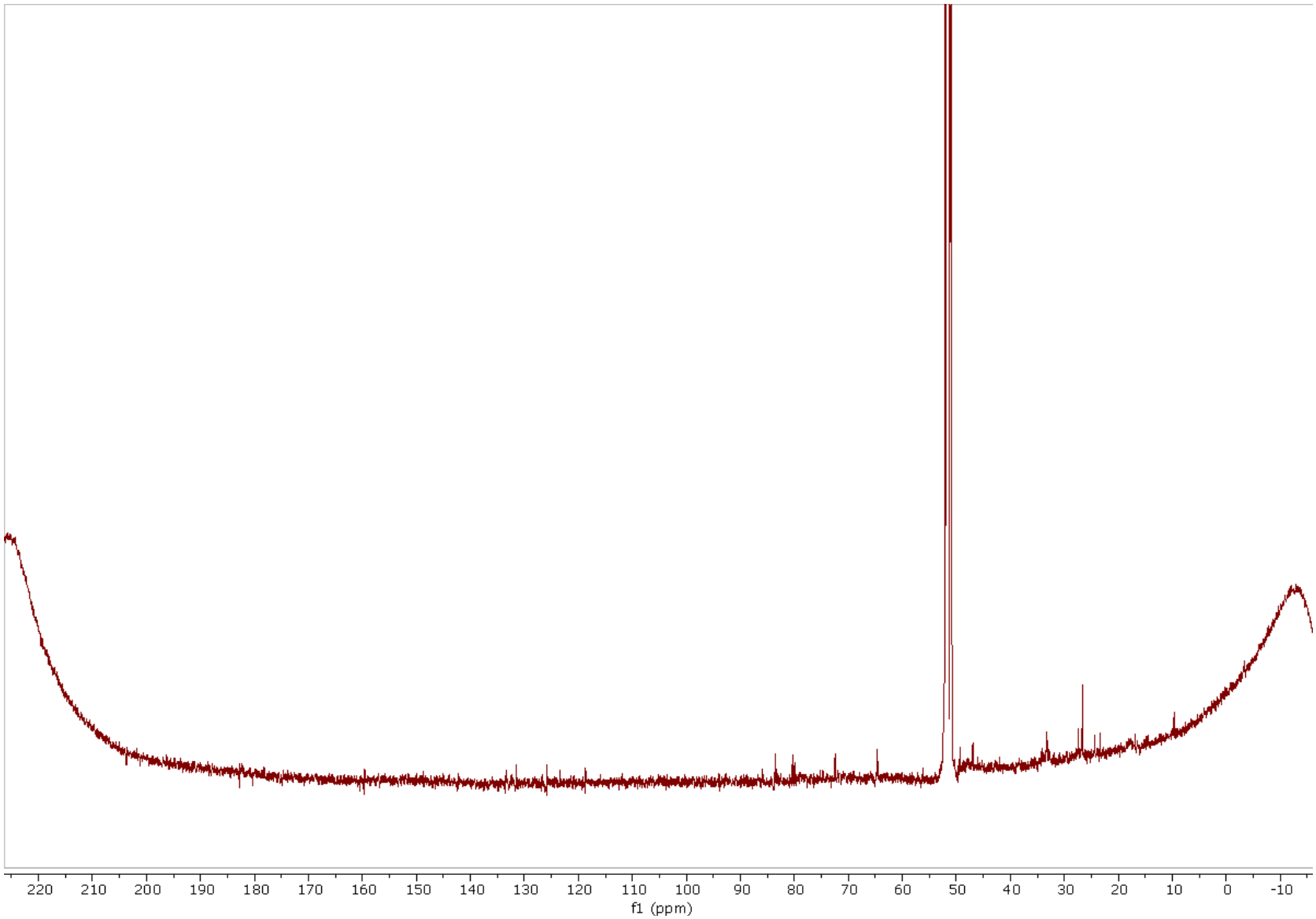
NMR spectra of 11-hydroxyptilosarcenone (6)



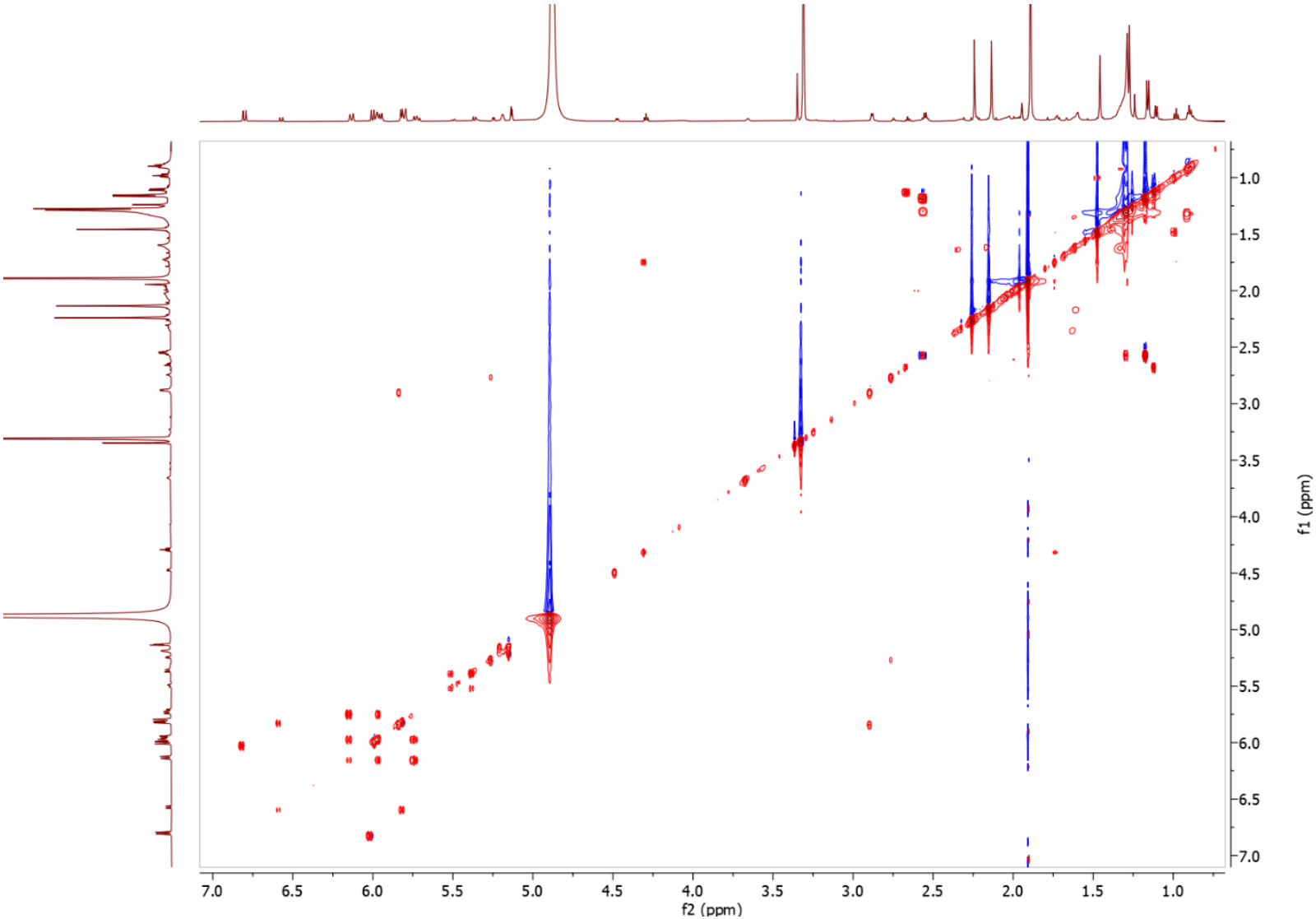
^{1}H NMR spectrum of 11-hydroxyptilosarcenone (**6**) in $CDCl_3$ at 600 MHz.



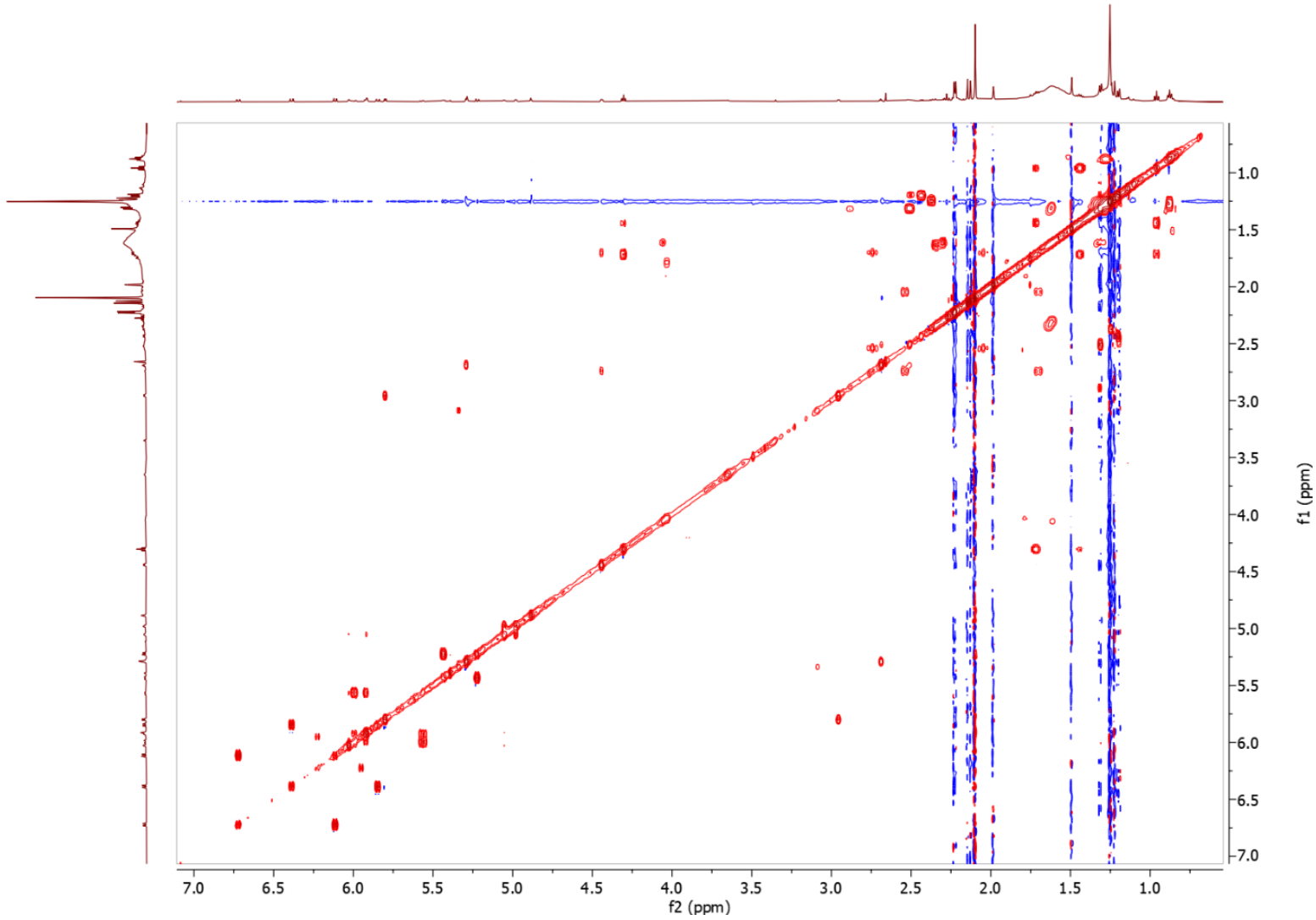
^{1}H NMR spectrum of 11-hydroxyptilosarcenone (**6**) in CD_3OD at 600 MHz.



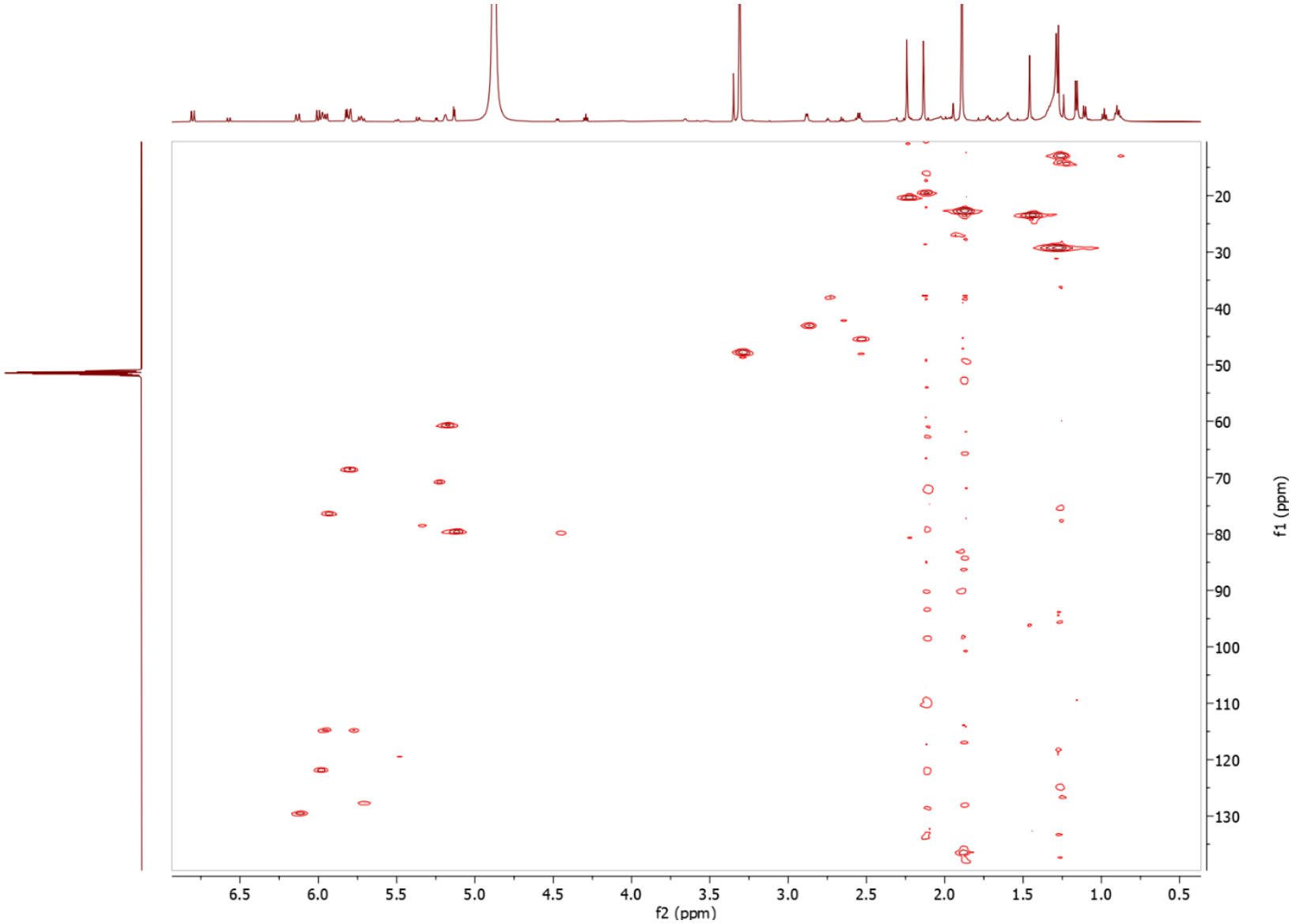
^{13}C NMR spectrum of 11-hydroxyptilosarcenone (**6**) in CD_3OD at 125 MHz.



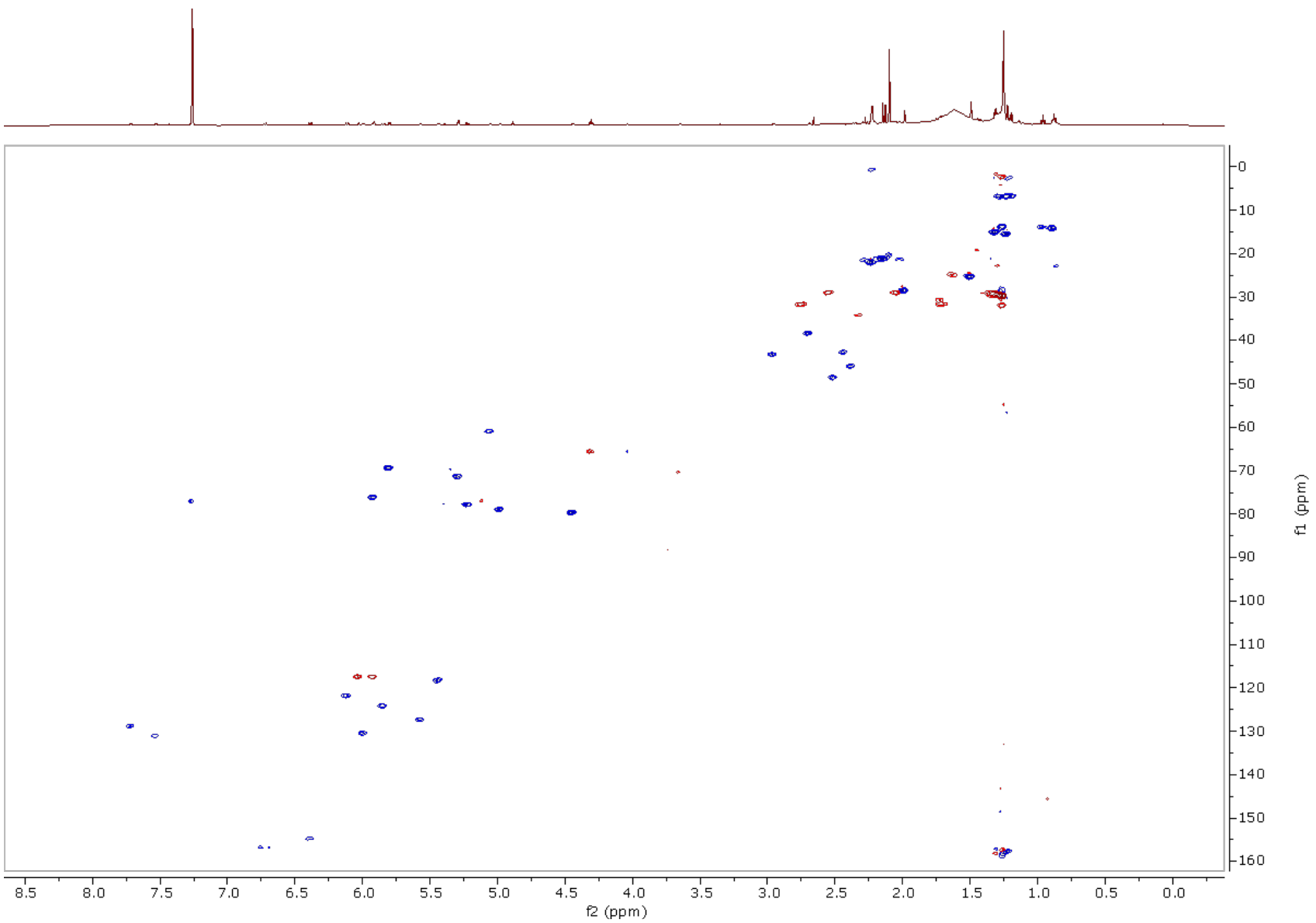
^{1}H-^{1}H COSY spectrum of 11-hydroxyptilosarcenone (**6**) in CD_3OD at 600 MHz.



