ChemBioChem

Supporting Information

Heterocyclic and Alkyne Terpenoids by Terpene Synthase-Mediated Biotransformation of Non-Natural Prenyl Diphosphates: Access to New Fragrances and Probes

Benjamin Weigel[+], Jeanette Ludwig[+], Roman A. Weber[+], Steve Ludwig, Claudia Lennicke, Paul Schrank, Mehdi D. Davari, Mohamed Nagia, and Ludger A. Wessjohann*

Table of Contents

Supporting results

Fragmentation of monoterpene-analogous products resulting from substrates 1, 2, and 3

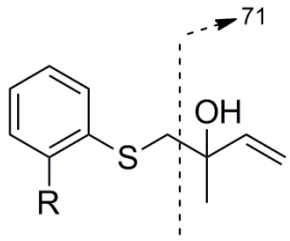


Figure S1: α-cleavage of the aromatic products **1d**, **2b,** and **3b** leading to the formation of the common fragment ion with *m/z* = 71.

Biotransformation of GPP by TEAS

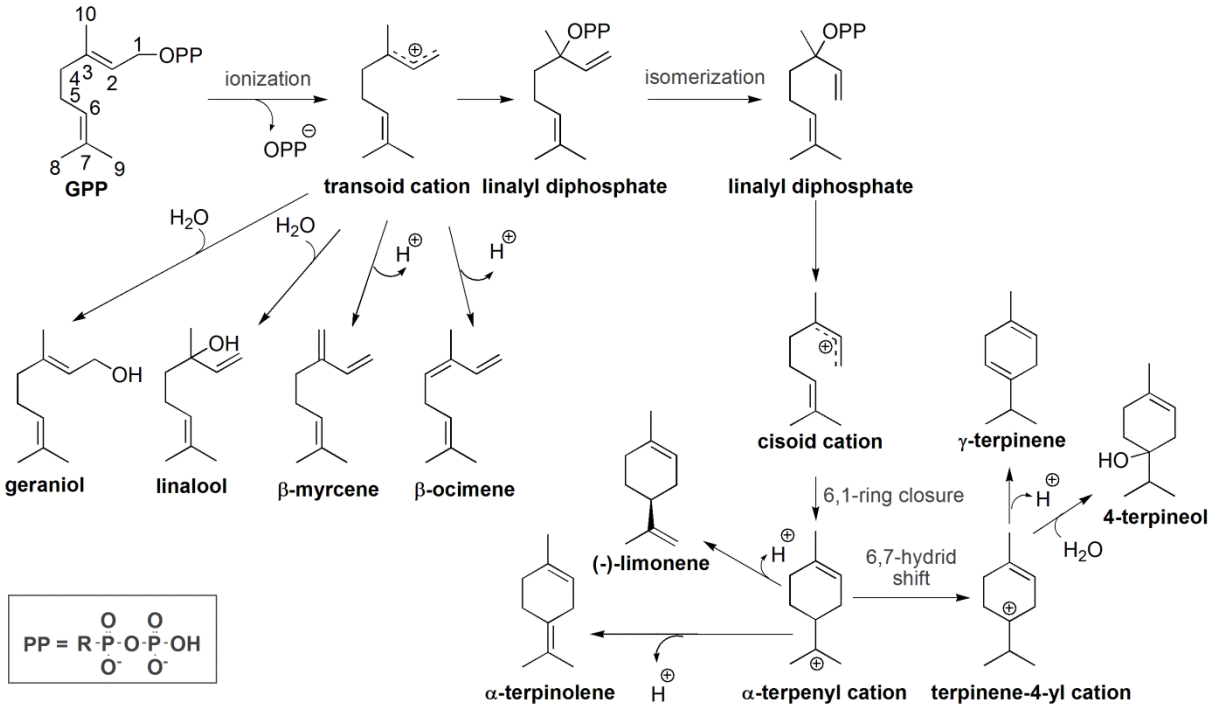


Figure S2: Proposed mechanism for the formation of cyclic and acyclic monoterpenes by TEAS via the transoid and cisoid cations.

