Electronic Supplementary Material (ESI) for Catalysis Science & Technology.
This journal is © The Royal Society of Chemistry 2022

Supporting Information

Rational reprogramming of the sesquiterpene synthase BcBOT2 yields new terpenes with presilphiperfolane skeleton

Vanessa Harms,[a] Jenny Irwan,[a] Trang Nguyen,[a] Jörg Fohrer,[b] Philipp Elbers[c], Paul Schrank[c], Mehdi D. Davari,[c*] Andreas Kirschning[a*]

[a] Institute of Organic Chemistry, Leibniz University Hannover, Schneiderberg 1b, 30167 Hannover, Germany

[b] Department of Chemistry, Technical University Darmstadt, Alarich-Weiss-Straße 4, 64287 Darmstadt, Germany.

[c]Department of Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry Weinberg 3, 06120 Halle, Germany

The SI contains additional information (graphics, tables, figures, etc.) not covered in the materials and methods section of the main manuscript.

1.	General information about the devices and equipment used	S4
2.	Biological and chemical section	S6
3.	Computational section	S17
4.	References (Supporting information)	S36
5.	Attachments (A. NMR spectra, B. GC chromatograms, C. MS spectra)	S38

Table of Contents:

Figure or Table	page
Table S1: Enzymes and corresponding gene bank accession numbers used in this study	**S6**
Table S2: *E. coli* stocks used for transformations	**S6**
Table S3: Forward and reverse primers used for site-directed mutagenesis	**S6**
Table S4: Primer sequences for DNA sequencing	**S7**
Table S5: Components of the PCR sample	**S7**
Table S6: Temperature program for gradient PCR	**S7**
Figure S1: Agarose gels for the identification of successful amplification of mutant *bcbot2*-pET28a(+) plasmid DNA and of optimal annealing temperature	**S8/9**
Table S7: Solutions for plasmid isolation	**S9**
Table S8: Composition of the separating and collecting gels used	**S9**
Figure S2: SDS-Page for purification of WT BcBOT2	**S10**
Figure S3: SDS-Page for purification of BcBOT2 variants	**S10/11**
Figure S4: Chromatogram of the biotransformation of WT BcBOT2 with FPP	**S12**
Table S9: Additional biotransformation products that could not be identified by RI or MS-spectroscopy	**S12**
Table S10: Semi-quantitative determination in percent [%] of the detected compounds for the wild type (WT) and all variants	**S12//13**
Figure S5: Analysis of thermostability for BcBOT2 variants Y211S, Y211F, T242N, and T242S	**S14**
Figure S6: Analysis of the thermostability for BcBOT2-W118Q variant	**S14**
Figure S7: Analysis of thermostability for BcBOT2-F138V and yields of the five main products **2**, **4**, and **25**-**27**	**S14**
Table S11: Data on 2D NMR and 1d-NOE experiments of product 25	**S15**
Table S12: Data of 2D NMR and 1d-NOE experiments of product 27	**S16**
Table S13: Evaluation of the five BcBOT2 models predicted by AlphaFold2 program	**S18**
Figure S8: Multiple structural alignment of models for BcBOT2 build using AlphaFold2	**S19**
Figure S9: Structural model of BcBOT2 constructed using AlphaFold2 (Model Number 5)	**S20**
Figure S10: Comparing secondary structure prediction of N – terminal region (amino acids 1 – 80) for BcBOT2 sequence using a variety of prediction servers	**S21**
Figure S11: Comparing secondary structure of N – terminal region (amino acids 1 – 80) of BcBOT2 structural models build by AlphaFold2	**S22**
Figure S12: Comparing the build secondary structure of the Alphafold2 generated model and the general consensus from prediction servers for whole BcBOT2 sequence	**S23**

Figure S13: The amino acids sequence of BcBOT2 colored by evolutionary importance **S24**

Figure S14: The amino acid sequence of BcBOT2 residue 1-399 colored by conservation **S25**

Figure S15: Structural alignment of BcBOT2 model and crystal structure of TEAS (5IKA) **S26**

Figure S16: The structural model and active pocket of BcBOT2 **S28**

Table S14: List of amino acid residues that form the active pocket **S29**

Figure S17: Structural alignment of BcBOT2 model and crystal structure of Chain A of SdS (4OKM) **S30**

Figure S18: Docking of intermediates of Tantillo route in wild-type BcBOT2 structure **S32**

Figure S19: Docking of intermediate 8 and 11 into wild-type BcBOT2 and the variants W118F and W118Q **S33**

Figure S20: Investigation of chair- and boat-like conformations for Intermediate **11** **S35**

1. General information about the devices and equipment used

^{1}H NMR spectra were recorded at 400 MHz with a BRUKER Avance-400 and BRUKER Ascend-400 as well as at 500 MHz with a BRUKER DRX-500 spectrometer or at 600 MHz with an BRUKER Ascend-600 or BRUKER Avance III-600 at 298K or 253K for low temperature NMR measurements. ^{13}C NMR spectra were recorded at 100 MHz with a BRUKER Avance-400 and BRUKER Ascend-400 or at 125 MHz with a BRUKER DRX-500 or at 151 MHz with a BRUKER Ascend-600/ Avance III-600 instrument. ^{31}P NMR spectra were recorded at 162 MHz with a BRUKER Avance-400 and BRUKER Ascend-400 or at 202 MHz with a BRUKER DRX-500 instrument. Multiplicities are described using the following abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, sex = sextet, m = multiplet, b = broad. Chemical shift values of ^{1}H and ^{13}C NMR spectra are commonly reported in ppm relative to residual solvent signal as internal standard. The multiplicities refer to the resonances in the off-resonance decoupled spectra and were elucidated using phase-sensitive HSQC experiments. Mass spectra were obtained with a lockspray dual ion source in combination with a WATERS Alliance 2695 LC system, or with a type Q-TOF premier (MICROMASS) spectrometer (ESI mode) in combination with a WATERS Acquity UPLC system equipped with a WATERS Acquity UPLC BEH C18 1.7 µm (SN 01473711315545) column (solvent A: water + 0.1 % (v/v) formic acid, solvent B: MeCN or MeOH (given in experimental part) + 0.1 % {v/v} formic acid; flow rate = 0.4 mL/min; gradient {t [min]/solvent B [%]}: {0/5} {2.5/95} {6.5/95} {6.6/5} {8/5}; retention times {r_t} given in the experimental part). Ion mass signals (*m/z*) are reported as values in atomic mass units.

GC-O analyses were carried out with an Agilent GC 7890B chromatograph with Gerstel CIS4 Cold Injector, coupled to a Gerstel OPD 3 sniffer. Samples were analysed on a Zebron ZB-FFAP (7KG-G009-11) column (60 m x 0.25 mm i.d. x film thickness 0.25 µm). Carrier gas, He; flow 1.4 mL/min; injector temp.: 50°C to 260 °C at 12 °C/s; split ratio 1:5; temp. program: 50 °C to 230 °C at 8 °C/min held isothermal for 22.5 min. Transfer to Gerstel ODP 3 is made through 1 m x 150 µm column without stationary phase at 280 °C. Further GC-O analyses were carried out with an Agilent GC 7890B chromatograph with Gerstel CIS4 Cold Injector, coupled to a Gerstel OPD 3 sniffer. Samples were analysed on a Zebron ZB-1 column (60 m x 0.25 mm i.d.; ilm thickness 0.25 µm). Carrier gas, He; flow 11.4 mL/min; injector temp.: 50°C to 260 °C at 12 °C/s; split ratio 1:5 up to 1:25; temp. program: 50 °C to 250 °C at 8 °C/min held isothermal for 22.5 min. Transfer to Gerstel ODP 3 is made through 1 m x 150 µm column without stationary phase at 280°C.

GC/MS analyses were carried out on with an Agilent 7890B GC with 5977B GC/MSD and Gerstel MPS Robotic XL with KAS 4C injector. Samples were analysed on an Optima 5HT column, (30 m x 250 µm i.d. x film thickness 0.25 µm). Carrier gas, He; injector temp., 60°C to 300°C at 12°C/min, splitless; temp. program: 50°C (isothermal 1 min) to 300°C, at 20 °C/min and held isothermal for 6.5 min at 300°C; FID: 300°C , H_2: 30 mL/min, N_2: 25 mL/min, MSD: ion source: EI 70 eV, 230 °C; detector: quadrupole, EI mass spectra were acquired over the mass range of 30 –650 amu. Further GC/MS analyses were carried out with an Agilent GC 7890B chromatograph with Gerstel CIS4 Cold Injector. Samples were analysed on a Zebron ZB-FFAP (7KG-G009-11) column (60 m x 0.25 mm i.d. x film thickness 0.25 µm). Carrier gas, He; flow 1.4 mL/min; injector temp.: 50°C to 260 °C at 12 °C/s; split ratio 1:5; temp. program: 50 °C to 230 °C at 8 °C/min held isothermal for 22.5 min. MSD: ion source: EI 70 eV, 230 °C; detector: quadrupole.

HR-GC/MS analyses were carried out on a Waters GCT Premier mass spectrometer coupled with an Agilent 6890n GC with CTC CombiPAL sampler. Samples were analysed on an Optima 5HT column, 30 m x 250 µm i.d. x film thickness 0.25 µm). Carrier gas, He; injector temp. 300°C, split ratio 1:40; temp. program: 50°C (isothermal 1 min) to 300°C, at 20 °C/min and held isothermal for 6.5 min at

300°C; FID: 300°C, H_2: 30 mL/min, N_2: 25 mL/min, GCT-Premier: ion source: EI 70 eV, 250 °C; detector-voltage: 2500 V, EI mass spectra were acquired over the mass range of 20 –800 amu.

Preparative GC (pGC) with a nonpolar ZB-1 column (Phenomenex, 30 m x 0.53 mm i.d. x 3 µm film thickness, serial no. 628795) was carried out on a HP 6890 chromatograph with HP 7683 autosampler and hot injection. New terpenoids were isolated with a Gerstel PFC. The temperature gradient was optimized for each isolation, carrier gas: H_2 (5 mL/min), injector temp.: 250 °C, splitless, injection volume: 1 µL, detector: FID, temperature: 250°C, H_2: 40 mL/min, N_2: 45 mL/min.

Analytical thin-layer chromatography was performed using precoated silica gel plates (Macherey-Nagel, Düren) and the spots were visualized with UV light at 254 nm or alternatively by staining with ninhydrin, permanganate, anisaldehyde or 4-methoxybenzaldehyde solutions commercially available reagents, chromatography type or dry solvents were used as received or purified by standard techniques according to the literature.[1] The preparation of 3,3-dimethylallyl bromide is reported in reference [2] and the synthesis of tris(tetra-*n*-butylammonium)hydrogenpyrophosphate trihydrate in reference [3, 4] Isoprene and geraniol are commercially available. The chemical synthesis of FPP and the general description of diphosphate synthesis via the chloride was reported in reference [5, 6].

2. Biological and chemical experiments

2.1 Genes, strains, and plasmids

Table S1: Enzymes and corresponding gene bank accession numbers used in this study.

enzyme	gene bank accession number	reference
BcBOT2	Q6WP50	[7]

2.2 Microbiological methods

Table S2: *E. coli* stocks used for transformations.

organism	genotype
***E. coli* TOP10**	F$^-$ mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galK16 rpsL(Str^R) endA1 λ^-
***E. coli* BL21 (DE3)**	F$^-$ ompT gal dcm Ion hsd Sʙ (rʙ- mʙ-) λ(DE3 [lacI lacUV5-T7 gene 1 ind 1 sam7 nin5])

2.3 Primer sequences

Table S3: Forward and reverse primers used for site-directed mutageneses. The exchanged amino acid is shown in blue.

	forward primer (5'□3')	reverse primer(5'□3')
R373H	CCGGACCCAGATAGTGGCCGGTCTGAAAA	TTTTCAGACCGGCCACTATCTGGGTCCGG
R373G	CGGACCCAGATAGCCGCCGGTCTGAAAAC	GTTTTCAGACCGGCGGCTATCTGGGTCCG
F138V	AAACTGGTCATCGAACAGCACCACCCAGTGGTTCCAATC	GATTGGAACCACTGGGTGGTGCTGTTCGATGACCAGTTT
F138Y	CAAACTGGTCATCGAACAGATACACCCAGTGGTTCCAATC	GATTGGAACCACTGGGTGTATCTGTTCGATGACCAGTTTG
F138A	CAAACTGGTCATCGAACAGCGCCACCCAGTGGTTCCAATC C	GGATTGGAACCACTGGGTGGCGCTGTTCGATGACCAGTTTG
F138M	AAACTGGTCATCGAACAGCATCACCCAGTGGTTCCAATC	GATTGGAACCACTGGGTGATGCTGTTCGATGACCAGTTT
F138N	AACTGGTCATCGAACAGGTTCACCCAGTGGTTCCAATC	GATTGGAACCACTGGGTGAACCTGTTCGATGACCAGTT
F138Q	AACCACTGGGTGCAGCTGTTCGATGAC	GAACAGCTGCACCCAGTGGTTCCAATC
Y211S	TAACAGCTGATCGAAGGAGCGCTTGTGCTGGTC	GACCAGCACAAGCGCTCCTTCGATCAGCTGTTA
Y211F	TAACAGCTGATCGAAGAAGCGCTTGTGCTGGTC	GACCAGCACAAGCGCTTCTTCGATCAGCTGTTA
T242S	CACACCAATGGAACCGCGGCGCAGGTC	GACCTGCGCCGCGGTTCCATTGGTGTG
T242N	GGATACACACCAATGTTACCGCGGCGCAGG	CCTGCGCCGCGGTAACATTGGTGTGTATCC
T282S	TCGTTCACCAGGGACACCAGATCGGCG	CGCCGATCTGGTGTCCCTGGTGAACGA
W118Q	GTGCATCCGGTGCCTGCATACTGGCCAGAAAGCA	TGCTTTCTGGCCAGTATGCAGGCACCGGATGCAC
W118F	TGCATCCGGTGCAAACATACTGGCCAGAAAGCACAGG	CCTGTGCTTTCTGGCCAGTATGTTTGCACCGGATGCA
W118	GTGCATCCGGTGCCATCATACTGGCCAGAAAGCA	TGCTTTCTGGCCAGTATGATGGCACCGGATGCAC

M		
W118H	CCGGTGCATCCGGTGCATGCATACTGGCCAGAAAGCAC	GTGCTTTCTGGCCAGTATGCATGCACCGGATGCACCGG
W118N	CGGTGCATCCGGTGCGTTCATACTGGCCAGAAAGCA	TGCTTTCTGGCCAGTATGAACGCACCGGATGCACCG
W118R	GGTGCATCCGGTGCACGCATACTGGCCAGAAAGC	GCTTTCTGGCCAGTATGCGTGCACCGGATGCACC
W367A	GGTAACCTGTACGCGAGTTTTCAGACC	GGTCTGAAAACTCGCGTACAGGTTACC
W367F	CAGGGTAACCTGTACTTTAGTTTTCAGACCGGCC	GGCCGGTCTGAAAACTAAAGTACAGGTTACCCTG
W367H	CCAGGGTAACCTGTACCATAGTTTTCAGACCGGCC	GGCCGGTCTGAAAACTATGGTACAGGTTACCCTGG
W367K	CAGGGTAACCTGTACAAGAGTTTTCAGACCGGC	GCCGGTCTGAAAACTCTTGTACAGGTTACCCTG
W367N	GGGTAACCTGTACAACAGTTTTCAGACCG	GAAAACTGTTGTACAGGTTACCCTGGGC
W367T	CCAGGGTAACCTGTACACGAGTTTTCAGACCGGC	GCCGGTCTGAAAACTCGTGTACAGGTTACCCTGG
W367V	CAGGGTAACCTGTACGTGAGTTTTCAGACCG	CGGTCTGAAAACTCACGTACAGGTTACCCTG

Table S4: Primer sequences for DNA sequencing.

primer	sequence
T7 minus primer	5‘ TAATACGACTCACTATAGGG 3‘
RP primer	5‘ CTAGTTATTGCTCAGCGG 3‘

2.4 Molecular biological methods

Table S5: Components of the PCR sample.

components	concentration	PCR sample [µL]
5x HF buffer	-	4.0[A]
dNTPs [mM]	10.0	0.4[A]
starting DNA [µg·µL^{-1}]	5.4	0.15[A]
forward primer [µM]	10.0	1.0[B]
reverse primer [µM]	10.0	1.0[B]
Phusion™ polymerase	-	0.2[B]
DMSO	-	0.5[A]
H_2O	-	12.7[A,B]
total volume [µL]	-	20.0

[A] master mix 1: [B] master mix 2

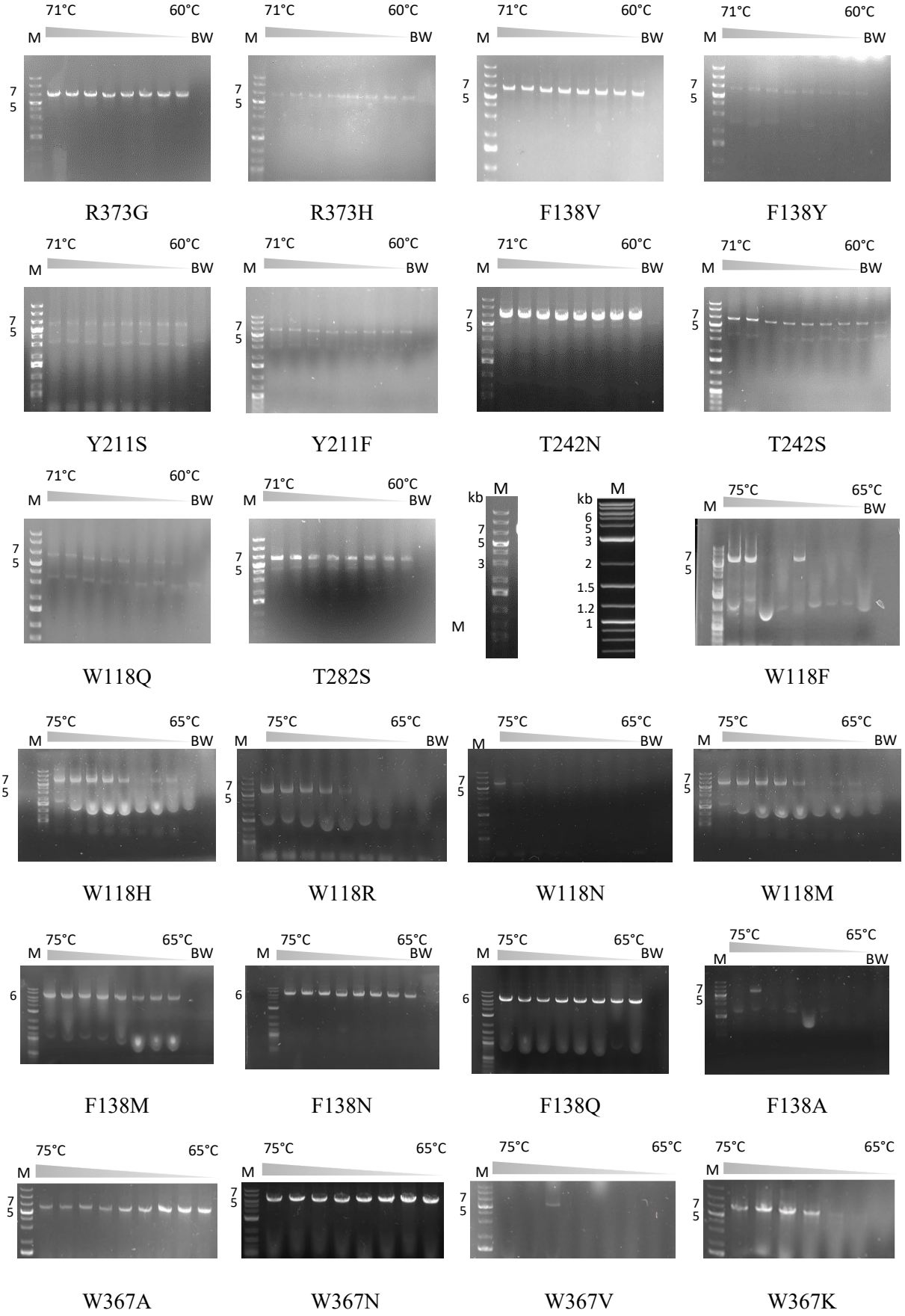
Table S6: Temperature program for gradient PCR.

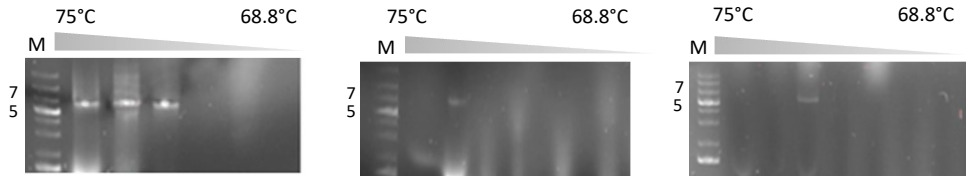
	step	temperature	duration
I	Initial denaturation	98°C	30 sec
II	denaturation	98°C	10 sec
III	Annealing	gradient	15 sec
IV	Elongation	72°C	30 min
V	Final elongation	72°C	10 min

15

71.0°C
70.2°C
68.8°C
66.7°C
64.2°C
62.2°C
60.8°C
60.0°C

2.5 Agarose electrophoresis of *bcbot2*-pET28a(+) mutants





W367F W367H W367T

Figure S1: Agarose gels for the identification of successful amplification of mutant *bcbot2*-pET28a(+) plasmid DNA and of optimal annealing temperature. M: size standard, BW: negative control (water instead of plasmid DNA).

2.6 Plasmid isolation

Table S7: Solutions for plasmid isolation.

solution 1	solution 2	solution 3
50 mM glucose **25 mM Tris base** **10 mM EDTA**	200 mM NaOH 1% [w/v] SDS	3M KCH_3COO
pH 8.0	-	pH 5.5

2.7 Protein biochemical methods

Table S8: Composition of the separating and collecting gels used.

	15% separating gel (10 mL)	5% collecting gel (5 mL)
30% Acrylic-Bisacrylic Mix	5 mL	0.83 mL
1.5 M Tris base	2.5 mL (pH 8.8)	0.63 mL (pH 6.8)
SDS	0.1 %	0.05 %
APS	0.1 %	0.05 %
TEMED	4 µL	5 µL

2.8 Protein overexpression and purification

For the overexpression, a preculture (3 mL and 3 µL kanamycin (50 µg/mL)) was inoculated and incubated for 4 - 5 h at 37 °C. This was set to inoculate the main culture, which was composed of 1 mL preculture, 50 mL 2TY- medium, and 50 µL kanamycin (50 µg/mL). First, the main culture was incubated for a period of about 1.5 h at 37 °C and 200 rpm shaking. The optical density (OD_{600}) was guarded. When an OD_{600} value of 0.5- 0.8 was reached, 0.5 mM IPTG (isopropyl-β-D-thiogalactopyranoside) was added. After the IPTG induction, the main cultures were incubated for 22 h at 16 °C and shaking at 180 rpm. Finally, the cells were centrifuged at 5000 xg for 10 min after incubation. The cells were disrupted using an ultrasonic sonotrode. For the disruption, an amplitude of 34-45 %, a time between 7-10 min,

and a pulse of 10 s digestion with 10 s pause were used. The cell lysate was then separated from insoluble cell debris by centrifugation at 4 °C for 20 min at 10000 gx. A nickel- NTA agarose column was used to purify the proteins. The purification was performed at ascending imidazole concentrations from 25 mM to 500 mM. To remove the salts, the proteins were changed to 50 mM HEPES, 5 mM DTT buffer (pH 7.5) via a PD10 desalting column.

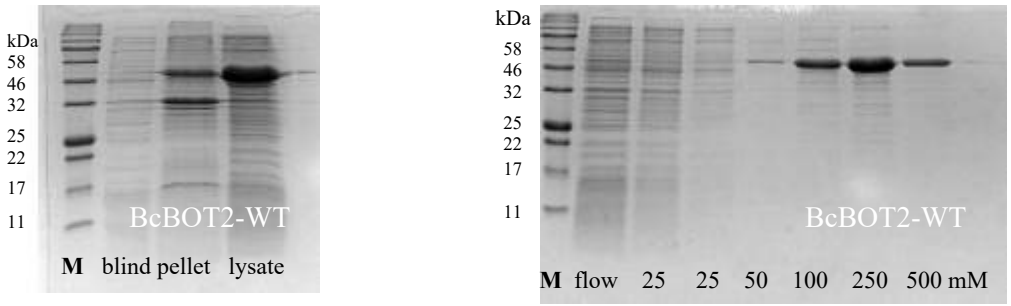
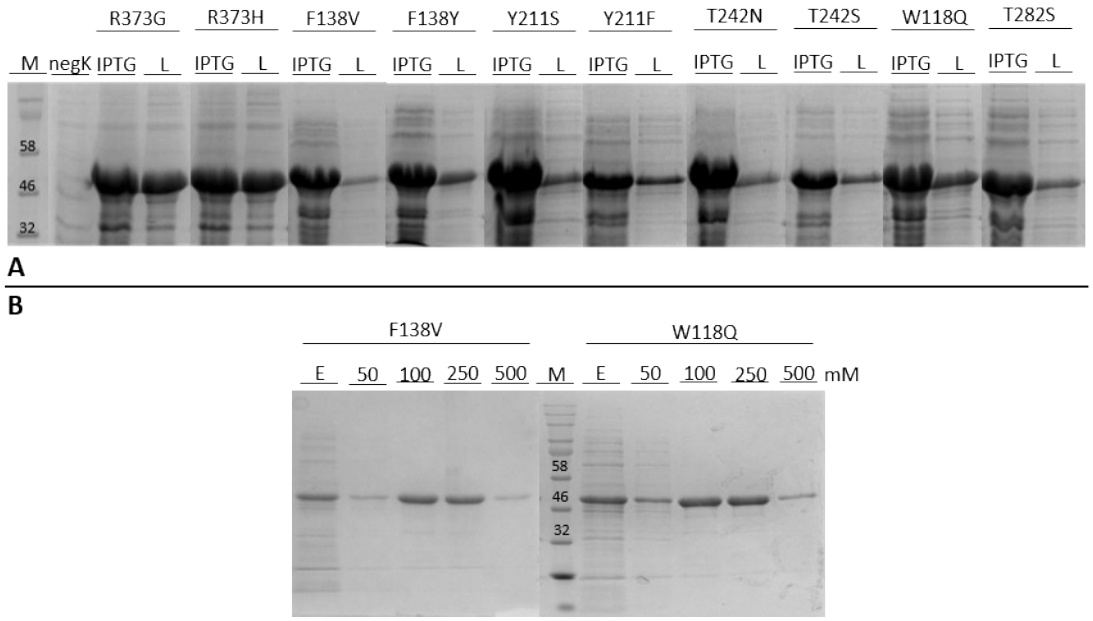


Figure S2: SDS-Page for purification of WT BcBOT2. Purification was done by discontinous imidazol gradient (25 mM up to 500 mM). Colour Prestained Protein Standard, Broad Range 10-20% Tris-glycine was used as marker. Selected marker bands were labelled with kb size.