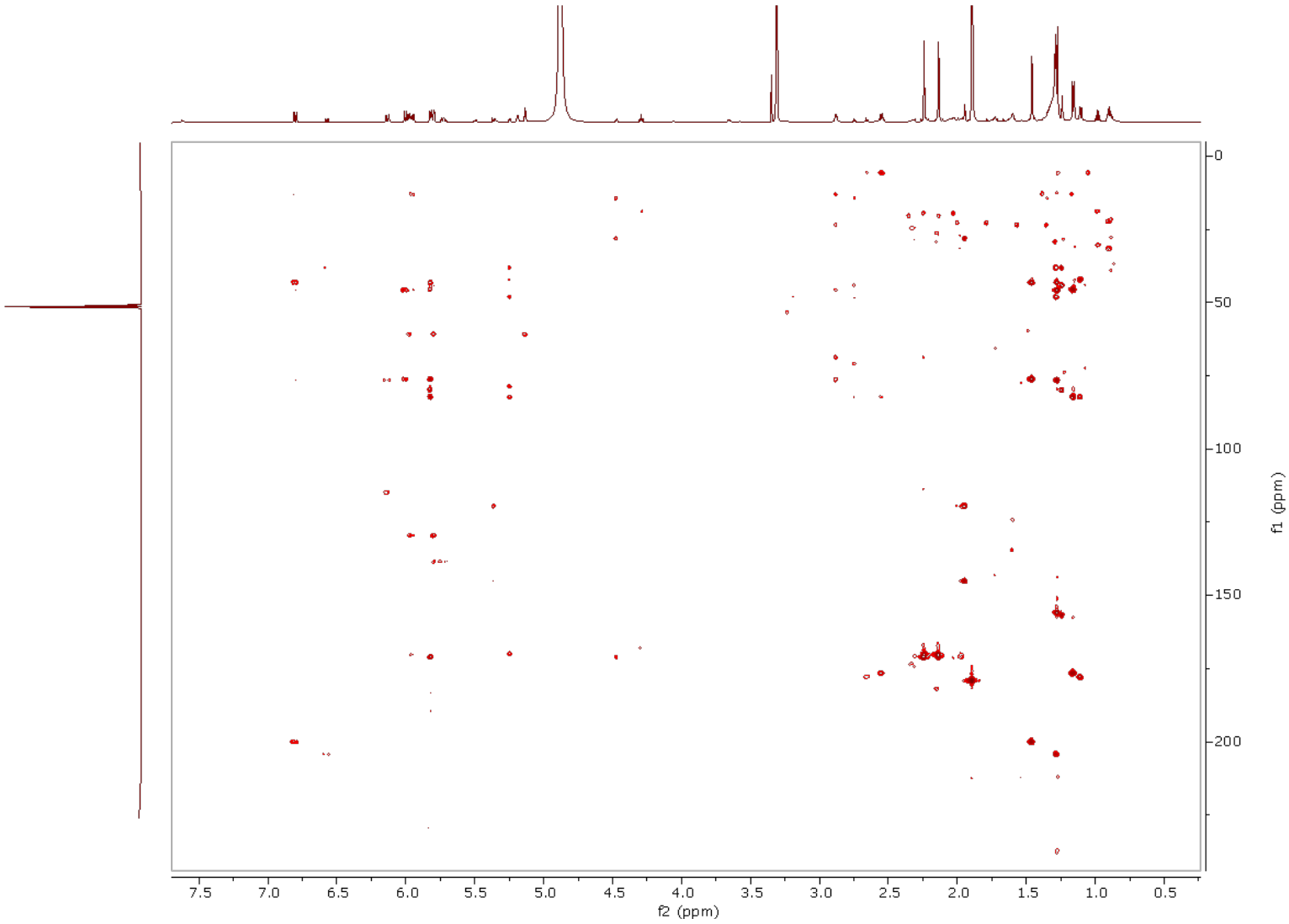
^{1}H-^{1}H COSY spectrum of 11-hydroxyptilosarcenone (**6**) in $CDCl_3$ at 600 MHz.



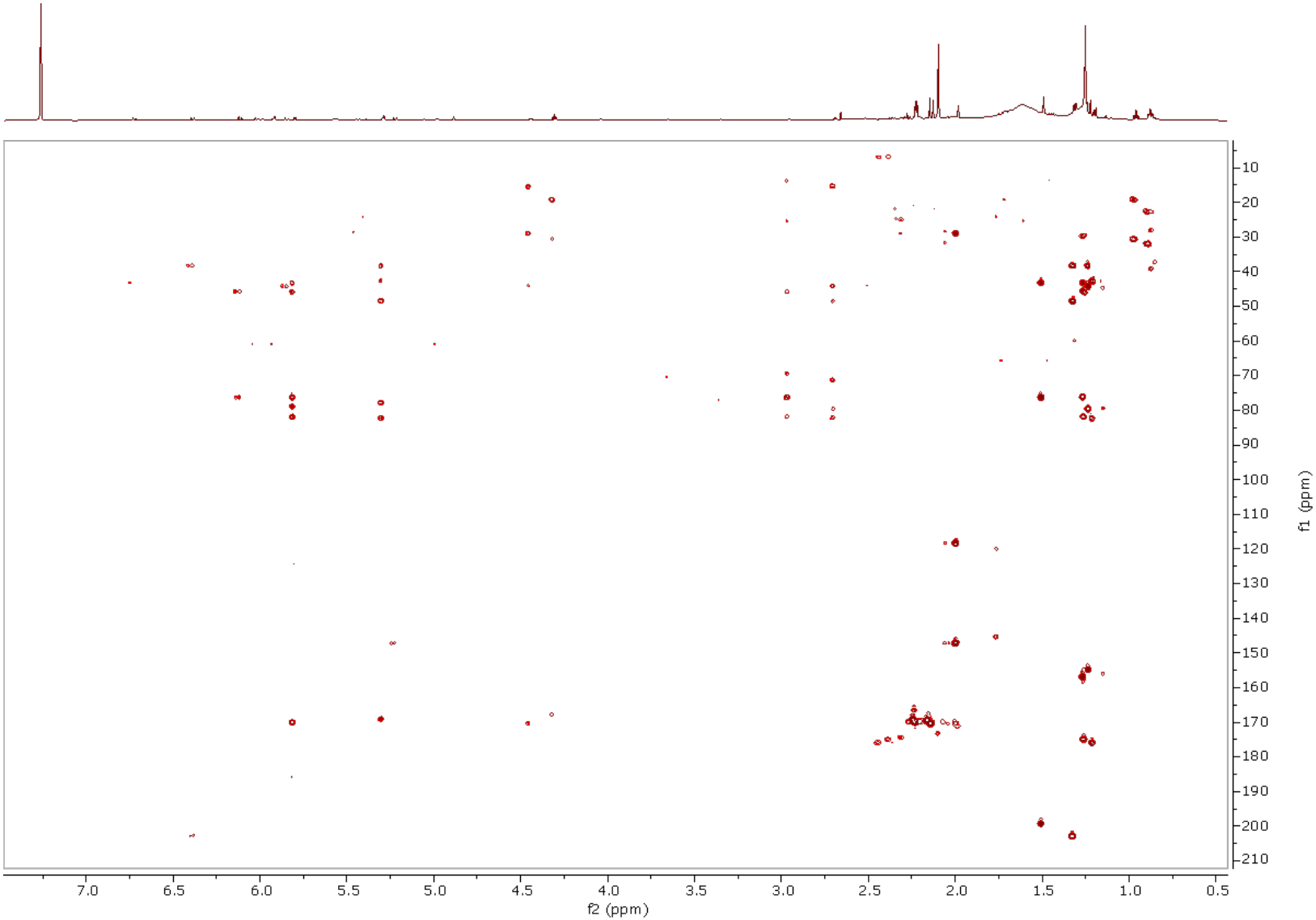
^{1}H-^{13}C HSQC spectrum of 11-hydroxyptilosarcenone in (**6**) CD_3OD at 600 MHz.



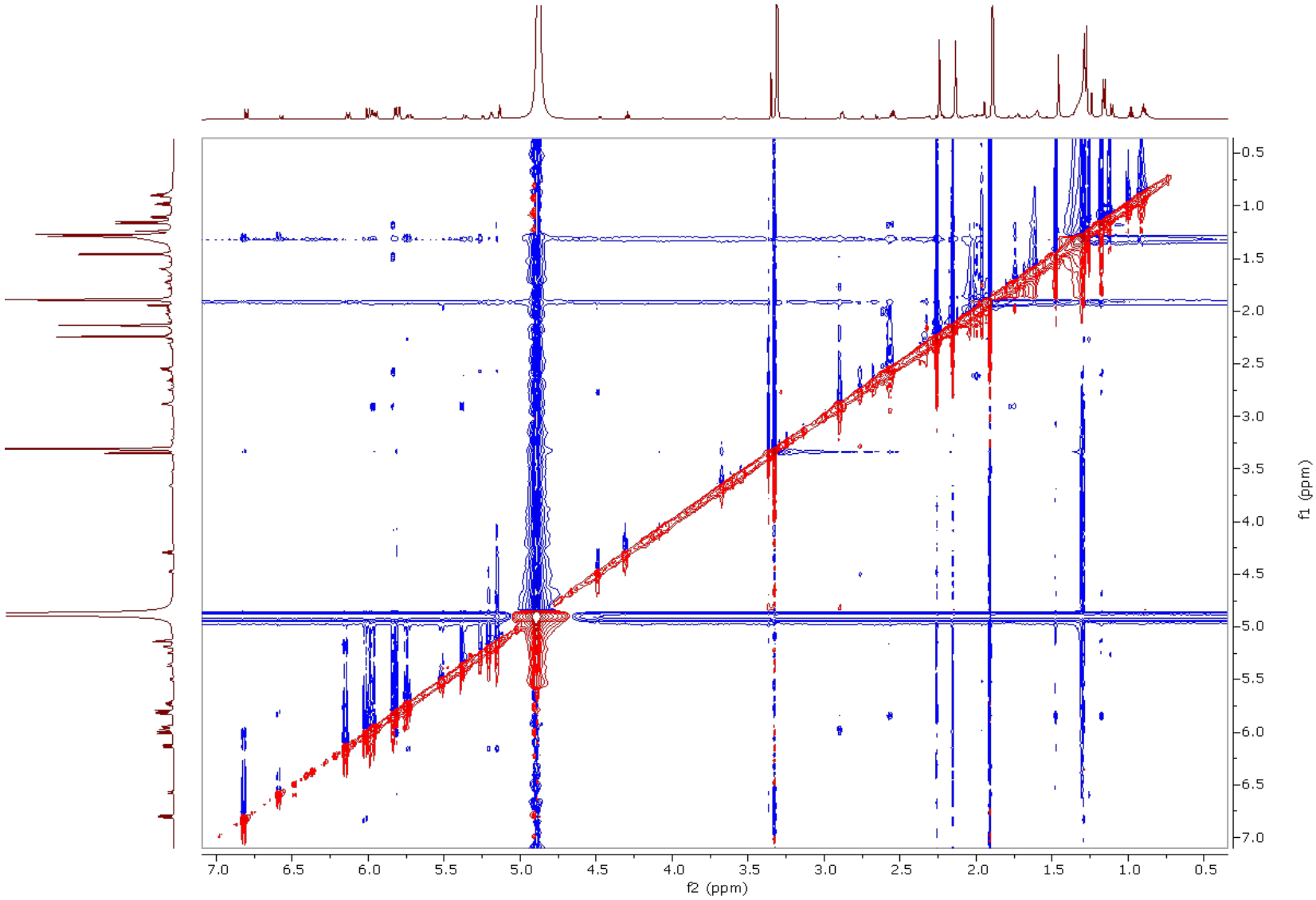
^{1}H-^{13}C HSQC spectrum of 11-hydroxyptilosarcenone (**6**) in $CDCl_3$ at 600 MHz.



^{1}H-^{13}C HMBC spectrum of 11-hydroxyptilosarcenone (**6**) in CD_3OD at 600 MHz.

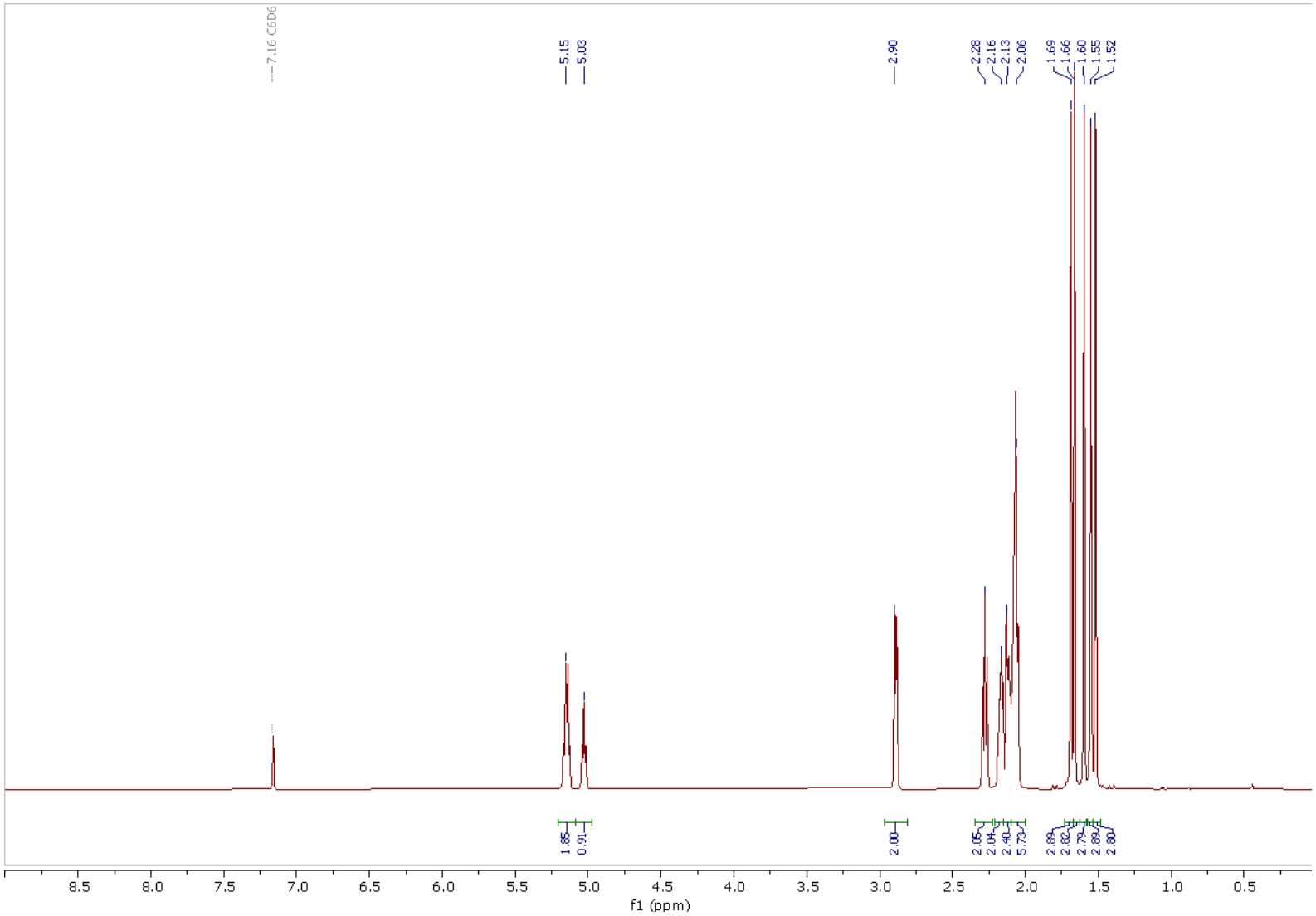


^{1}H-^{13}C HMBC spectrum of 11-hydroxyptilosarcenone (**6**) in $CDCl_3$ at 600 MHz.

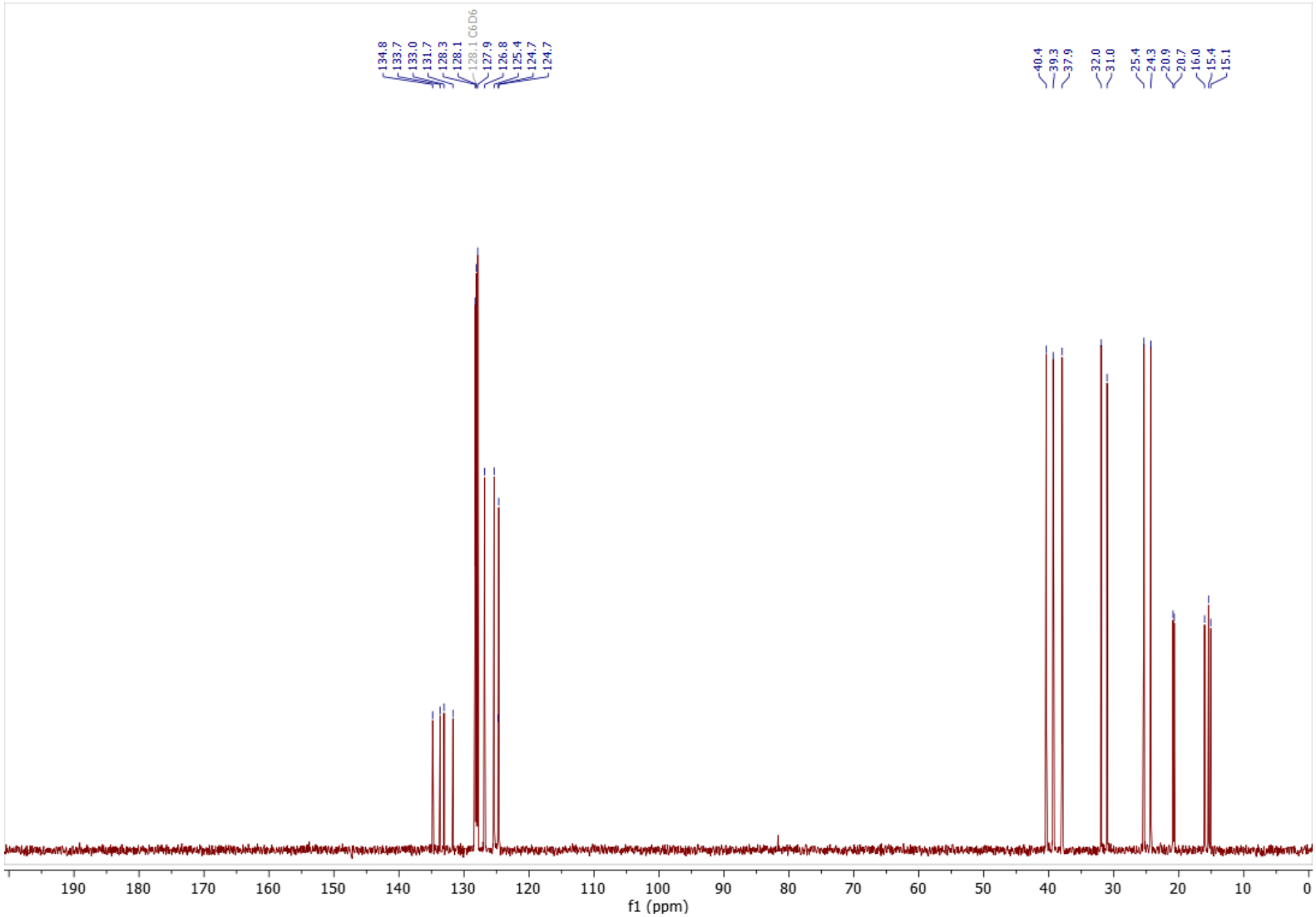


H-^{1}H NOESY spectrum of 11-hydroxyptilosarcenone (**6**) in CD_3OD at 600 MHz.

NMR spectra of cembrene B (10).

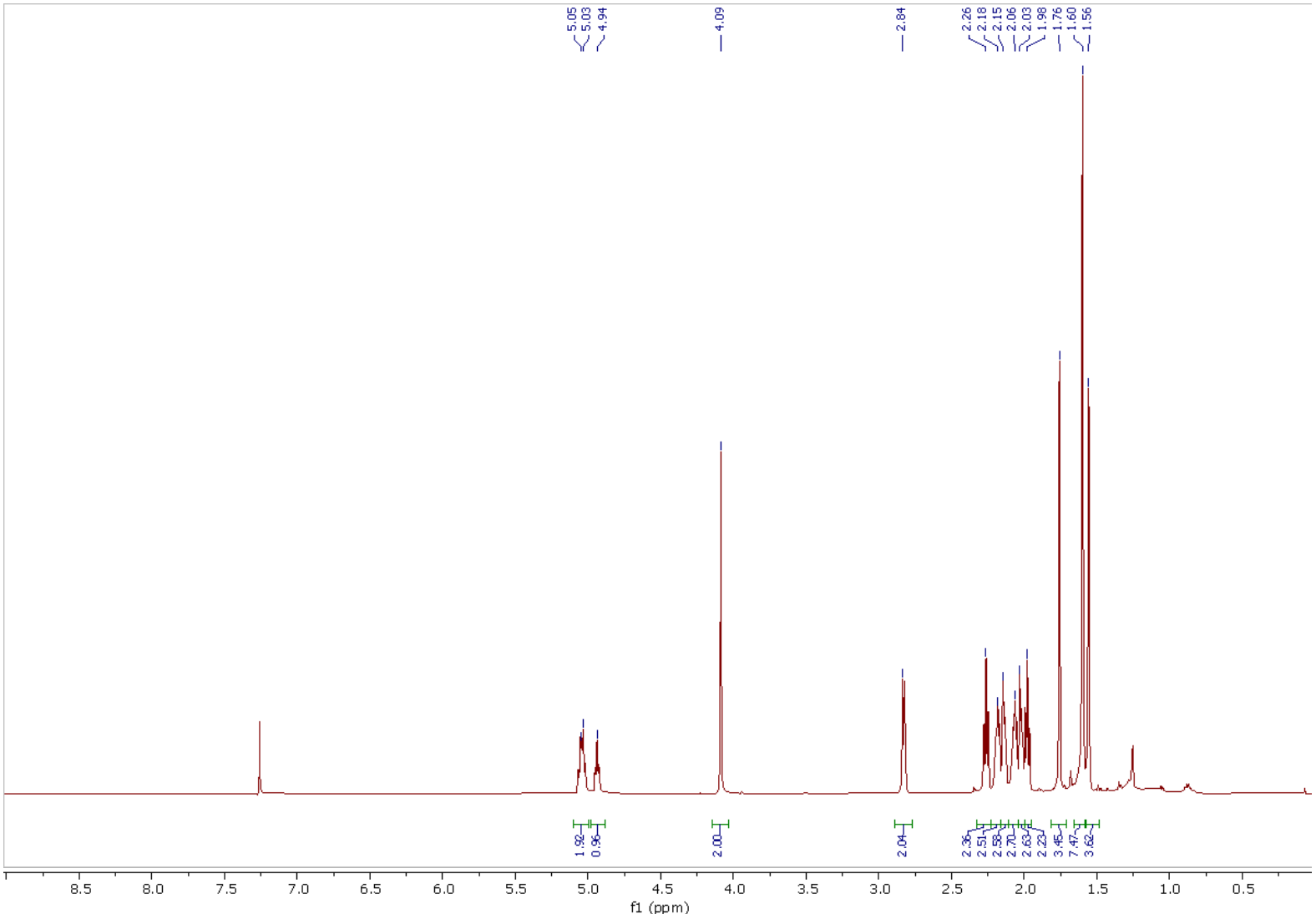


^{1}H NMR spectrum of **10** in benzene-d_6 at 500 MHz.