Chiral GC-MS of the enzymatic products

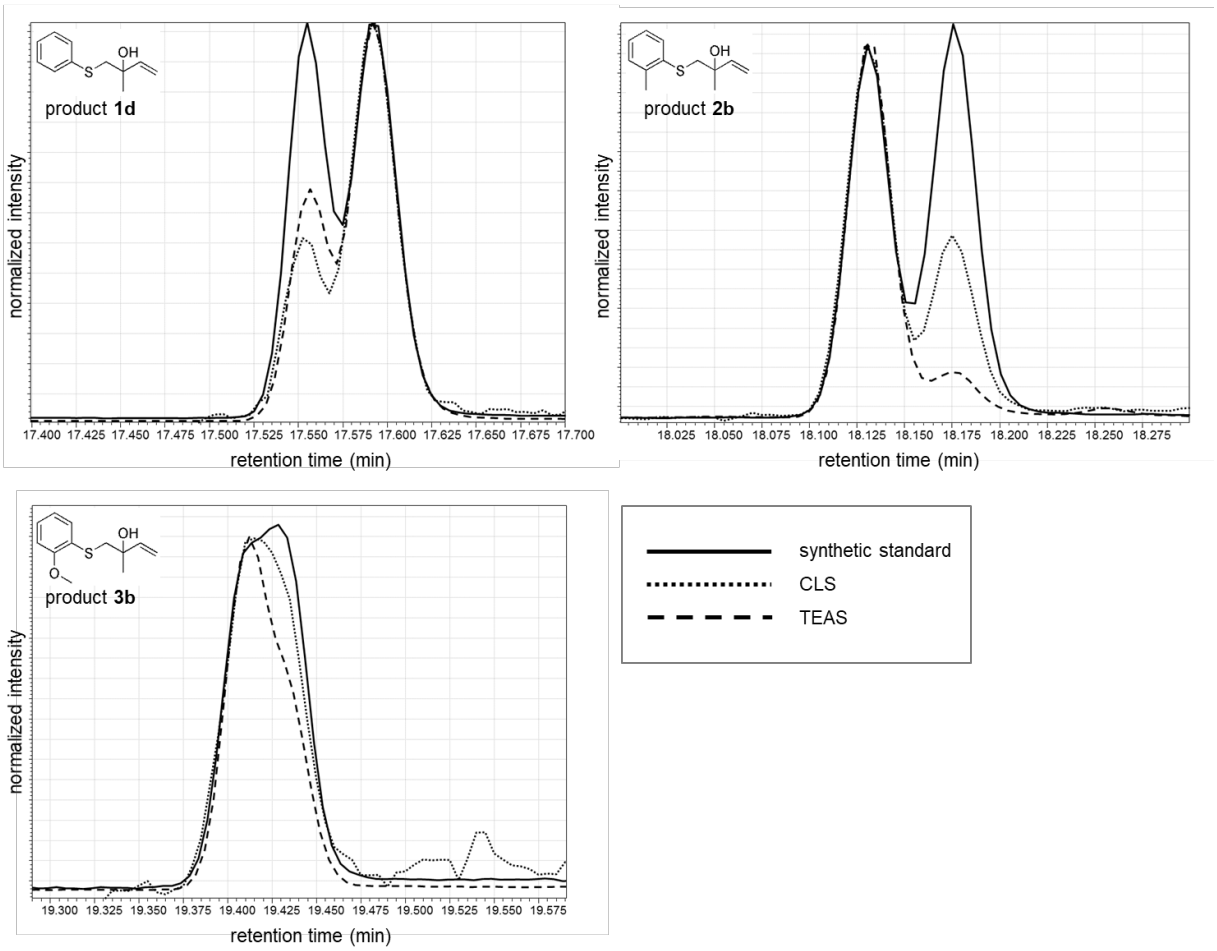


Figure S3: Chiral column GC/MS chromatograms of the identified products **1d**, **2b,** and **3b** in comparison to the synthetic racemic standards. Shown are only segments of the TIC-chromatograms with relevant peaks of the identified main product. For **3b**, the enantiomers are not well separated and just show a much-broadened peak. However, the enzymatic products show a sharp peak at the front end, clearly indicating the non-racemic product formation.