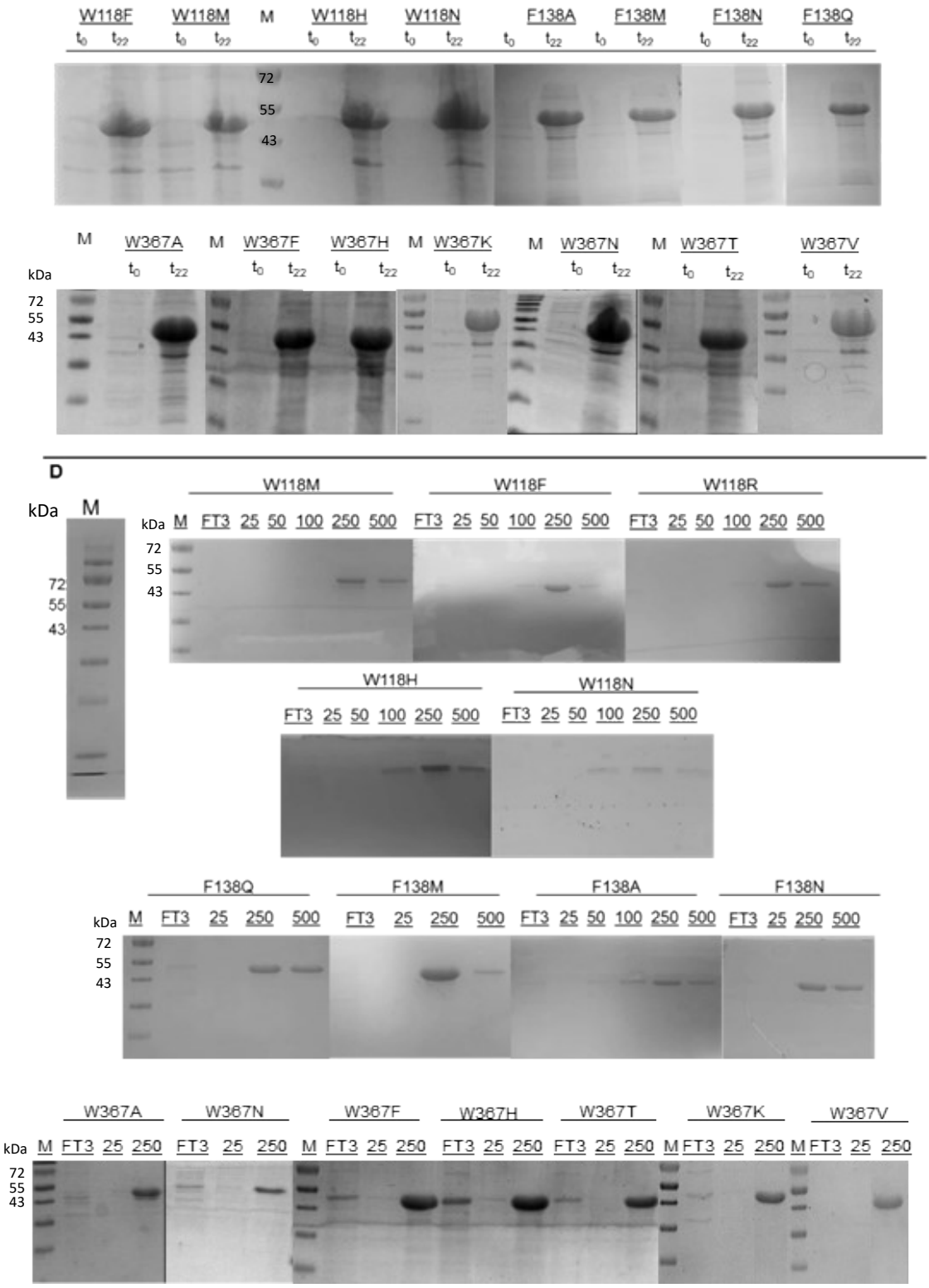


Figure S3: A: SDS-PAGE of *E. coli*-BL21 (DE3) major cultures before (negK) and after induction with 0.5 mM IPTG and the plotting of the lysate (L) to demonstrate protein solubility for the BcBOT2 variants. **B**: SDS page of protein purification by nickel affinity chromatography for BcBOT2 variants F138V and W118Q. The elution steps with 50 mM - 500 mM imidazole are shown next to the flow rate (E). **C**: SDS page of E.coli BL21(DE3) main cultures before (t_0) and after induction (t_{22}) with IPTG. **D**: SDS page of protein purification by nickel affinity chromatography for BcBOT2 variants W118F, W118M, W118H, W118N, W118R, F138Q, F138M, F138A , F138N, W367A, W367F, W367H, W367K, W367N, W367T, W367V.

2.9 Biotransformations

Biotransformation of WT BcBOT2 with FPP

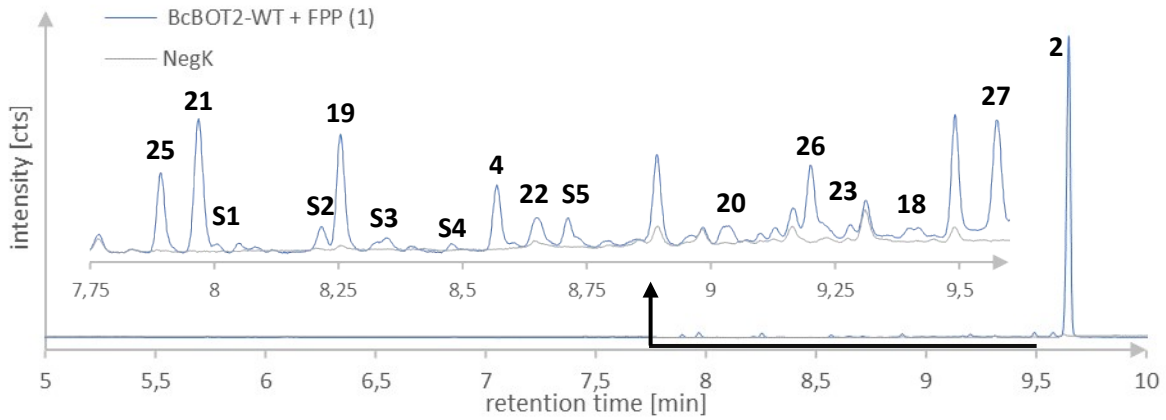


Figure S4: Chromatogram of the biotransformation of WT BcBOT2 with FPP.

Table S9: Additional biotransformation products that could not be identified by RI or MS-spectroscopy. The table contains the RI (DB5-column) and die area percentages (A [%]) of the compounds.

	25	**S1**	**S2**	**S3**	**S4**	**S5**	**26**	**27**
RI_{DB5}	1345	1368	1392	1412	1433	1470	1550	1612
A [%]	0.9	0.1	0.3	0.3	0.1	0.3	0.9	1.3

Table S10: Semi-quantitative determination in percent [%] of the detected compounds for the wild type (WT) and all variants. If a compound from the wild type could not be detected in the variants, it is marked with a horizontal bar (-). Yields indicated by "<" were not calculated semi-quantitatively but were derived from the calculated yields of the other byproducts. All calculations are based on a triplicate determination of the biotransformations.

product	**WT**	**R373G**	**R373H**	**F138V**	**F138Y**	**F138A**	**F138M**	**F138N**	**F138Q**
25	0.4 ± 0.1	< 0.1	< 0.1	3.2 ± 0.2	< 0.2	0.46 ± 0.07	1.33 ± 0.14	0.53 ± 0.1	0.6 ± 0.07
21	0.9 ± 0.1	-	-	0.2 ± 0.1	-	-	-	-	-
S1	< 0.1	-	-	< 0.1	-	-	-	-	-
S2	< 0.2	-	-	-	-	-	-	-	-
19	0.5 ± 0.1	-	-	0.3 ± 0.1	-	-	-	-	-
S3	< 0.2	-	-	0.4 ± 0.1	-	-	-	-	-
S4	< 0.1	-	-	-	-	-	-	-	-
4	0.4 ± 0.1	0.2 ± 0.1	< 0.1	6.8 ± 0.4	0.2 ± 0.1	2.3 ± 0.23	1.39 ± 0.09	5.92 ± 0.82	6.52 ± 0.72
22	0.3 ± 0.1	-	-	0.6 ± 0.1	0.7 ± 0.1	9.52 ± 1.16	9.27 ± 0.86	1.93 ± 0.38	6.02 ± 0.65
S5	< 0.2	-	-	-	-	-	-	-	-
20	<0.2	-	-	-	-	-	-	-	-
26	0.5 ± 0.1	-	-	4.7 ± 0.3	0.3 ± 0.1	1.96 ± 0.28	4.73 ± 0.8	0.29 ± 0.05	1.15 ± 0.1
23	< 0.1	-	-	-	-	-	-	-	-
18	< 0.1	-	-		-	-	-	-	-
28	-	-	-	-	-	-	-	-	-
27	0.8 ± 0.1	< 0.1	< 0.1	3.4 ± 0.2	0.3 ± 0.1	2.8 ± 0.37	1.7 ± 0.28	6.08 ± 0.79	4.3 ± 0.33
2	52.6 ± 4.2	-	-	4.5 ± 0.3	0.7 ± 0.1	0.82 ± 0.14	0.67 ± 0.05	0.58 ± 0.06	0.59 ± 0.04

24	-	-	-	-	-	-	-	-	-

product	Y211S	Y211F	T242N	T242S	W118Q	T282S	W118F	W118M	W118H	W118N
25	0.4 ± 0.1	0.4 ± 0.1	0.5 ± 0.1	0.4 ± 0.1	-	0.5 ± 0.1	0.65 ± 0.02	-	-	-
21	1.2 ± 0.1	1.1 ± 0.2	0.8 ± 0.1	0.9 ± 0.1	-	0.7 ± 0.2	-	-	-	-
S1	< 0.1	< 0.1	< 0.2	< 0.1	-	< 0.2	-	-	-	-
S2	0.2 ± 0.1	< 0.2	< 0.1	< 0.2	-	< 0.2	-	-	-	-
19	0.5 ± 0.1	0.5 ± 0.1	0.7 ± 0.1	0.6 ± 0.1	-	0.6 ± 0.1	-	-	-	-
S3	< 0.1	< 0.1	< 0.2	< 0.2	-	< 0.2	-	-	-	-
S4	1.3 ± 0.1	< 0.1	< 0.2	< 0.2	-	< 0.2	-	-	-	-
4	< 0.1	0.7 ± 0.1	0.5 ± 0.1	0.4 ± 0.1	< 0.2	< 0.2	1.02 ± 0.13	0.64 ± 0.08	0.55 ± 0.04	0.32 ± 0.05
22	< 0.1	< 0.1	0.5 ± 0.1	0.4 ± 0.1	-	< 0.2				
S5	< 0.1	< 0.1	< 0.2	< 0.2	-	< 0.2	-	-	-	-
20	< 0.1	< 0.1	< 0.2	< 0.2	< 0.1	< 0.2	-	-	-	-
26	1.1 ± 0.1	0.4 ± 0.1	0.5 ± 0.1	< 0.2	-	0.6 ± 0.1	8.17 ± 0.52	3.78 ± 0.25	-	-
23	< 0.1	< 0.1	< 0.2	0.5 ± 0.1	-	< 0.2	-	-	-	-
18	< 0.1	< 0.1	< 0.1	< 0.2	-	< 0.2	-	-	-	-
28	-	-	-	-	-	-	-	-	-	-
27	0.5 ± 0.1	0.7 ± 0.1	0.8 ± 0.1	0.6 ± 0.1	0.4 ± 0.1	0.8 ± 0.1	1.5 ± 0.09	1.47 ± 0.12	0.69 ± 0.07	0.54 ± 0.07
2	50.1 ± 3.3	49.4 ± 7.1	49.2 ± 1.3	50.6 ± 5.3	0.7 ± 0.1	43.9 ± 0.5	42.25 ±3.1	11.07 ± 0.8	3.14 ± 0.28	-
24	-	-	-	-	18.7 ± 0.8	-	9.18 ± 0.65	44.2 ± 2.69	41.58 ±2.13	20.29 ± 1.01

Continuation Table S10:

product	W367A	W367F	W367H	W367K	W367N	W367T	W367V
25	-	0.21 ± 0.06	< 0.2	-	-	< 0.2	< 0.1
21	-	3.15 ± 0.58	0.45 ± 0.09	-	-	0.24 ± 0.06	< 0.1
S1	-	1.17 ± 0.3	-	-	-	< 0.1	-
S2	-	-	-	-	-	< 0.2	-
19	-	-	-	-	-	< 0.2	-
S3	-	-	-	-	-		-
S4	-	-	-	-	-		-
4	-	< 0.2	0.5 ± 0.12	-	-	< 0.1	-
22	-	-	-	-	-	-	-
S5	-	-	-	-	-	-	-
20	-	-	-	-	-	-	-
26	0.45 ± 0.05	0.84 ± 0.18	0.43 ± 0.11	-	0.49 ± 0.11	0.27 ± 0.07	< 0.1
23	-	-	-	-	-	-	-
18	-	-	-	-	-	-	-
28	0.9 ±	-	-	-	-	-	-
27	0.09	0.27 ± 0.06	< 0.2	-	-	< 0.1	-
2	4.1 ± 0.42	24.47 ± 5.58	2.43 ± 0.65	-	2.74 ± 0.57	2.56 ± 0.21	< 0.2
24	-	-		-	-	-	-

2.10 Temperature optimization with BcBOT2-mutants and FPP

The analytical enzyme assays were conducted in analogy to the general procedure. For thermostability testing, the enzyme assays were incubated at various temperatures ranging from 20°C - 50°C and the yield of product was determined by addition of an internal standard. All biotransformations were done in triplicate. The optimal temperature for WT BcBOT2 was determined between 20-25°C.[6]

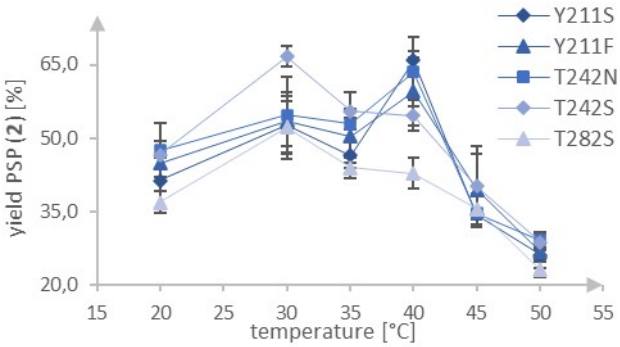


Figure S5: Analysis of thermostability for BcBOT2 mutants Y211S, Y211F, T242N, and T242S. The yields for **2** were determined using an added internal standard.

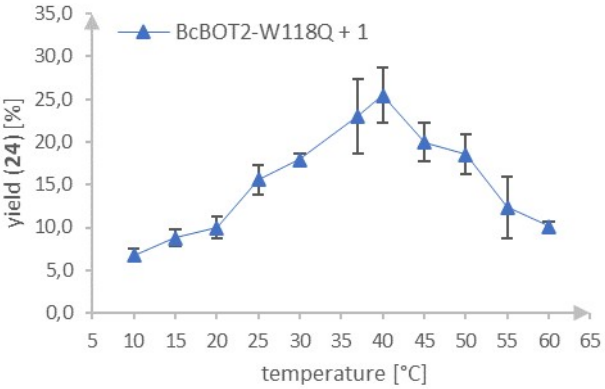


Figure S6: Analysis of the thermostability for BcBOT2-W118Q mutant. The determined thermostability agrees with that found by Arnold and coworkers.[8]

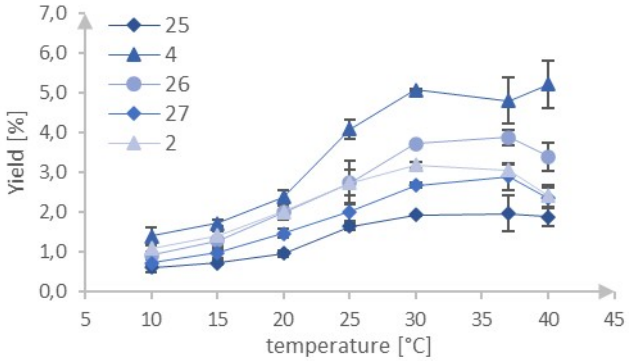


Figure S7: Analysis of thermostability for BcBOT2-F138V and yields of the five main products **2**, **4**, and **25-27**.

2.11 Upscaling of biotransformations of BcBOT2-mutants F138V and W118Q with FPP

Table S11. Data on 2D NMR and 1d-NOE experiments of product **25**.

No.	δ_c	δ_H	#H	Multi	J(Hz)	COSY	HMBC	1d-NOE
1	151.6						2, 3a, 3b, 7, 8, 9, 10b, 15	
2	113.8	4.95	1	t	2.3	3a, 3b, 8	3a, 3b, 8, 9	3a, 3b, 9, 15
3a	52.0	2.35	2	dt	2.5, 16.5	2,3b, 8, 14	2, 5a, 5b, 8, 14	2, 3b, 5a, 7, 12
3b		2.43		dd	1.8, 16.5	2, 3a		2, 3a, 14
4	43.5						2,3a, 3b, 8, 14	
5a	59.8	1.55	2	d	12.9	5b, 13	3a, 3b, 12, 13, 14	
5b		1.62		d	12.9	5a, 12, 14		
6	42.6						5a, 5b, 12, 13	
7	63.5	0.97	1	td	2.5, 12.0	8, 11a, 11b	5a, 5b, 11a, 11b, 12, 13	2, 3a, 5a, 10a, 11b
8	63.3	2.26	1	dd	2.6, 11.9	2, 3a, 7	2, 3a, 3b, 5a, 5b, 7, 9, 11a, 11b, 14	5b, 11a, 13, 14, 15
9	33.0	2.73	1	m		10a, 15	2, 10b, 11a, 11b, 15	2, 10a, 10b, 15
10a	36.9	1.53	2	m		10b, 9, 11a, 11b	7, 9, 11a, 11b, 15	
10b		1.62		m		10a, 11a, 11b, 15		
11a	22.4	1.31	2	qd	4.3, 12.3	7, 10a, 10b, 11b	8, 9, 10a	8, 11b, 13, 15
11b		1.45		ddt	2.5, 4.7, 12.3	7, 10a, 10b, 11a		7, 10a, 10b, 11a, 12, 13
12	28.1	0.96	3	s		5b, 13	5a, 13	5a, 11b, 13
13	21.6	0.89	3	s		5a, 12	5a, 5b, 7, 12	5b, 8, 11a, 12, 14
14	30.4	1.17	3	s		3a, 5b	3a, 3b, 5a, 5b, 8	3b, 5b, 8, 13
15	19.0	1.10	3	dd	0.3, 7.1	9,10b	3a, 3b, 9, 10a, 10b	2, 8, 9, 10b, 11a

Table 12. Data of 2D NMR and 1d-NOE experiments of product **27**.

N~o.	δ_c	δ_H	#H	Multi	J(Hz)	COSY	HMBC	1d-NOE
1a **1b**	37.5	1.48 1.64	2	M T	 10.0	1b,2 1a, 2, 13	2, 12, 13	 1a, 4a, 12,15
2	52.6	1.44	1	M		1a, 1b, 10	1a, 1b, 4a, 4b, 10, 15	13
3	73.2						1b, 4a, 4b, 5a, 5b, 10, 15	
4a **4b**	41.7	1.40 1.42	2	M M		4b, 5a, 5b 4a, 5a, 5b	5a, 5b, 15	1b, 6, 10, 15 5b
5a **5b**	23.4	2.00 2.10	2	M M		4a, 4b, 5b, 6 4a, 4b, 5a, 6	4b, 6	 4a, 6, 14
6	123.5	5.37	1	T	7.9	5a, 5b, 14	4a, 4b, 5a, 5b, 8a, 8b, 14	4a, 5b, 8a, 10
7	135.8						5a, 5b, 8a, 8b, 9b, 14	
8a **8b**	40.7	1.84 1.89	2	Qd Dt	12.3, 3.9 7.3, 3.6	8b, 9a, 9b 8a, 9a, 9b	6, 10, 14	6, 9b, 10 6, 9a, 9b, 14
9a **9b**	30.6	1.18 1.32	2	Dddd M	14.5, 12.8, 10.7, 4.0	8a, 8b, 9b, 10 8a, 8b, 9a	8a, 8b, 13	8a, 9b, 13 5b, 8b, 9a, 12, 13
10	46.8	1.59	1	M		2, 9a	1b, 2, 8a, 8b, 12, 13	6, 8a, 9b, 12
11	31.8						1a, 1b, 10, 12, 13	
12	30.2	0.96	3	S		13	1a, 1b, 10, 13	1b, 9b, 10, 13
13	23.8	0.87	3	S		1b, 12	1a, 1b, 10, 12	1a, 2, 9a, 9b, 12
14	16.4	1.47	3	M		6	6, 8a, 8b	5b, 8b, 13, H15
15	31.1	0.82	3	S		-		1b, 4a, 4b
16	*hydroxyl group*							

3 Computational section

3.1 Protein modelling and substrate docking

The structural model of *Bc*BOT2 was predicted using the AlphaFold2 program [9], with the sequence of *Bc*BOT2 FASTA format as input (fetched from UniProt – Entry: **Q6WP50**). Models were built exclusively as monomers, as this is proposed to be sufficient organization for catalytic functionality in *Bc*BOOT2 and other sesquiterpenecyclases [10-12]. Models were then scored by calculating various geometrical, energetical and structural descriptors (QMEAN4 [13], QMEANDisCO [13], Energy Z score [14], MolProbitity score [15], Clash score [15]). Additionally, coverage of models was checked, and the validity of secondary structure for each model was analyzed respectively, by comparing them to the secondary structure predictions for *Bc*BOT2 sequence using different serves (Porter 5.0 [16], PSIPRED [17], PSSPred [18], Jpred4 [19], NetSurfP [20]). In total, five models were evaluated in this manner. The final model was selected based on the best performance during the evaluation process (See Table S13). All models generated by AlphaFold2 appear to predict an apparently disordered loop in the N-terminal region between positions 1-45 (See Figs. S8, S9). Additional analysis to define this region was performed to test whether the predicted disordered region was due to the sequence data or the lack of structural information. Secondary structure was determined based on the protein sequence using five different secondary structure prediction servers (see Figs. S10-S12). A comparison of the predicted and modeled secondary structure for the BcBOT2 sequence showed a high degree of resemblance, especially the existence of the disordered loop in the N-terminal region (1-45) (see Figs. S10-S12). In addition, the N-terminal region exhibits a low degree of evolutionary conservation and, arguably, concomitant reduced biological significance (see Figs. S13, S14).

Based on the final apoenzyme model for BcBOT2, a second holoenzyme model with trinuclear Mg^{2+} ions was built by using an existing crystal structure with bound trinuclear Mg^{2+} ions of a similar STC as a template. The crystal structure of TEAS from *Nicotiana tabacum* with bound PP_i and trinuclear Mg^{2+} ions (PDB ID: 5IKA[21]) was used as a template and aligned with the BcBOT2 model through the PyMOL2[22] "align" command, producing a structural alignment with a global RMSD of 2.46 Å. Further, visual inspection revealed high structural similarities of the active site of both enzymes, especially when comparing the metal binding DDXXD – and DTE/NSE – motifs (Figure S15). Therefore, the positions of the trinuclear Mg^{2+} ions of the TEAS crystal structure were deemed sufficient for the BcBOT2 model and directly extracted as heteroatoms to the BcBOT2 model via the PyMOL2[22] "copy" command. The generated holoenzyme model of BcBOT2 with trinuclear Mg^{2+} was exported, energy minimized, and used for docking of FPP substrate.

Table S13: Evaluation of the five *Bc*BOT2 models predicted by AlphaFold2 program. Generation of structural/statistical descriptors QMEAN4 and QMEANDisCo (as Z scores) was done using "QMEAN" – tool of SWISS-MODEL server (https://swissmodel.expasy.org/) [13] . Normalized energy Z – score for the models was calculated using ANOLEA server (Melolab.org/anolea) [14]. Geometrical/structural evaluation through Clash and MolProbitity scoring was done using MolProbitity server (http://molprobity.biochem.duke.edu/) [15]. The first numeric values in these two bracket describe the crystallographic resolution they expect the model to be in for their observed structural/geometrical quality of the model and the second percentile value describes how the model ranked quality wise in comparison to other models in similar resolution range [15]. Good, mediocre and bad scoring for each descriptor, except QMEANDisCo, are highlighted in green, yellow and red, respectively. For QMEANDisCo the score is colored representatively to their ranking based on a scale 0 (bad – orange) to 1 (good – blue).

	Number	***QMEANDisCo***	***QMEAN4***	***Energy Z-score***	***MolProbity score*** *(normalized geometry)*	***Clash score*** *(atomic contacts)*	***Oligomerization***
Alphafold2 models of *Bc*BOT2	*1*	0,64	- 0,72	1,64	2,72 (98th %)	1,65 (91th %)	Monomer
	2	0,65	- 0,49	1,51	1,92 (99th %)	1,46 (96th %)	Monomer
	3	0,64	- 1,09	1,63	1,44 (99th %)	1,32 (98th %)	Monomer
	4	0,65	- 0,16	1,57	1,12 (99th %)	1,37 (98th %)	Monomer
	5	0,65	- 0,99	1,73	1,92 (99th %)	1,35 (98th %)	Monomer

Legend

QMeanDisCo
0 (bad) 1 (good)
0.0 0.5 1.0

QMean4 / Energy Z-score
good | mediocre | bad
|Score|<1 | 1<|Score|<2 | |Score|>2

Clash score / MolProbity score
good | mediocre | bad
Perc. ≥ 66 | 66 > Perc. ≥ 33 | Perc. < 33

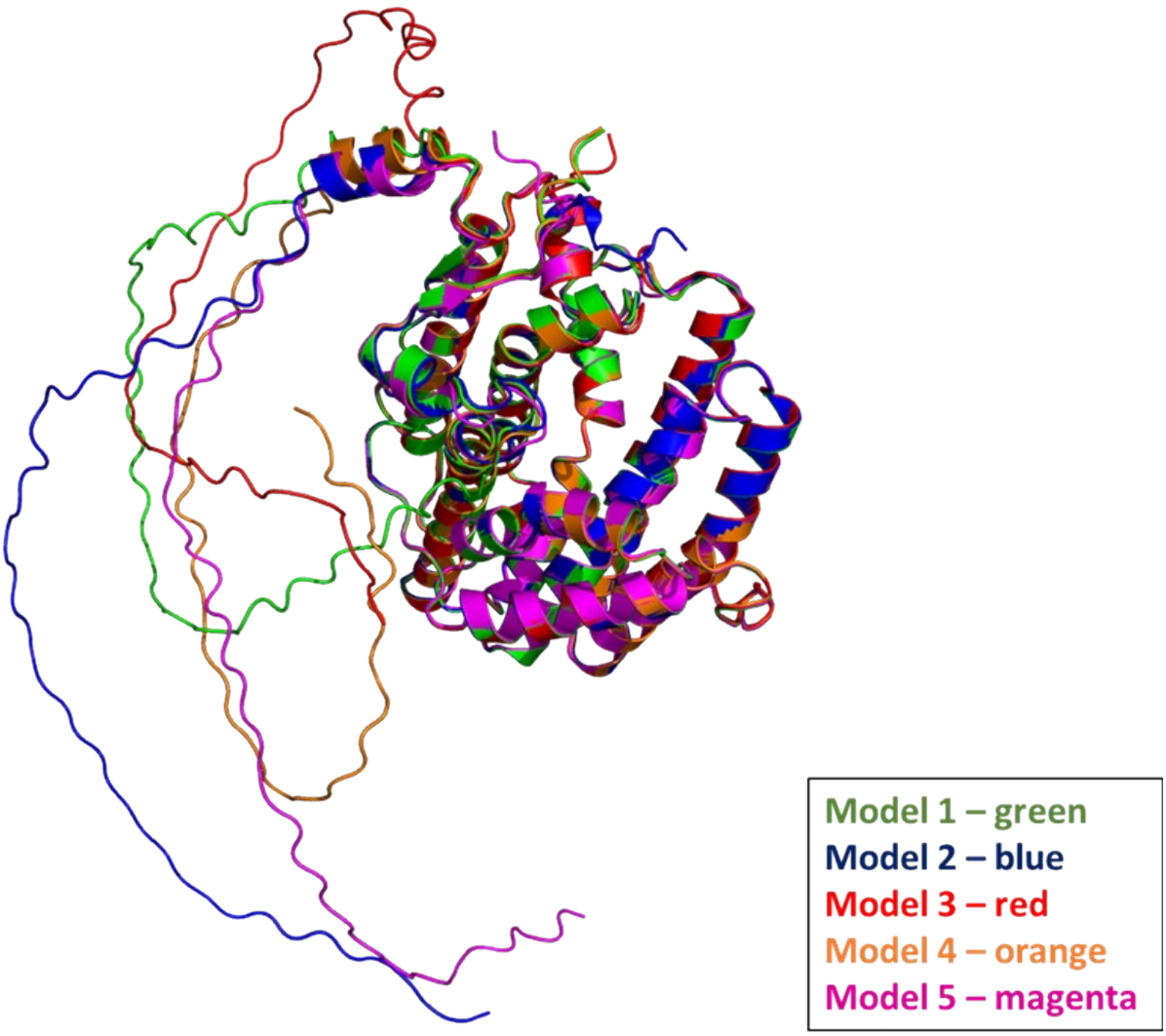


Figure S8: Multiple structural alignment of models for *Bc*BOT2 build using AlphaFold2. The models are shown as cartoon representations in respected colors (model 1 in green, model 2 in blue, model 3 in red, model 4 in orange, model 5 in magenta). N – terminal part (1 – 45) of each model was mostly unstructured, thus showing indications for an intrinsically disordered region.