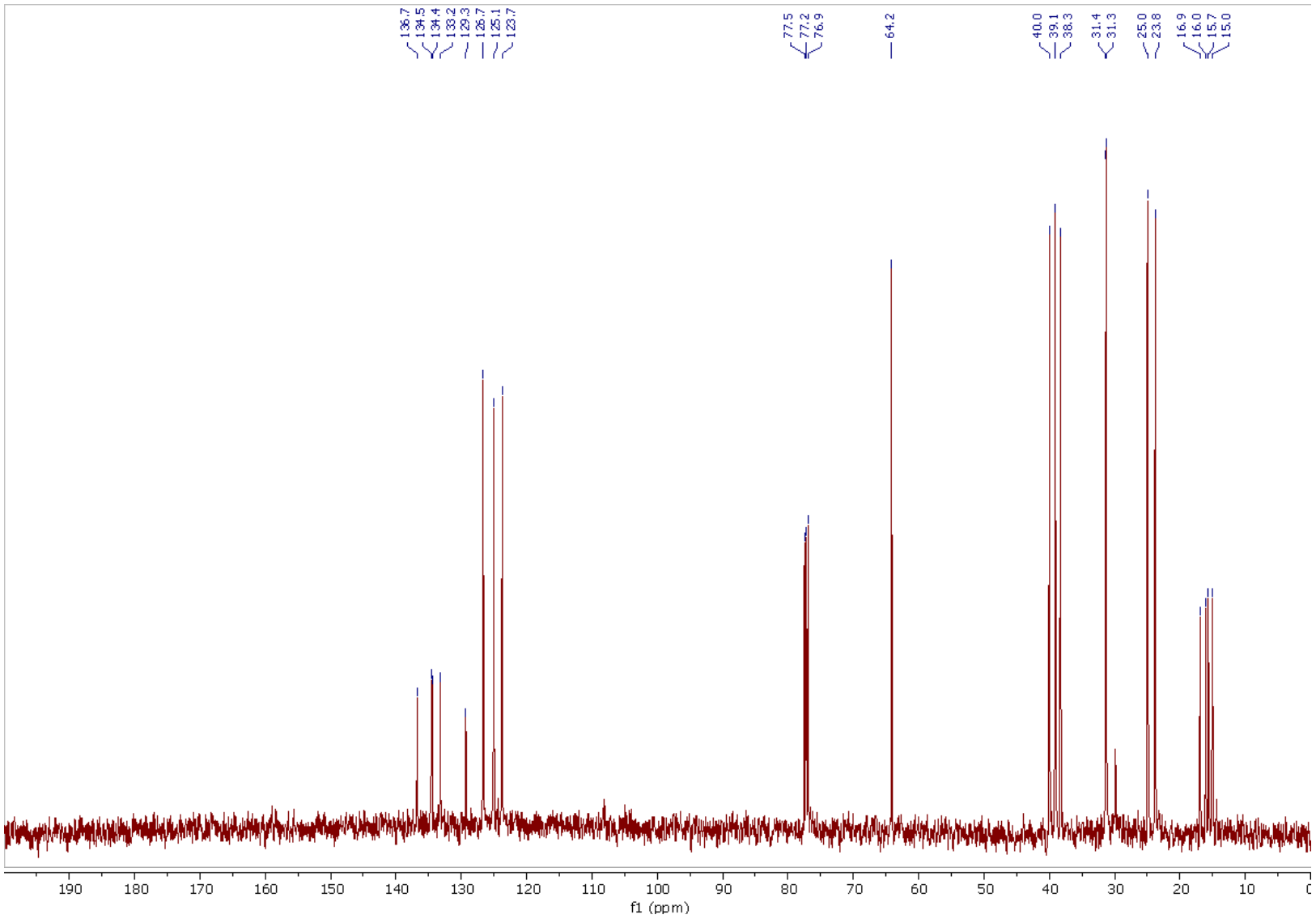


^{13}C NMR spectrum of **10** in benzene-d_6 at 125 MHz.

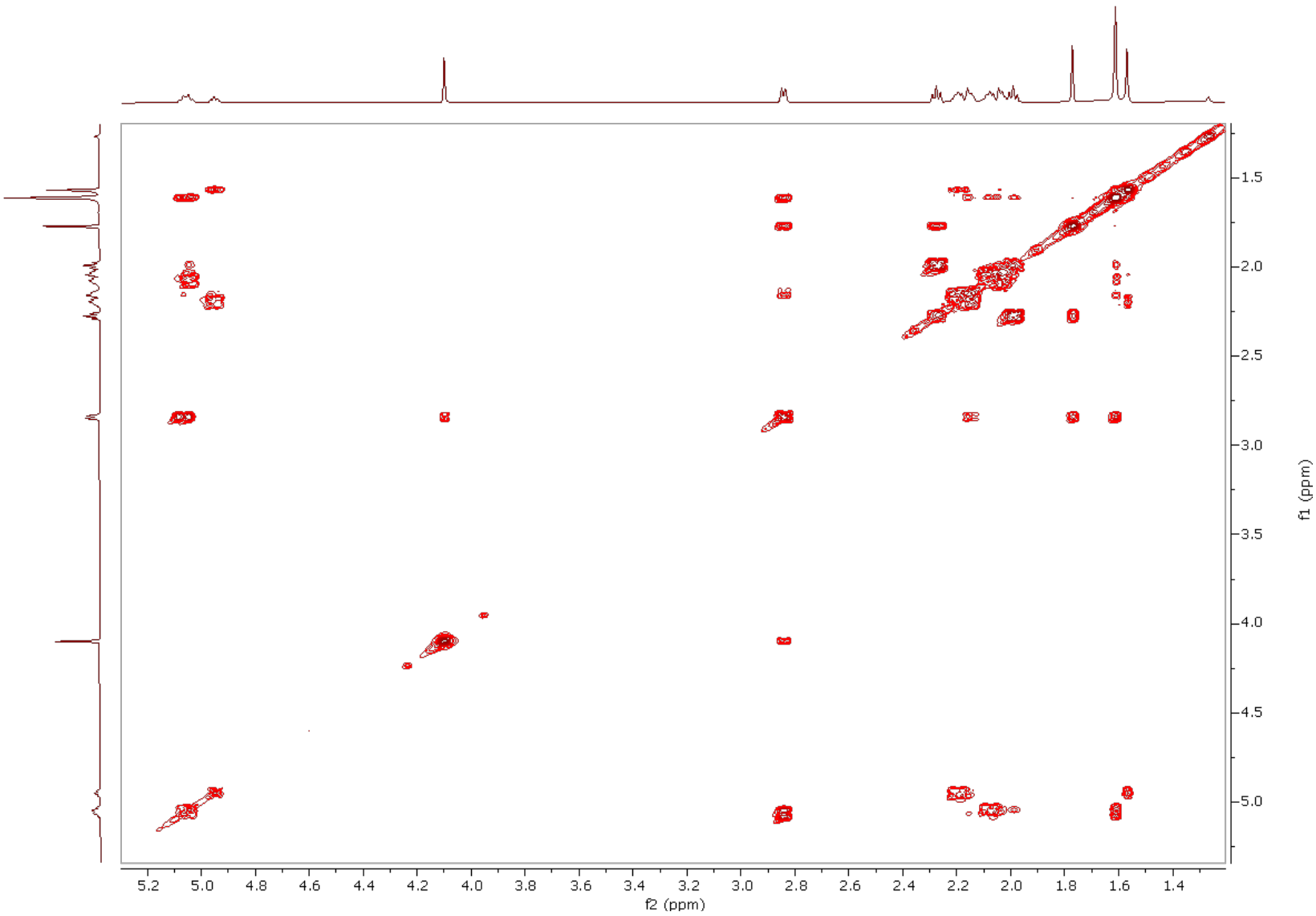
NMR spectra of 19-hydroxycembrene B (11).



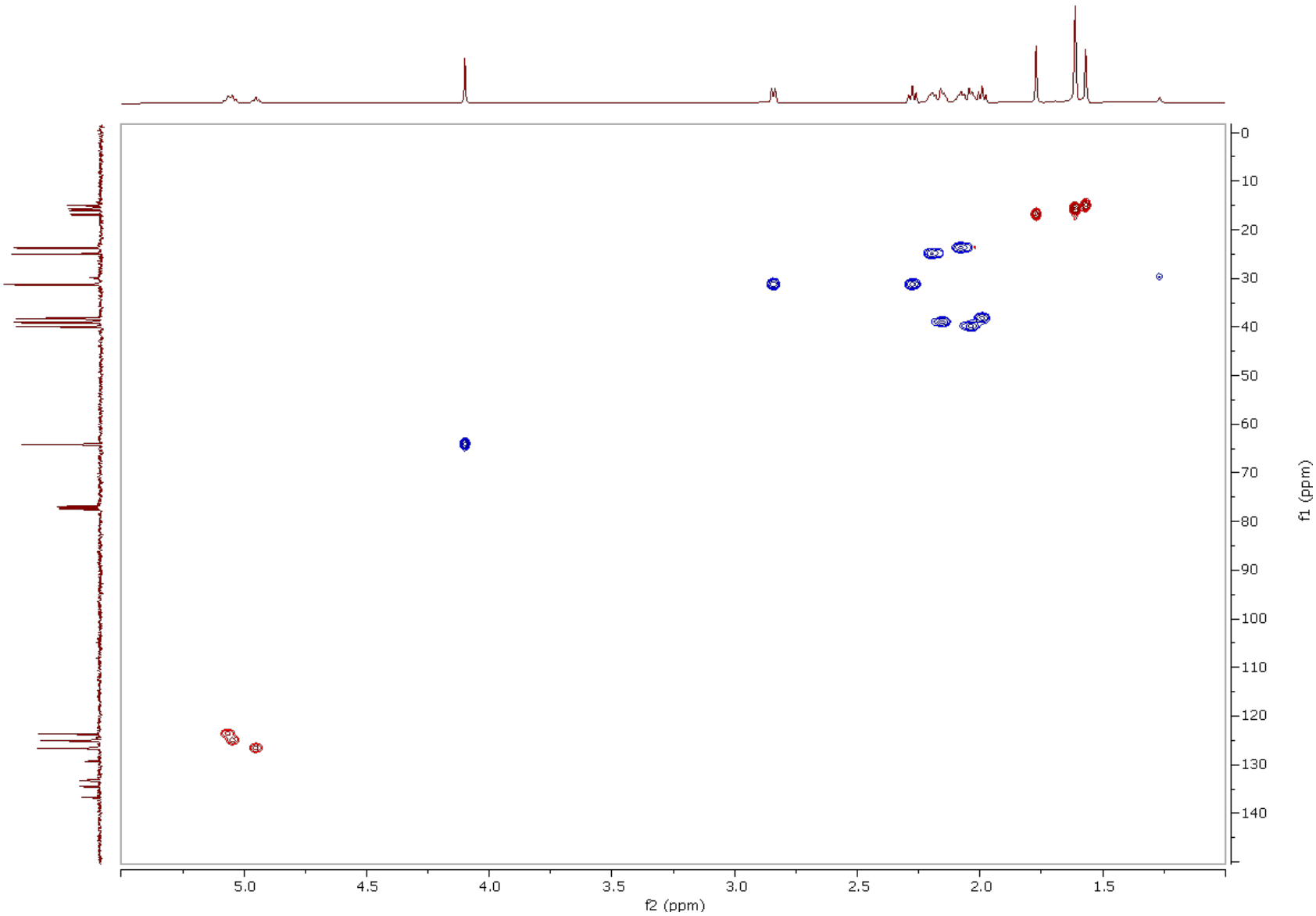
^{1}H NMR spectrum of **11** in $CDCl_3$ at 500 MHz.



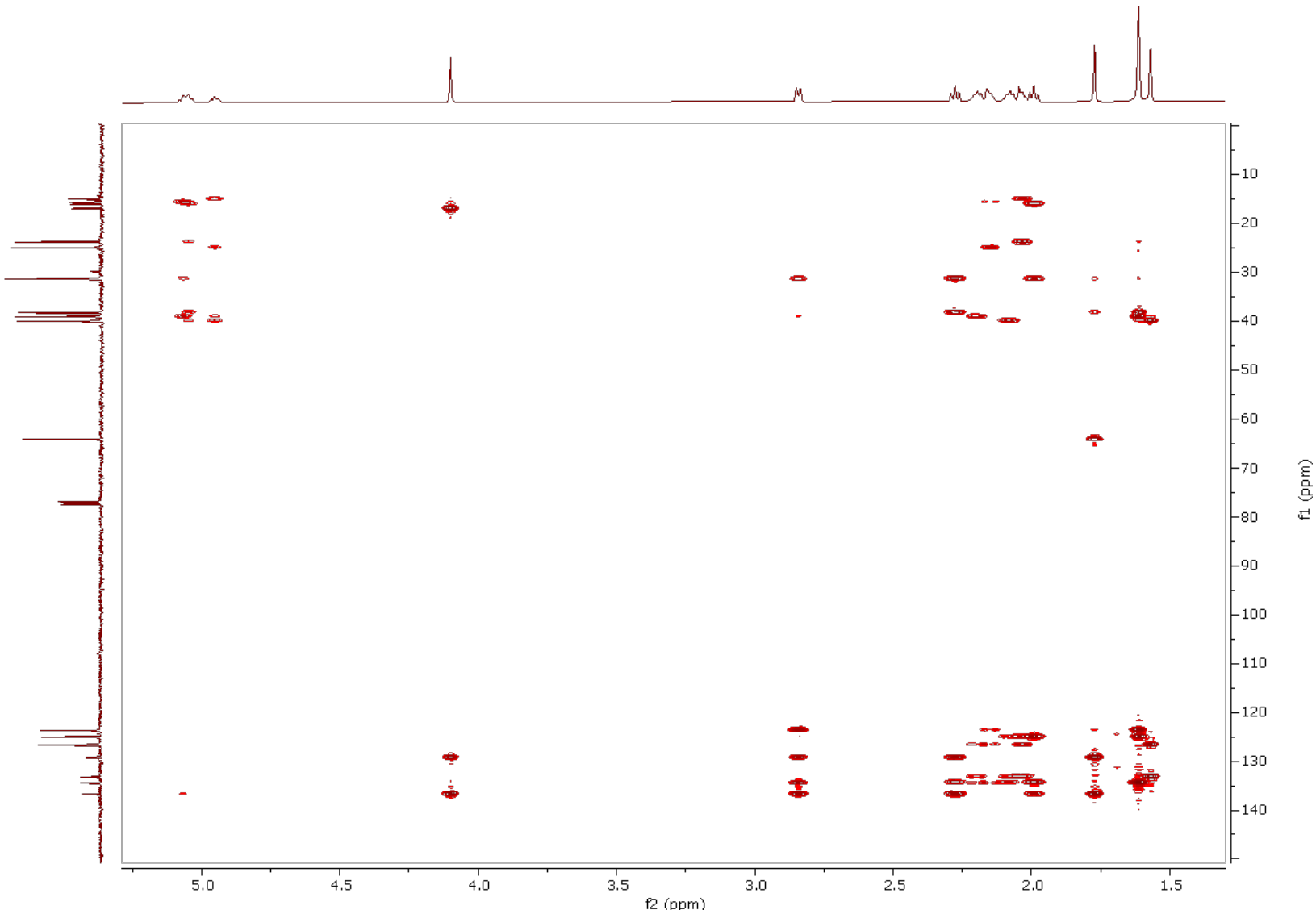
^{13}C NMR spectrum of **11** in $CDCl_3$ at 125 MHz.



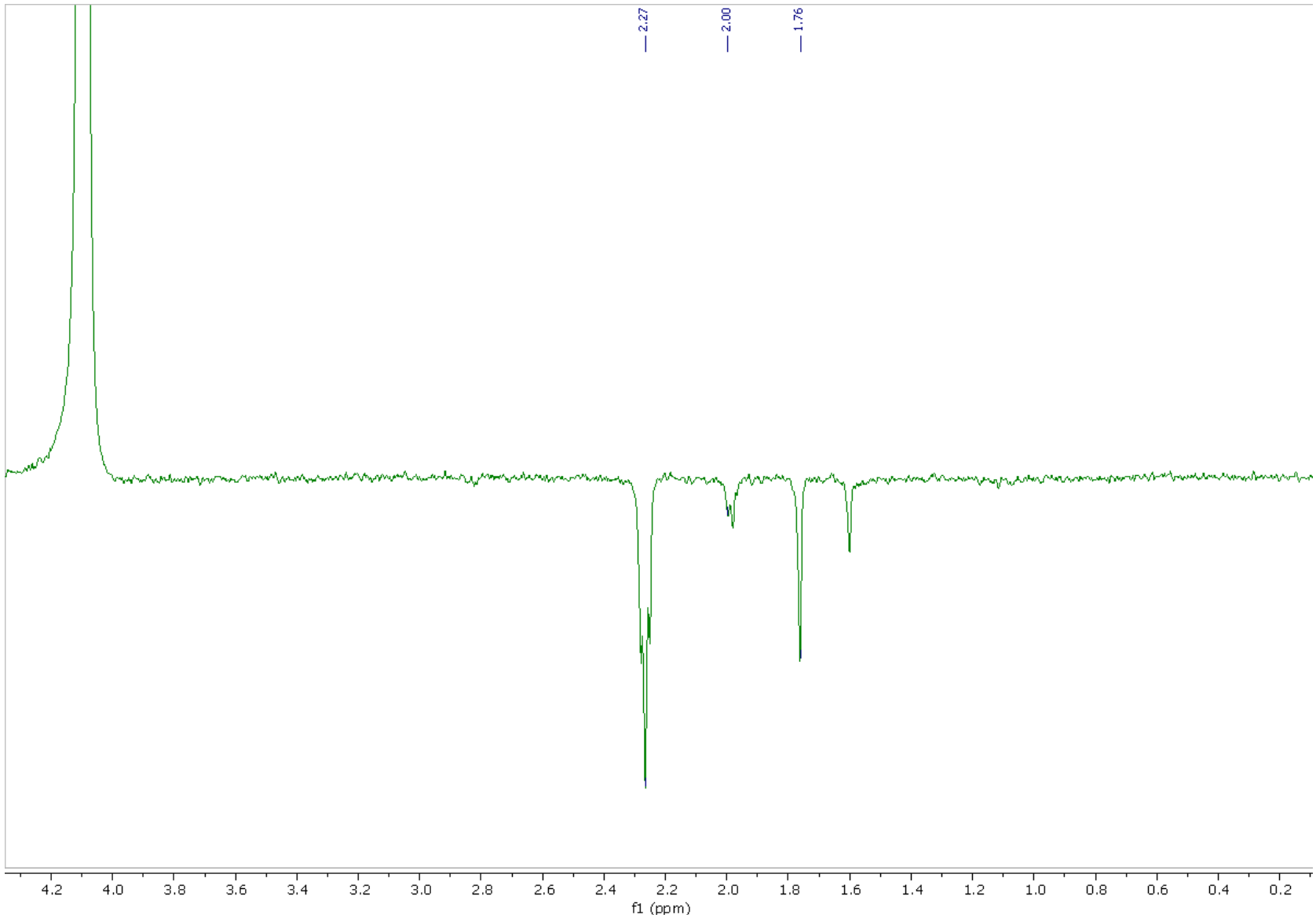
^{1}H-^{1}H COSY spectrum of **11** in $CDCl_3$ at 500 MHz.