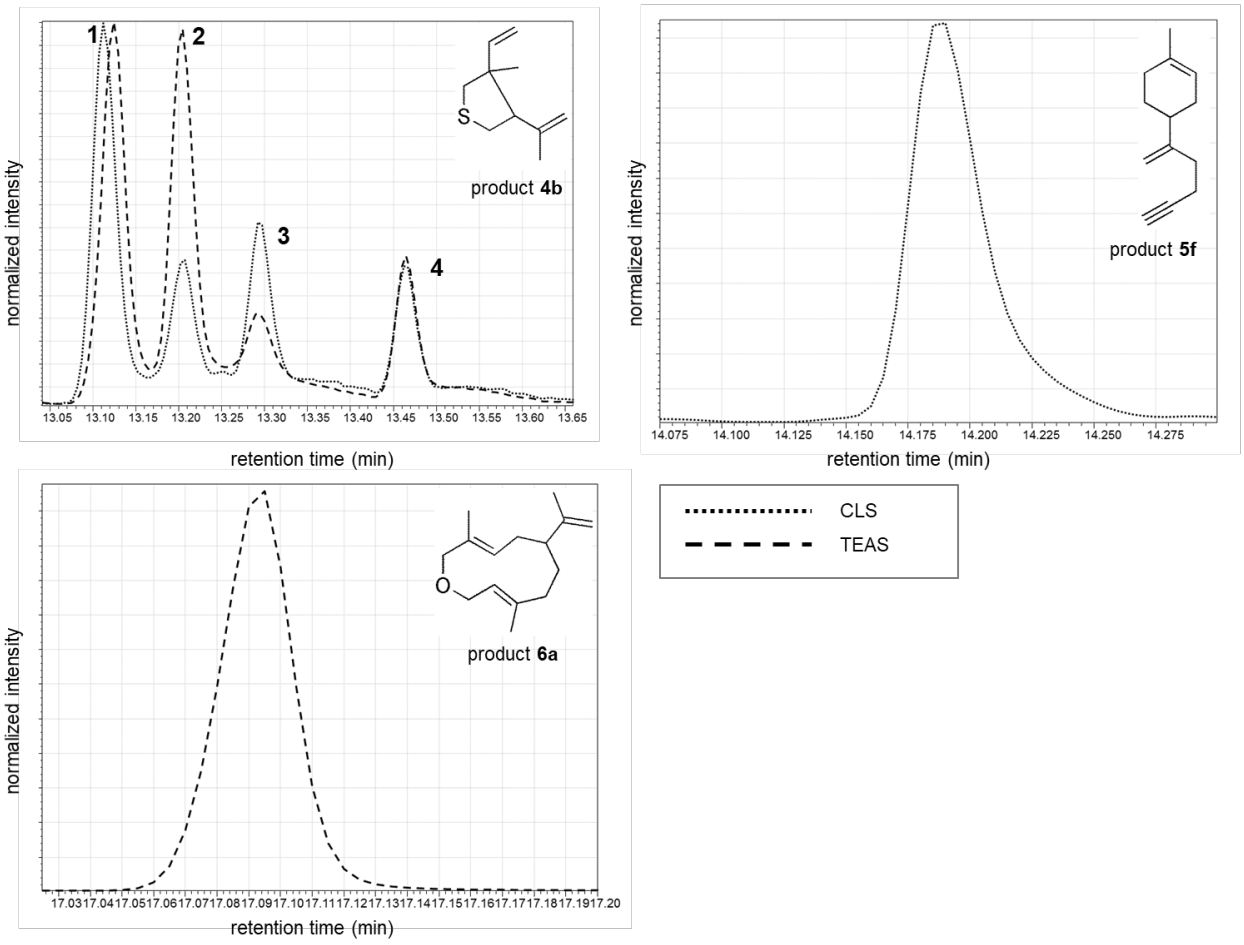
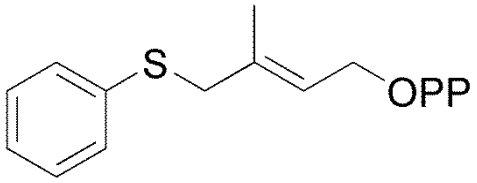