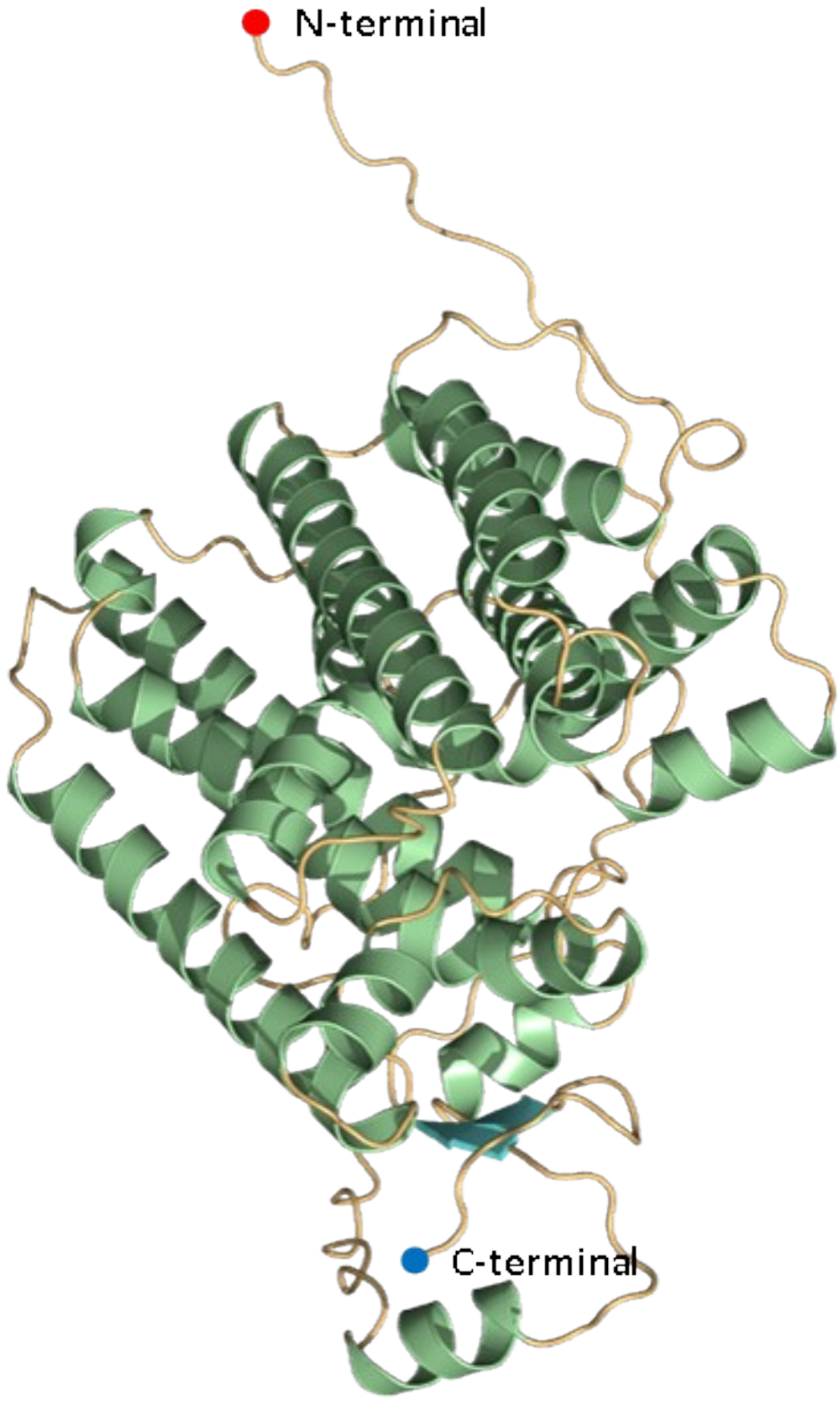


Figure S9: Structural model of *Bc*BOT2 constructed using AlphaFold2 (Model Number 5). The model is shown in cartoon representation. Alpha helices are colored green. Loops are shown orange. Beta sheets are colored cyan. N – terminal is circled in red, c - terminal in blue. It is mostly composed of alpha helices interconnected by loops, with two short beta strands inserted near the C – terminal region. The N – terminal region (1 – 45) was mostly unstructured, thus showing indications for intrinsic disordering. Geometrical, energetical and structural descriptors showed, similar to the other four models, entirely good to sufficient scoring, rendering it an adequate candidate for further analysis.

	Sequence Positions:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Servers	Amino Acids:	M	A	I	P	A	L	E	P	Q	L	H	D	A	D	T	S	S	N	N	M
Jpred 4	3 - class pred. (Q3)	C	C	C	C	C	H	C	C	C	C	C	C	C	C	C	C	C	C	C	C
NetSurfP	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Porter Ver. 5.0	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
PSIPRED	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
PSSpred	3 - class pred. (Q3)	C	C	C	C	C	C	C	H	H	H	C	C	C	C	C	C	C	C	C	C

	Sequence Positions:	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Servers	Amino Acids:	S	S	N	S	T	D	S	G	Y	D	T	N	S	T	T	P	L	E	K	S
Jpred 4	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
NetSurfP	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Porter Ver. 5.0	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
PSIPRED	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
PSSpred	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C

	Sequence Positions:	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Servers	Amino Acids:	E	K	P	N	T	Q	E	L	K	Q	Q	Q	L	D	P	K	R	P	P	F
Jpred 4	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	E
NetSurfP	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Porter Ver. 5.0	3 - class pred. (Q3)	C	C	C	C	C	C	H	H	H	C	C	C	C	C	C	C	C	C	C	C
PSIPRED	3 - class pred. (Q3)	C	C	C	C	H	H	H	H	H	H	H	H	H	C	C	C	C	C	C	E
PSSpred	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	H	H	H	H	H	C	C	C	C	C	C	E

	Sequence Positions:	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
Servers	Amino Acids:	V	R	V	P	D	L	F	G	S	I	M	S	T	K	P	V	V	N	P	N
Jpred 4	3 - class pred. (Q3)	E	E	E	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	H	H
NetSurfP	3 - class pred. (Q3)	E	E	C	C	C	C	C	H	H	C	C	C	C	C	C	C	C	C	C	C
Porter Ver. 5.0	3 - class pred. (Q3)	E	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
PSIPRED	3 - class pred. (Q3)	E	E	C	C	C	C	H	H	H	C	C	C	C	C	C	C	C	C	C	C
PSSpred	3 - class pred. (Q3)	E	E	E	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	H

C	— loop/coil
H	— α helix
E	— β strand

Figure S10: Comparing secondary structure prediction of N – terminal region (amino acids 1 – 80) for *Bc*BOT2 sequence using a variety of prediction servers. Predictions were performed according to the algorithms of the respected server (Jpred4, NetSurfP, Porter Ver. 5.0, PSIPRED, PSSPred) with default parameters. For each prediction, Q3 secondary structure classification was carried out, allowing differentiation between alpha helices (colored green), beta strands (colored blue) and loops/coils (colored orange).

	Sequence Positions:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Models	Amino Acids:	M	A	I	P	A	L	E	P	Q	L	H	D	A	D	T	S	S	N	N	M
Rank 1	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Rank 2	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Rank 3	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Rank 4	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Rank 5	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C

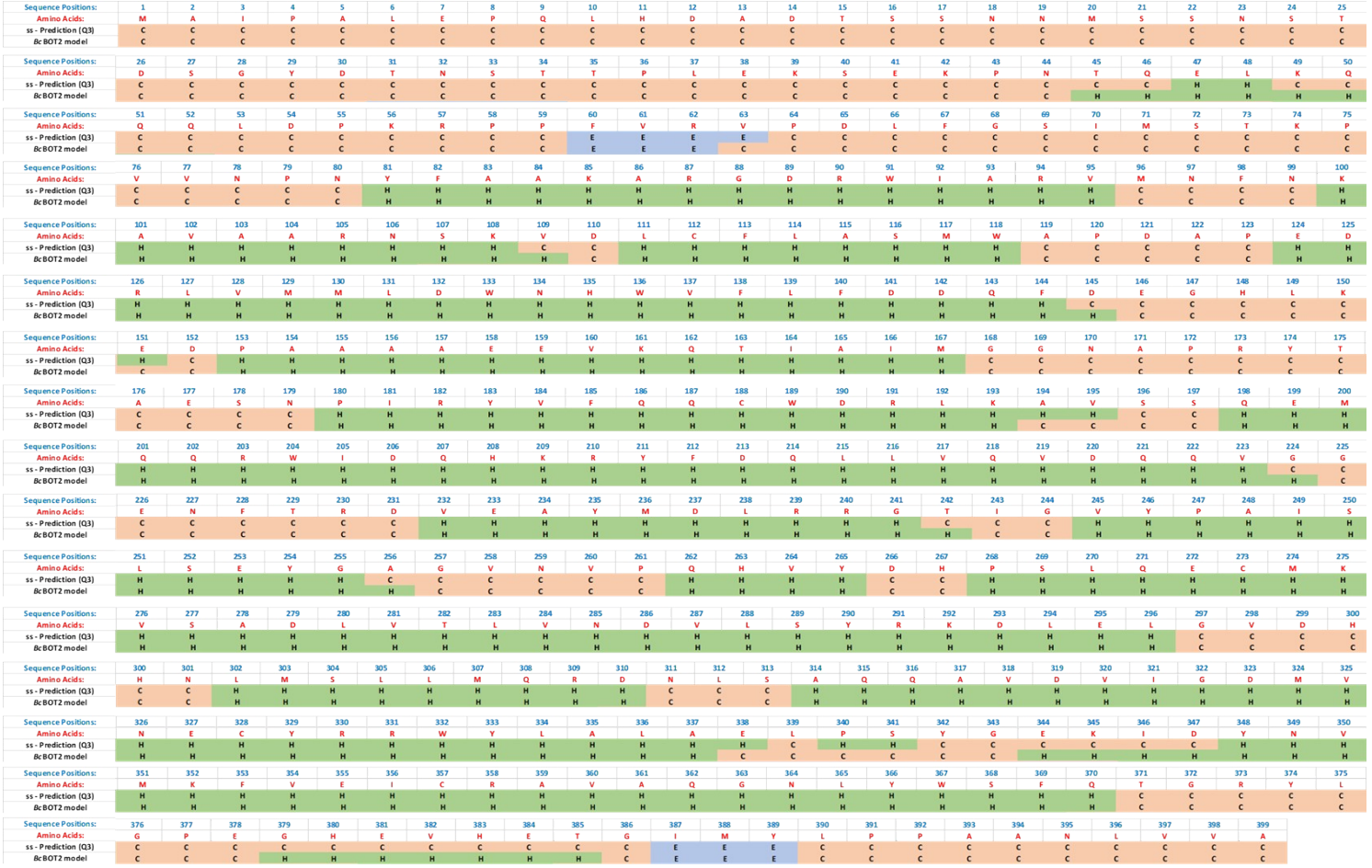
	Sequence Positions:	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Models	Amino Acids:	S	S	N	S	T	D	S	G	Y	D	T	N	S	T	T	P	L	E	K	S
Rank 1	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Rank 2	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Rank 3	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Rank 4	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Rank 5	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C

	Sequence Positions:	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Models	Amino Acids:	E	K	P	N	T	Q	E	L	K	Q	Q	Q	L	D	P	K	R	P	P	F
Rank 1	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	E
Rank 2	3 - class pred. (Q3)	C	C	C	C	H	H	H	H	H	H	H	C	C	C	C	C	C	C	C	E
Rank 3	3 - class pred. (Q3)	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	E
Rank 4	3 - class pred. (Q3)	C	C	C	C	H	H	H	H	H	H	C	C	C	C	C	C	C	C	C	E
Rank 5	3 - class pred. (Q3)	C	C	C	C	H	H	H	H	H	H	H	C	C	C	C	C	C	C	C	E

	Sequence Positions:	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
Models	Amino Acids:	V	R	V	P	D	L	F	G	S	I	M	S	T	K	P	V	V	N	P	N
Rank 1	3 - class pred. (Q3)	E	E	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Rank 2	3 - class pred. (Q3)	E	E	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Rank 3	3 - class pred. (Q3)	E	E	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Rank 4	3 - class pred. (Q3)	E	E	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Rank 5	3 - class pred. (Q3)	E	E	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C

C	— loop/coil
H	— α helix
E	— β strand

Figure S11: Comparing the secondary structure of N – terminal region (amino acids 1 – 80) of *Bc*BOT2 structural models built by AlphaFold2. Structural information of each model was directly extracted by loading the pdb – files into JALVIEW Ver.2.11.2.3 and exporting the structural annotation in a CVS – file format. For each model, Q3 secondary structure classification was carried out, allowing differentiation between alpha helices (colored green), beta strands (colored blue) and loops/coils (colored orange).



C	— loop/coil
H	— α helix
E	— β strand

Figure S12: Comparing the secondary structure of the Alphafold2 generated model and the general consensus from prediction servers for the whole *Bc*BOT2 sequence. Structural information of the model was directly extracted by loading the pdb – files into JALVIEW Ver.2.11.2.3 and exporting the structural annotation in a CVS – file format.[23] The consensus was generated based on what the majority of the five prediction servers proposed for each amino acid position. Again, Q3 secondary structure classification was carried out, allowing differentiation between alpha helices (colored green), beta strands (colored blue), and loops/coils (colored orange).

3.2 Evolutionary Trace Analysis

The biological importance of BcBOT2 residues was analyzed by using the Evolutionary Trace Server with default settings (evolution.lichtargelab.org).[24, 25]

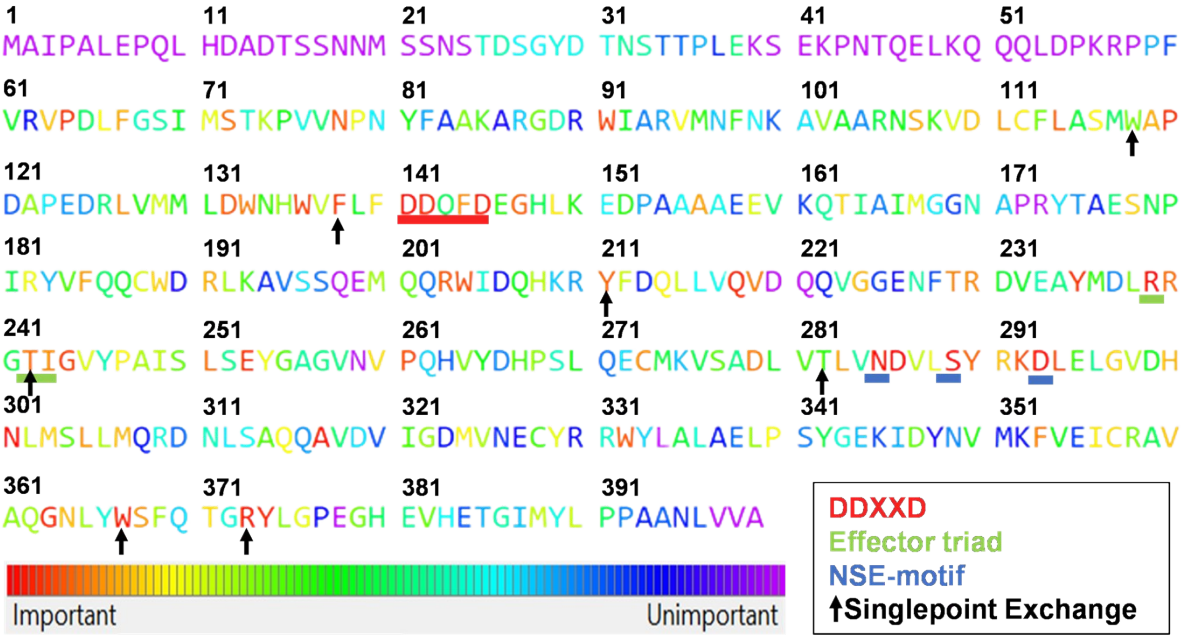


Figure S13: The amino acid sequence of BcBOT2 colored by evolutionary importance. Importance can be catalytical or structural. The DDxxD-motif is underlined red, the effector triad is underlined green, the NSE-motif is underlined blue, and the positions for substitution are marked with a black arrow. The color scheme (based on Gobblestopper) for mapping the scale of evolutionary importance to the respected amino acid positions is shown below the sequence map.

3.3 Evolutionary Conservation Analysis of BcBOT2

The evolutionary conservation of BcBOT2 amino acid residues was analyzed by using the ConSurf Server (https://consurf.tau.ac.il/) [26-29] with the BcBOT2 sequence as input and default settings.

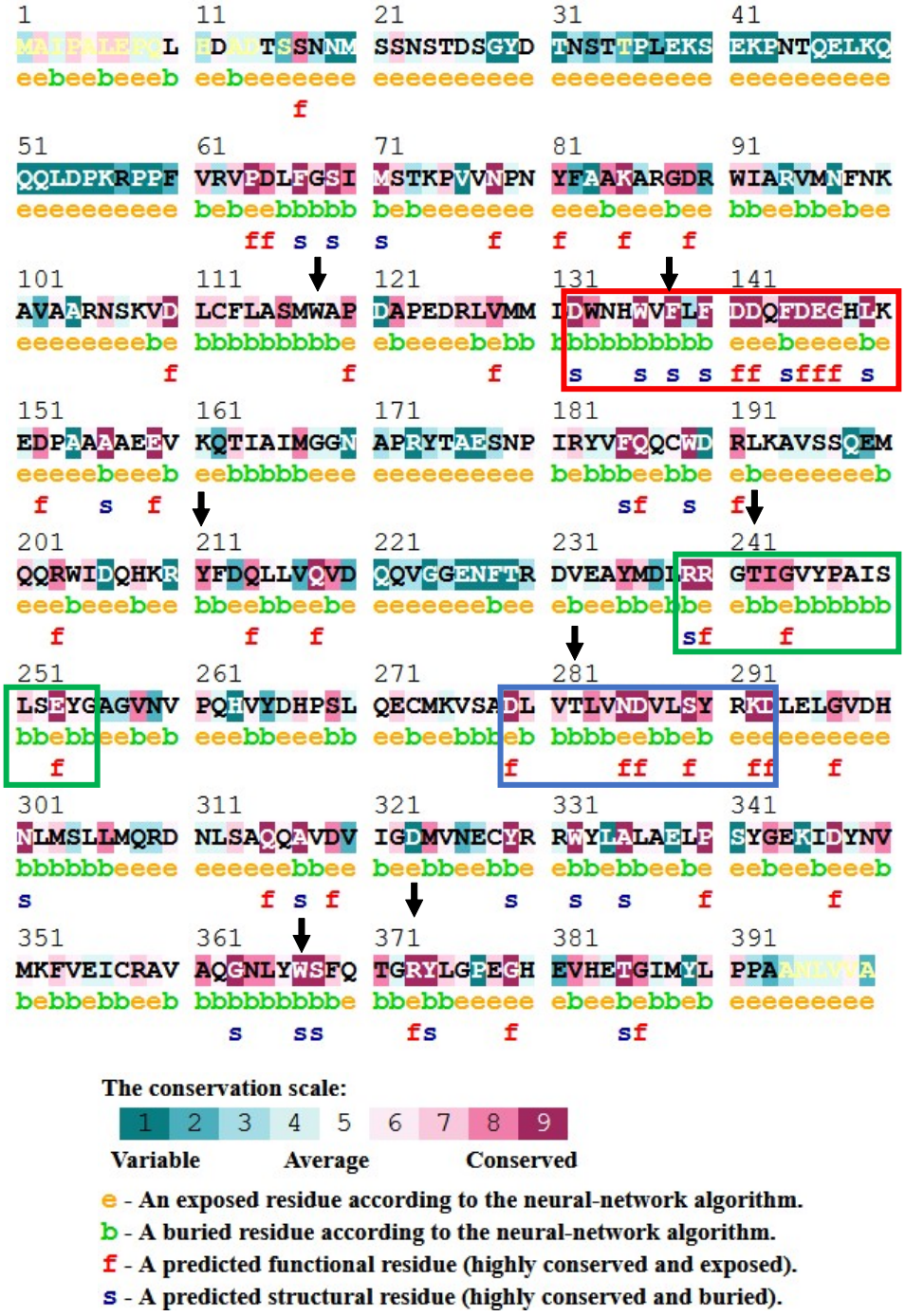


Figure S14: The amino acid sequence of BcBOT2 residue 1-399 colored by conservation. Substituted residues are marked with black arrows, the DDxxD-motif (DDQFD) is marked with a red cuboid, the effector triad is marked with a green cuboid, and the NSE-motif is marked with a blue cuboid. The scale for the conservation and the key to the abbreviations are shown below.

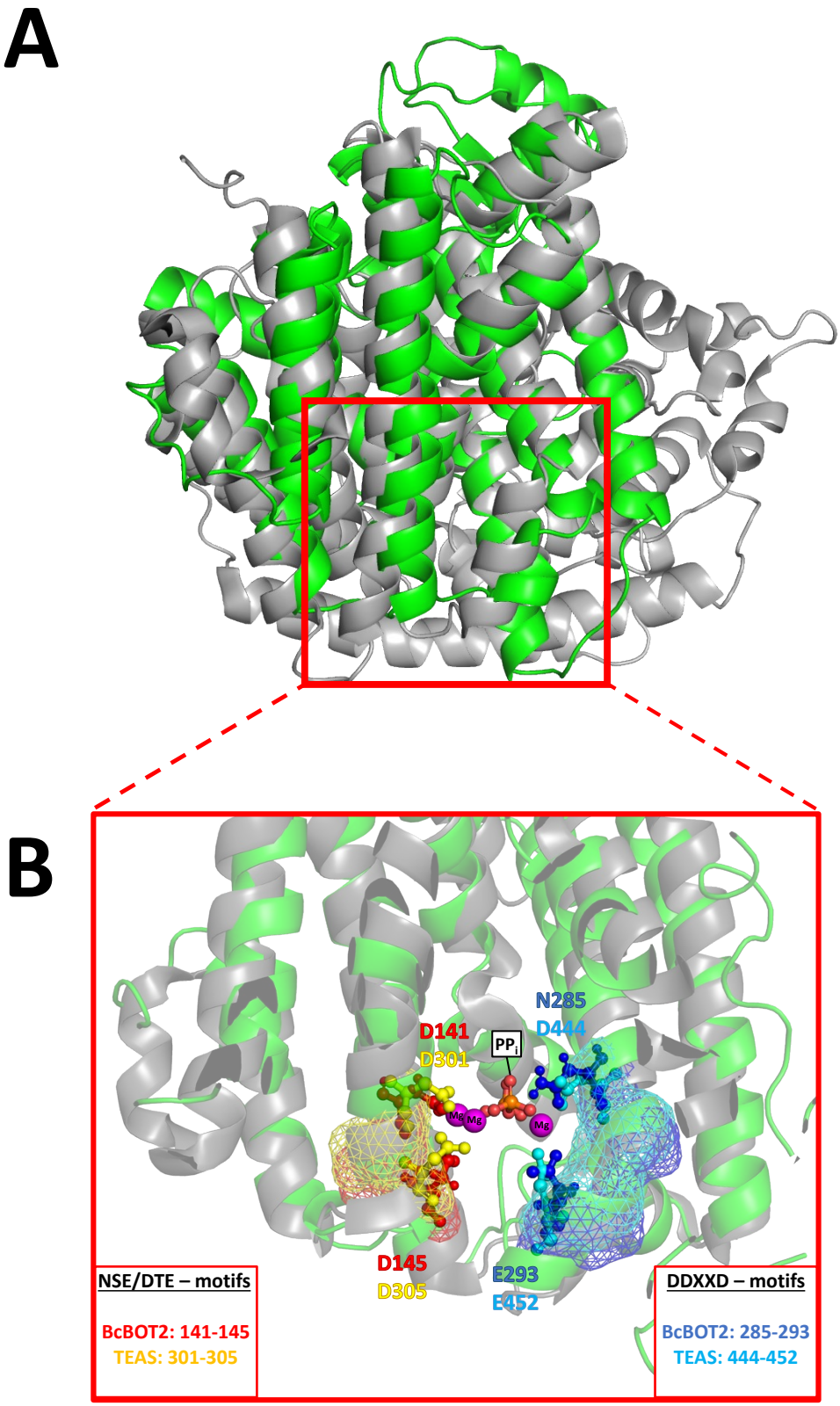


Figure S15: Structural alignment of BcBOT2 model and crystal structure of TEAS (5IKA). A: Alignment of BcBOT2 apoenzyme model (green) and TEAS crystal structure (grey), with bound PP_i and trinuclear Mg^{2+} ions, through PyMOL2[22] "align" command (global RMSD = 2.46 Å). **B:** Comparison of active sites of the aligned structures. The DDXD binding sites (BcBOT2 – red, TEAS – yellow) and NSE/DTE binding sites (BcBOT2 – blue, TEAS – cyan) and conserved residues in both structures are highlighted as mesh and ball – stick representations, respectively.

3.4 Cavity analysis of BcBOT2

Identification and analysis of actives Pockets for wildtype and single point mutants was carried out using the local fPocket tool provided by GitHub (clone: https://github.com/Discngine/fpocket.git). [30] For identification of active site pocket in wildtype *Bc*BOT2 was energy minimized based on steepest decent using AMBER10 forcefield. Then, the default fPocket tool with customized settings (m: 3 Å; M: 5.5 Å; n: 10; v: 5000) was used on the minimized structure for pocket detection. Based on mechanistic and structural knowledge of the enzyme, the most promising detected pocket was then picked, thus yielding the final active site pocket. This active site pocket is in proximity to the metal binding sites, the effector triad and stands out through its high hydrophobicity. The chosen active site has a druggability score of 0.910, a volume of 567.8 $Å^3$, a hydrophobicity score of 36.478, and a flexibility of 0.956. The active site is formed by 23 amino acids, including the DDxxD motif, the effector triad, and the N285 of the NSE-motif. Molecular docking of FPP in this active pocket was carried out, as described in "Molecular Docking of substrate and intermediates", leading to identification of binding position of natural substrate in *wild-type Bc*BOT2.

For identification of homologous active site pocket for the *in silico* generated single – substitutions mutants, structures were again energy minimized based on steepest decent using AMBER10 as forcefield. FPocket detection with customized parameters (m: 3 Å; M: 5.5 Å; n: 10; v: 5000) was carried out, with the addition of filtering and combining pockets, which had their center of mass in sufficient distance (≤4.5 Å) to any of the FPP atoms in binding position, to one explicit active pocket (using the Dpocket mode).[30] Thus, leading to the identification of the respected active site pockets for each of the variants.