^{1}H-^{13}C HSQC spectrum of **11** in $CDCl_3$ at 500 MHz.

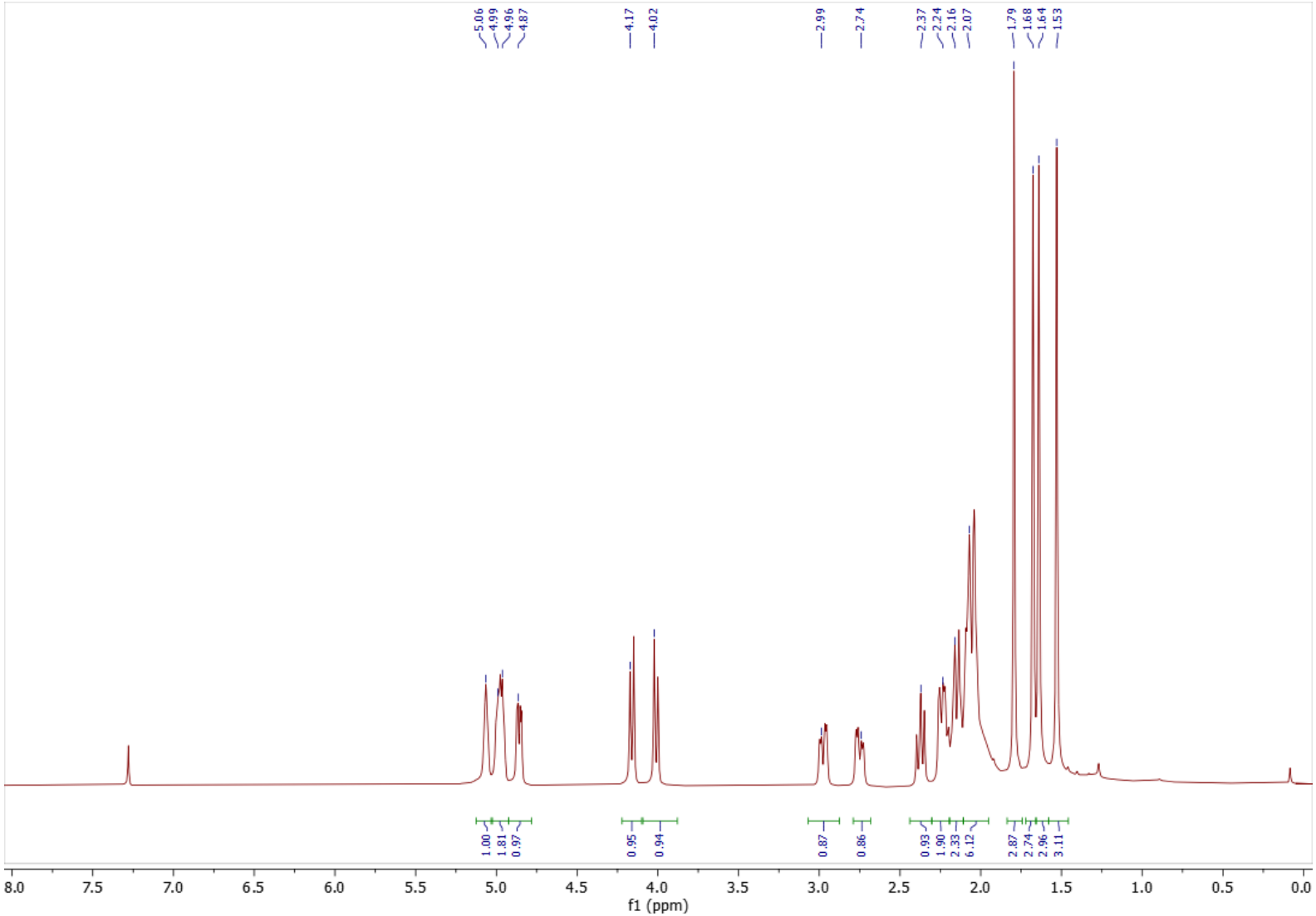


^{1}H-^{13}C HMBC spectrum of **11** in $CDCl_3$ at 500 MHz.

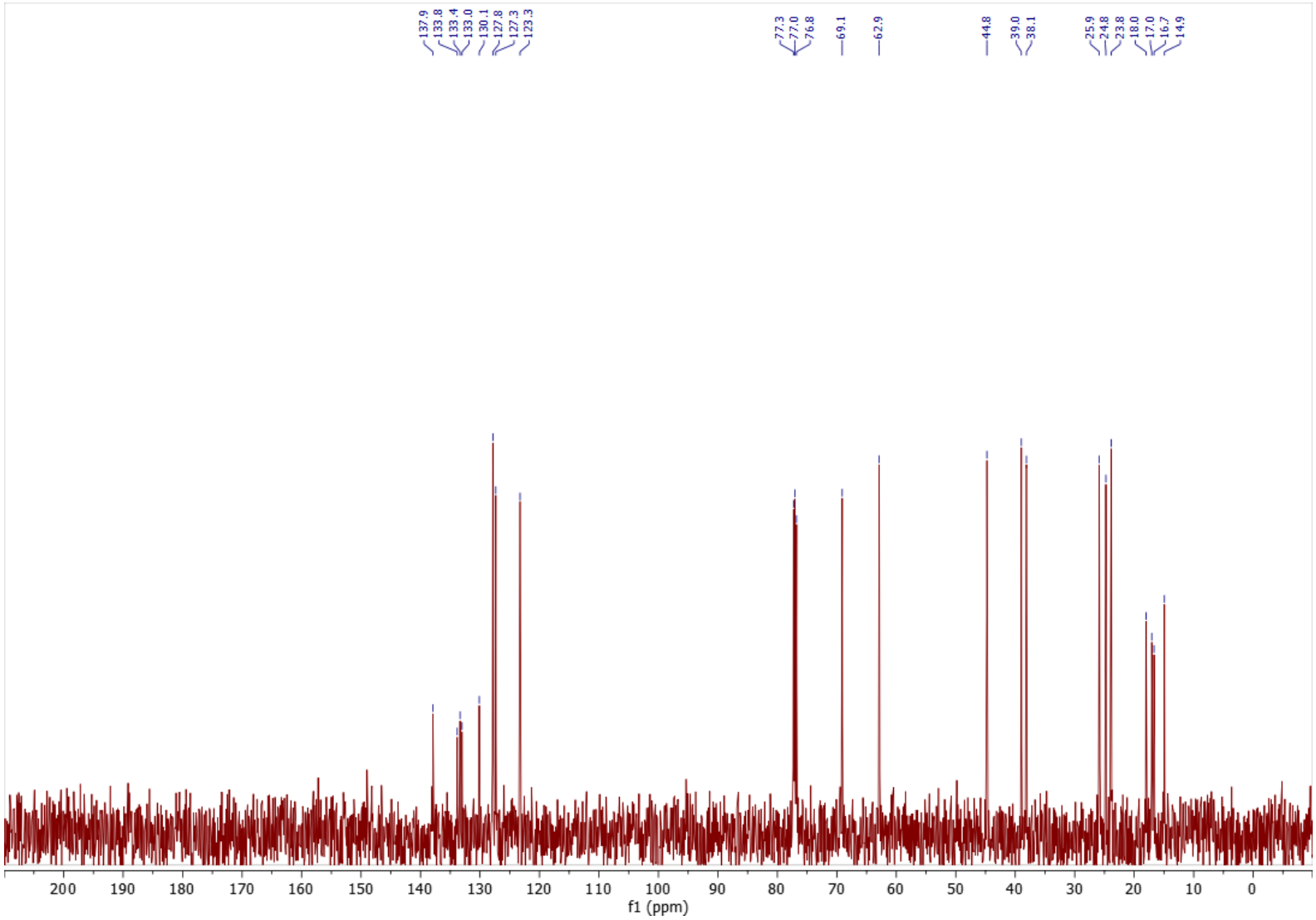


^{1}H 1D DPFGSE NOESY spectrum of **11** in $CDCl_3$ at 500 MHz with selective excitation at 4.09 ppm.

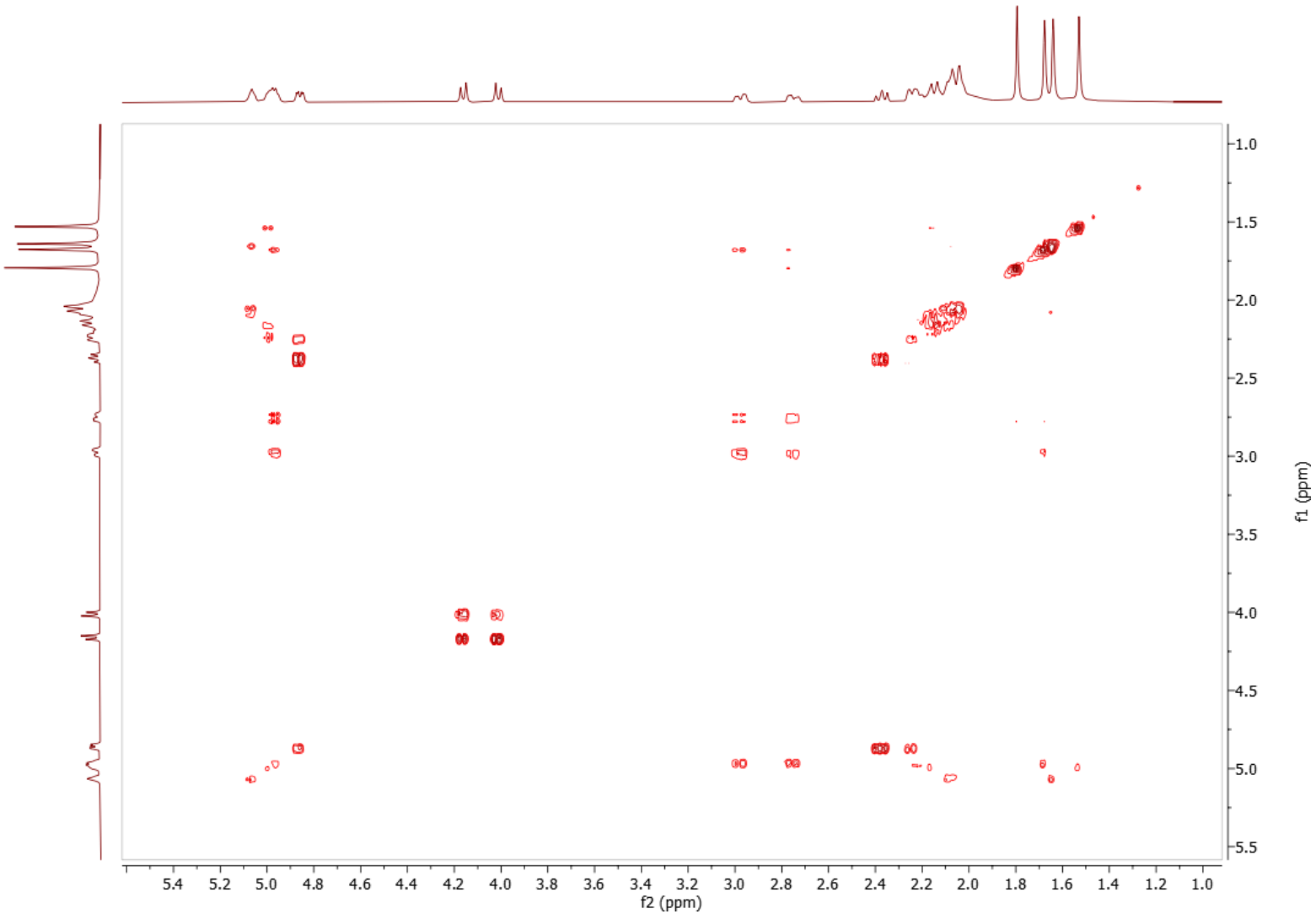
NMR spectra of dihydroxycembrene B (12).



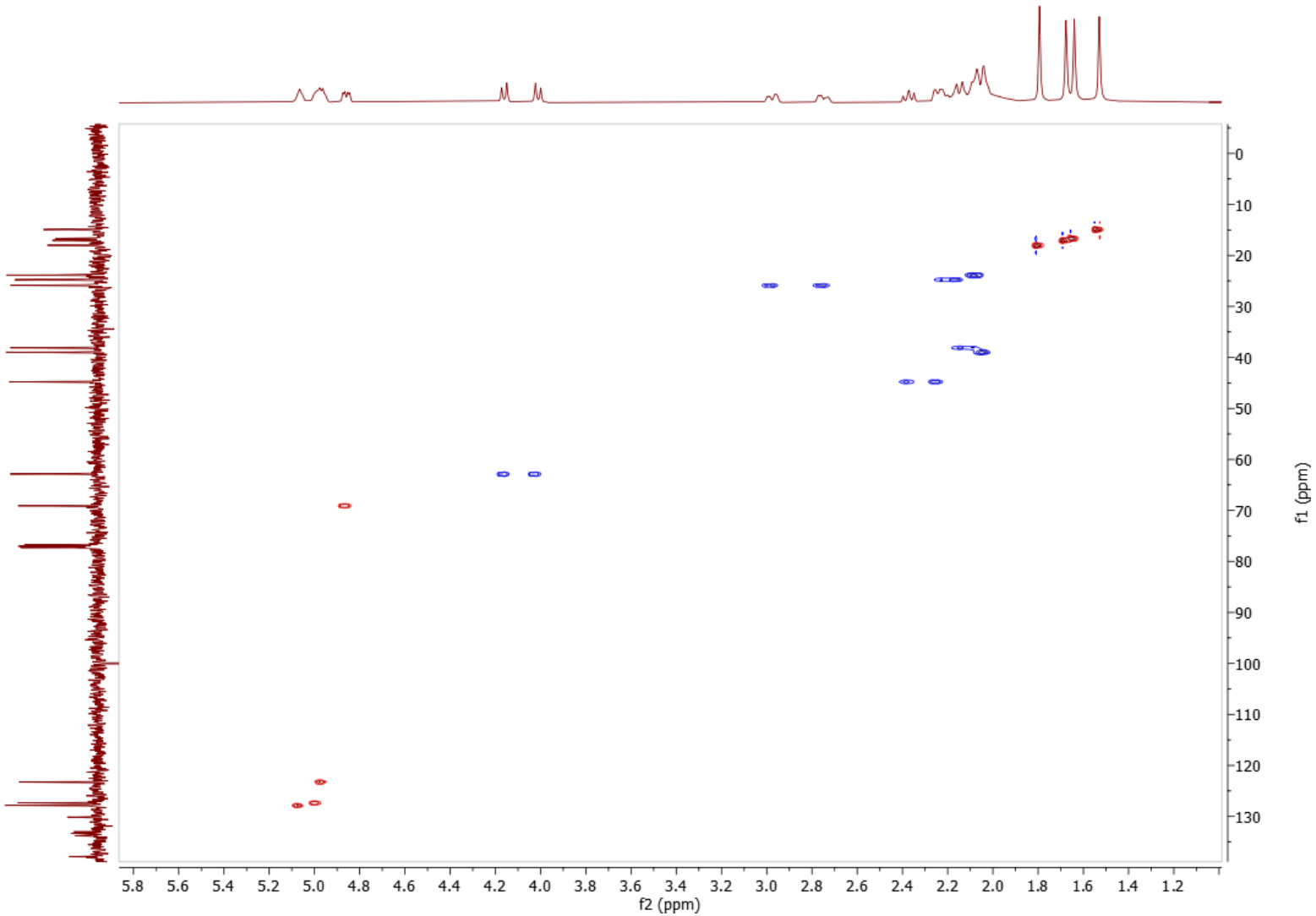
^{1}H NMR spectrum of **12** in $CDCl_3$ at 500 MHz.



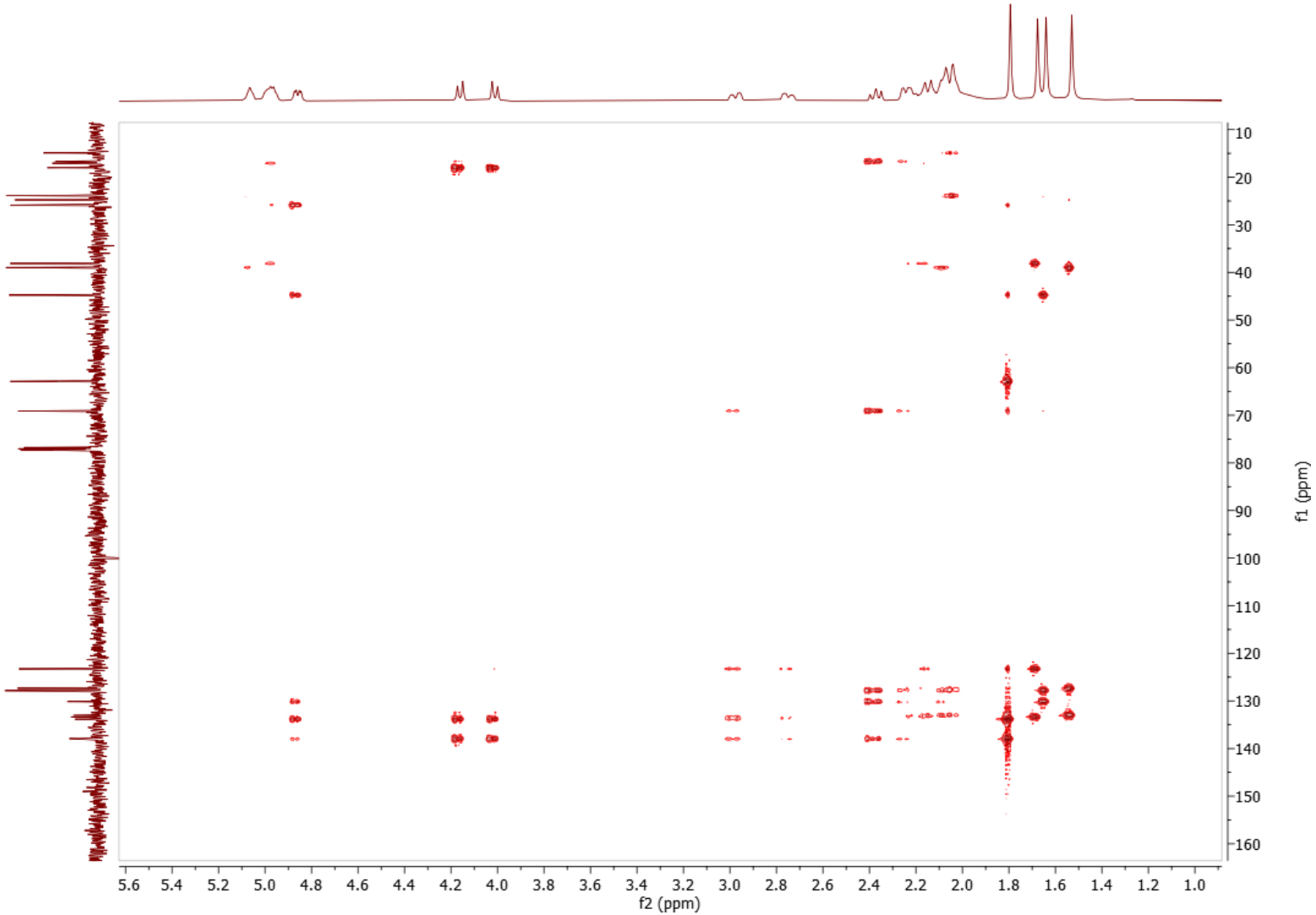
^{13}C NMR spectrum of **12** in $CDCl_3$ at 125 MHz.



^{1}H-^{1}H COSY spectrum of **12** in $CDCl_3$ at 500 MHz.