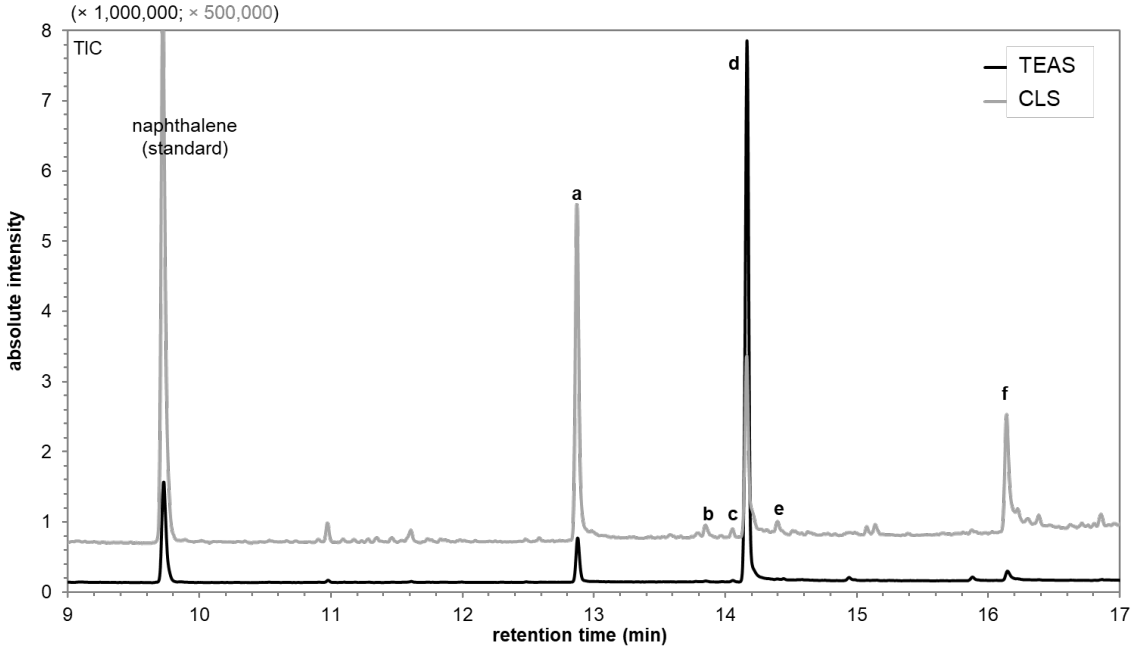
Figure S4: Chiral column GC/MS chromatograms of the identified products **4b**, **5f,** and **6a**. Shown are only segments of the TIC-chromatograms with relevant peaks of the identified main products of substrates **4**, **5,** and **6**. The formation of four different diastereomers of product **4b** (1-4) by CLS and three diastereomers (2-4) by TEAS could be confirmed. The first depicted peak of TEAS-products of substrate **4** is not corresponding to diastereomers of product **4b** due to its different EI-fragmentation pattern. Product **5f** likely is only one enantiomer as only one sharp peak occurs in chiral GC/MS measurements. The peak of **6a** corresponds to (-)-oxa-germacrene.

GC-MS chromatograms and spectra from the product mixtures of successful biotransformations with non-natural and natural substrates.

Table S1. GC/MS chromatogram of the single vial assays of CLS and TEAS with corresponding EI mass spectra for each product peak with substrate **1**. RT, retention time; RI, retention index