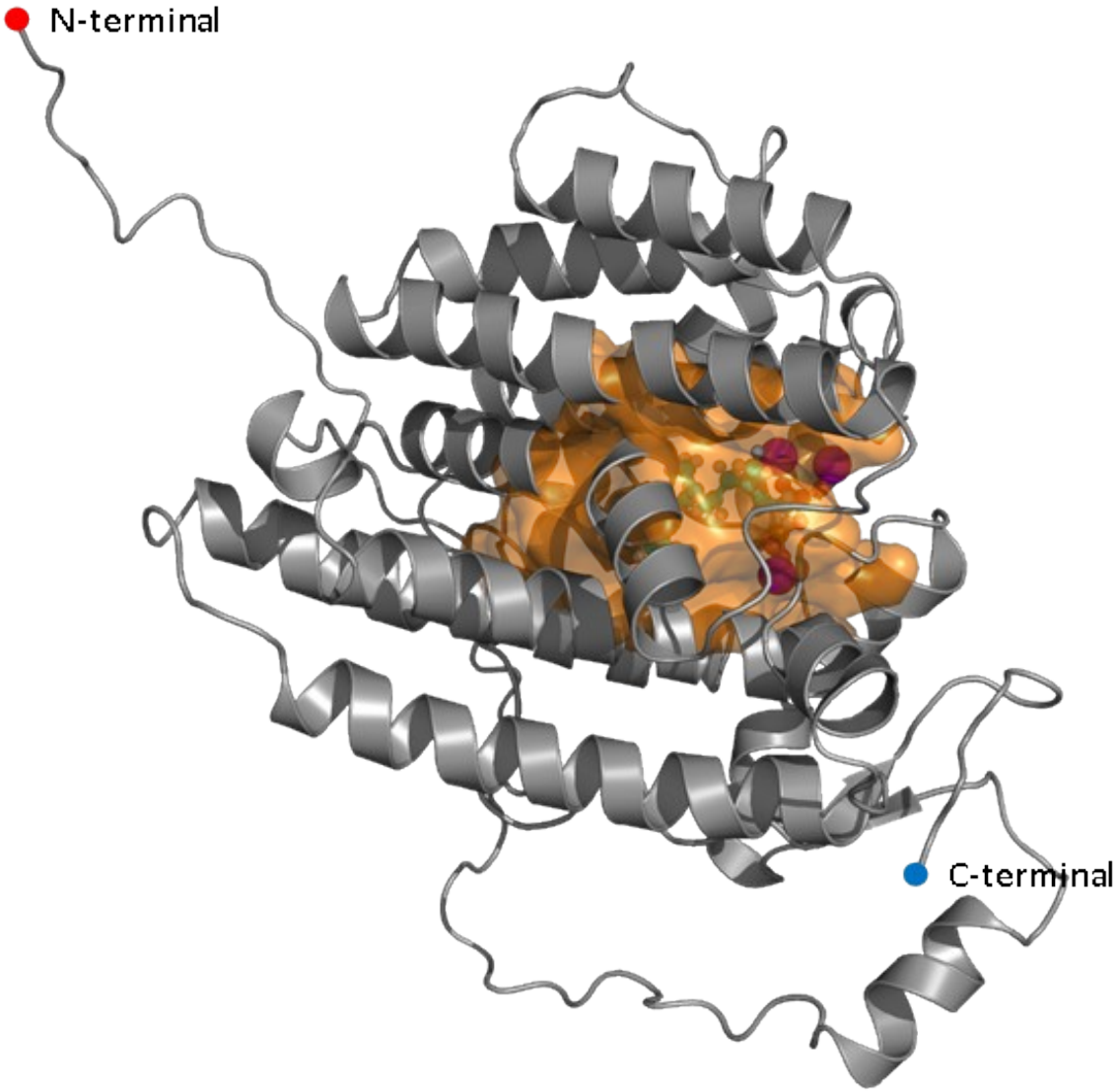


Figure S16: The structural model and active pocket of *Bc*BOT2. The model is shown in grey cartoon representation. Natural substrate FPP is displayed in a cyan ball-stick representation. Mg^{2+} ions are shown as magenta-colored spheres. The active pocket for *Bc*BOT2 (identified using the fPocket tool) is presented in orange surface representation.

Table S14: List of amino acid residues that form the active pocket. The active site is composed of the 23 amino acids listed below. The DDxxD motive is colored red, the effector triad green, and the N285 of the NSE motif is colored blue.

Amio acid residue	Location	Amio acid residue	Location
L114	**2. shell**	**I243**	**1. shell, effector triad**
W118	**1. shell, substitution**	G244	**1. shell**
W133	**2. shell**	V245	**2. shell**
N134	**2. shell**	A248	**1. shell**
V137	**1. shell**	V281	**2. shell**
F138	**1. shell, substitution**	**N285**	**1. shell, NSE-motif**
D141	**1. shell, DDxxD-motif**	V360	**2. shell**
D142	**2. shell, DDxxD-motif**	N364	**2. shell**
D145	**1. shell, DDxxD-motif**	W367	**2. shell , substitution**
Y211	**2. shell , substitution**	R373	**1. shell, substitution**
R239	**1. shell, effector triad**	Y374	**2. shell**
T242	**1. shell, effector triad, substitution**		

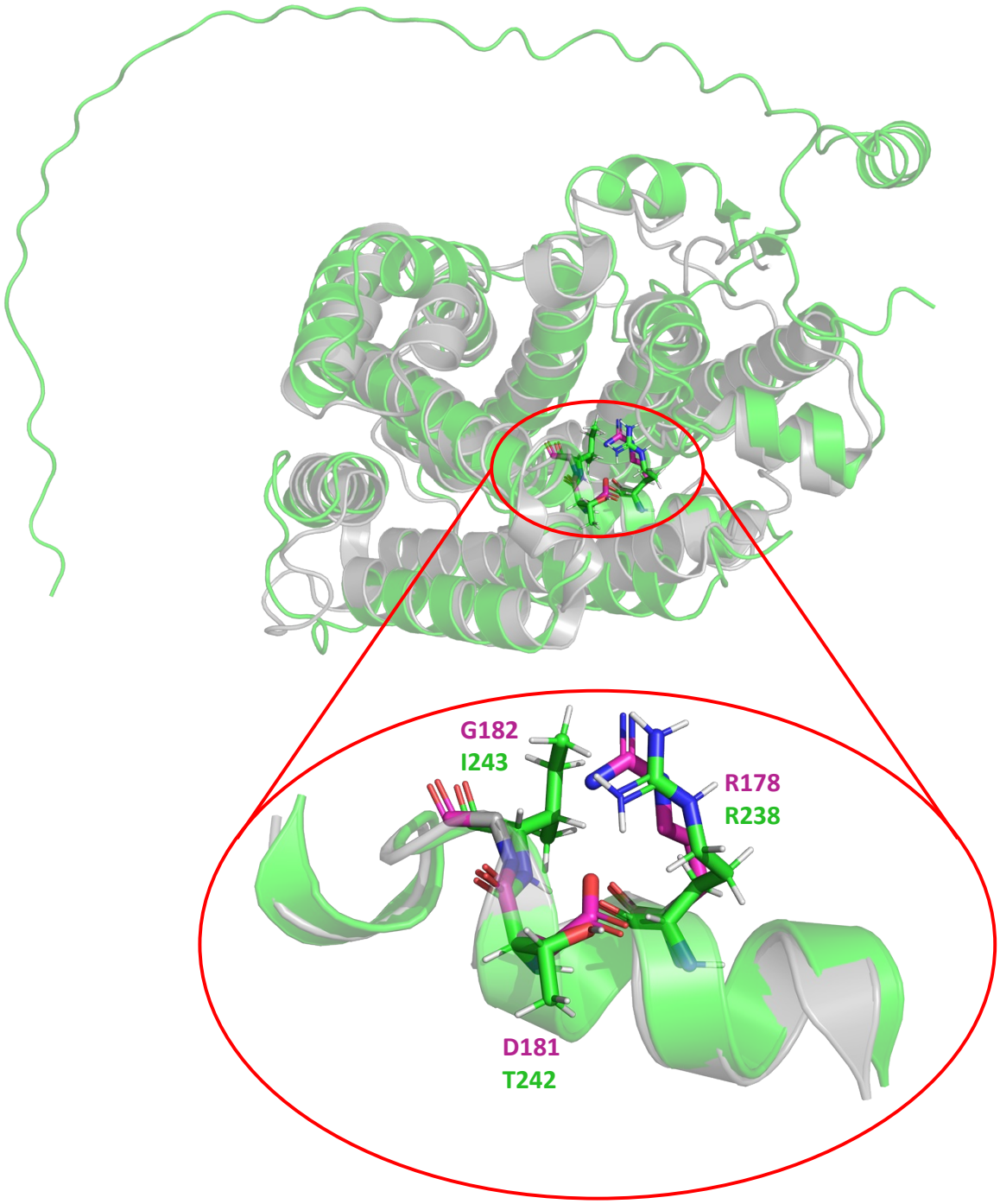


Figure S17: Structural alignment of BcBOT2 model and crystal structure of Chain A of SdS (4OKM). Alignment of BcBOT2 model (green) and Chain A of SdS crystal structure (grey) through PyMOL2[22] "align" command (global RMSD = 1.58 Å). The effector triad is red encircled and zoomed in. The zoom-in only shows a closeup of the active site structure including the effector triad and 10 surrounding amino acids. The sidechain of the effector triads are depicted in stick representation and are colored green for BcBOT2 and purple for SdS.

3.5 *In silico* mutagenesis of BcBOT2

Structural models of BcBOT2 variants (W118F, W118H, W118Q, F138V) were constructed using YASARA Version 21.12.19 [31] employing the FoldX plugin suit 4 [32]. Starting point for the mutagenesis was the previously repaired BcBOT2 model of AlphaFold2. After the substitution, the structures were again repaired. The resulting mutants were solvated in a 5 A° cube around all atoms in TIP3P water [33] and the energy was minimized using the AMBER10 force field [34].

3.6 Molecular Docking of substrate and intermediates

Molecular mechanics docking was performed using MOE 2022.02 software [35]. The docking procedure was as follows:

A ligand database was prepared (containing structures **1-2** and **6-14**) by drawing a 2D representation using ChemDraw Professional 20.1 (PerkinElmer Informatics) and importing them directly into the mdb file. Additionally, protein models with or without Mg^{2+} ions for docking were prepared according to MOEs standard "Quick Prep" protocol, firstly protonating the structure (using the Protonate 3D tool) and secondly energy minimizing it based on steepest descent (using AMBER10/EHT forcefield and allowing RMS gradient of 0.1 kcal/mol/Å^2). Lastly, the binding site for docking was identified using the SiteFinder tool using/with default parameters.

After the preparation, the docking was performed using the "general" docking tool. For each variant and wildtype, the whole array of substrates and intermediates in the database was docked. "Triangle Matcher" was used for placement, and "Induced Fit" method was used for refinement of ligand poses, (allowing flexible sidechains during docking). At least 300 docking poses were generated for each structure and scored based on London dG + GBVI/WSA dG values. Then, five best-scored poses were further analyzed for the selection of final docking pose. For docking of substrates, the BcBOT2 model with trinuclear Mg^{2+} ions was used, during which the ions where treated as a flexible part of the receptor, allowing them to move during ligand placement and refinement. For docking the reaction intermediates the BcBOT2 model without any Mg^{2+} ions was used.

Final docking poses for each intermediate in the respected variants were selected based on proposed catalytic mechanism for FPP in *Bc*BOT2 and general STCs.[11] For FPP (**1**) and the nerolidyl diphosphate (**12**) logical positioning of the diphosphate group into the Mg^{2+} coordination shell (compromised by the DDxxD –, NSE/DTE – motif and the effector triad), and sufficient orientation of the carbon moiety into the hydrophobic part of active pocket was considered. For all other structures (**6-11**, **13-14**), reasonable orientation in the hydrophobic part of the active pocket and, if needed, sufficient distance between the carbons used for cyclization was contemplated. Additionally, the scoring of each pose for the respected structures were examined. Binding Affinity and solvation for the final poses were calculated automatically after the steepest descent energy minimization of the respected enzyme-substrate/intermediate complexes using the "Ligand Properties" tool. Most important docking results are shown in Figure S18-S19.

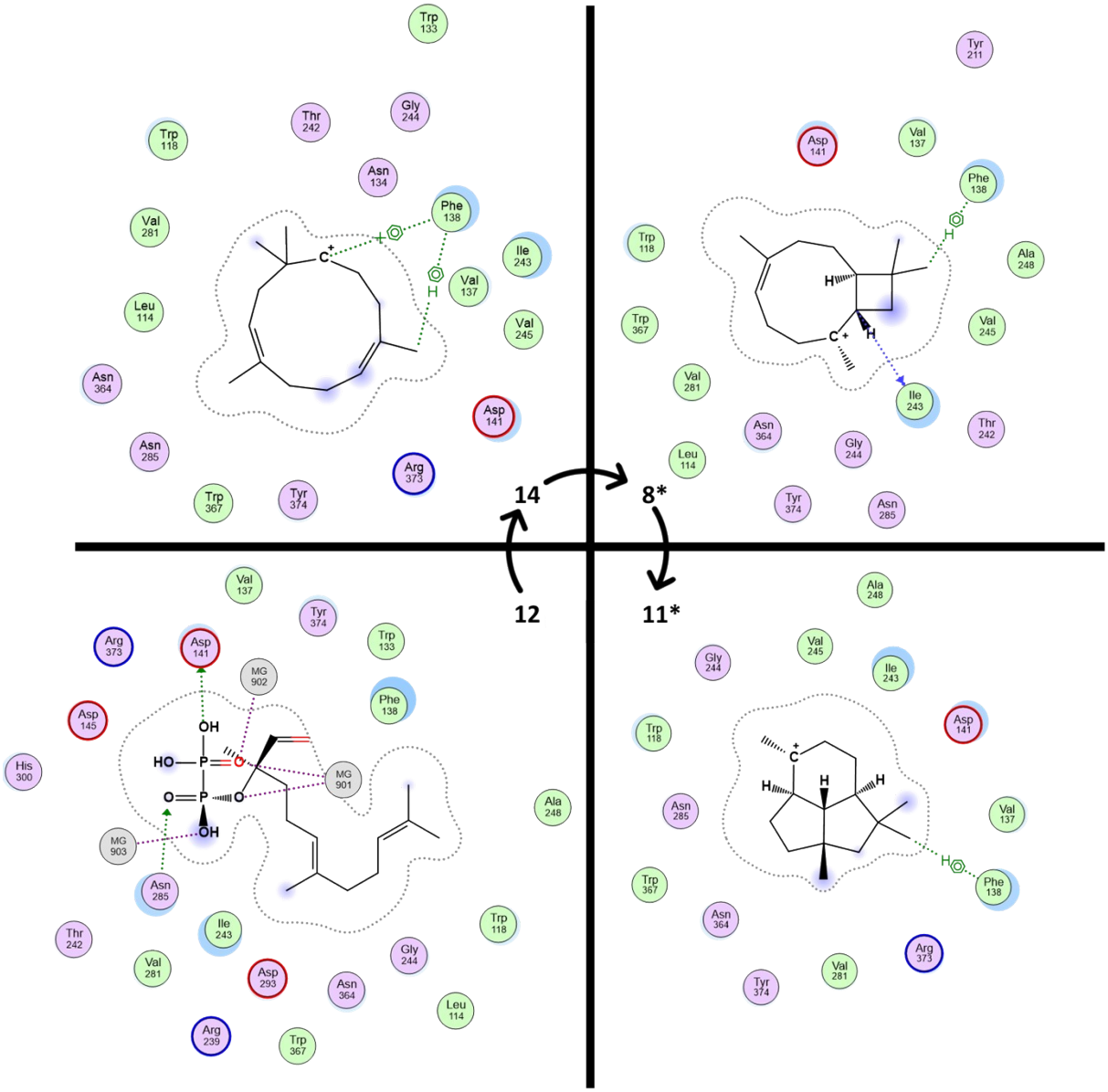


Figure S18: Docking of Intermediates of Tantillo route in wild-type BcBOT2 structure.[34] The bottom left shows the constitutional isomer of FPP, which is the starting point for the Tantillo route. The top left shows the intermediate **14**, stabilized by a cation-arene interaction and a hydrogen-aren interaction from F138. The top right shows the intermediate **8** with a hydrogen-arene interaction from F138 and I243 interacts as a backbone donor. The bottom right shows the last intermediate **11** that is stabilized by a hydrogen-arene interaction.

*Deception of configuration and stereochemistry in the ligand interaction representations for the intermediates **8** and **11** can differ from the actual arrangement. The figures are produced directly in MOE by using a tool that takes a fitting perspective from the 3D view of the docked structures and flattens it to an uncluttered 2D representation. During the flattening the structures can be read out in altered configurations and with missing or wrongly added stereo centers. Thus, the representation of **8** contains a wrongly configured double bond (*Z* instead of *E*) and in **8** and **11** a stereo center was added at the carbocation.

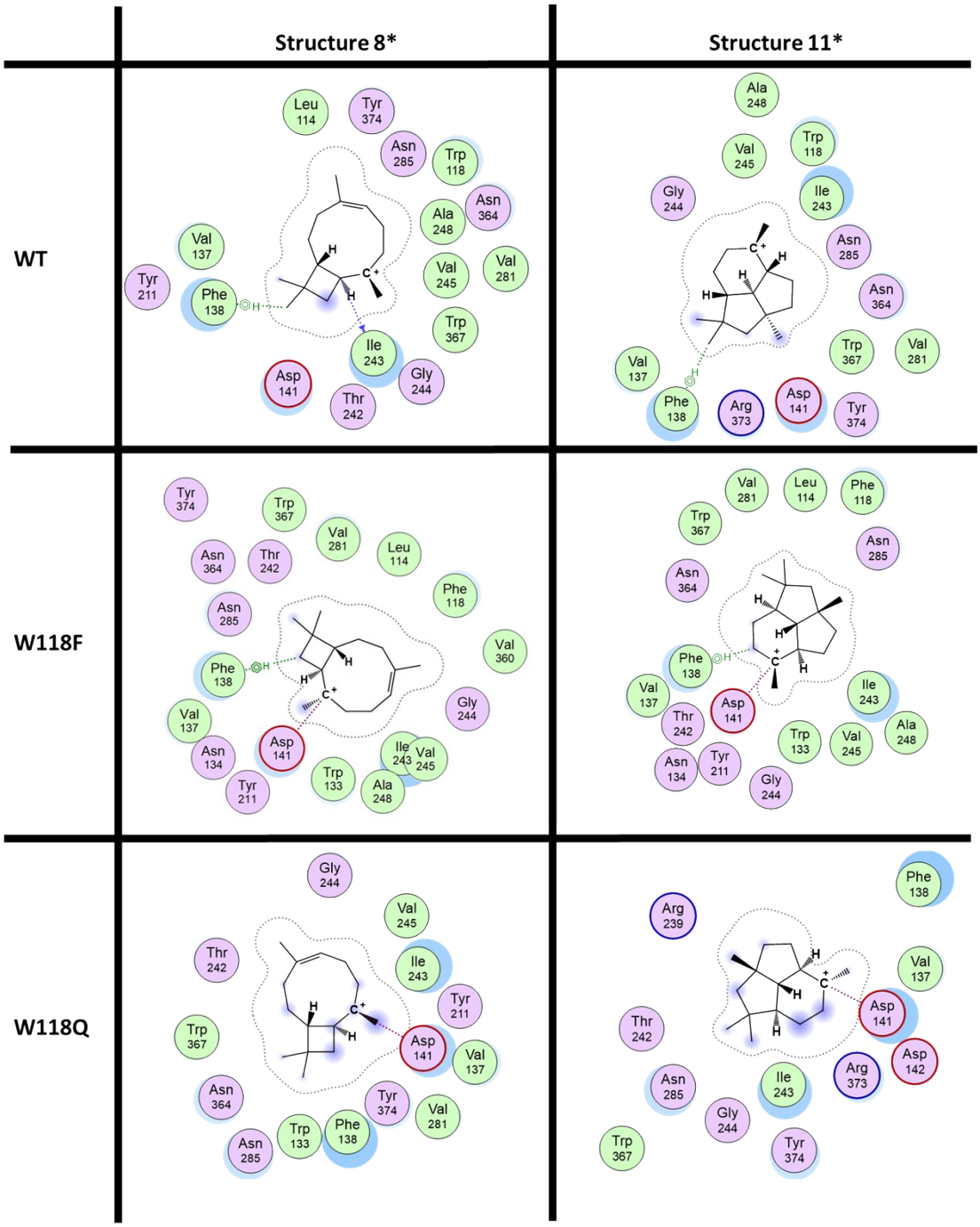


Figure S19: Docking of intermediate 8 and 11 into wild-type BcBOT2 and the variants W118F and W118Q. The wild-type BcBOT2 shows no ionic interaction of D141, with neither of the two intermediates. It shows hydrogen- arene interaction of F138 for both intermediates and a backbone binding from I243 for intermediate **8**. The intermediates in W118F variant are stabilized by the aromatic stabilization of F138 and by the ionic interactions of D141. Lastly, the intermediates in W118Q only show ionic interaction with D141.

*Deception of configuration and stereochemistry in the ligand interaction representations for the intermediates **8** and **11** can differ from the actual arrangement. The figures are produced directly in MOE by using a tool that takes a fitting perspective from the 3D view of the docked structures and flattens it to an uncluttered 2D representation. During the flattening the structures can be read out in altered configurations and with missing or wrongly added stereo centers. Thus, the representation of **8** contains a wrongly configured double bond (*Z* instead of *E*) and in **8** and **11** a stereo center was added at the carbocation.

3.7 Generation and analysis of boat-, chair - and twist boat-like conformations of intermediate 11

Boat-, chair- and twist boat-like conformations of intermediate **11** were build, geometry optimized and analyzed using Chem3D software produced by PerkinElmer Informatics [36]. Boat- and chair-like conformation of the cyclohexane ring of **11** were generated manually using the "builder" tool of Chem3D. Geometry optimization and energy calculation of the conformations was carried out using the "energy minimization" tool with the MM2 forcefield[37]. To not alter/lose the designed conformations of **11**, the position of the carbon atoms of the cyclohexane ring were fixed during the geometry optimization step. Additionally, the chair- and boat-like conformations of **11** were energy minimized again using the MM2, but without fixing any atoms in the structures, to generate the conformation. The respected structures after minimization were fairly similar in conformation and energy and ultimately lead to the generation of the twist boat-like conformation for **11**, representing the energy minimal structure.

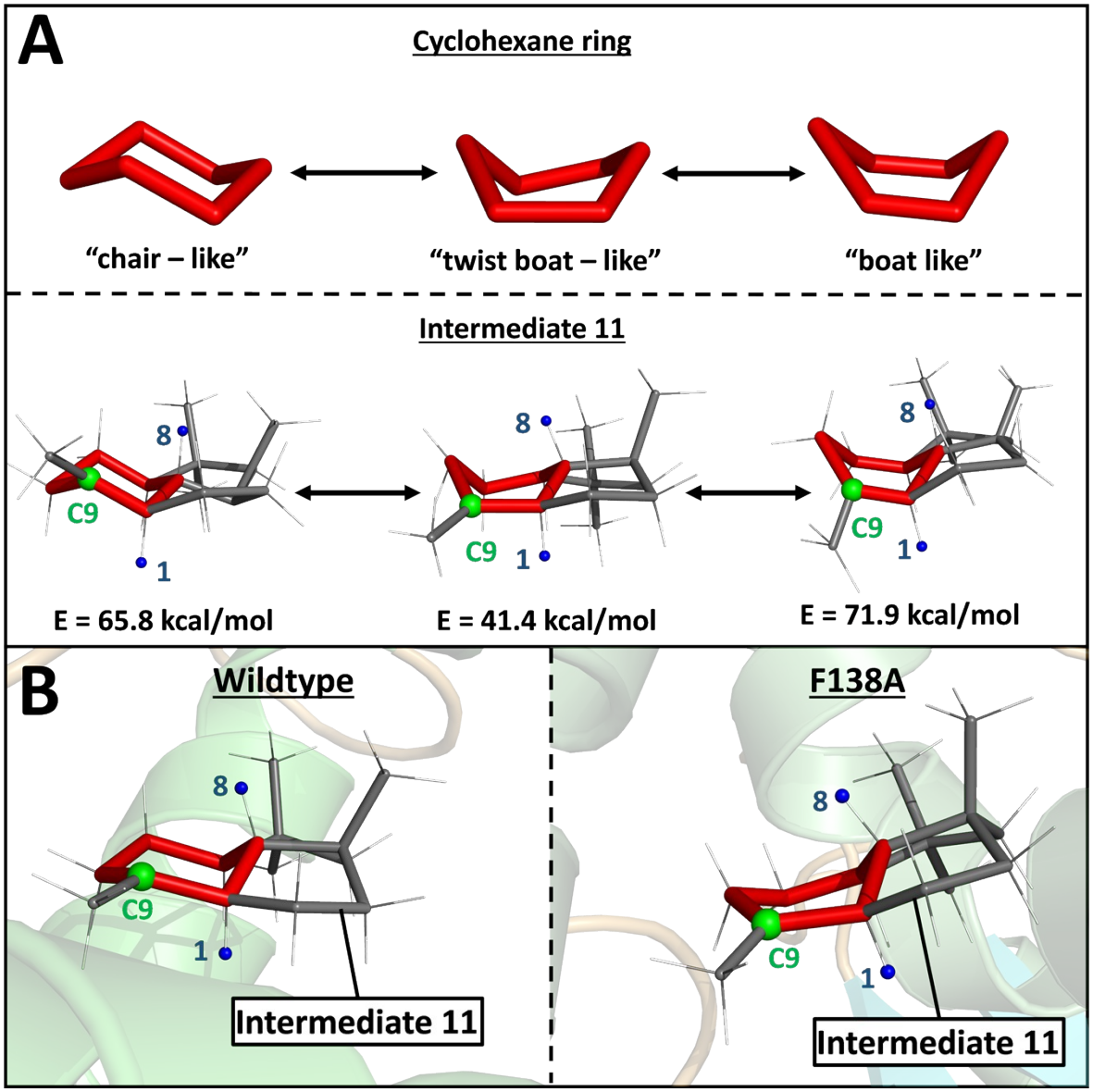


Figure S20: Investigation of chair- and boat-like conformations for Intermediate 11; A: Chair-, boat- and twist boat-like conformation of intermediate **11** shown in grey stick representation. The chair- and boat-like conformations were build, geometry optimized and analyzed using the Chem3D software [36]. The cyclohexane ring of the conformations is highlighted as red sticks. The C9 atoms and H atoms at positions 1 and 8 are shown in green and blue spheres, respectively. The E value stands for the potential energy of the conformations calculated after the energy minimization **B:** Comparison the conformations of docked intermediate **11** for BcBOT2 wild-type (left) and the variant F138A (right). The structures of **11** are shown in grey stick representations. The cyclohexane ring of the conformations is highlighted as red sticks. The C9 atoms and H atoms at positions 1 and 8 are shown in green and blue spheres, respectively.