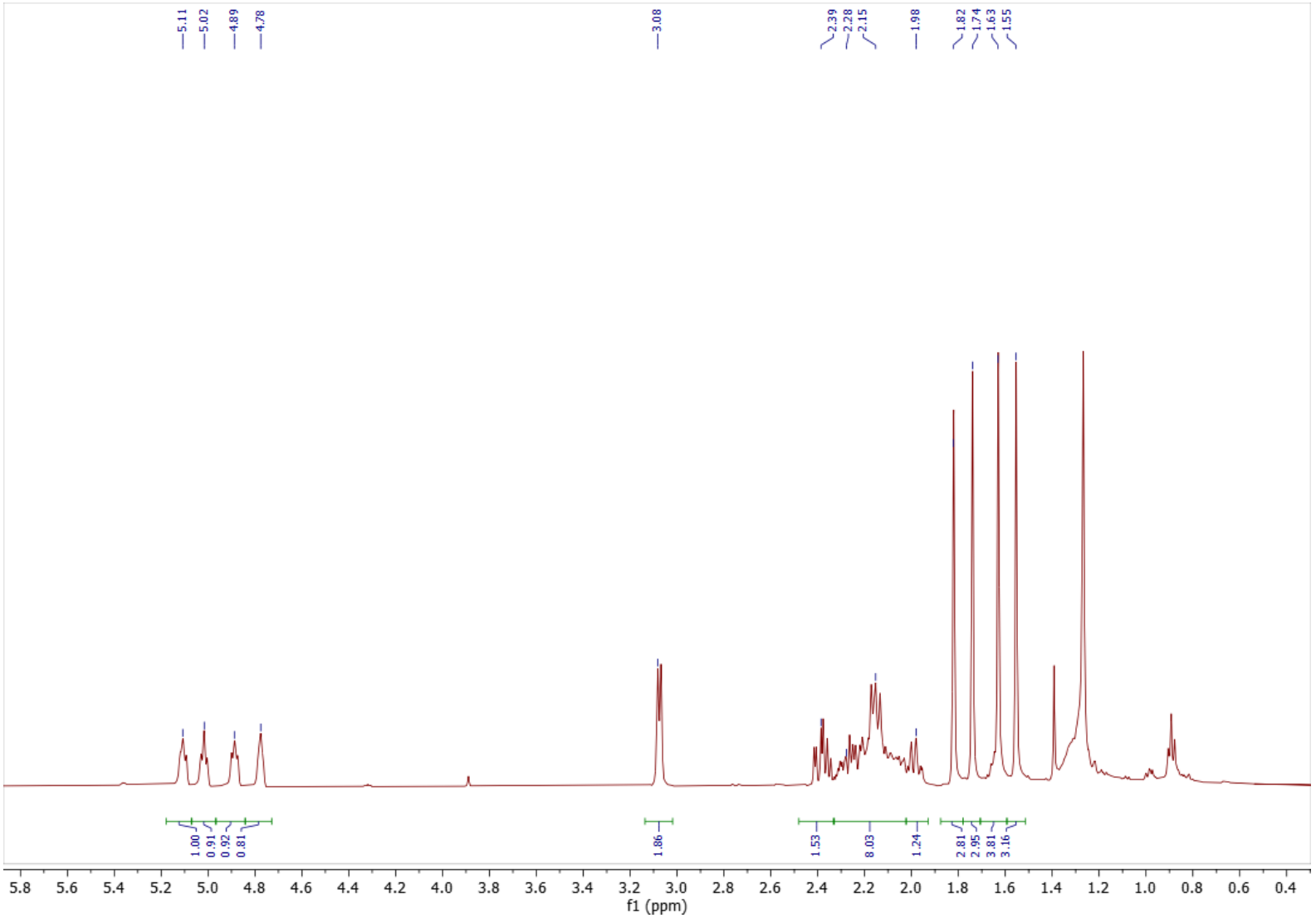


^{1}H-^{13}C HSQC spectrum of **12** in $CDCl_3$ at 500 MHz.

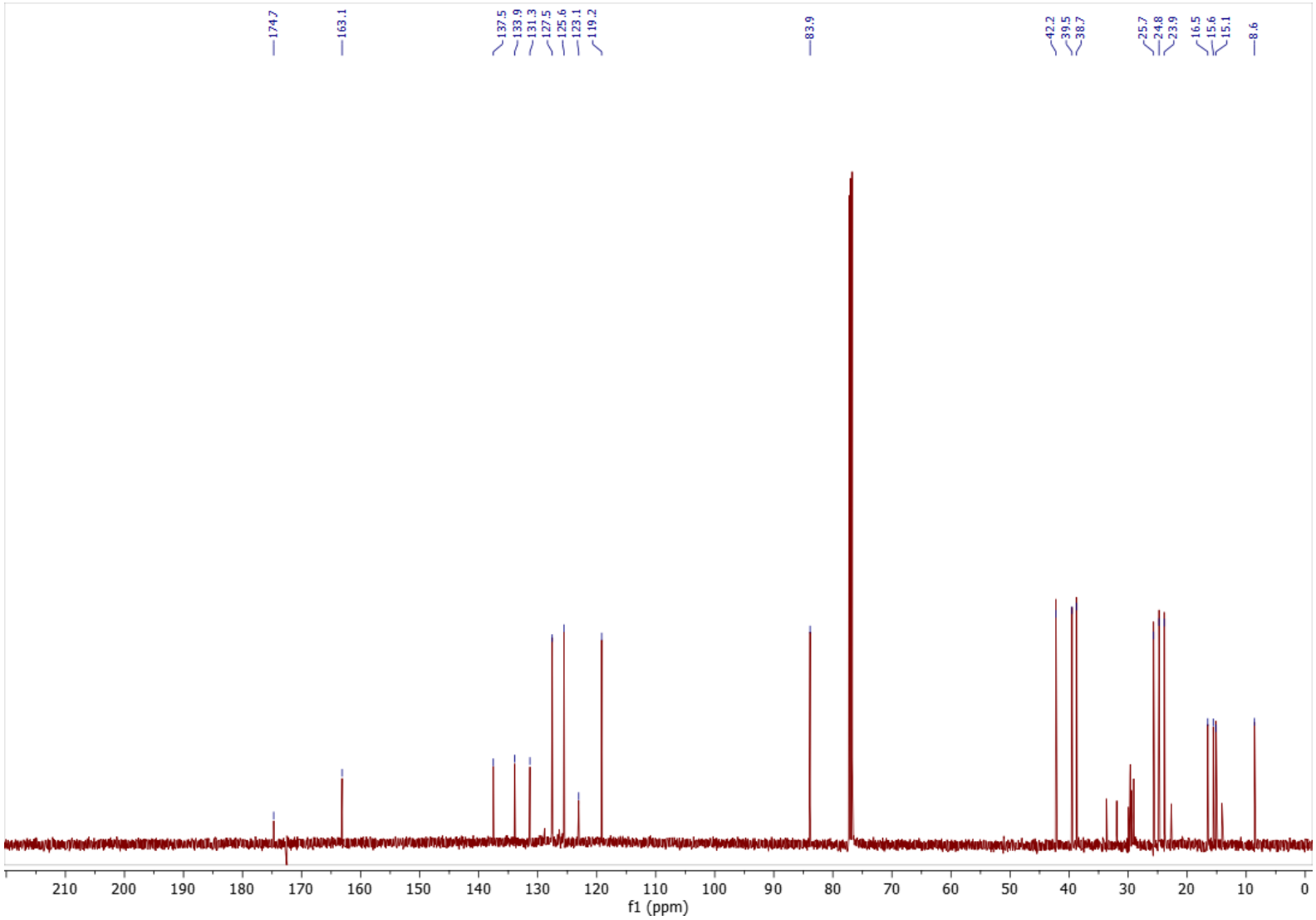


^{1}H-^{13}C HMBC spectrum of **12** in $CDCl_3$ at 500 MHz.

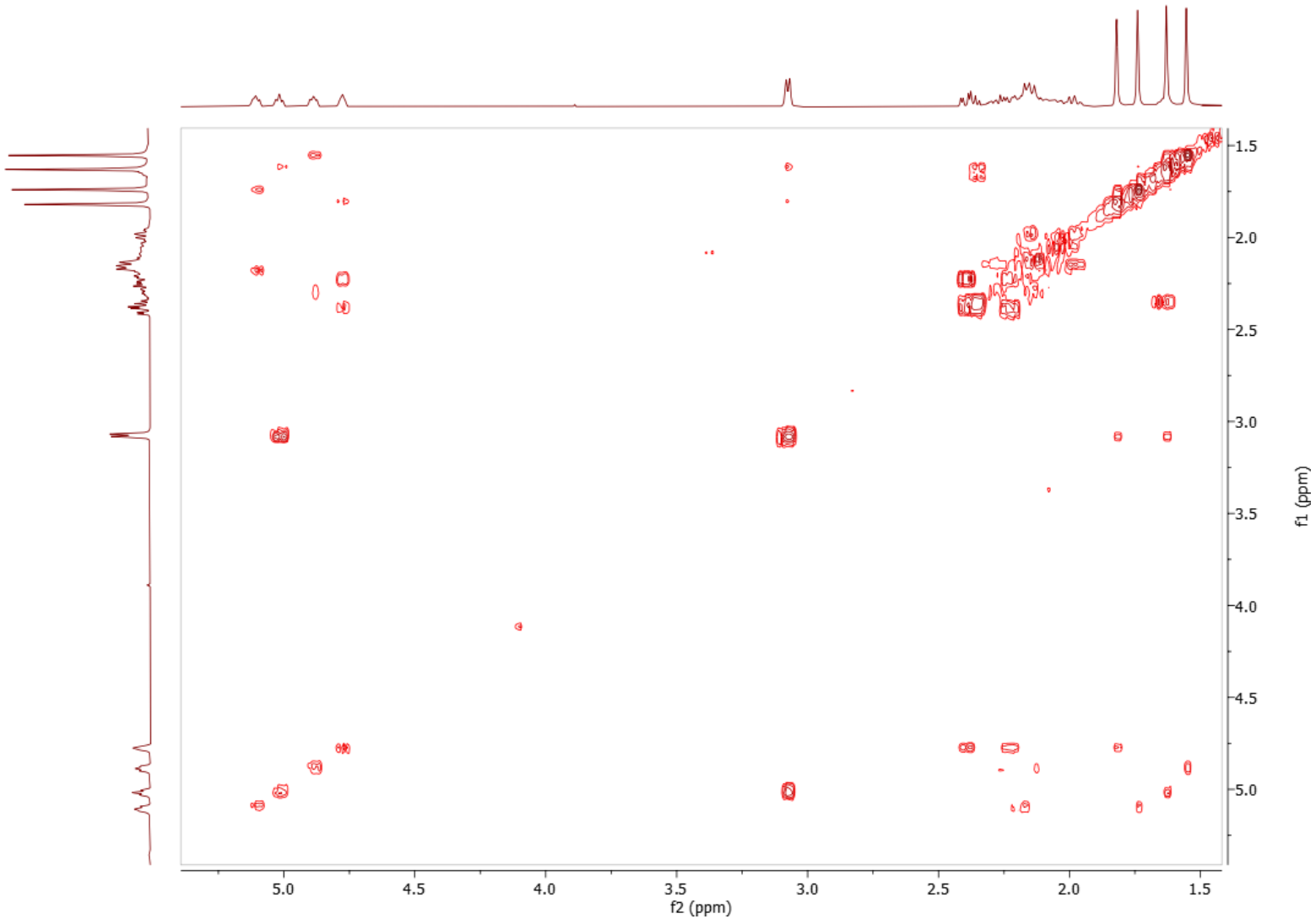
NMR spectra of cembrene B g-lactone (13).



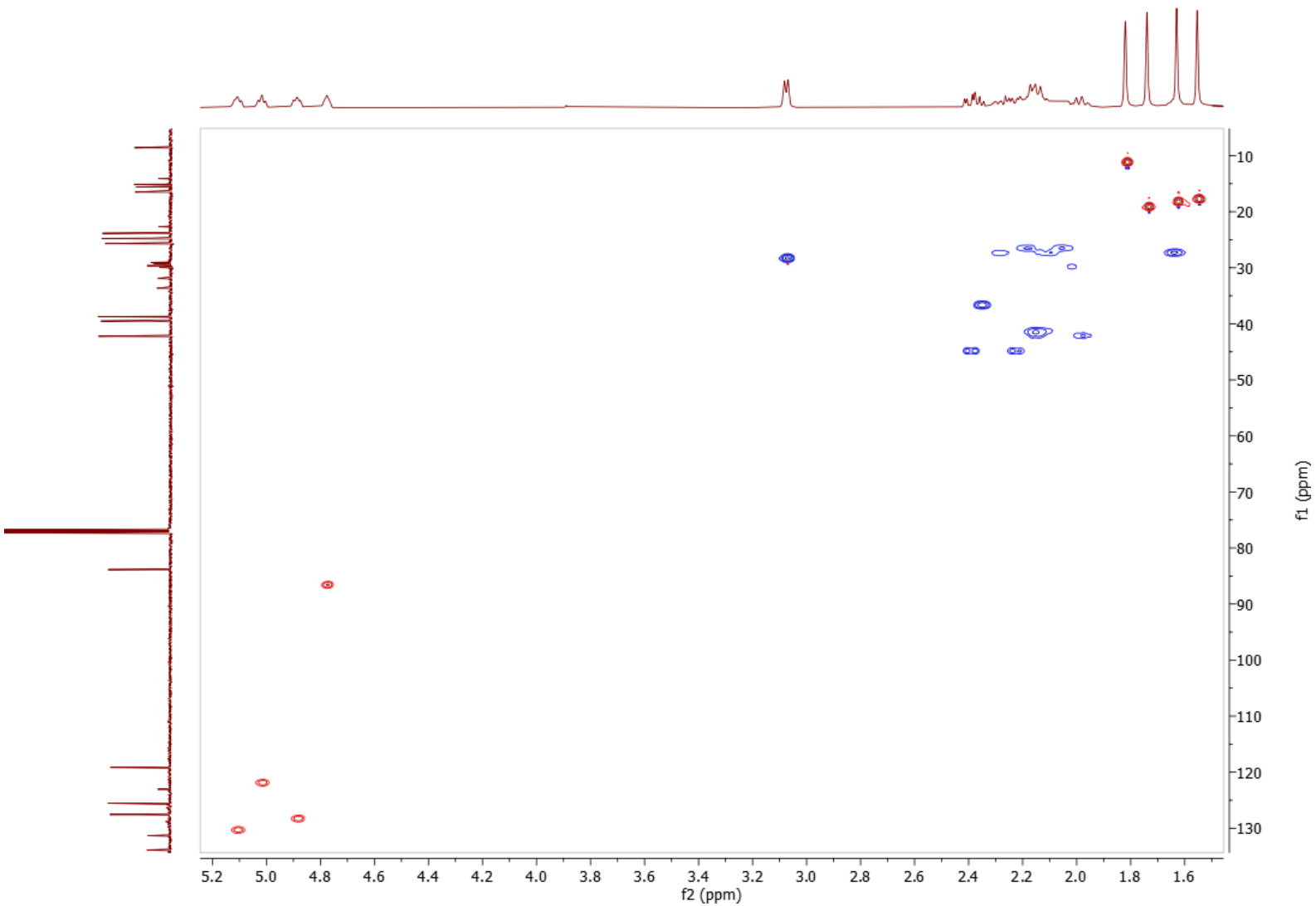
^{1}H NMR spectrum of **13** in $CDCl_3$ at 500 MHz.



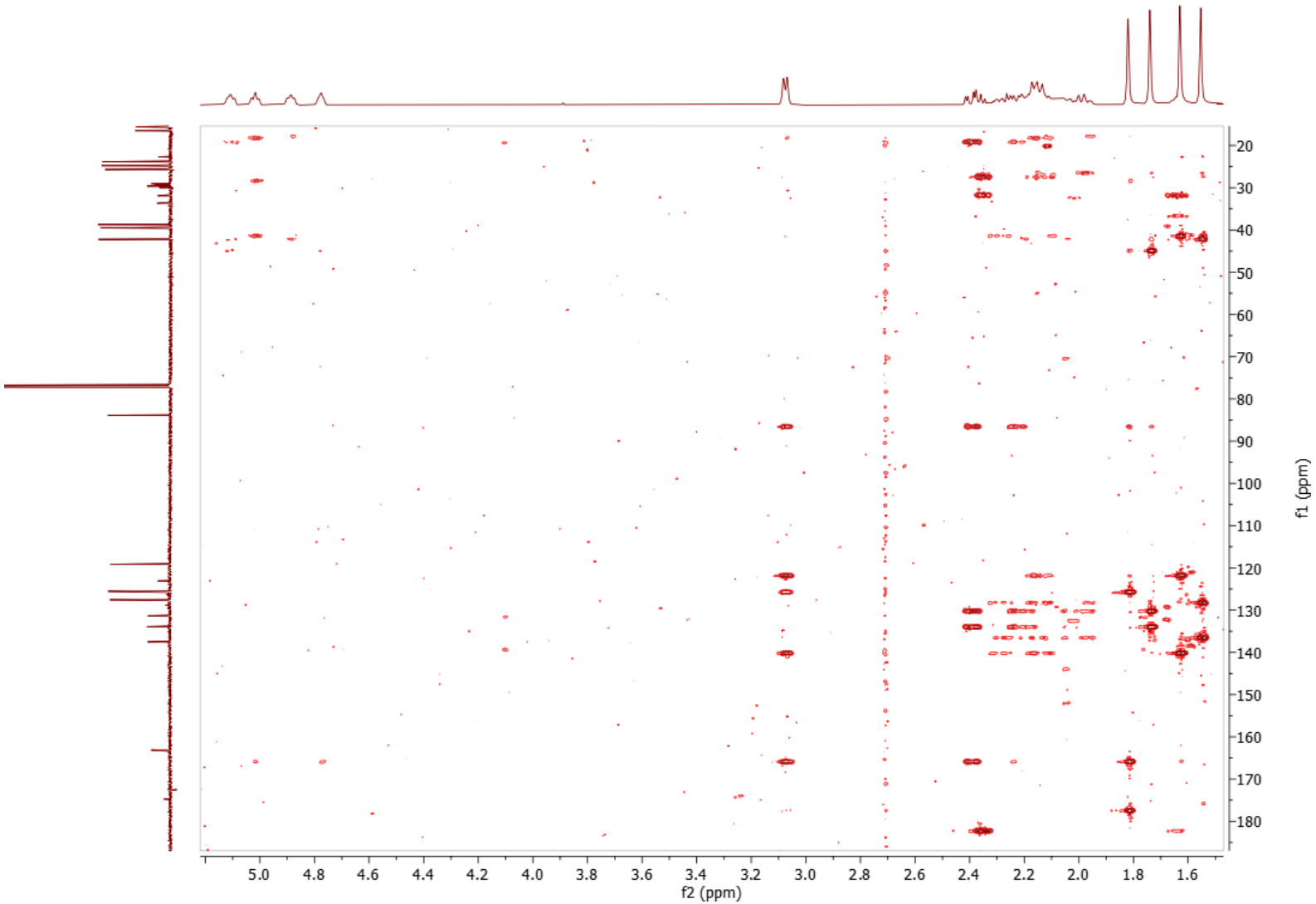
^{13}C NMR spectrum of **13** in $CDCl_3$ at 125 MHz.



^{1}H-^{1}H COSY spectrum of **13** in $CDCl_3$ at 500 MHz.

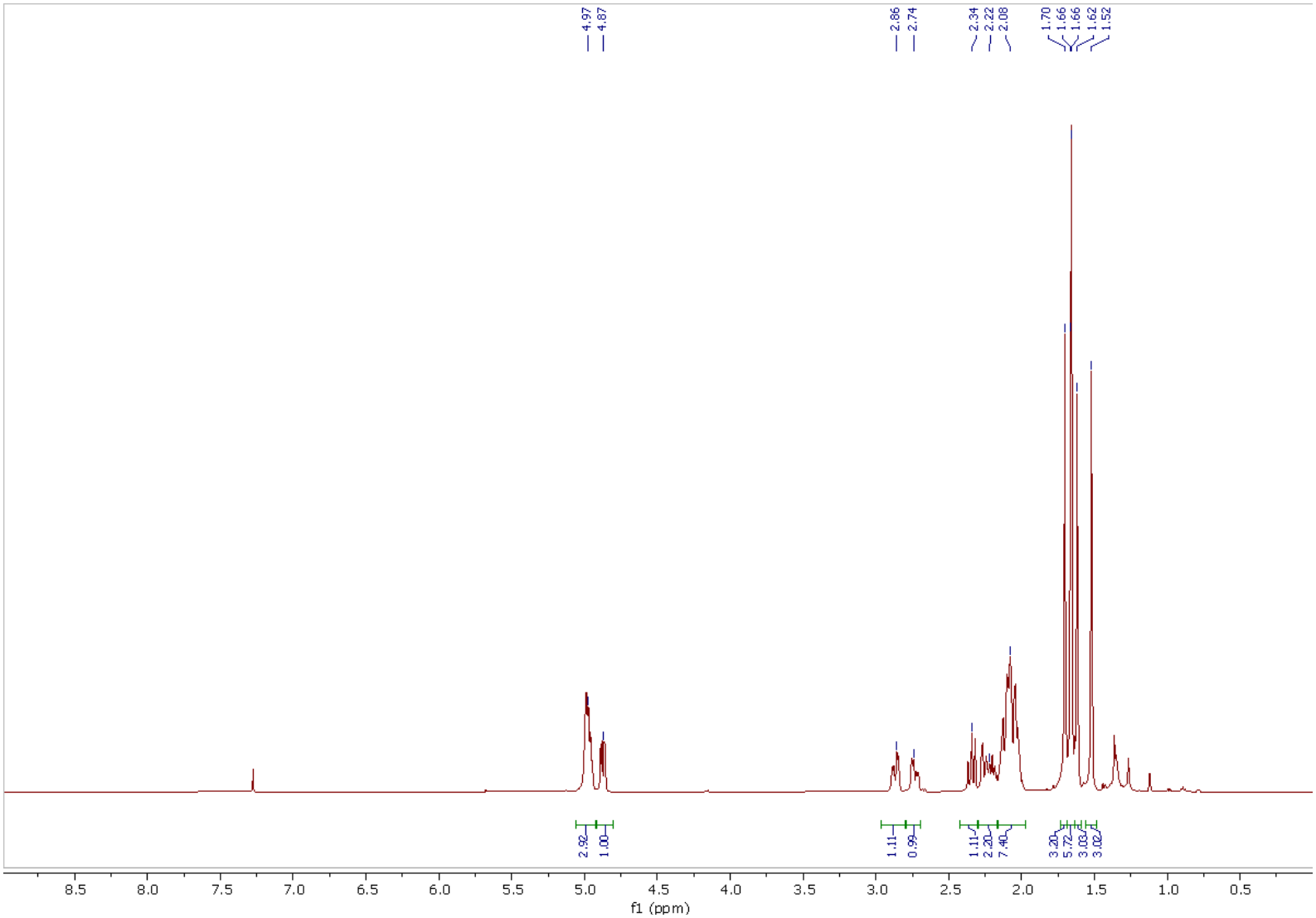


^{1}H-^{13}C HSQC spectrum of **13** in $CDCl_3$ at 500 MHz.