4 References (Supporting information)

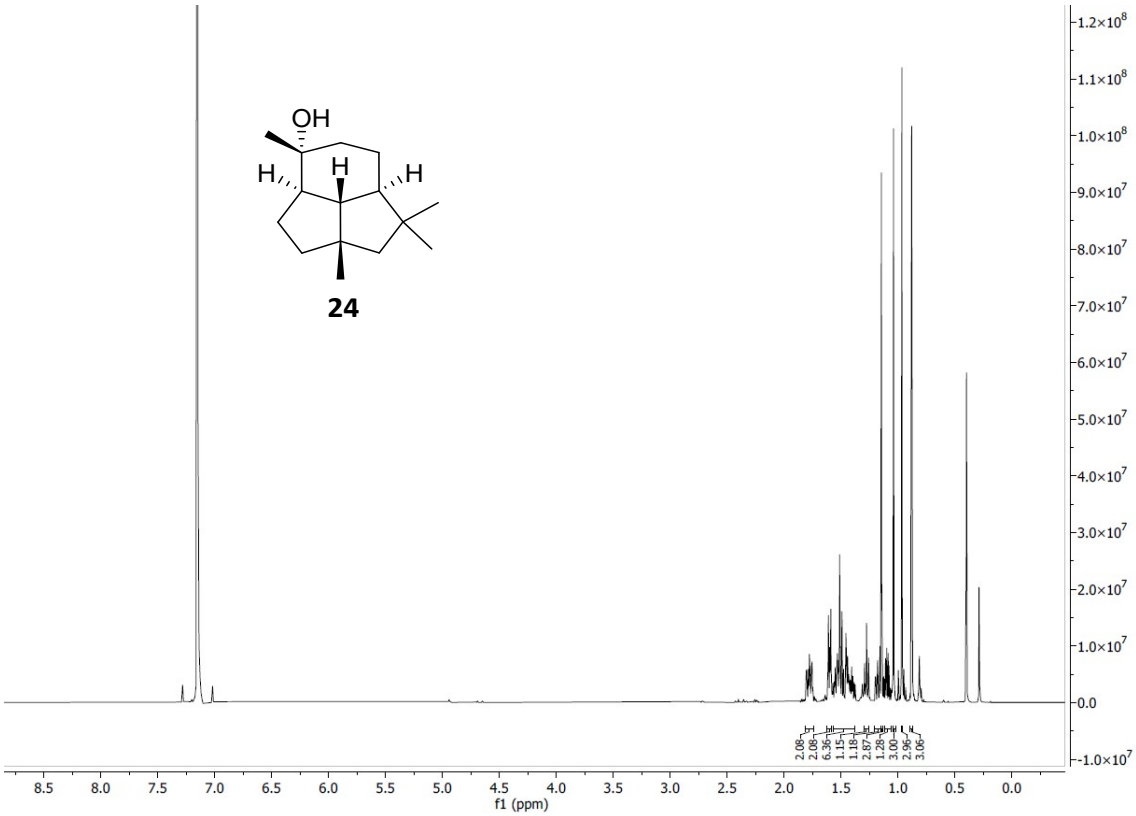
1. L. J., L. B. and P. G., *Praxis in der oranischen Chemie*, VCH, Weinheim, 1996.
2. P. B. Shrestha-Dawadi and J. Lugtenburg, *Eur. J. Org. Chem.*, 2003, **2003**, 4654-4663.
3. V. J. Davisson, A. B. Woodside, T. R. Neal, K. E. Stremler, M. Muehlbacher and C. D. Poulter, *J. Org. Chem.*, 1986, **51**, 4768-4779.
4. W. A. B., Z. Huang and C. D. Poulter, *Org. Synth.*, 1988, **66**.
5. V. Harms, B. Schröder, C. Oberhauser, C. D. Tran, S. Winkler, G. Dräger and A. Kirschning, *Organic Letters*, 2020, **22**, 4360-4365.
6. C. Oberhauser, V. Harms, K. Seidel, B. Schröder, K. Ekramzadeh, S. Beutel, S. Winkler, L. Lauterbach, J. S. Dickschat and A. Kirschning, *Angewandte Chemie International Edition*, 2018, **57**, 11802-11806.
7. W. C. M., H. R., L. X. and C. D. E., *J. Am. Chem. Soc.* , 2003, 8360-8361.
8. R. Lauchli, K. S. Rabe, K. Z. Kalbarczyk, A. Tata, T. Heel, R. Z. Kitto and F. H. Arnold, *Angew. Chem. Int. Ed.*, 2013, **52**.
9. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek and A. Potapenko, *Nature*, 2021, **596**, 583-589.
10. C. Pinedo, C.-M. Wang, J.-M. Pradier, B. Dalmais, M. Choquer, P. Le Pêcheur, G. Morgant, I. G. Collado, D. E. Cane and M. Viaud, *ACS Chem. Biol.*, 2008, **3**, 791-801.
11. D. W. Christianson, *Chem. Rev.*, 2017, **117**, 11570-11648.
12. C. Contreras-Martel, C. Dahout-Gonzalez, A. D. S. Martins, M. Kotnik and A. Dessen, *J. Mol. Biol.*, 2009, **387**, 899-909.
13. P. Benkert, M. Künzli and T. Schwede, *Nucleic Acids Res.*, 2009, **37**, W510-W514.
14. F. Melo, D. Devos, E. Depiereux and E. Feytmans, 1997.
15. I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall III, J. Snoeyink and J. S. Richardson, *Nucleic Acids Res.*, 2007, **35**, W375-W383.
16. G. Pollastri and A. McLysaght, *Bioinformatics*, 2005, **21**, 1719-1720.
17. L. J. McGuffin, K. Bryson and D. T. Jones, *Bioinformatics*, 2000, **16**, 404-405.
18. R. Yan, D. Xu, J. Yang, S. Walker and Y. Zhang, *Sci. Rep.*, 2013, **3**, 1-9.
19. A. Drozdetskiy, C. Cole, J. Procter and G. J. Barton, *Nucleic Acids Res.*, 2015, **43**, W389-W394.
20. B. Petersen, T. N. Petersen, P. Andersen, M. Nielsen and C. Lundegaard, *BMC Struct. Biol.*, 2009, **9**, 1-10.
21. H. J. Koo, C. R. Vickery, Y. Xu, G. V. Louie, P. E. O'maille, M. Bowman, C. M. Nartey, M. D. Burkart and J. P. Noel, *The Journal of Antibiotics*, 2016, **69**, 524-533.
22. The PyMOL Molecular Graphics System, Version 2.0, Schrodinger, LLC.
23. A. M. Waterhouse, J. B. Procter, D. M. Martin, M. Clamp and G. J. Barton, *Bioinformatics*, 2009, **25**, 1189-1191.
24. O. Lichtarge, H. R. Bourne and F. E. Cohen, *J. Mol. Biol.*, 1996, **257**, 342-358.
25. I. Mihalek, I. Reš and O. Lichtarge, *J. Mol. Biol.*, 2004, **336**, 1265-1282.
26. H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko and N. Ben-Tal, *Nucleic Acids Res.*, 2016, **44**, W344-W350.
27. G. Celniker, G. Nimrod, H. Ashkenazy, F. Glaser, E. Martz, I. Mayrose, T. Pupko and N. Ben-Tal, *Isr. J. Chem.*, 2013, **53**, 199-206.
28. H. Ashkenazy, E. Erez, E. Martz, T. Pupko and N. Ben-Tal, *Nucleic Acids Res.*, 2010, **38**, W529-W533.
29. C. Berezin, F. Glaser, J. Rosenberg, I. Paz, T. Pupko, P. Fariselli, R. Casadio and N. Ben-Tal, *Bioinformatics*, 2004, **20**, 1322-1324.
30. V. Le Guilloux, P. Schmidtke and P. Tuffery, *BMC Bioinform.*, 2009, **10**, 1-11.
31. E. Krieger, G. Koraimann and G. Vriend, *Proteins*, 2002, **47**, 393-402.
32. R. Guerois, J. E. Nielsen and L. Serrano, *J. Mol. Biol.*, 2002, **320**, 369-387.
33. S. Miyamoto and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 952-962.
34. D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, M. Crowley, R. C. Walker and W. Zhang, Amber 10, University of California, 2008.

35. Molecular Operating Environment (MOE), 2022.02 Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2022.
36. Chem3D software, *PerkinElmer Informatics*.
37. N. L. Allinger, *J. Am. Chem. Soc.*, 1977, **99**, 8127-8134.

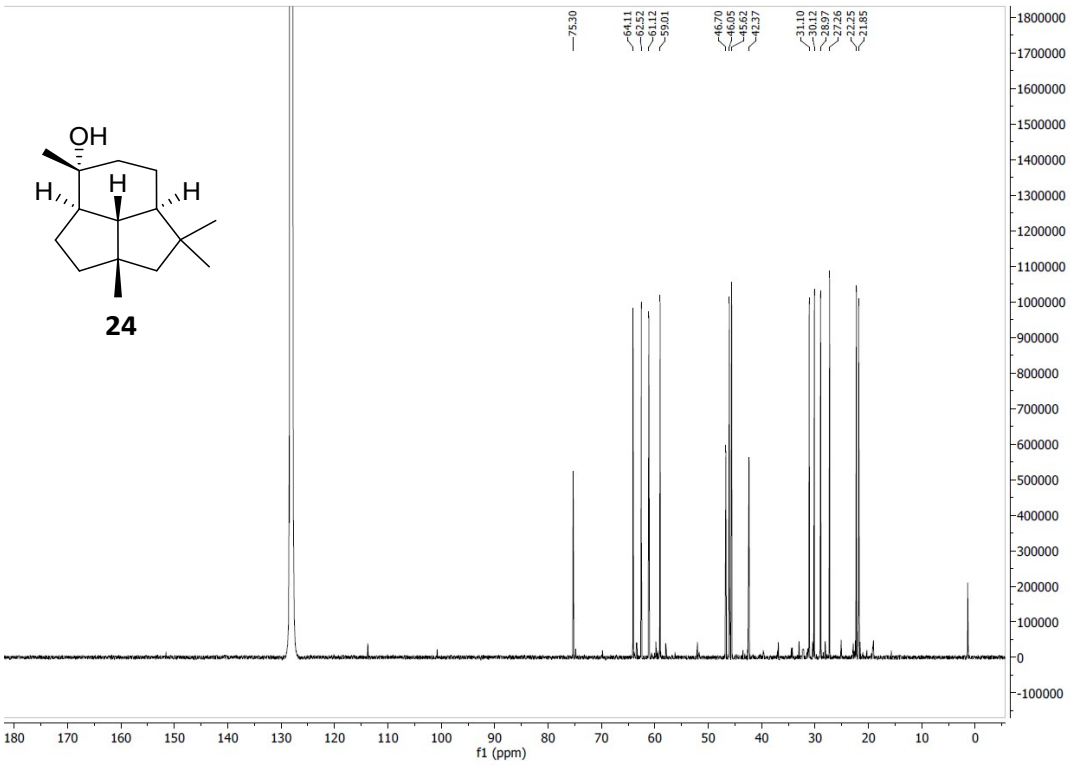
Attachments (A. NMR spectra, B. GC chromatograms, C. MS spectra)

A. NMR spectra

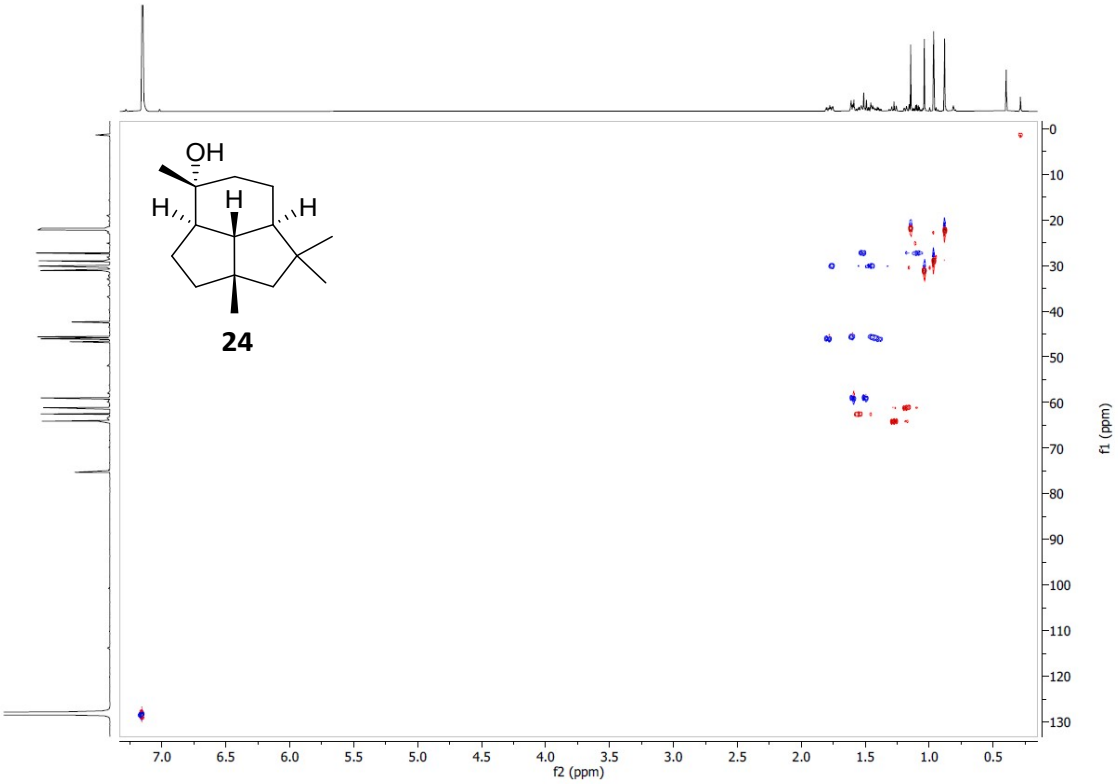
^{1}H-NMR (600 MHz, C_6D_6 δ= 7.16 ppm, **24**)



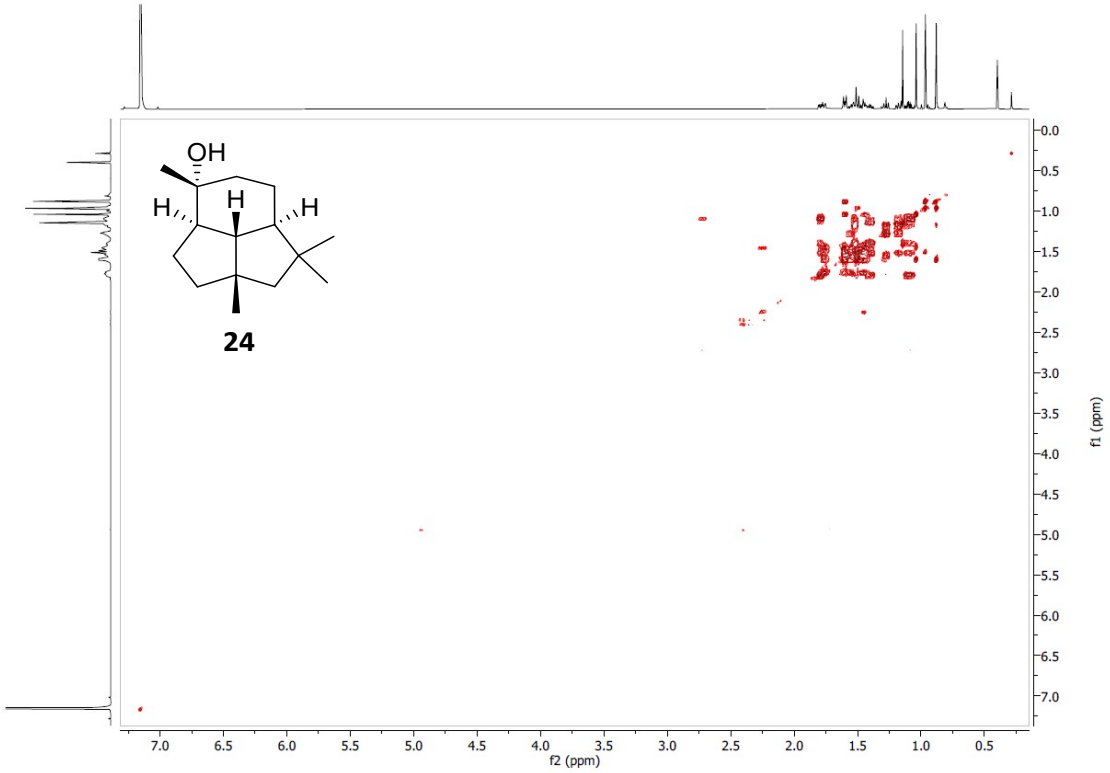
^{13}C-NMR (151 MHz, C_6D_6 δ= 128.06 ppm, **24**)



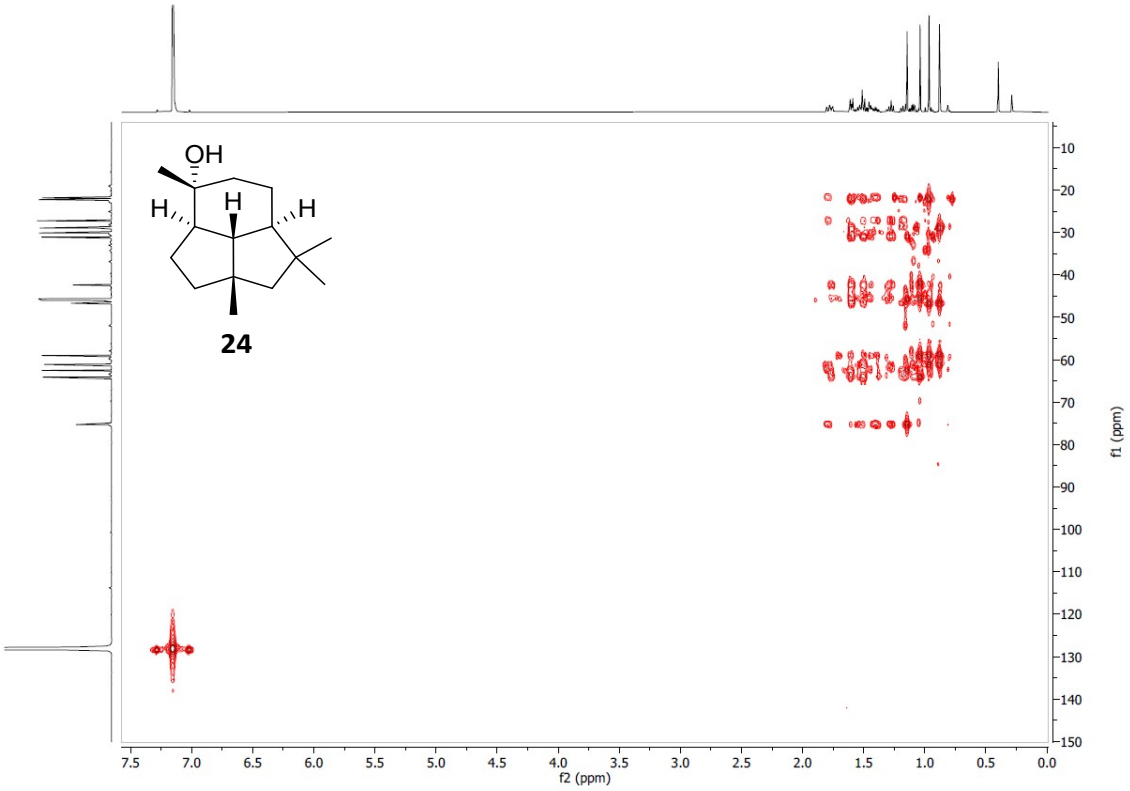
HSQC (**24**)



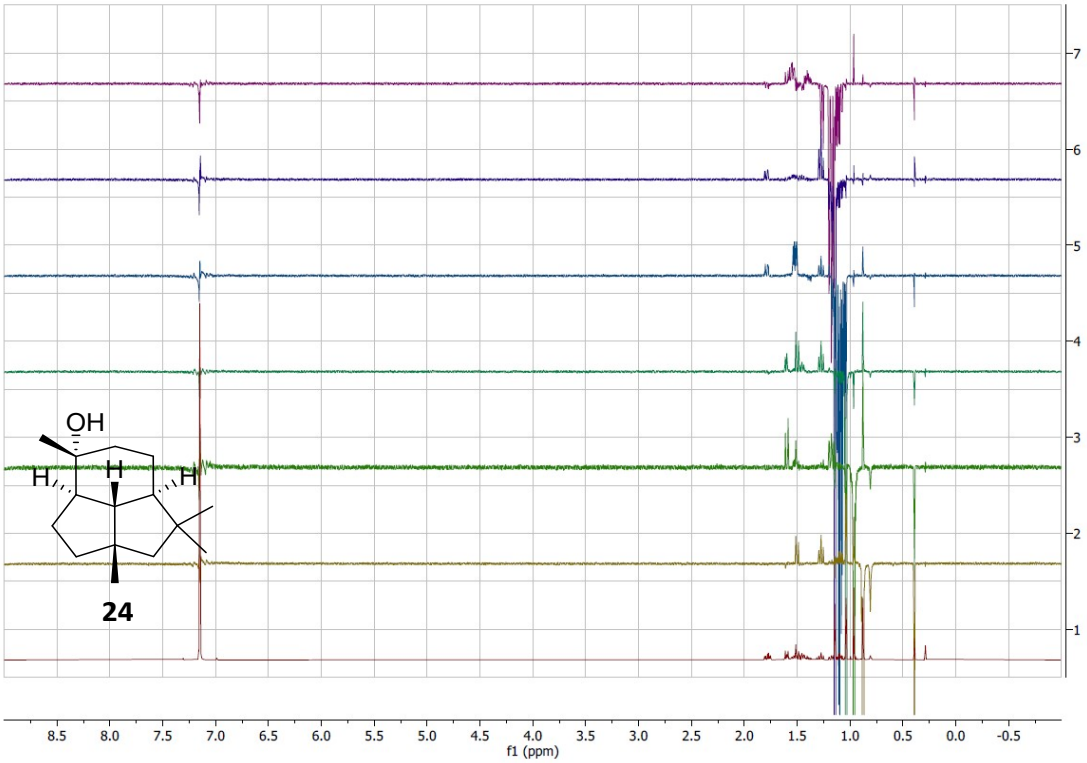
COSY (**24**)



HMBC (**24**)

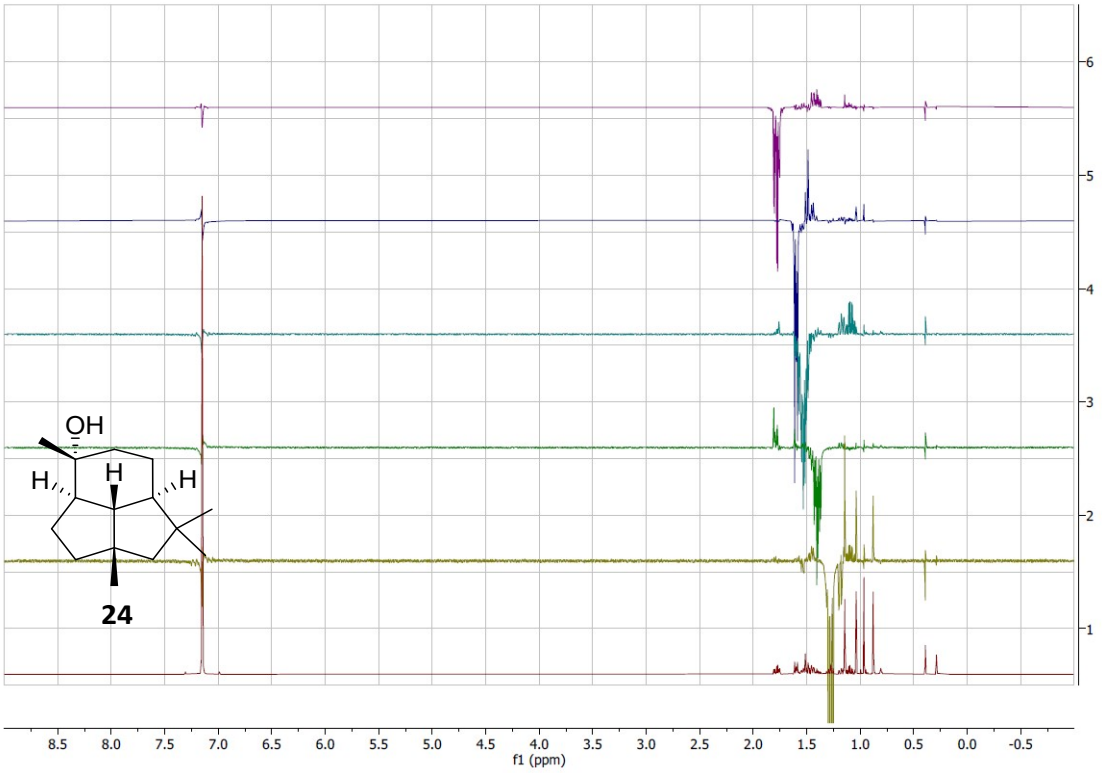


1d-NOE (**24**)



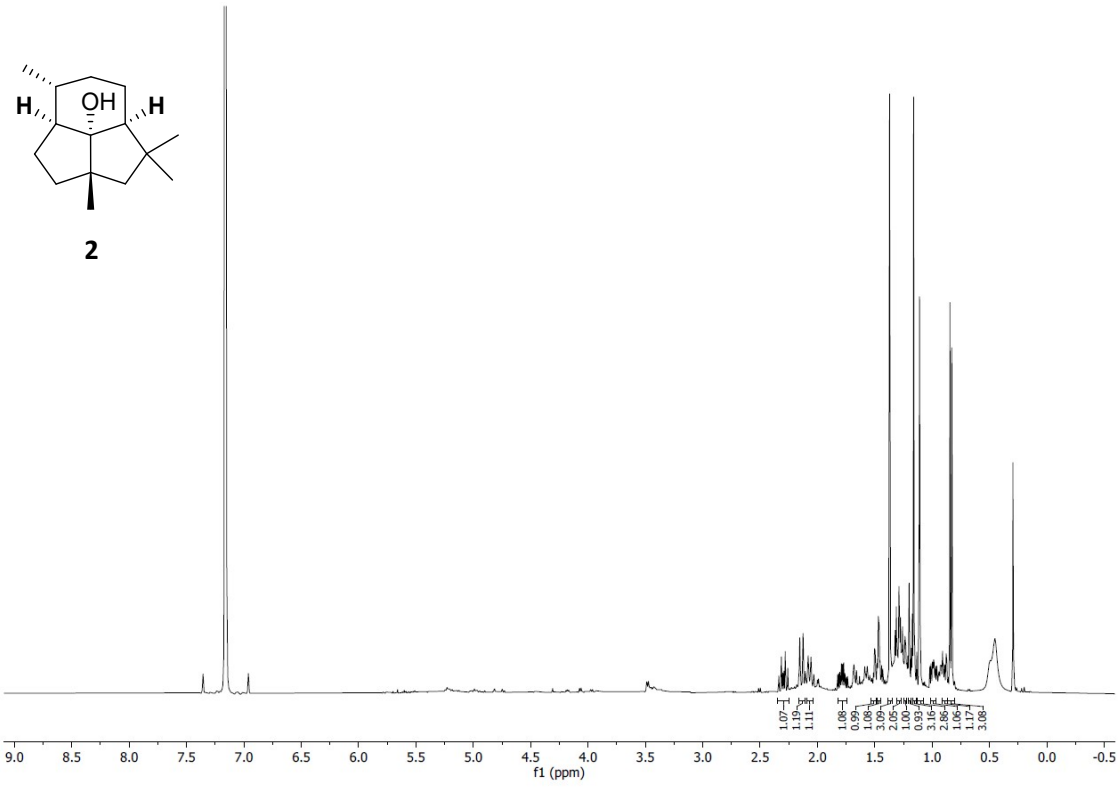
From bottom to top: ^{1}H-NMR, H-13, H-12, H-14, H-11a, H-15, H-7.

1d-NOE (**24**)

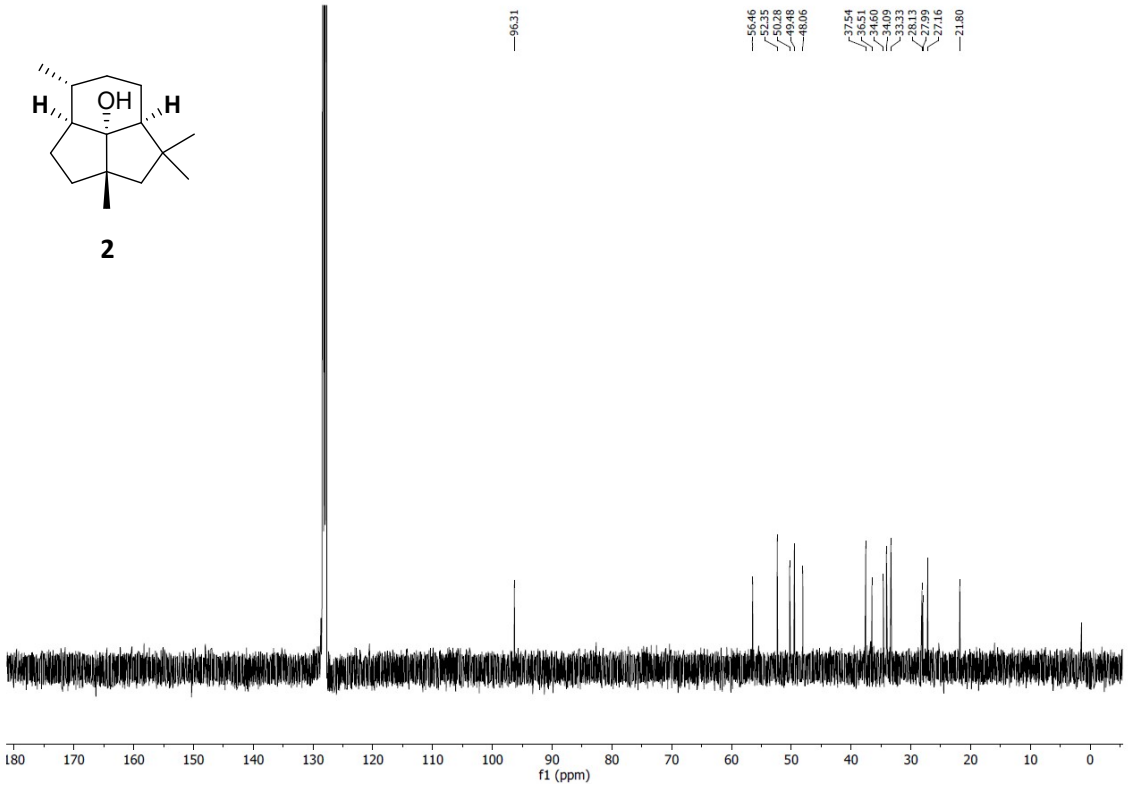


From bottom to top: ^{1}H-NMR, H-8, H-10a, H-11b, H-3b, H-10b.

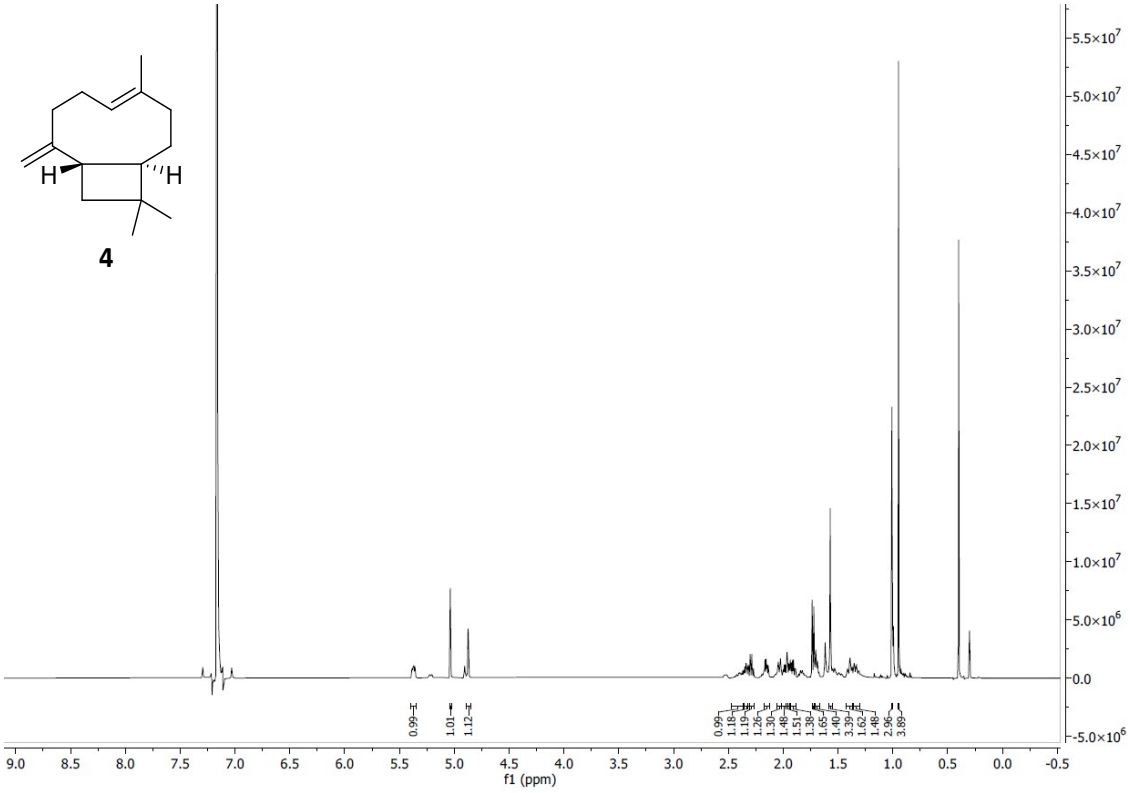
^{1}H-NMR (600 MHz, C_6D_6 δ= 7.16 ppm, **2**)



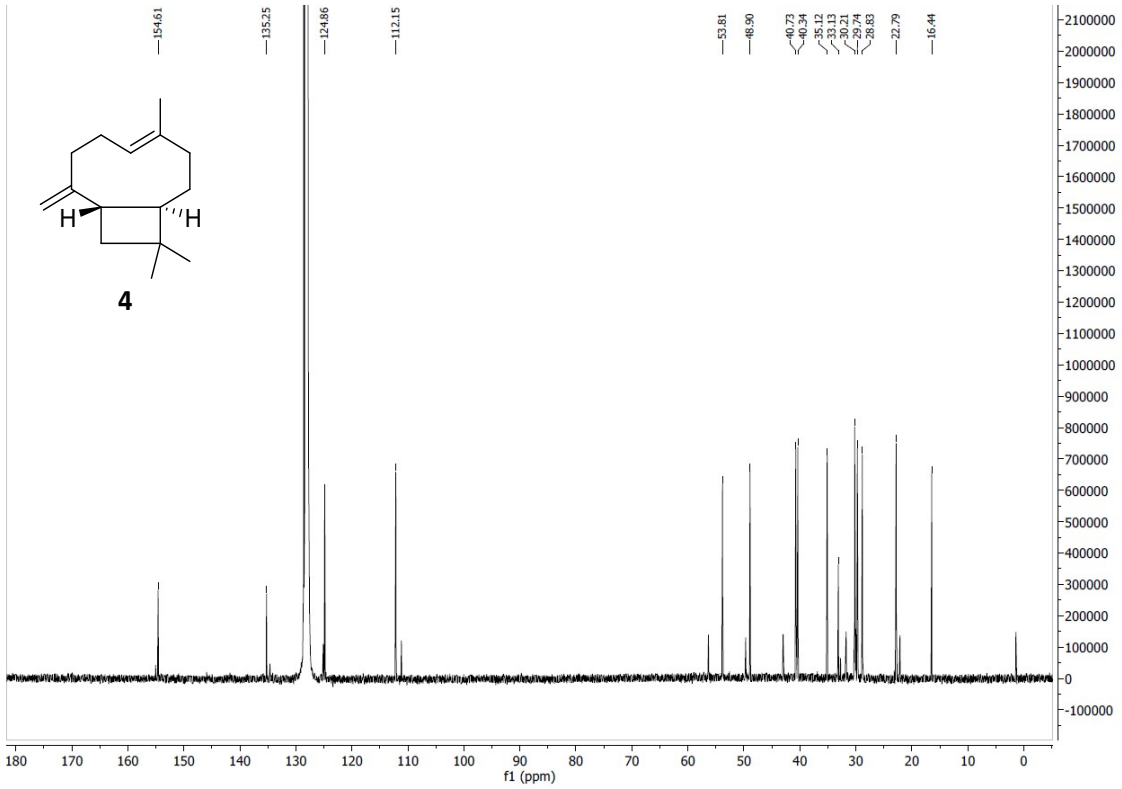
^{13}C-NMR (151 MHz, C_6D_6 δ= 128.06 ppm, **2**)



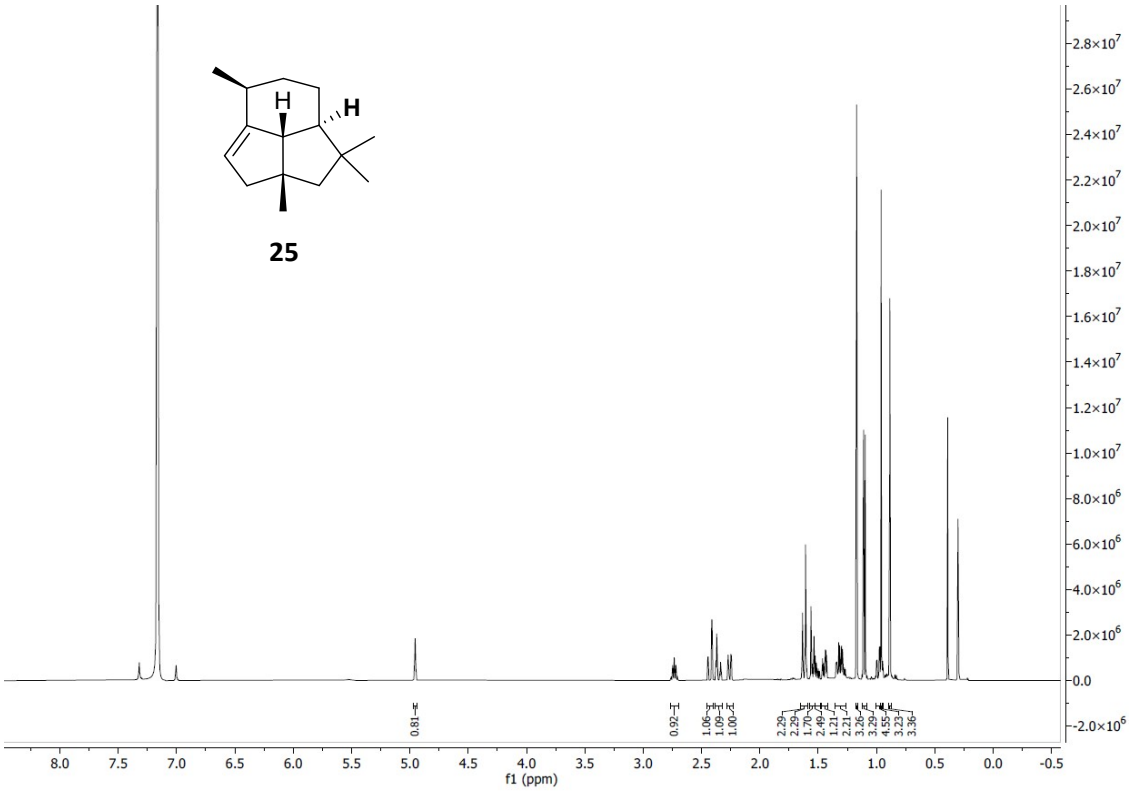
^{1}H-NMR (600 MHz, C_6D_6 δ= 7.16 ppm, **4**)



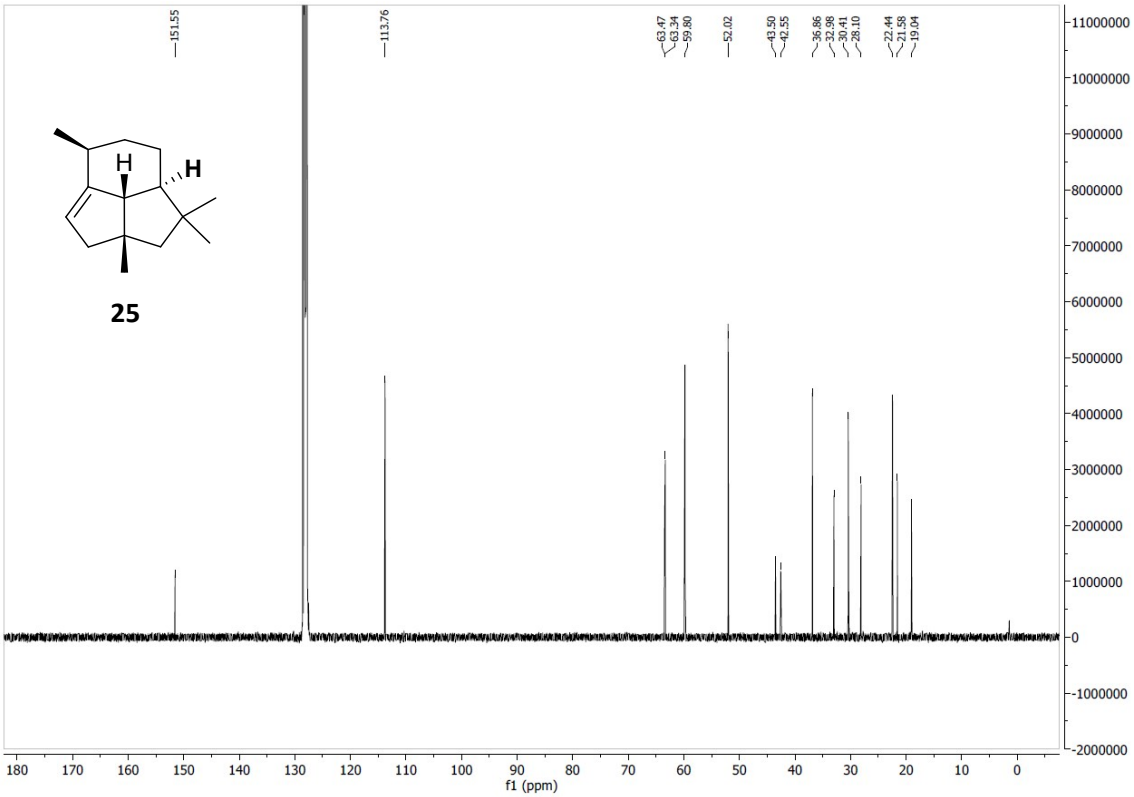
^{13}C-NMR (151 MHz, C_6D_6 δ= 128.06 ppm, **4**)



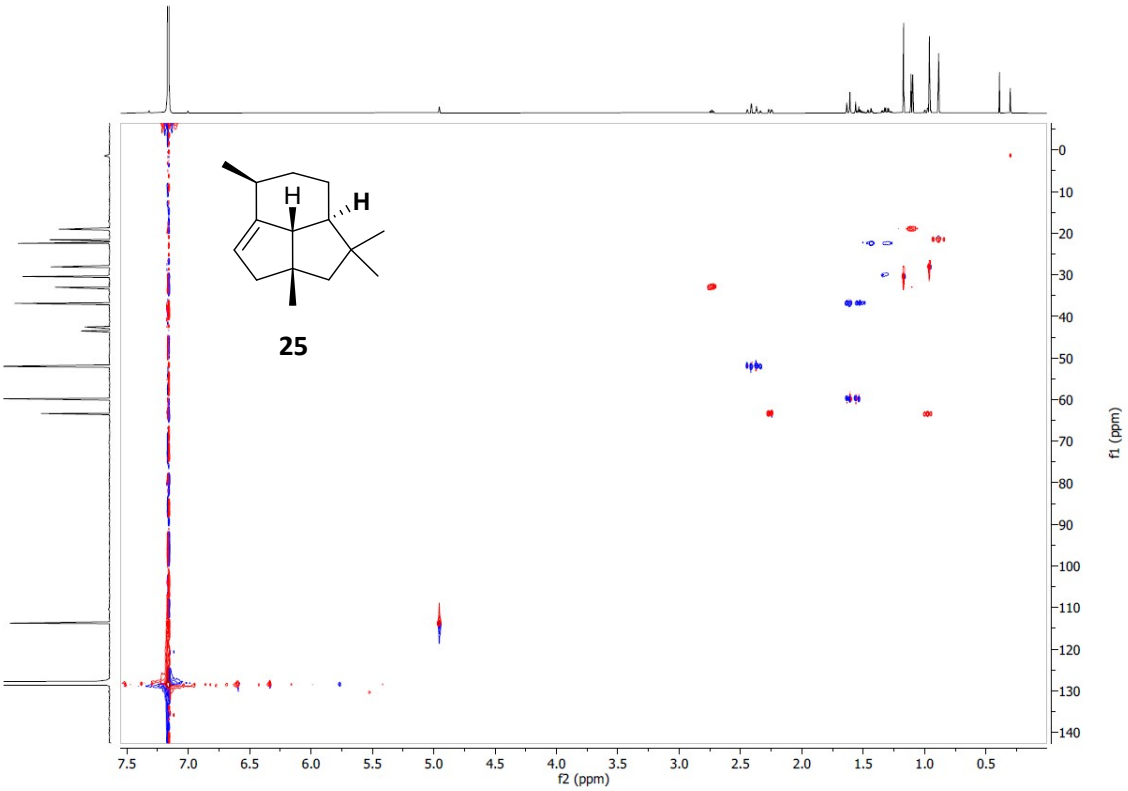
^{1}H-NMR (600 MHz, C_6D_6 δ= 7.16 ppm, **25**)



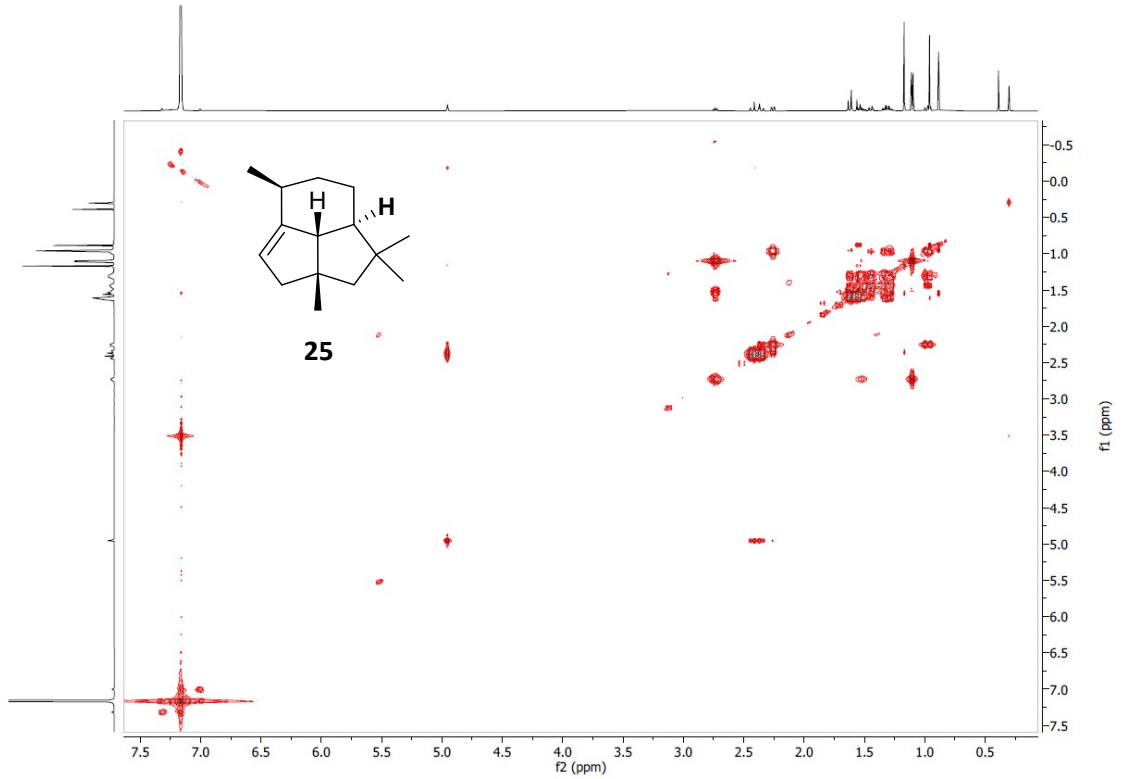
^{13}C-NMR (151 MHz, C_6D_6 δ= 128.06 ppm, **25**)



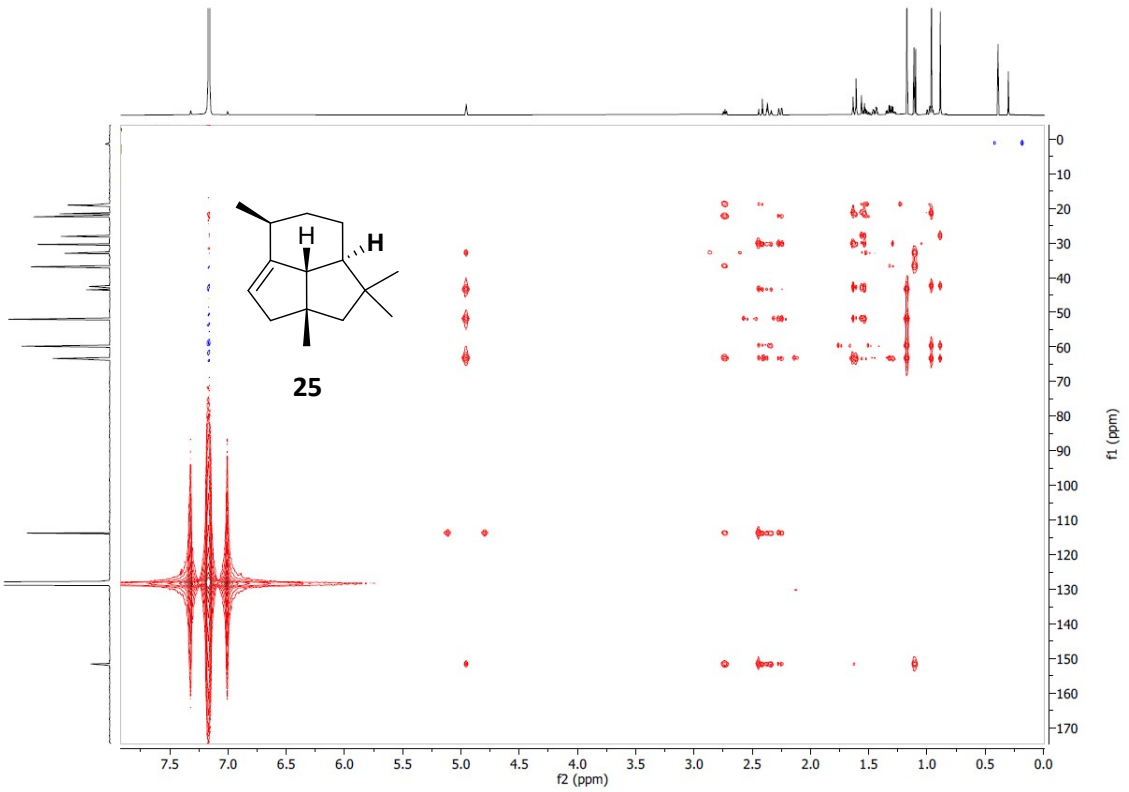
HSQC (**25**)



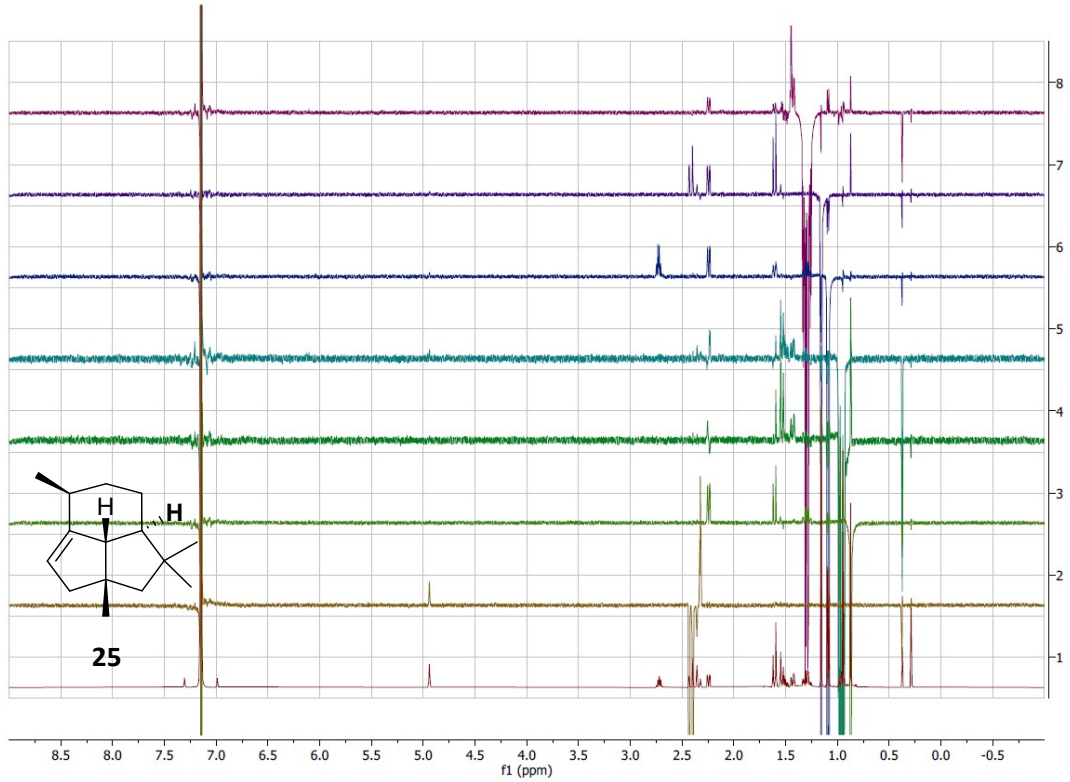
COSY (**25**)



HMBC (**25**)

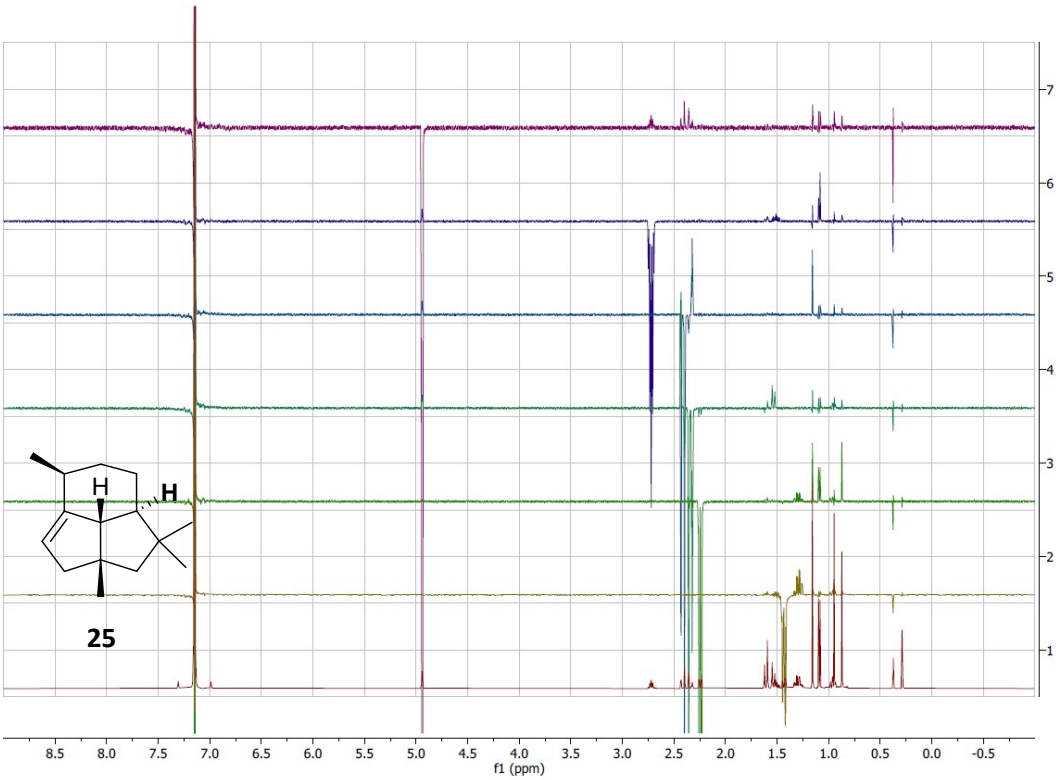


1d-NOE (**25**)



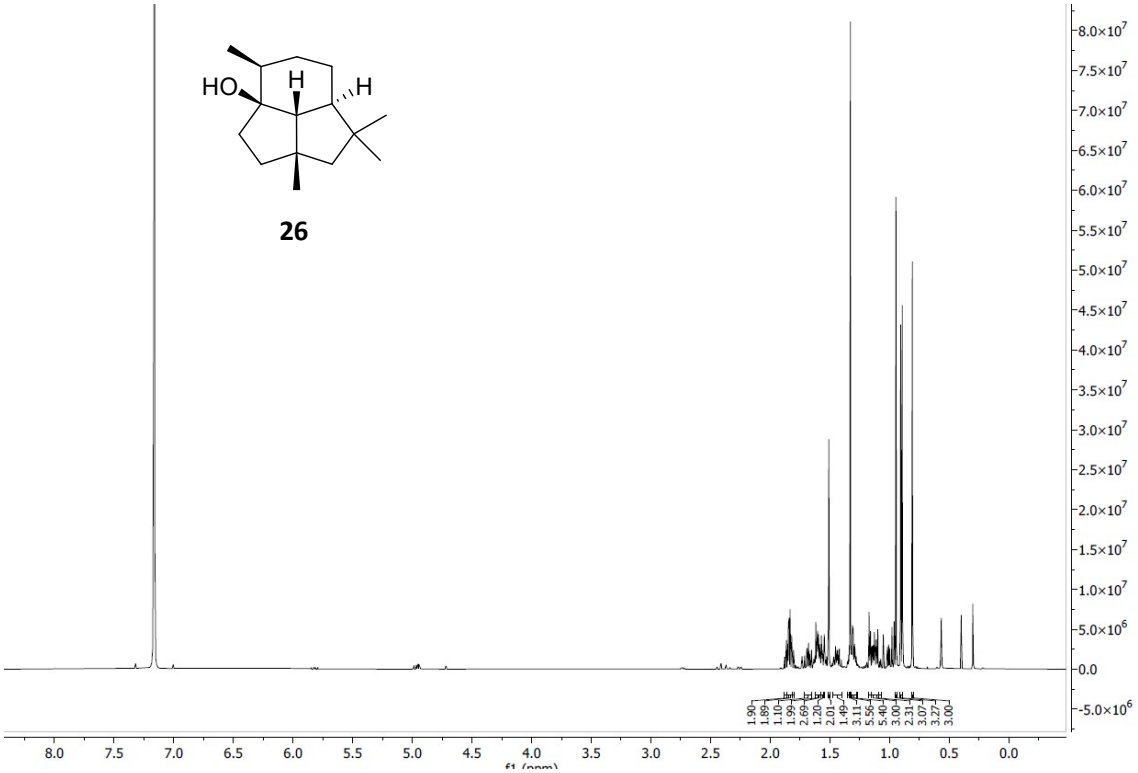
From bottom to top: ^{1}H-NMR, H-3b, H-13, H-12, H-7, H-15, H-14, H-11a.