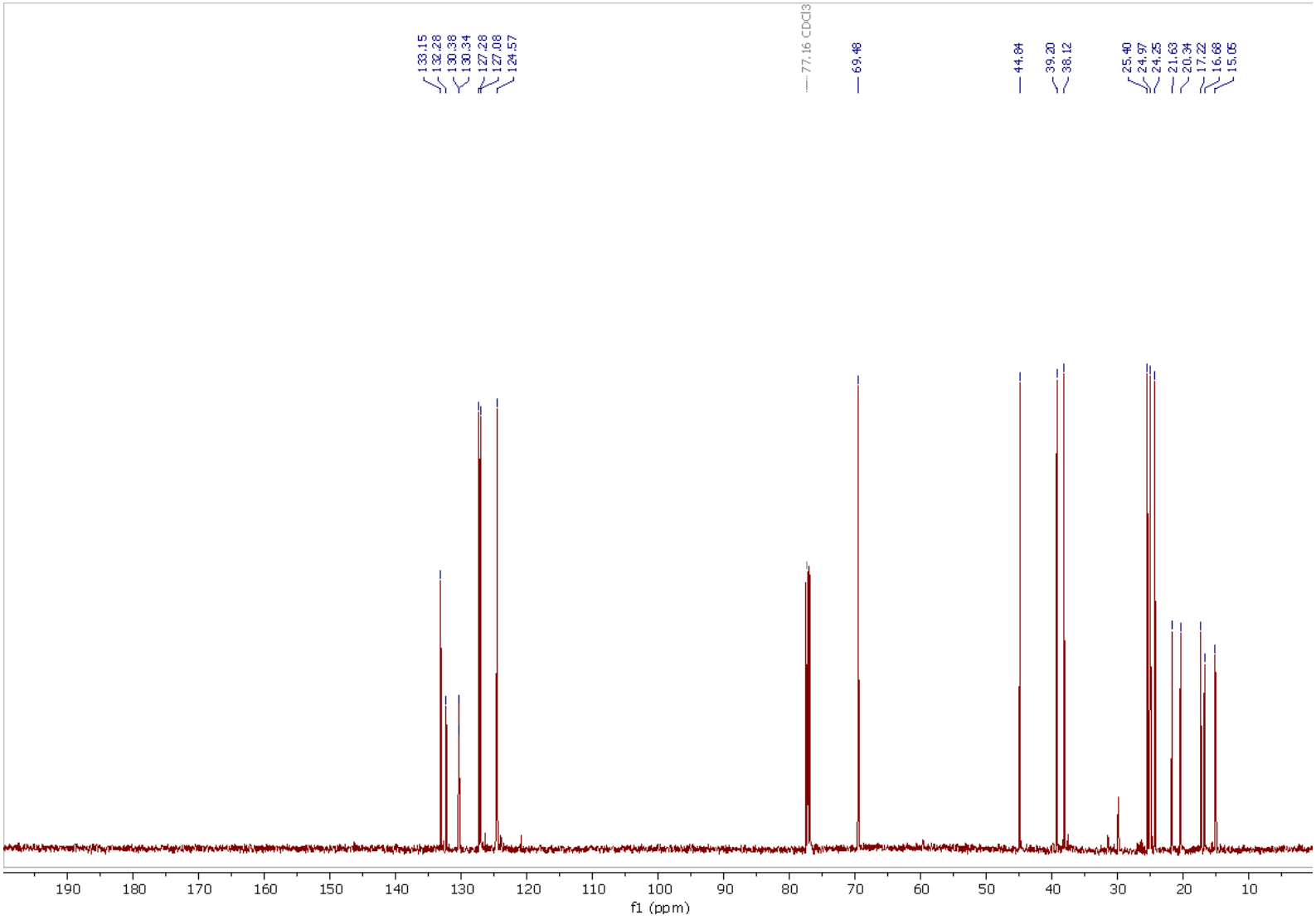


^{1}H-^{13}C HMBC spectrum of **13** in $CDCl_3$ at 500 MHz.

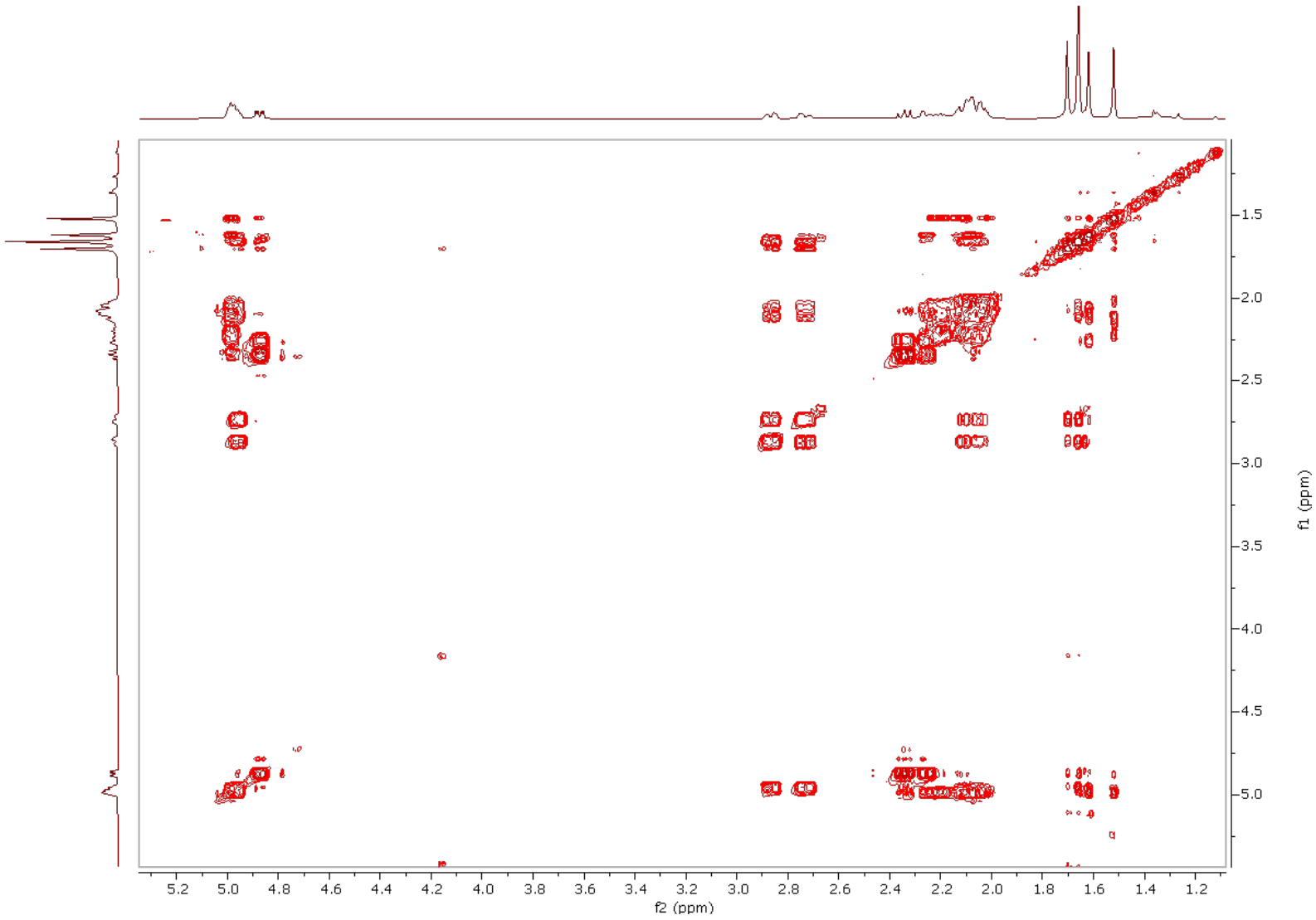
NMR spectra of 17-hydroxycembrene (S1).



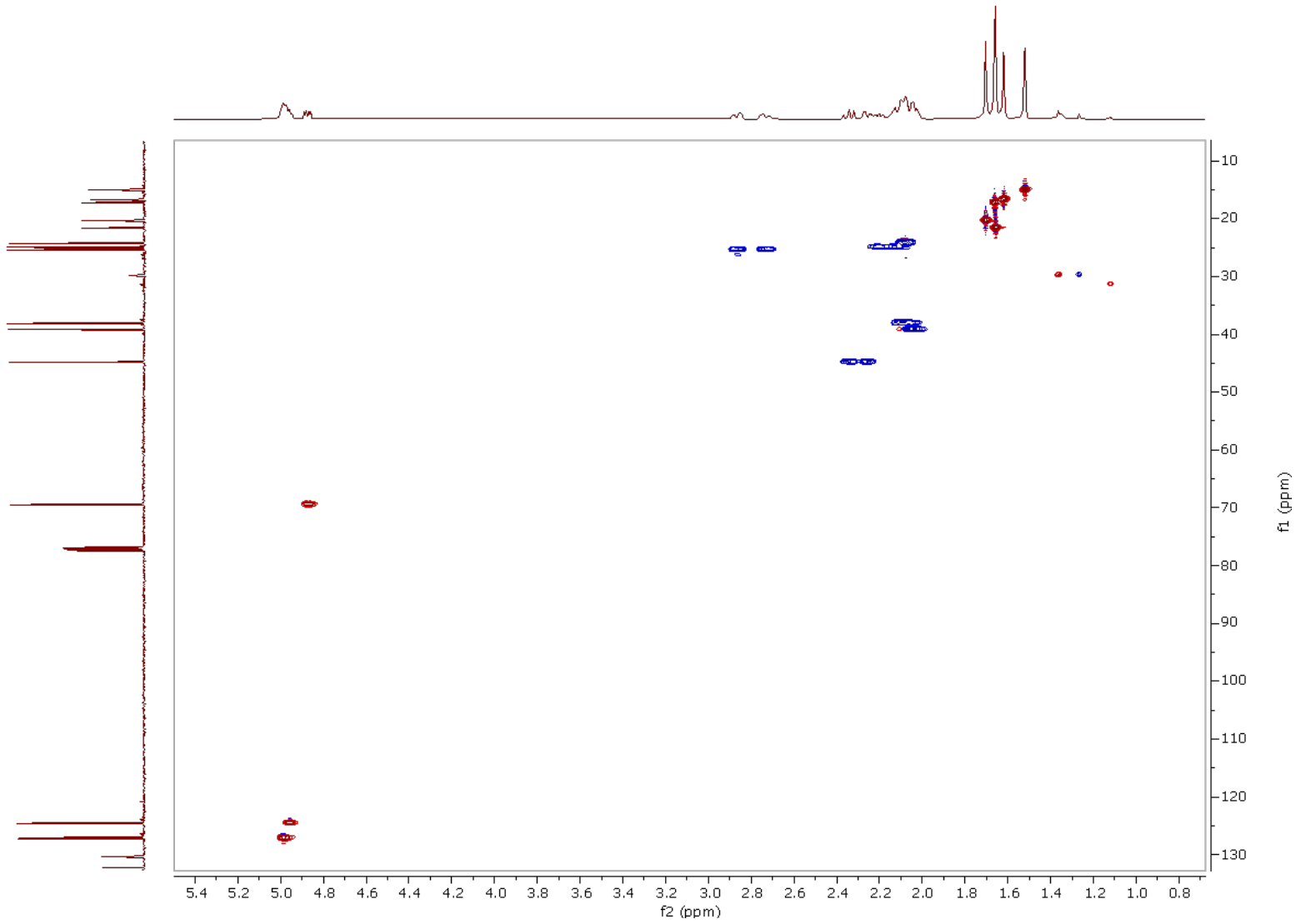
^{1}H NMR spectrum of **S1** in $CDCl_3$ at 500 MHz.



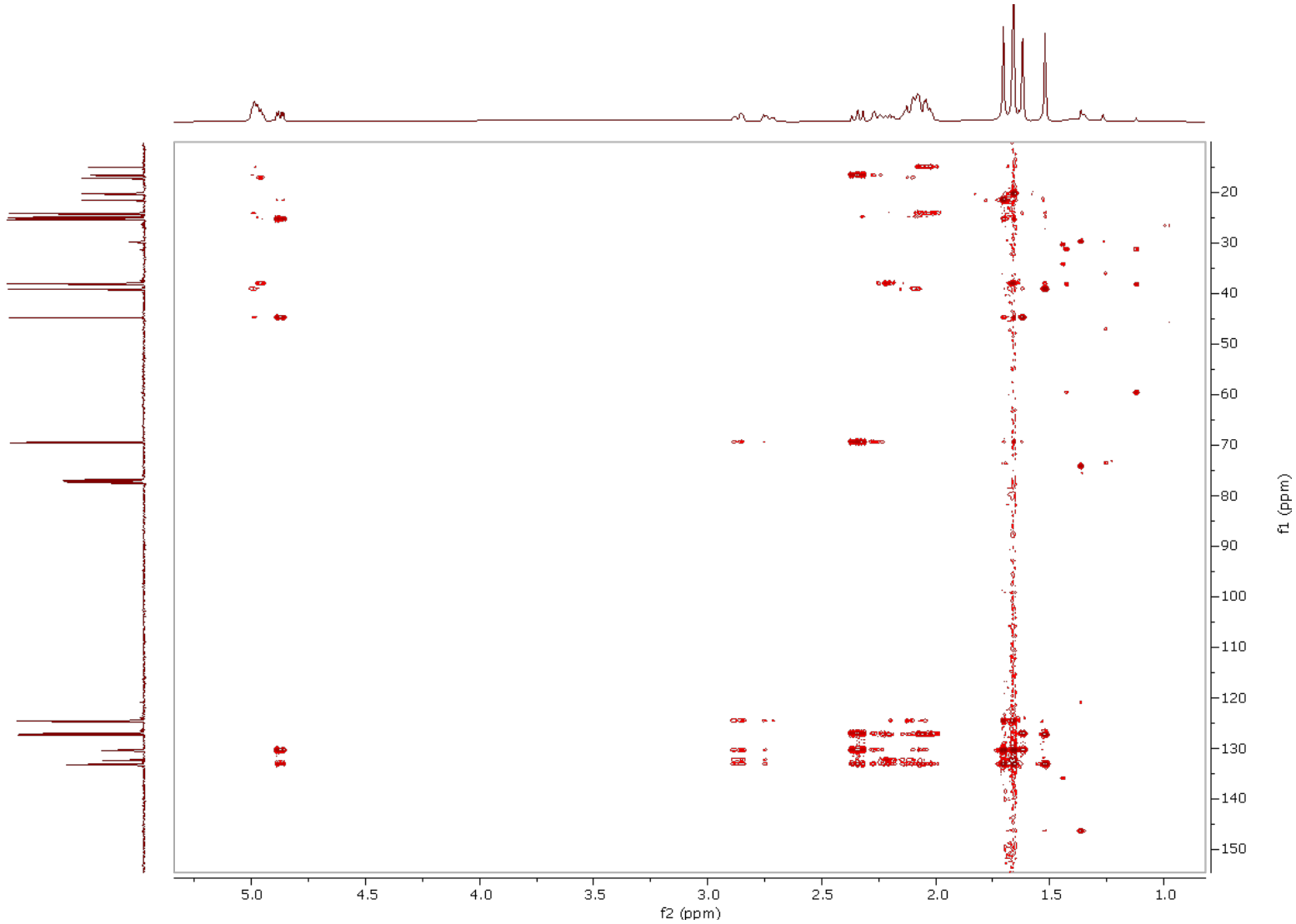
^{13}C NMR spectrum of **S1** in $CDCl_3$ at 125 MHz.



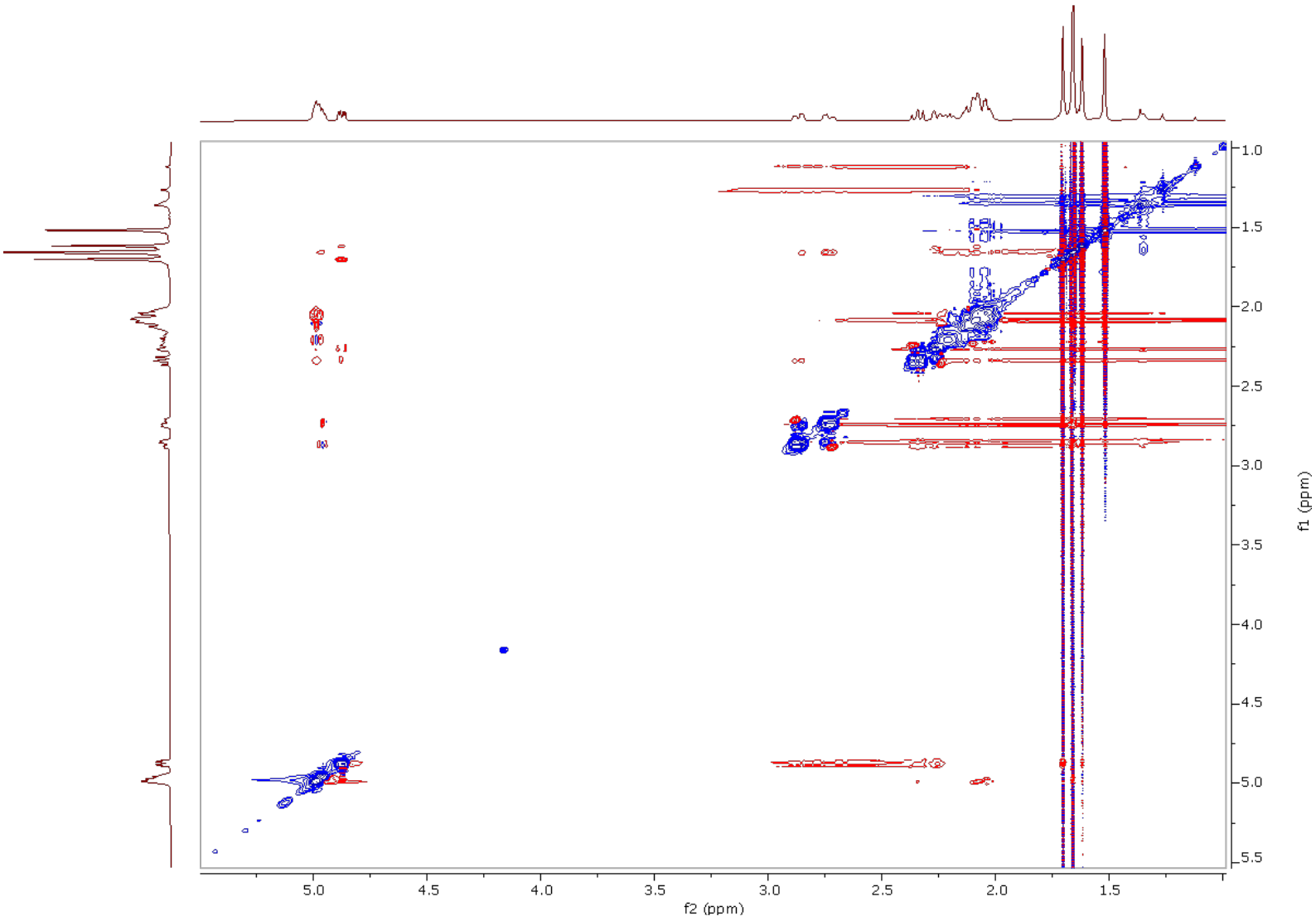
^{1}H-^{1}H COSY spectrum of **S1** in $CDCl_3$ at 500 MHz.



^{1}H-^{13}C HSQC spectrum of **S1** in $CDCl_3$ at 500 MHz.