1d-NOE (**25**)

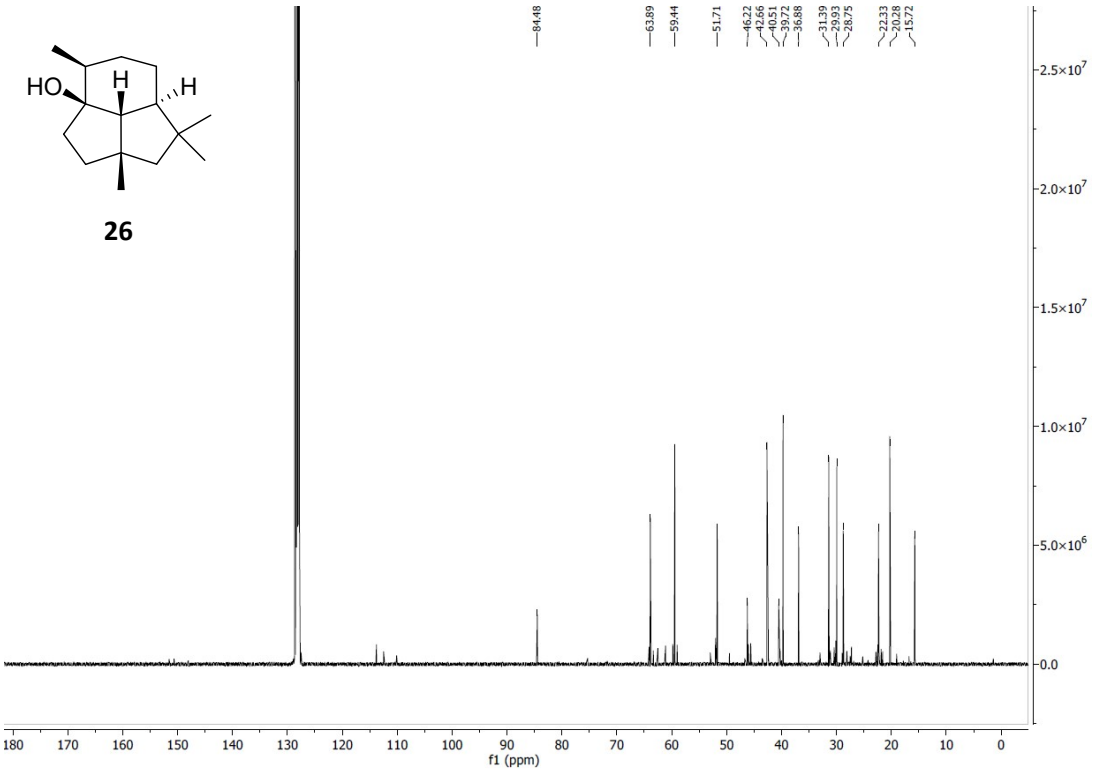


From bottom to top: ^{1}H-NMR, H-11b, H-8, H-3a, H-3b, H-9, H-2.

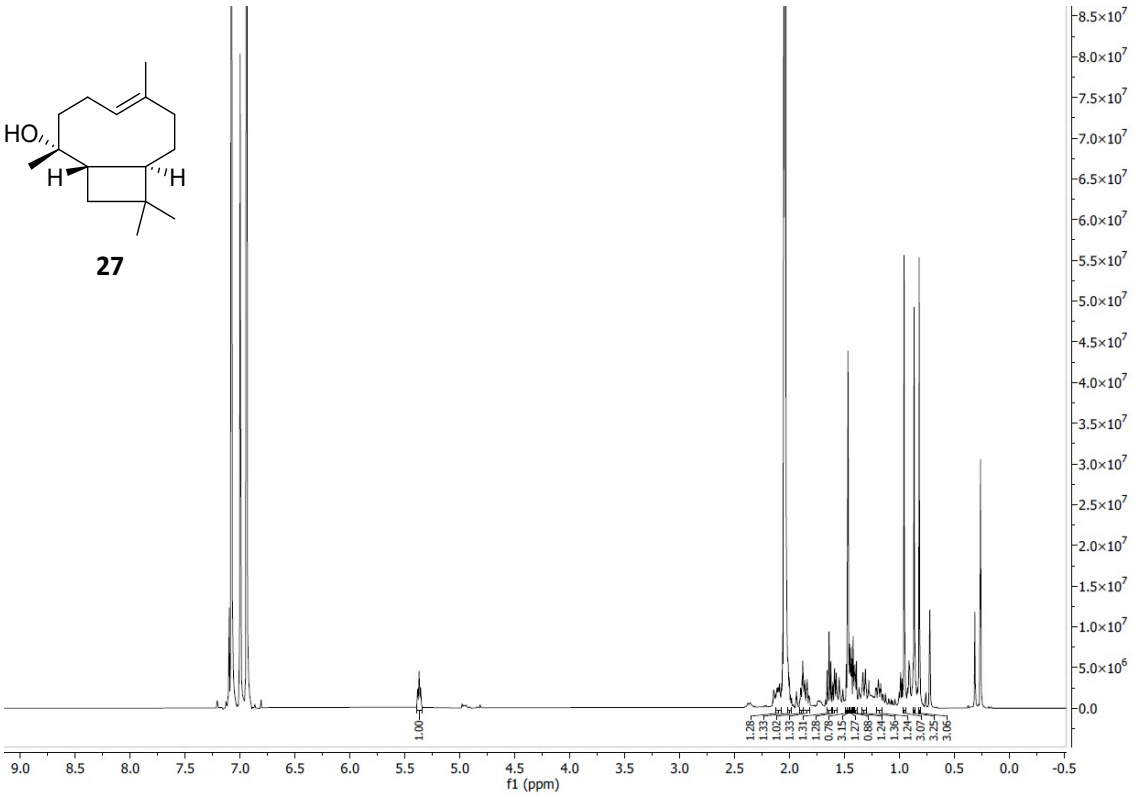
^{1}H-NMR (600 MHz, C_6D_6 δ= 7.16 ppm, **26**)



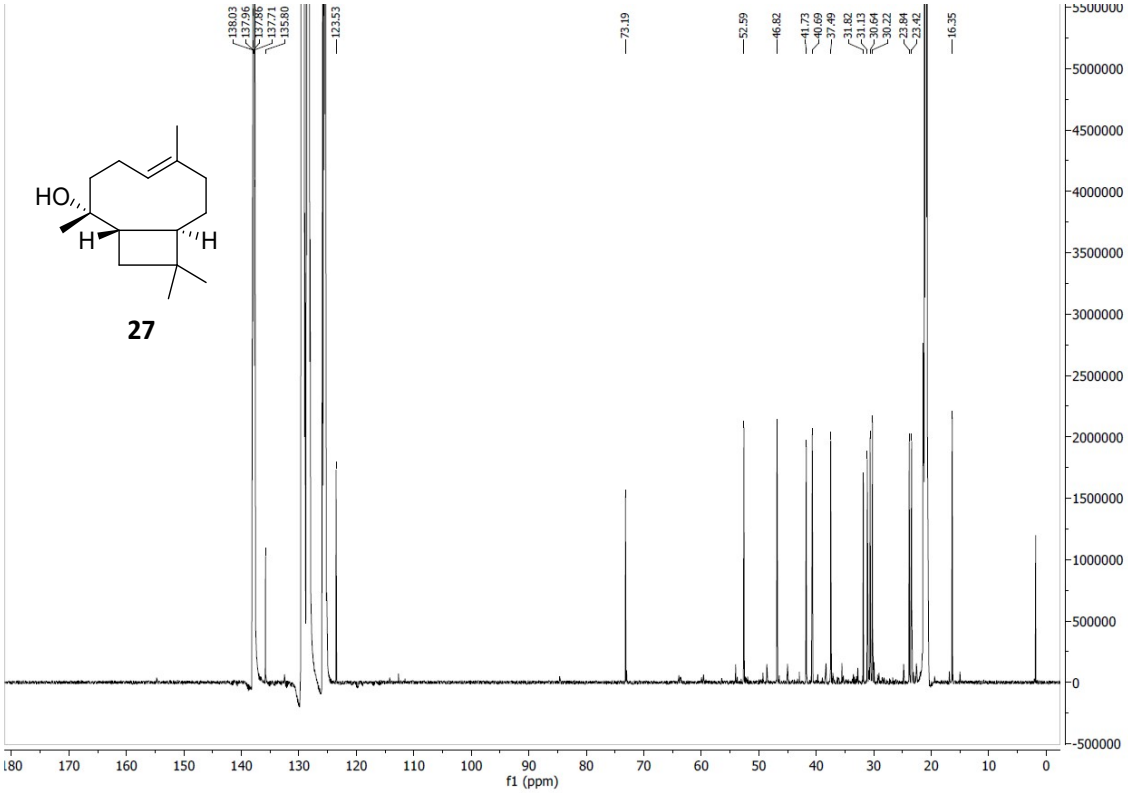
^{13}C-NMR (151 MHz, C_6D_6 δ= 128.06 ppm, **26**)



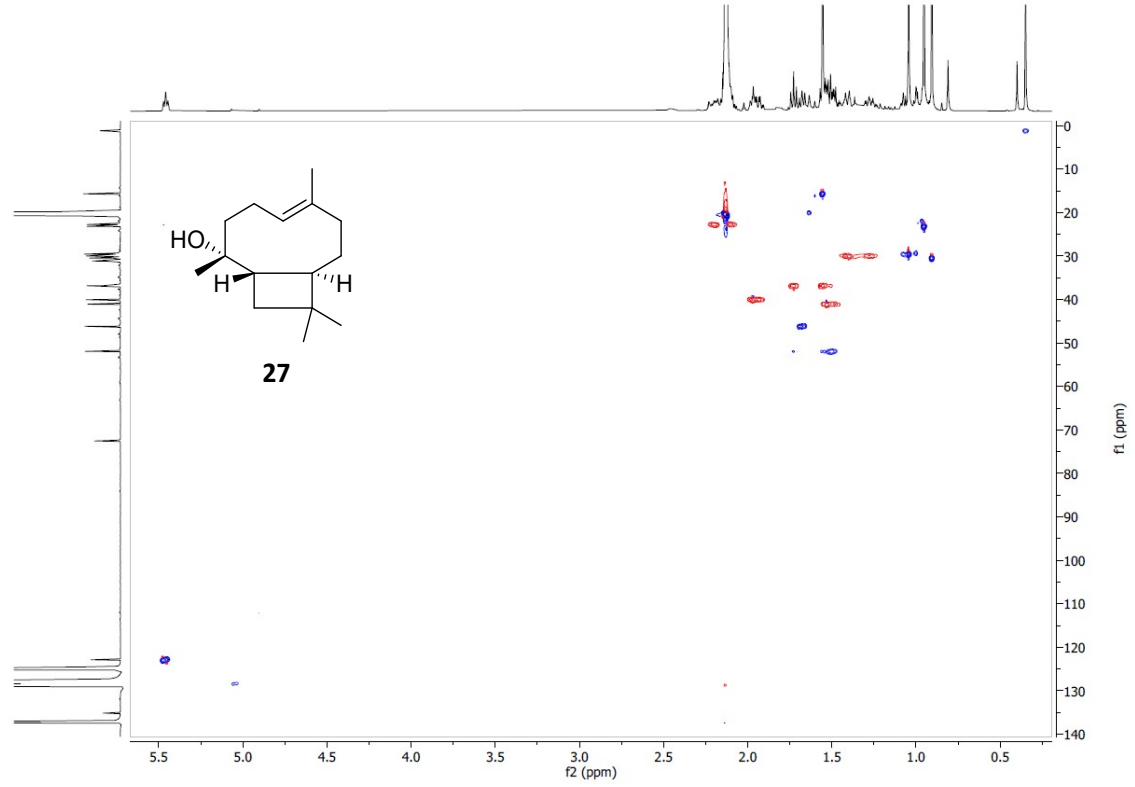
^{1}H-NMR (600 MHz, toluene-$d_{8,}$ δ= ppm, **27**)



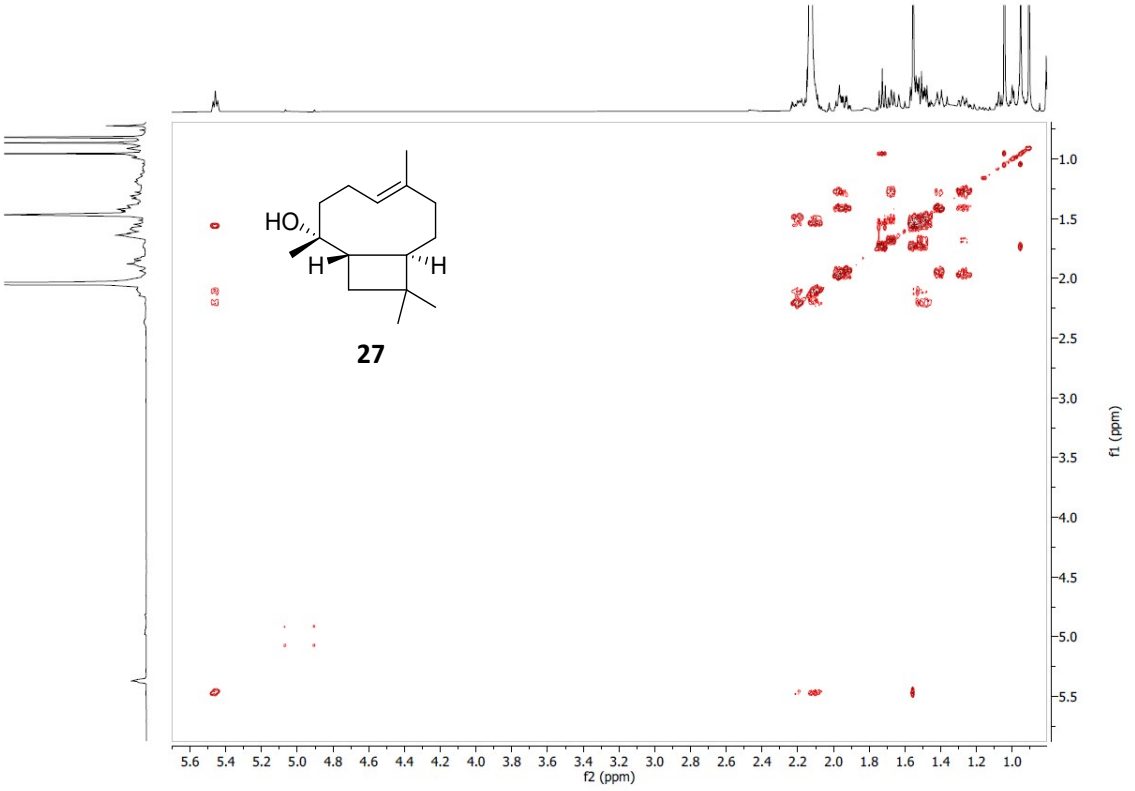
^{13}C-NMR (151 MHz, toluene-$d_{8,}$ δ= ppm, **27)**



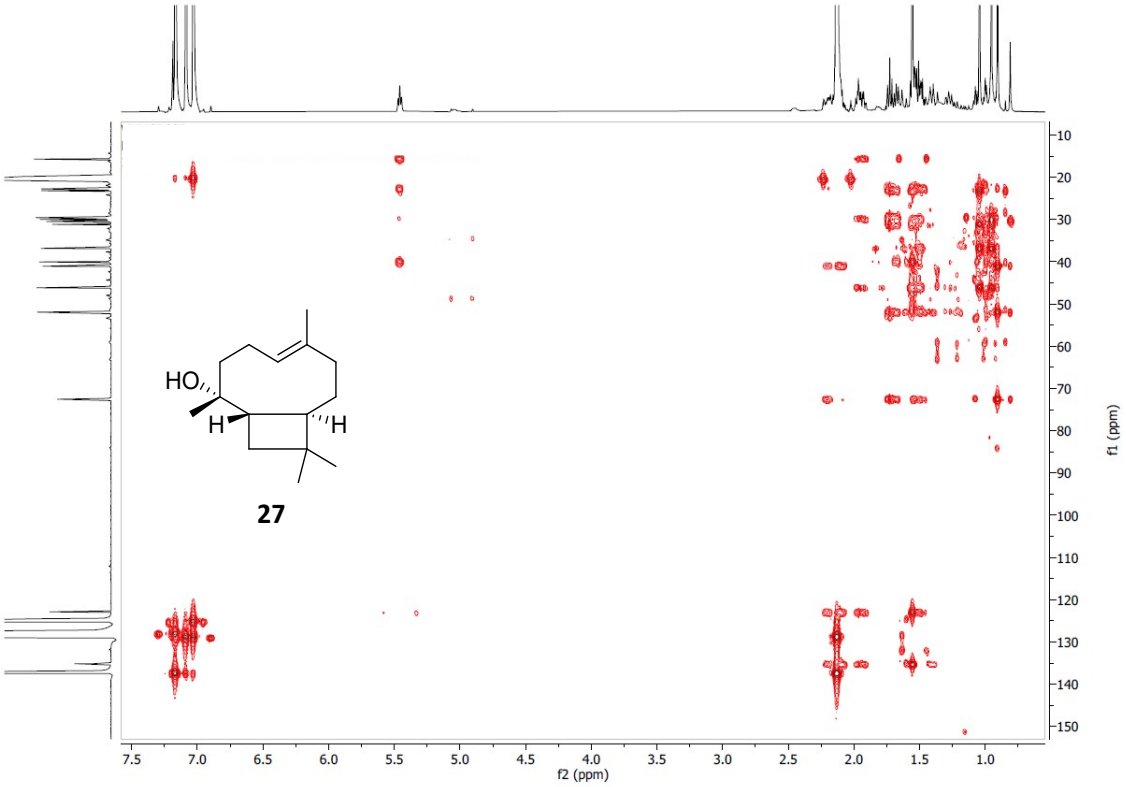
HSQC (**27**)



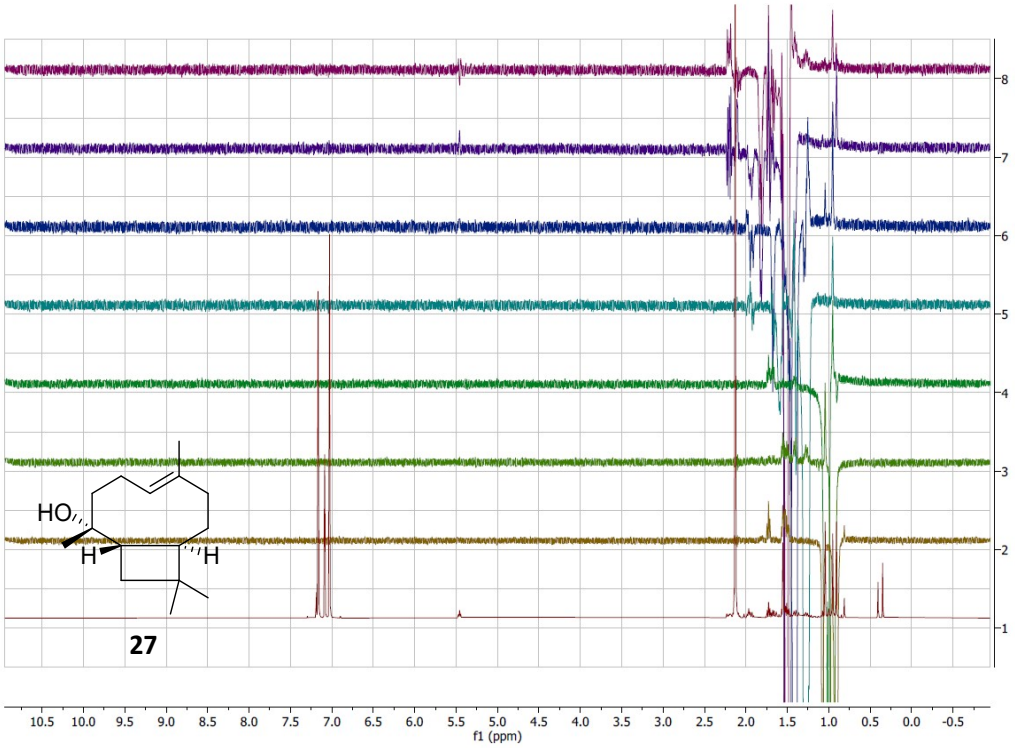
COSY (**27**)



HMBC (**27**)

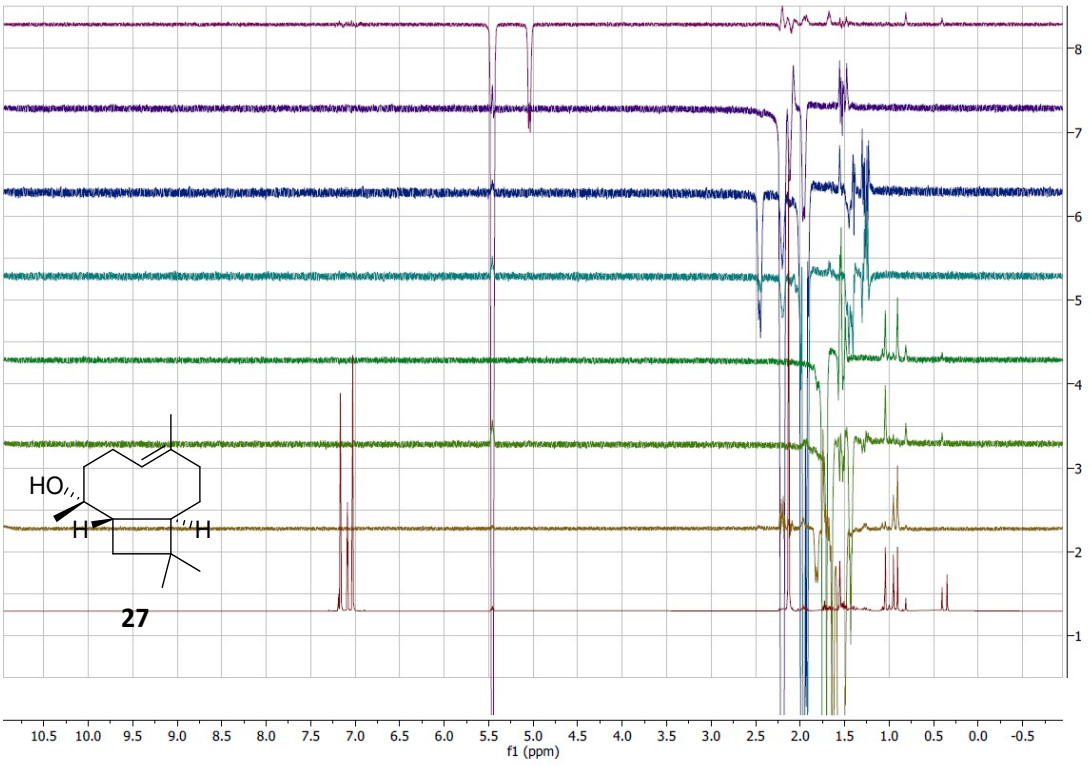


1d-NOE (**27**)



From bottom to top: ^{1}H-NMR, H-15, H-13, H-12, H-9a, H-9b, H-4a, H-4b.

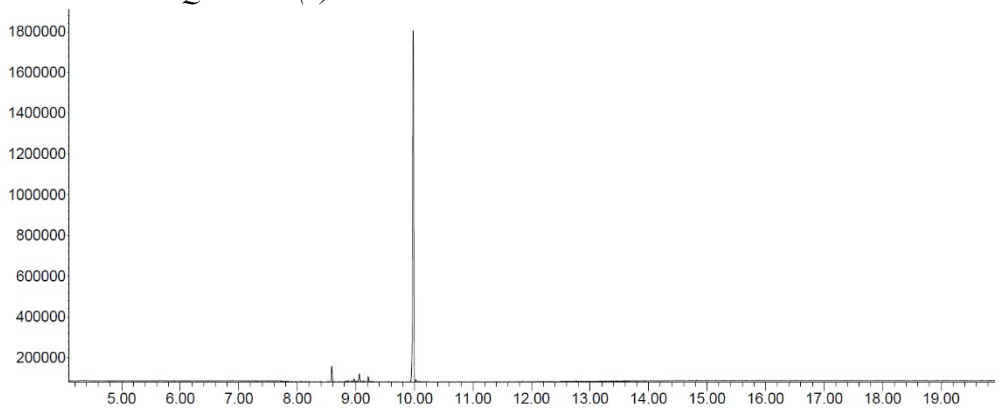
1d-NOE (**27**)



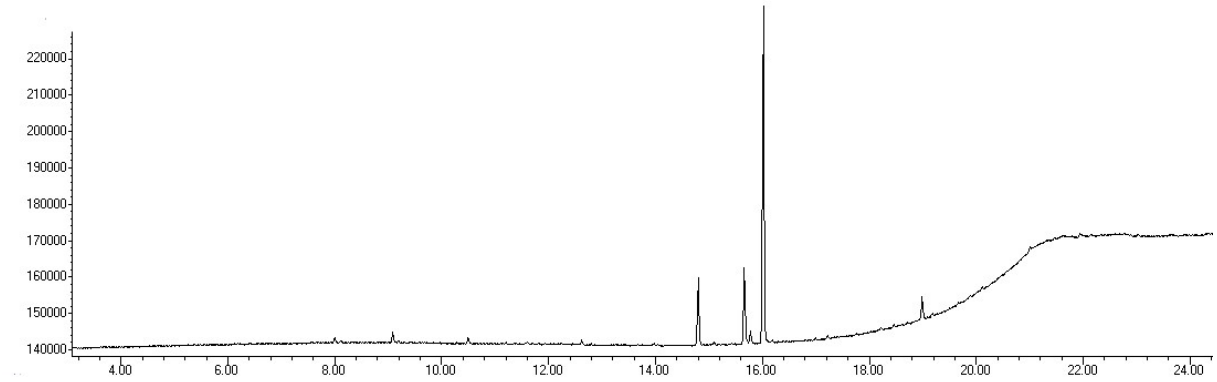
From bottom to top: ^{1}H-NMR, H-14, H-10, H-1b, H-8a, H-8b, H-5b, H-6.