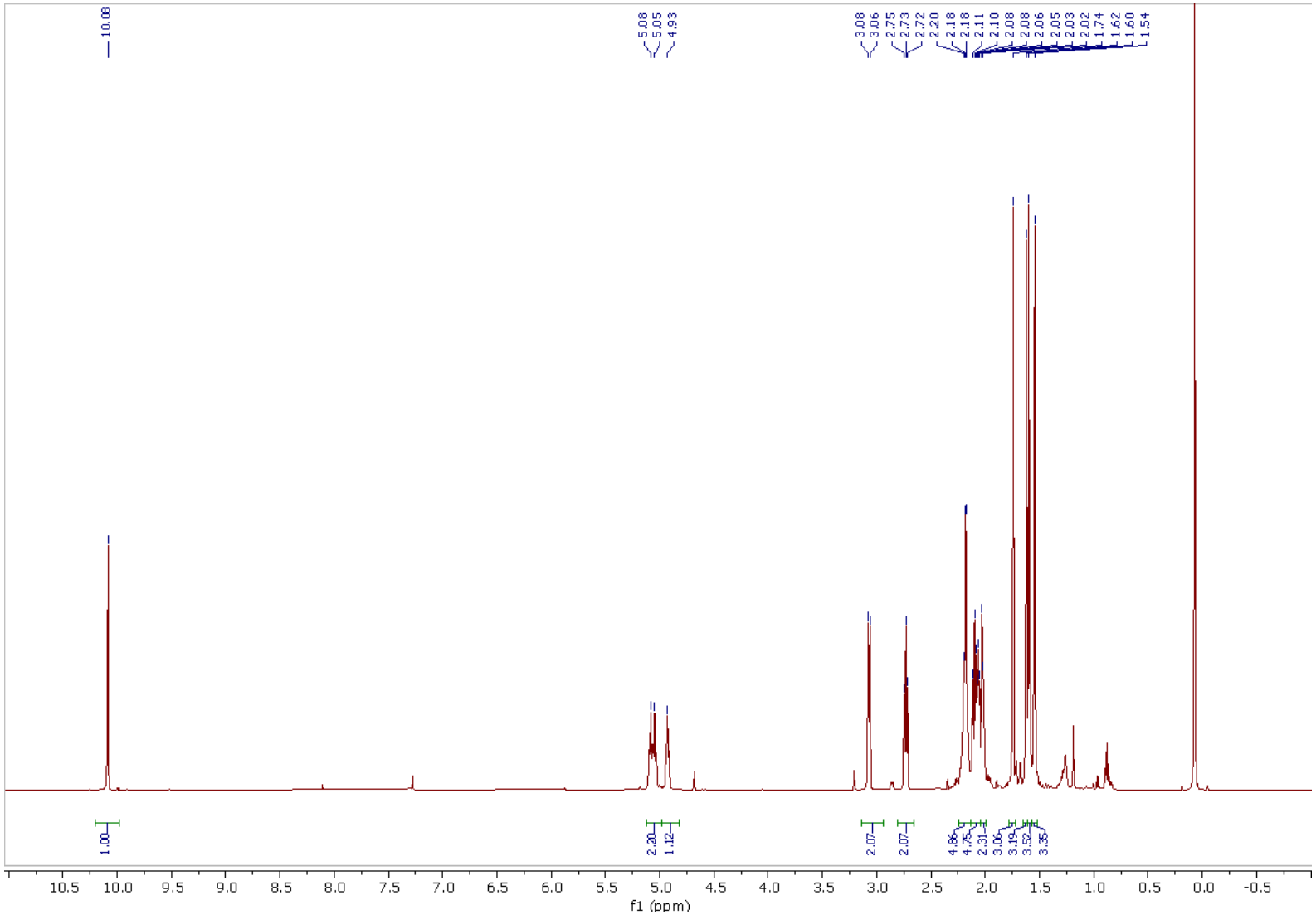


^{1}H-^{13}C HMBC spectrum of **S1** in $CDCl_3$ at 500 MHz.

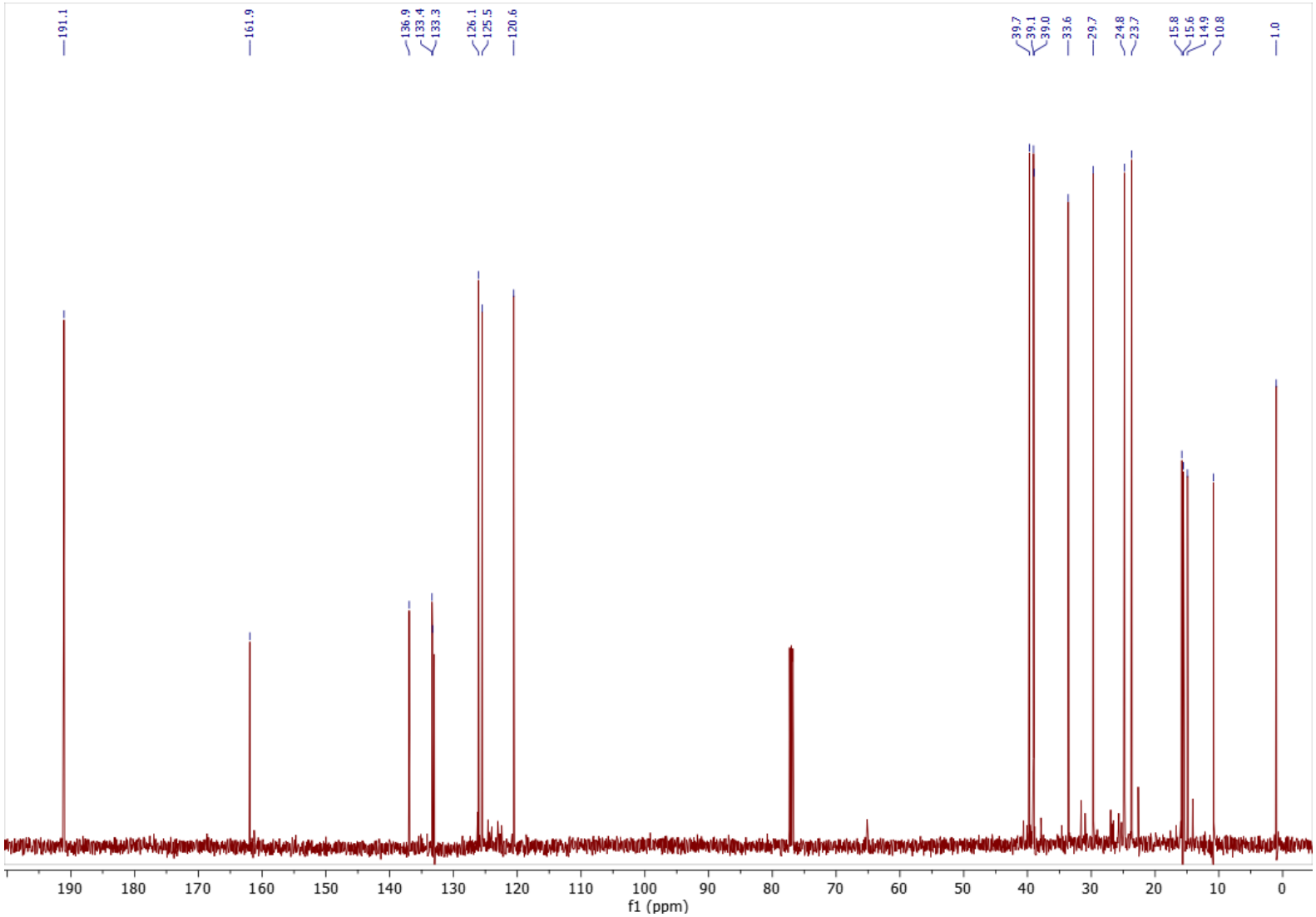


^{1}H-^{1}H NOESY spectrum of **S1** in $CDCl_3$ at 500 MHz.

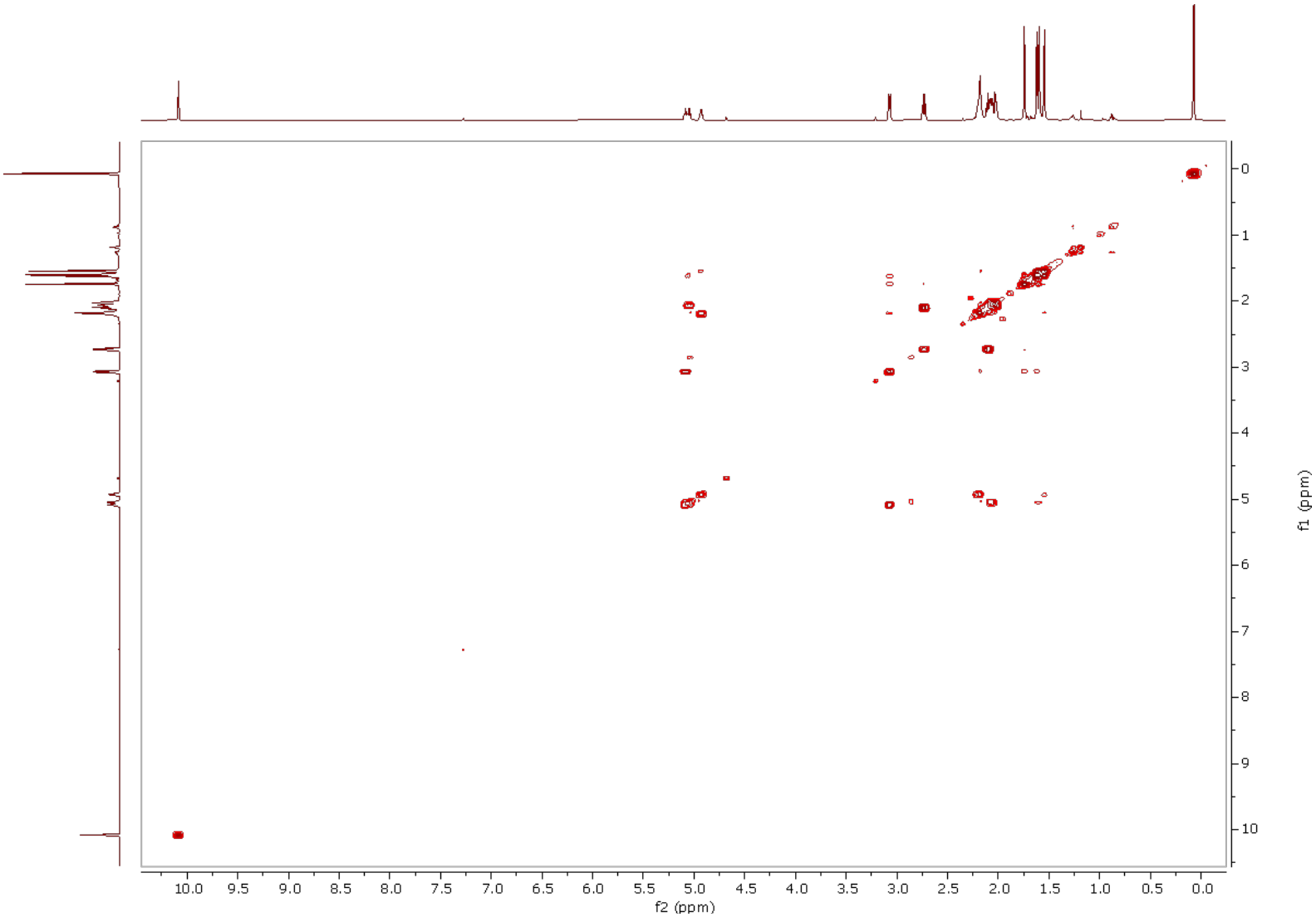
NMR spectra of cembrene B aldehyde (**S2**).



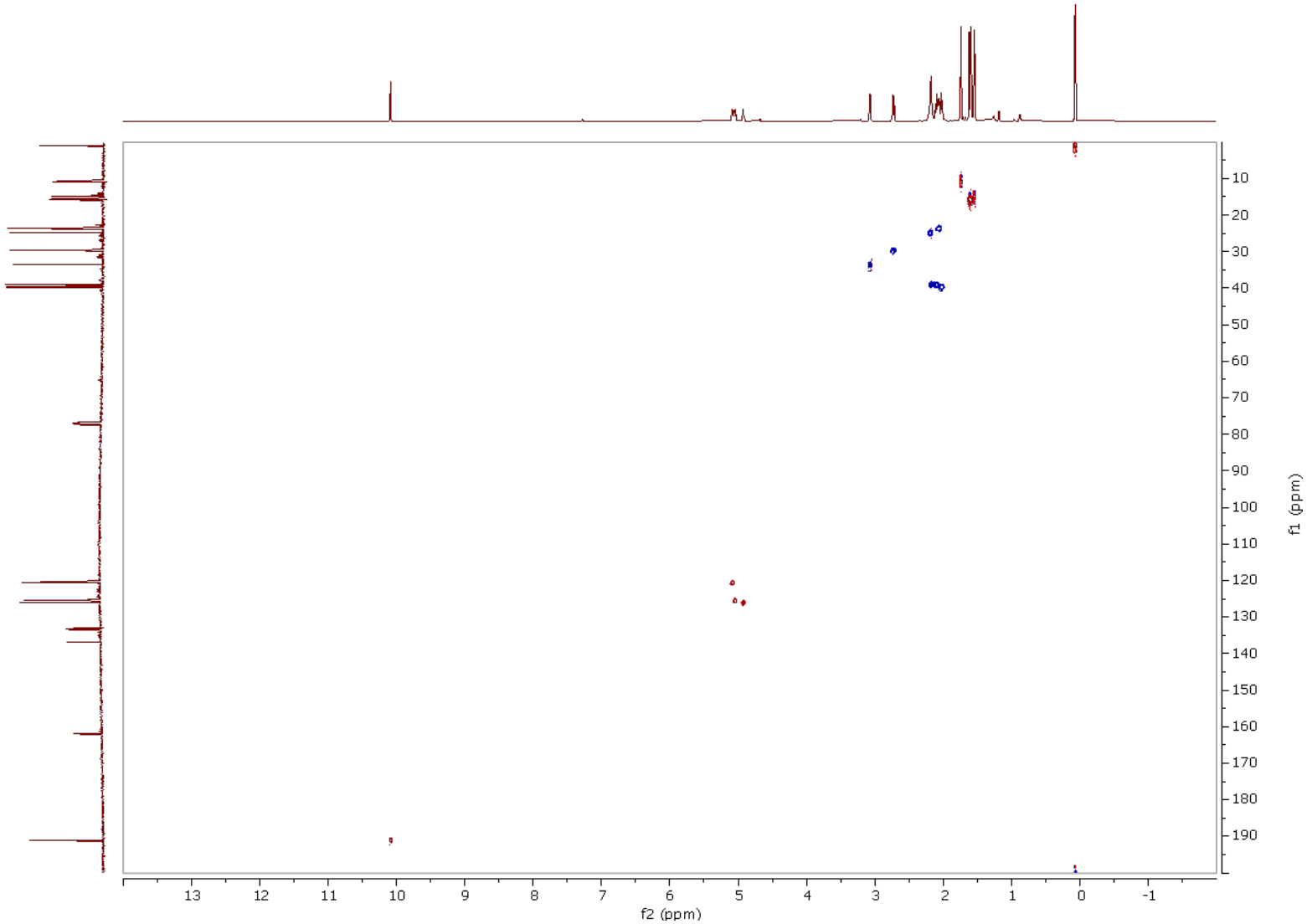
^{1}H NMR spectrum of **S2** in $CDCl_3$ at 500 MHz.



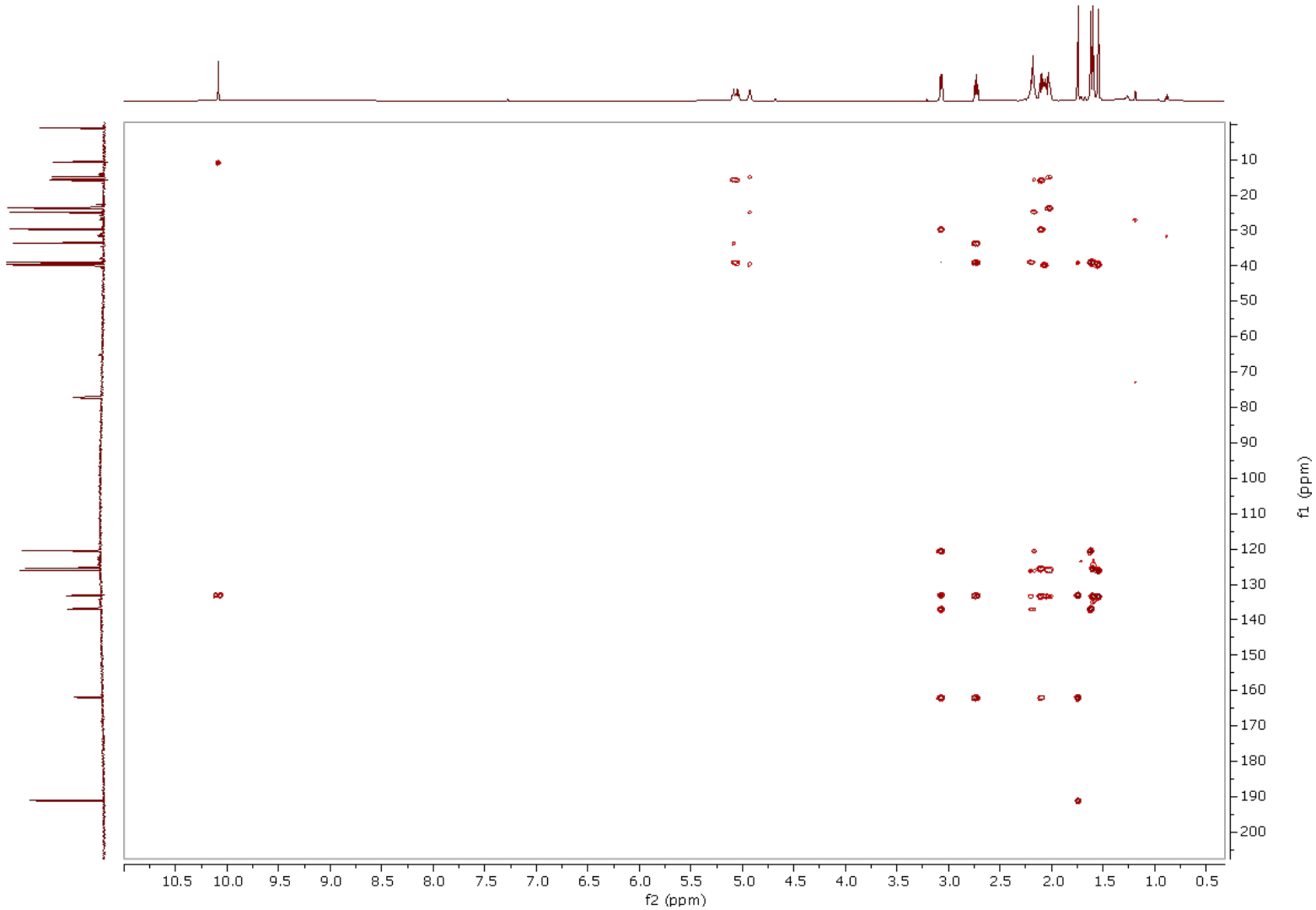
^{13}C NMR spectrum of **S2** in $CDCl_3$ at 125 MHz.