B. GC chromatograms

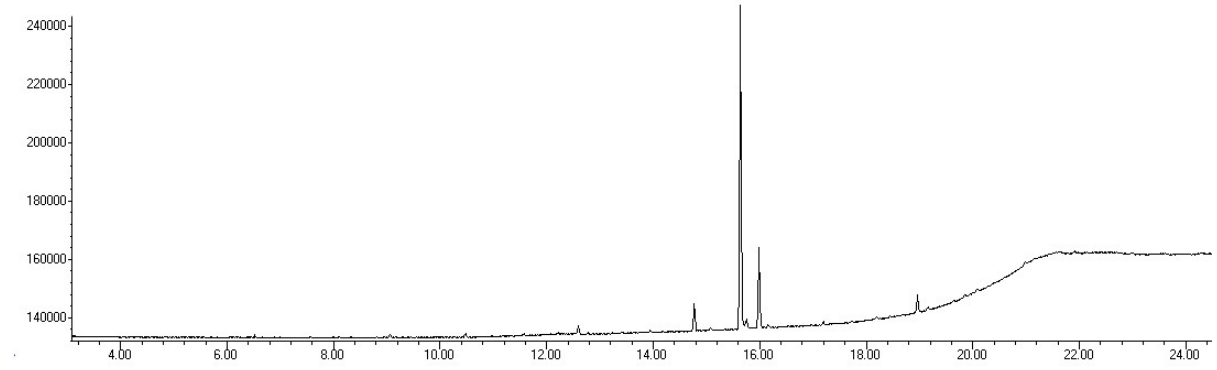
BcBOT2-W118Q + FPP (1)



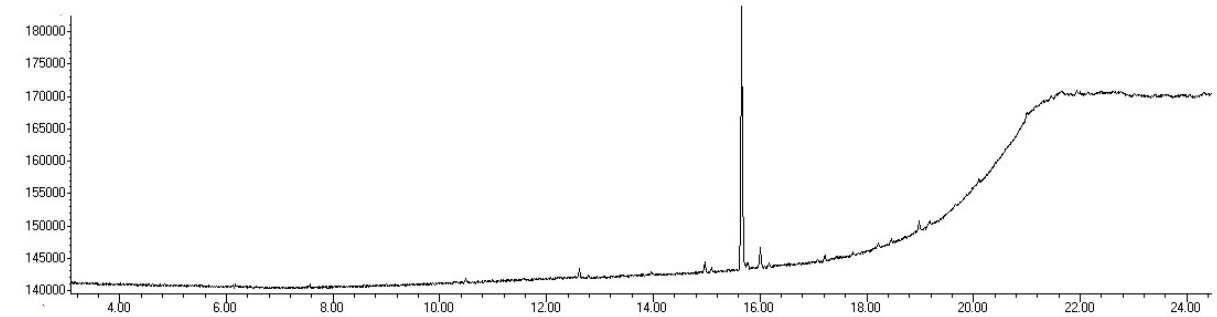
BcBOT2-W118F + FPP



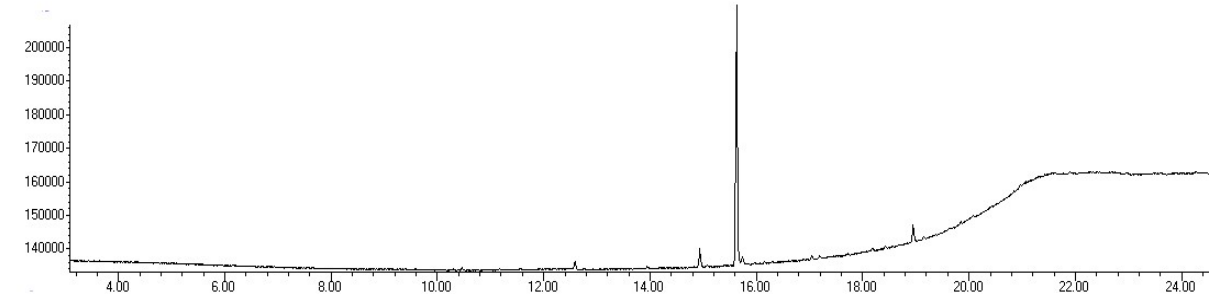
BcBOT2-W118M + FPP



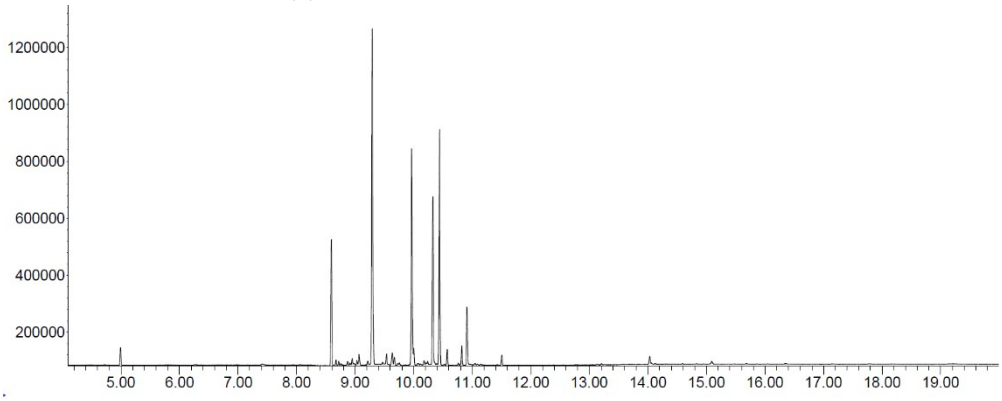
BcBOT2-W118H + FPP



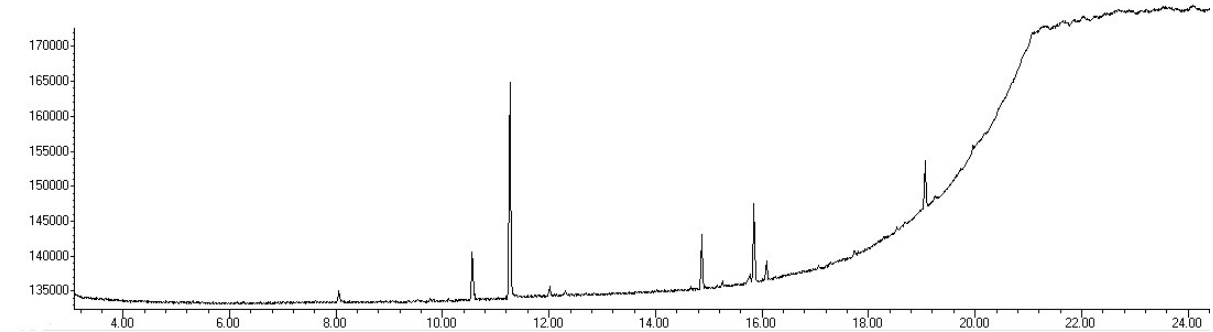
BcBOT2-W118N + FPP



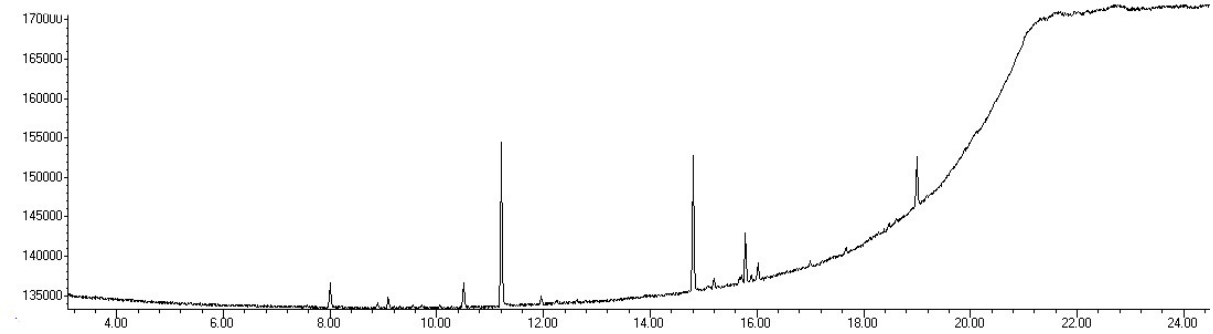
BcBOT2-F138V + FPP (1)



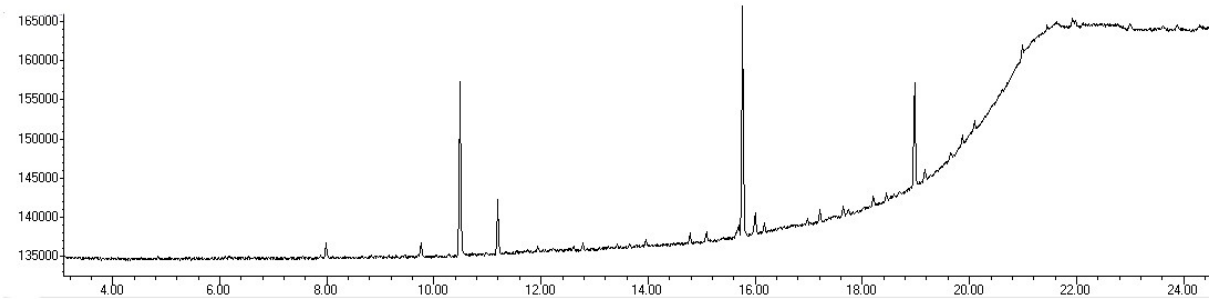
BcBOT2-F138A + FPP



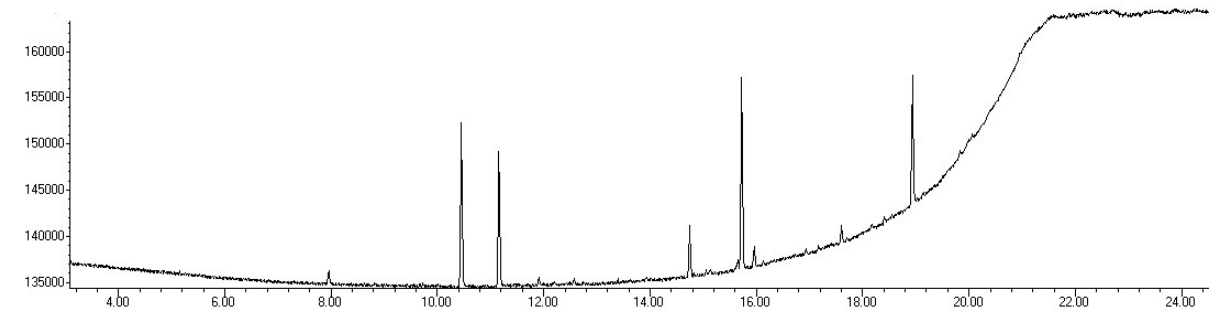
BcBOT2-F138M + FPP



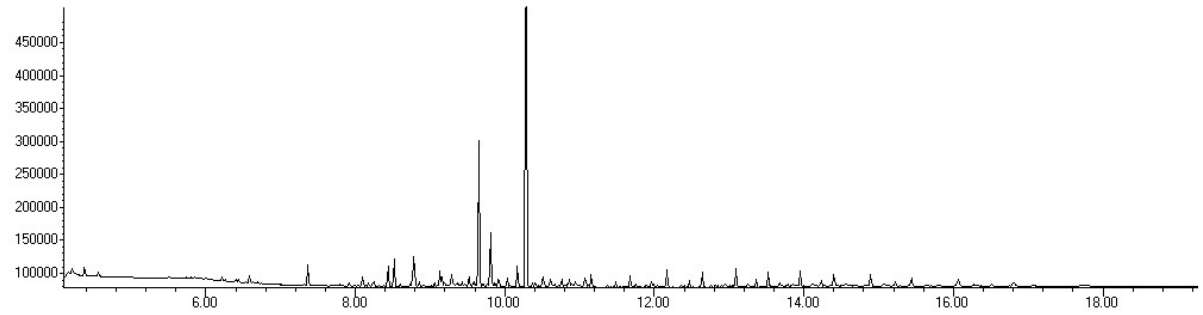
BcBOT2-F138N + FPP



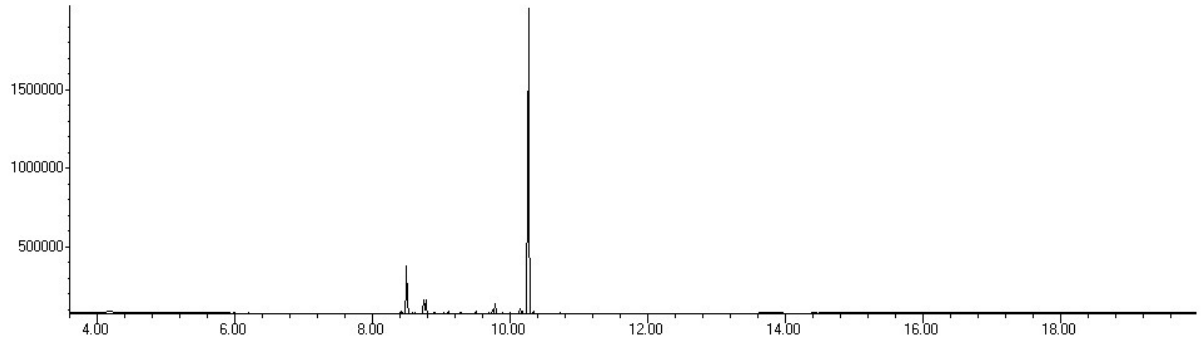
BcBOT2-F138Q + FPP



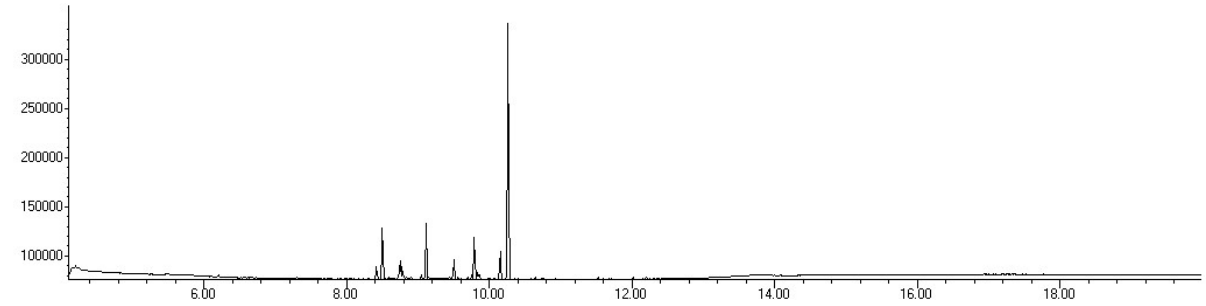
BcBOT2-W367A + FPP



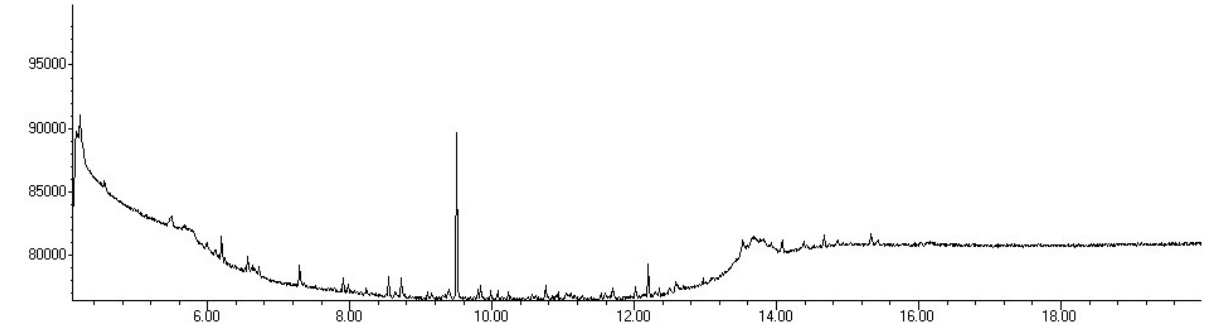
BcBOT2-W367F + FPP



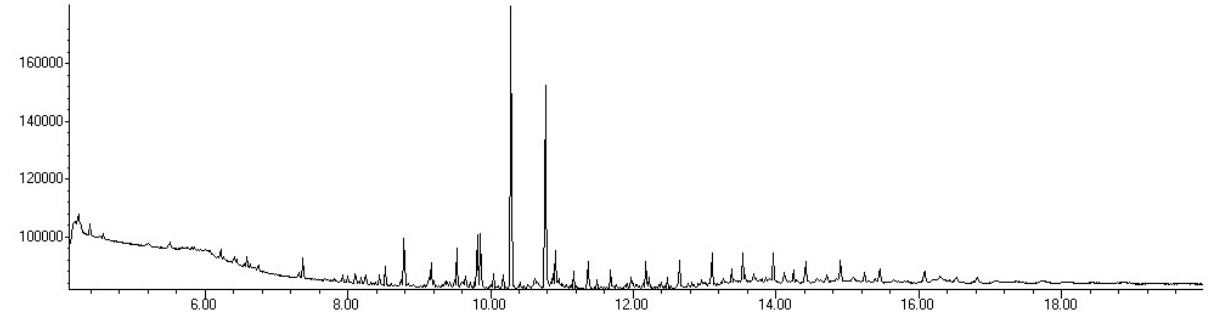
BcBOT2-W367H + FPP



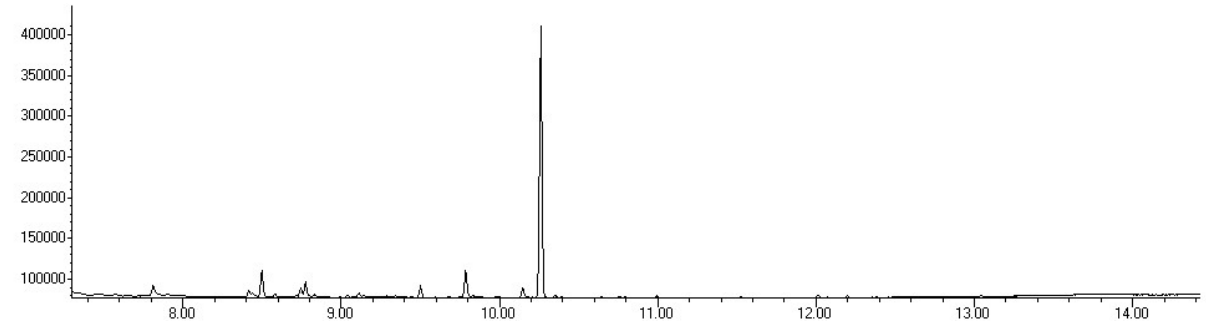
BcBOT2-W367K + FPP



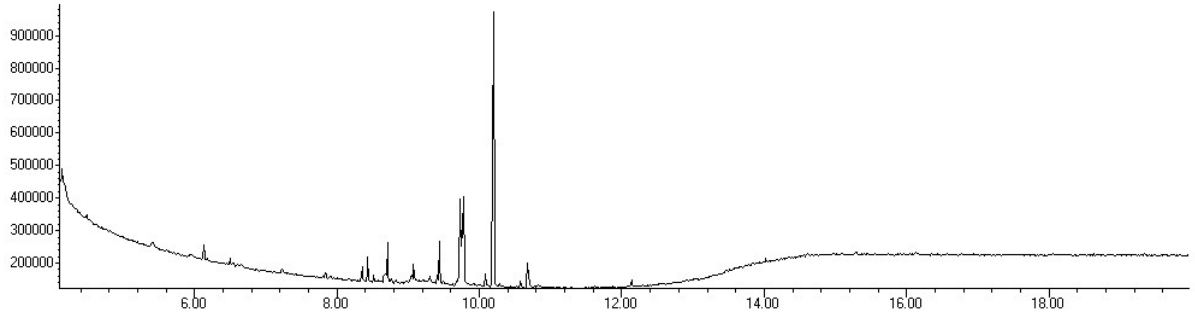
BcBOT2-W367N + FPP



BcBOT2-W367T + FPP

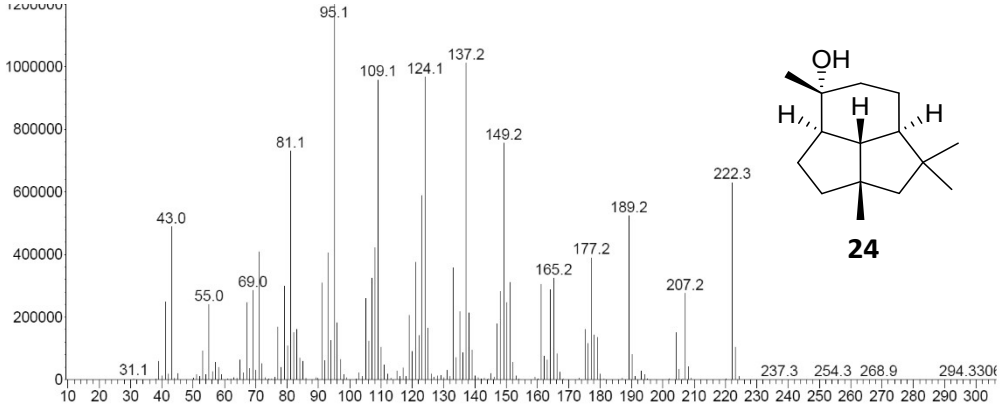


BcBOT2-W367V + FPP

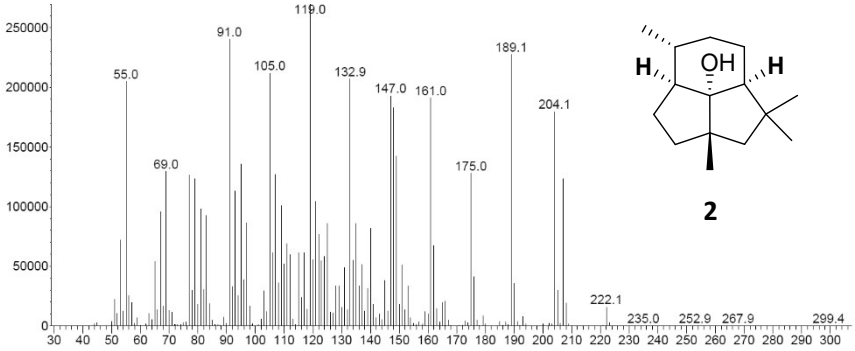


C. MS spectra

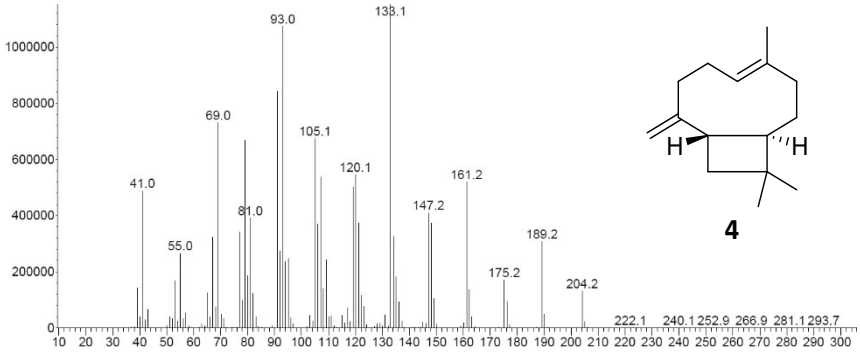
*Presilphiperfolan-9β-ol (**24**)*



*Presilphiperfolan-8β-ol (**2**)*

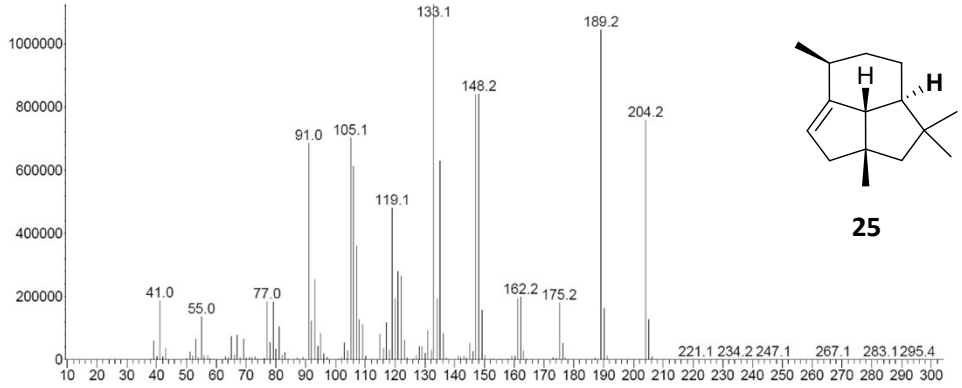


*B-Caryophyllene (**4**)*



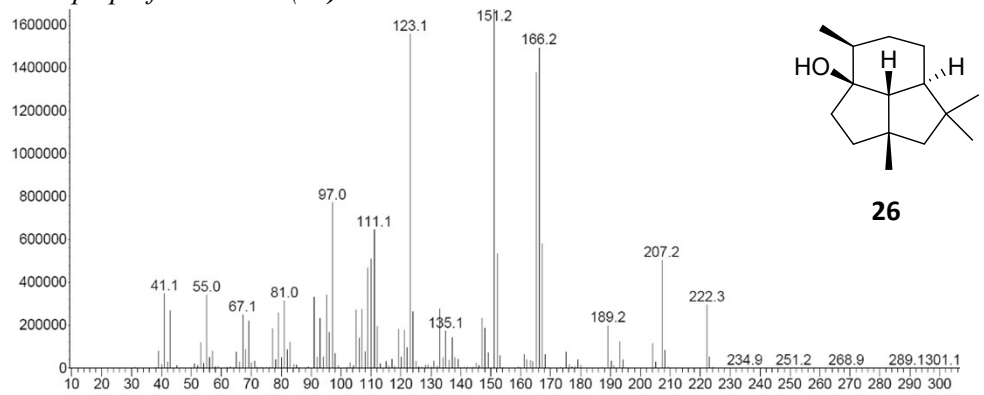
4

*Presilphiperfol-1-ene (**25**)*



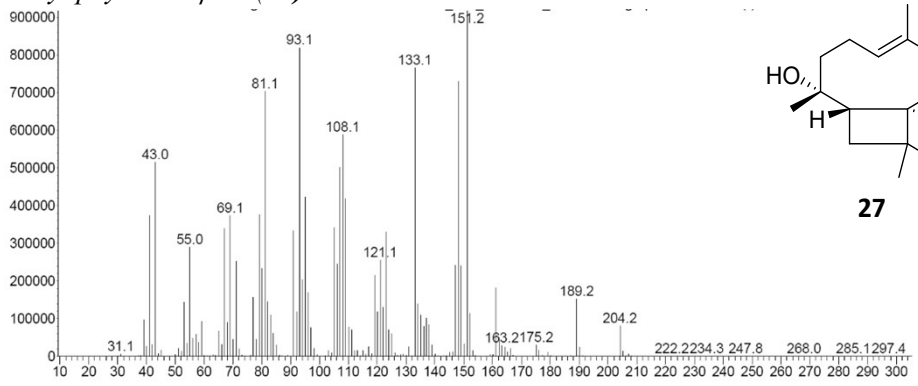
25

*Presilphiperfolan-1α-ol (**26**)*



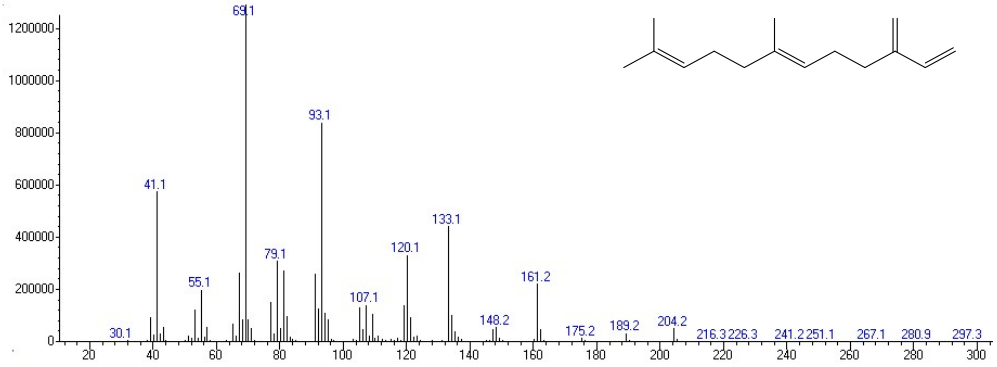
26

*Caryophyllene-8β-ol (**27**)*

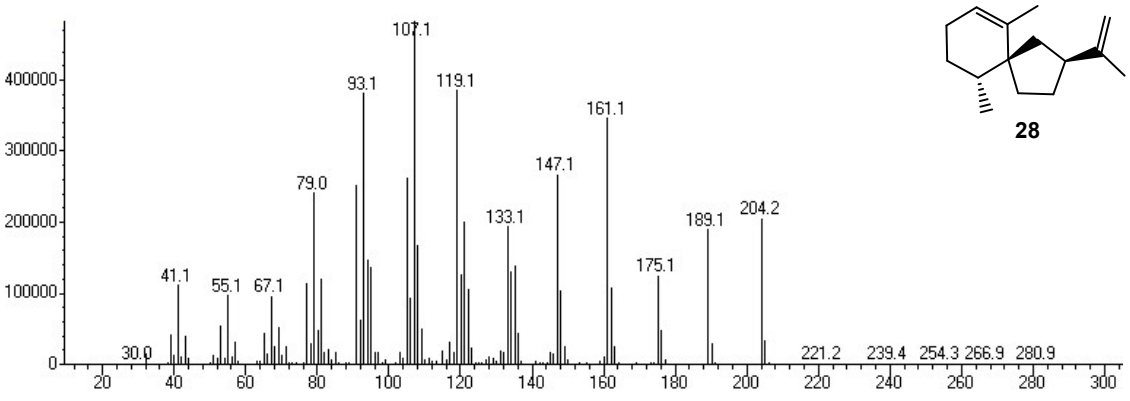


27

*(E)-β-farnesene (**22**)*



*Vetispiradiene (**28**)*



28