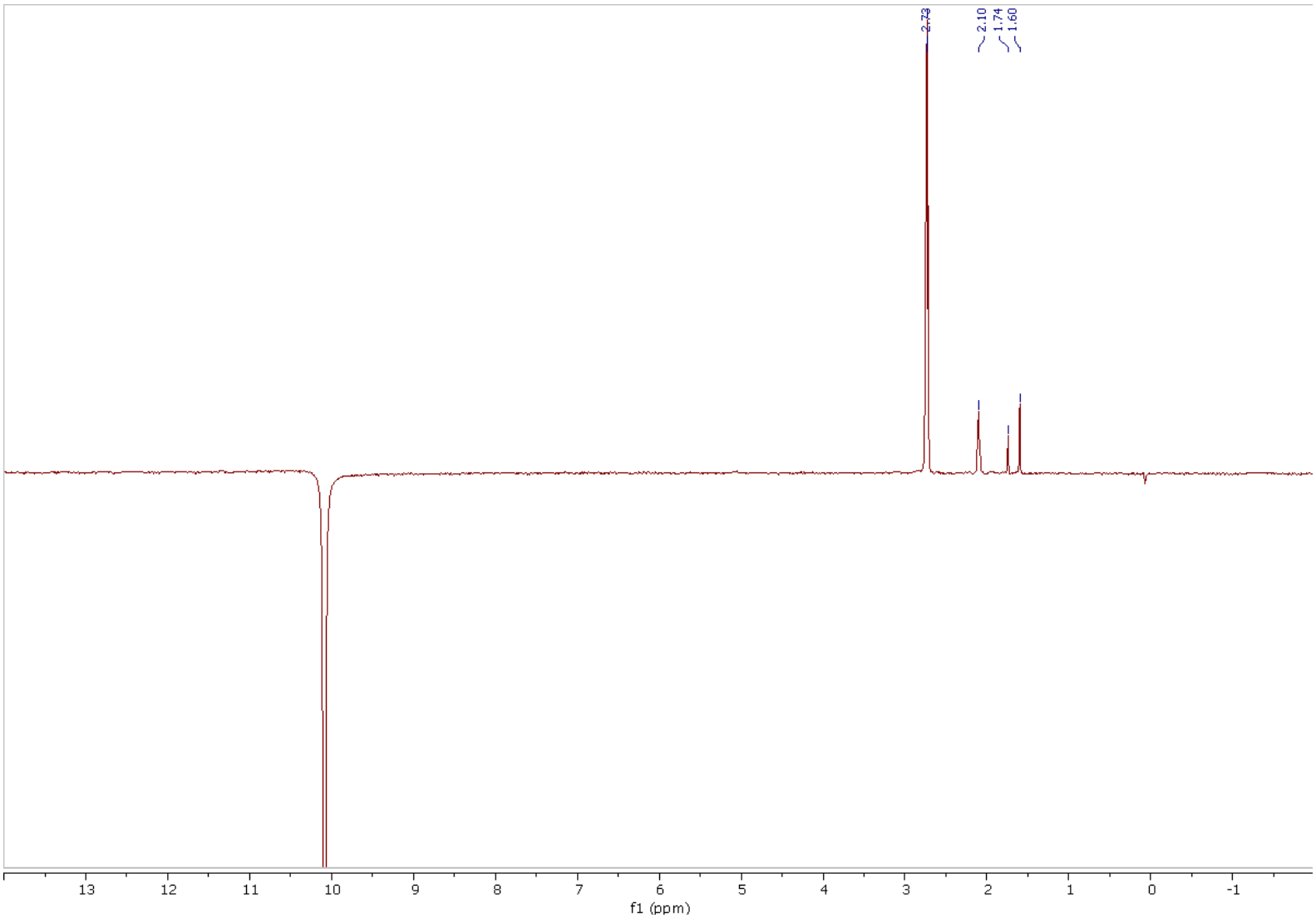
^{1}H-^{1}H COSY spectrum of **S2** in $CDCl_3$ at 500 MHz.



^{1}H-^{13}C HSQC spectrum of **S2** in $CDCl_3$ at 500 MHz.

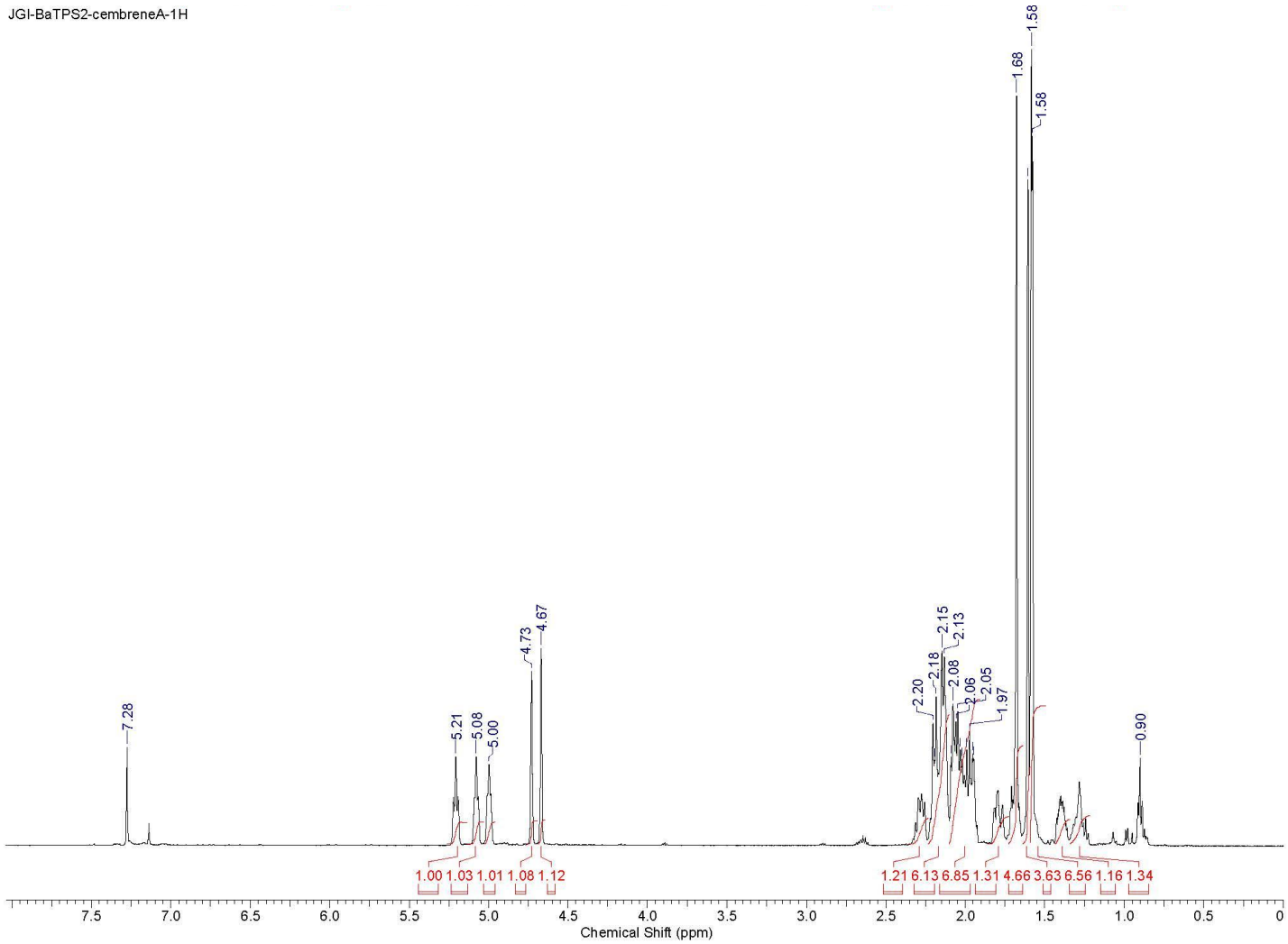


^{1}H-^{13}C HMBC spectrum of **S2** in $CDCl_3$ at 500 MHz.

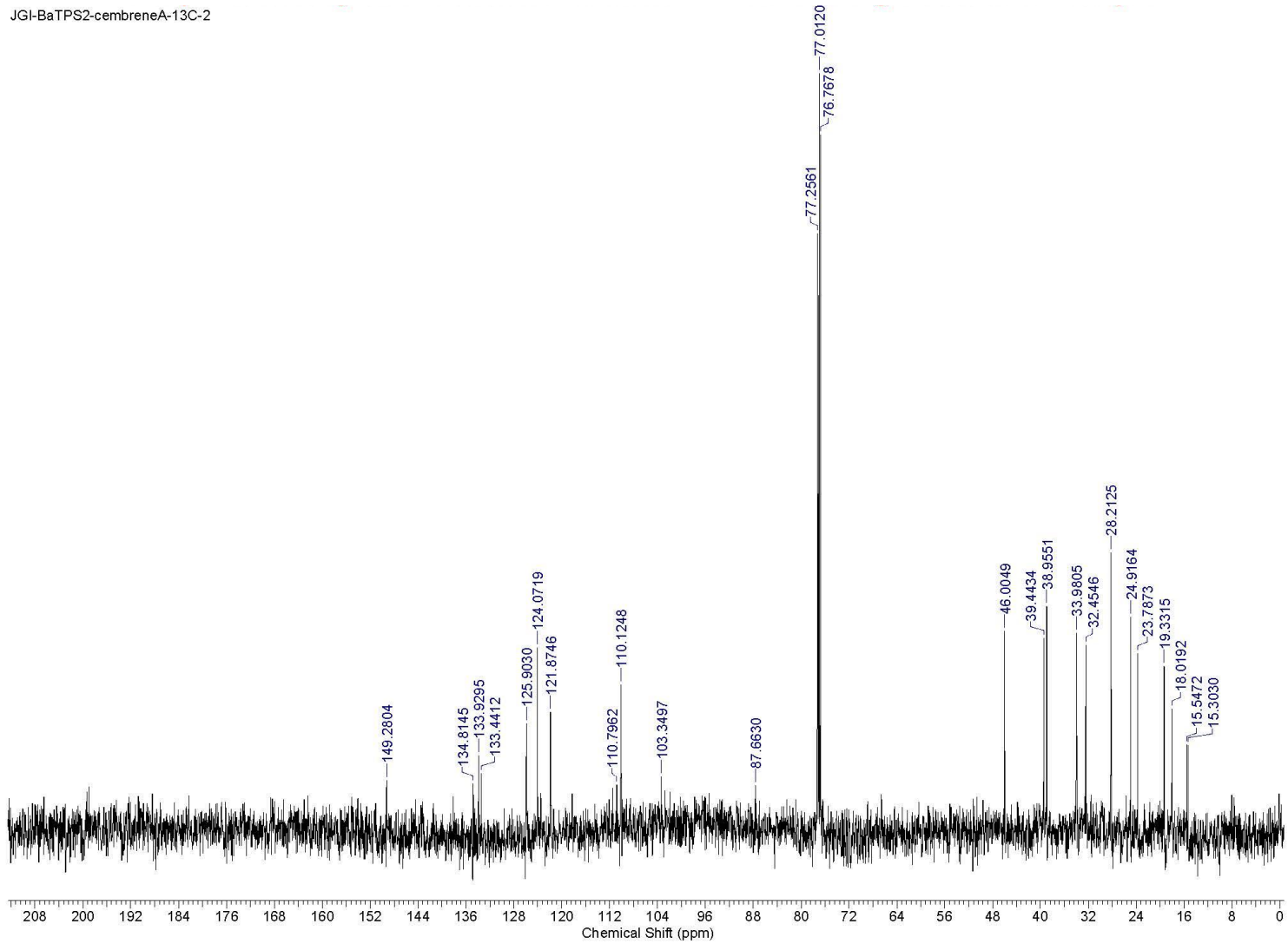


^{1}H 1D DPFGSE NOESY spectrum of **S2** in $CDCl_3$ at 500 MHz with selective excitation at 10.08 ppm.

NMR spectra of cembrene A.

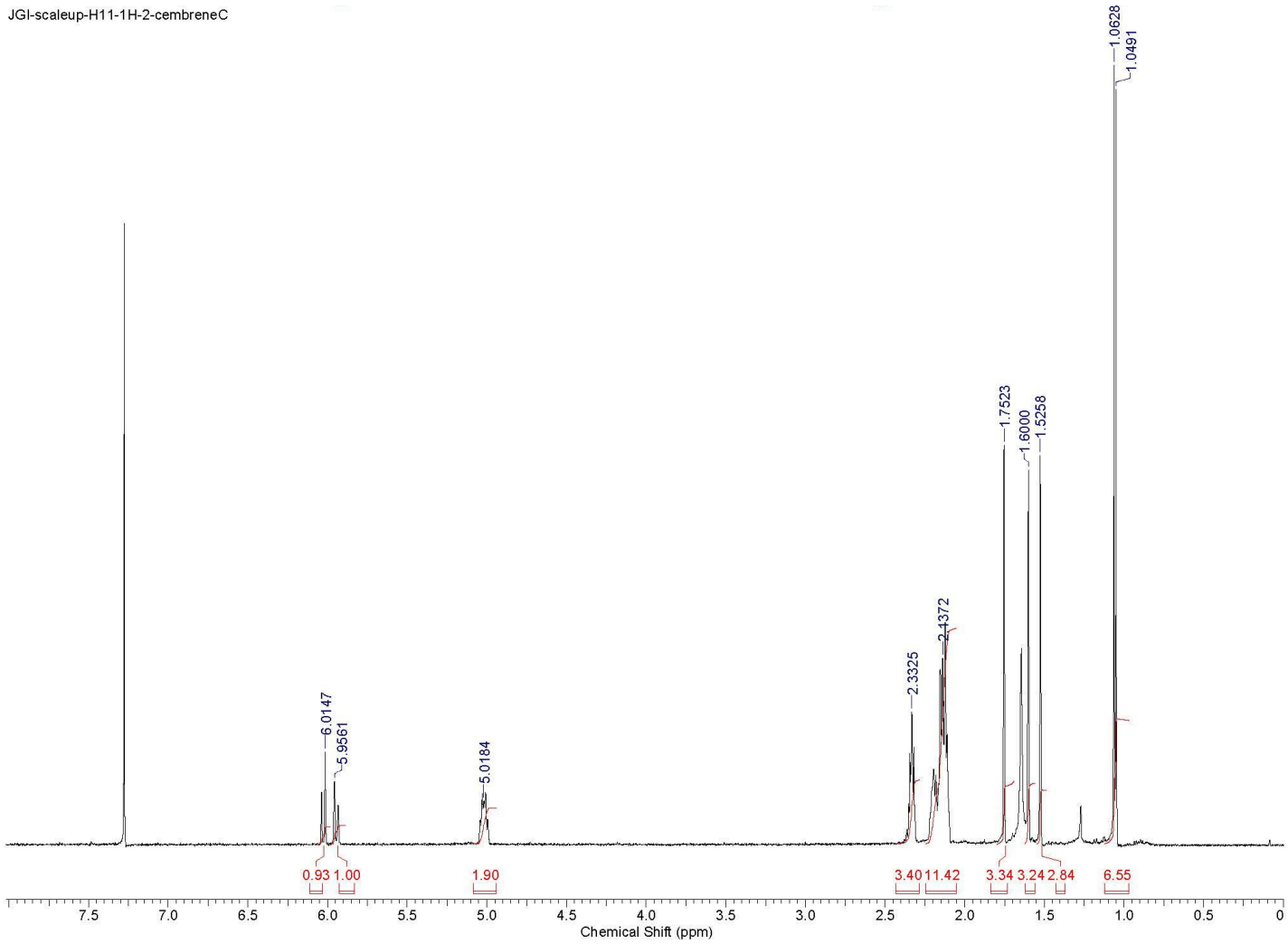


^{1}H NMR spectrum of cembrene A in $CDCl_3$ at 500 MHz



^{13}C NMR spectrum of cembrene A in $CDCl_3$ at 500 MHz

NMR spectra of cembrene C.



^{1}H NMR spectrum of cembrene C in $CDCl_3$ at 500 MHz

1. Williams, G. C. Living genera of sea pens (Coelenterata: Octocorallia: Pennatulacea): illustrated key and synopses. *Zool. J. Linn. Soc.* **113**, 93–140 (1995).
2. Burkhardt, I., Dürr, L., Grayson, N. E. & Moore, B. S. Chapter Thirteen - Methods for the discovery and characterization of octocoral terpene cyclases. in *Methods in Enzymology* (ed. Rudolf, J.) vol. 699 343–371 (Academic Press, 2024).
3. Scesa, P. D., Lin, Z. & Schmidt, E. W. Ancient defensive terpene biosynthetic gene clusters in the soft corals. *Nat. Chem. Biol.* **18**, 659–663 (2022).
4. Chen, S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta* **2**, e107 (2023).
5. Marçais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27**, 764–770 (2011).
6. Vurture, G. W. *et al.* GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics* **33**, 2202–2204 (2017).
7. De Coster, W. & Rademakers, R. NanoPack2: population-scale evaluation of long-read sequencing data. *Bioinformatics* **39**, btad311 (2023).
8. Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).
9. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
10. Vaser, R., Sović, I., Nagarajan, N. & Šikić, M. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* **27**, 737–746 (2017).

11. Walker, B. J. *et al.* Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLOS ONE* **9**, e112963 (2014).

12. Cheng, H. *et al.* Haplotype-resolved assembly of diploid genomes without parental data. *Nat. Biotechnol.* **40**, 1332–1335 (2022).

13. Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A. & Zdobnov, E. M. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Mol. Biol. Evol.* **38**, 4647–4654 (2021).

14. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).

15. Grabherr, M. G. *et al.* Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* **29**, 644–652 (2011).

16. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

17. Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A. & Korobeynikov, A. Using SPAdes De Novo Assembler. *Curr. Protoc. Bioinforma.* **70**, e102 (2020).

18. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).

19. Burkhardt, I., de Rond, T., Chen, P. Y.-T. & Moore, B. S. Ancient plant-like terpene biosynthesis in corals. *Nat. Chem. Biol.* **18**, 664–669 (2022).

20. Wang, J. *et al*. The conserved domain database in 2023. *Nucleic Acids Res*. **51**, D384–D388 (2022).

21. Solovyev, V., Kosarev, P., Seledsov, I. & Vorobyev, D. Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol*. **7**, S10 (2006).

22. Potter, S. C. *et al*. HMMER web server: 2018 update. *Nucleic Acids Res*. **46**, W200–W204 (2018).

23. Gilchrist, C. L. M. & Chooi, Y.-H. clinker & clustermap.js: automatic generation of gene cluster comparison figures. *Bioinformatics* **37**, 2473–2475 (2021).

24. Pankov, K. V. *et al*. The cytochrome P450 (CYP) superfamily in cnidarians. *Sci. Rep*. **11**, 9834 (2021).

25. Camacho, C. *et al*. BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).

26. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

27. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).

28. Lassmann, T., Frings, O. & Sonnhammer, E. L. L. Kalign2: high-performance multiple alignment of protein and nucleotide sequences allowing external features. *Nucleic Acids Res*. **37**, 858 (2009).

29. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).

30. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* **52**, W78–W82 (2024).

31. Chambers, M. C. *et al.* A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918–920 (2012).

32. Schmid, R. *et al.* Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat. Biotechnol.* **41**, 447–449 (2023).

33. Nothias, L.-F. *et al.* Feature-based molecular networking in the GNPS analysis environment. *Nat. Methods* **17**, 905–908 (2020).

34. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

35. Scesa, P. D. & Schmidt, E. W. Biomimetic Approach to Diverse Coral Diterpenes from a Biosynthetic Scaffold. *Angew. Chem. Int. Ed.* **62**, e202311406 (2023).

36. Alcaíno, J. *et al.* Functional Characterization of the Xanthophyllomyces dendrorhous Farnesyl Pyrophosphate Synthase and Geranylgeranyl Pyrophosphate Synthase Encoding Genes That Are Involved in the Synthesis of Isoprenoid Precursors. *PLOS ONE* **9**, e96626 (2014).

37. Ro, D.-K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006).

38. Sambrook, J. & Russell, D. W. Preparation and Transformation of Competent E. coli Using Calcium Chloride. *Cold Spring Harb. Protoc.* **2006**, pdb.prot3932 (2006).

39. Synthetic Complete (SC) Medium. *Cold Spring Harb. Protoc.* **2016**, pdb.rec090589 (2016).

40. Wratten, S. J., Faulkner, D. J., Hirotsu, K. & Clardy, J. Stylatulide, a sea pen toxin. *J. Am. Chem. Soc.* **99**, 2824–2825 (1977).

41. Meginley, R. J. *et al.* Briareolate Esters from the Gorgonian Briareum asbestinum. *Mar. Drugs* **10**, 1662–1670 (2012).

42. Fenical, W. & Pawlik, J. R. Defensive properties of secondary metabolites from the Caribbean gorgonian coral Erythropodium caribaeorum. *Mar. Ecol. Prog. Ser.* **75**, 1–8 (1991).

43. Ksebati, M. B. & Schmitz, F. J. Diterpenes from a Soft Coral, Minabea SP., from Truk Lagoon. *Bull. Sociétés Chim. Belg.* **95**, 835–851 (1986).

44. Li, C. *et al.* Bioactive 11,20-Epoxy-3,5(16)-diene Briarane Diterpenoids from the South China Sea Gorgonian Dichotella gemmacea. *J. Nat. Prod.* **74**, 1658–1662 (2011).

45. Nurco, D. J., Conklin, D. E., Shapiro, N. S. & Tran, E. The absolute structure of ptilosarcenone 2.5-hydrate, a diterpenoid briarane from the orange sea pen Ptilosarcus gurneyi (Gray). *Acta Crystallogr. Sect. E Struct. Rep. Online* **67**, o181–o182 (2010).

46. Hendrickson, R. L. & Cardellina, J. H. Structure and stereochemistry of insecticidal diterpenes from the sea penXXX. *Tetrahedron* **42**, 6565–6570 (1986).

47. Pordesimo, E. O., Schmitz, F. J., Ciereszko, L. S., Hossain, M. B. & Van der Helm, D. New briarein diterpenes from the Caribbean gorgonians Erythropodium caribaeorum and Briareum sp. *J. Org. Chem.* **56**, 2344–2357 (1991).