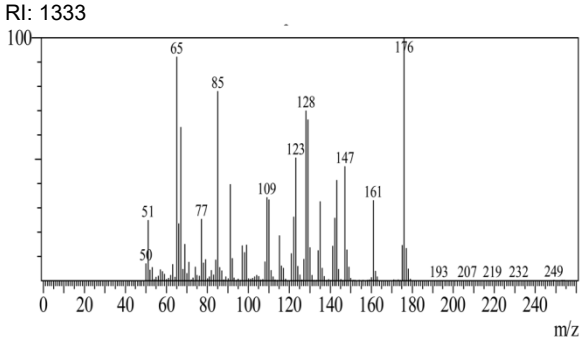


substrate **1**



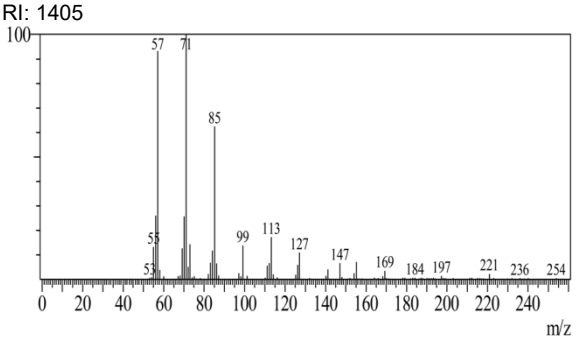
product a

RT: 12.872 min
RI: 1333



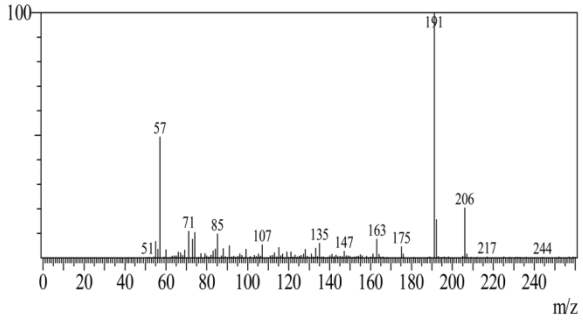
product b

RT: 13.849 min
RI: 1405



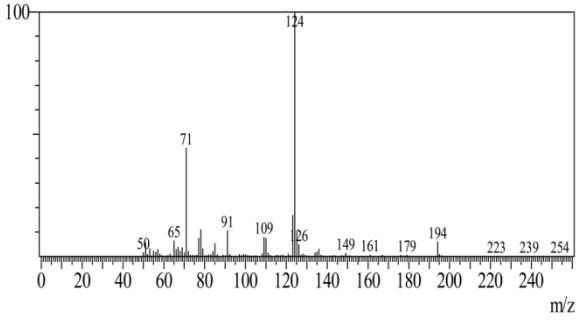
product c

RT: 14.055 min
RI: 1421



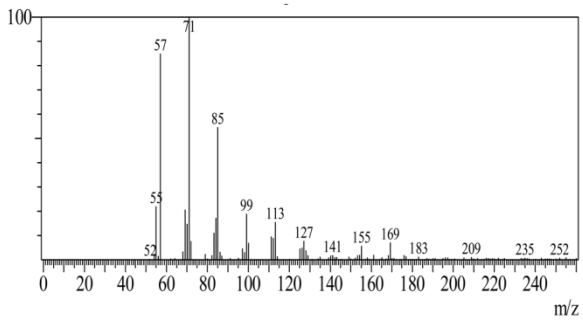
product d

RT: 14.163 min
RI: 1430



product e

RT: 14.397 min
RI: 1448



product f (3-methyl-4-(phenylthio)-(*E*)-2-buten-1-ol)

RT: 16.140 min
RI: 1588

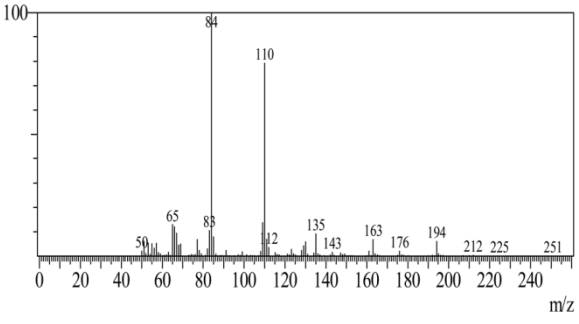
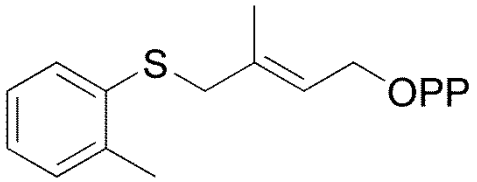
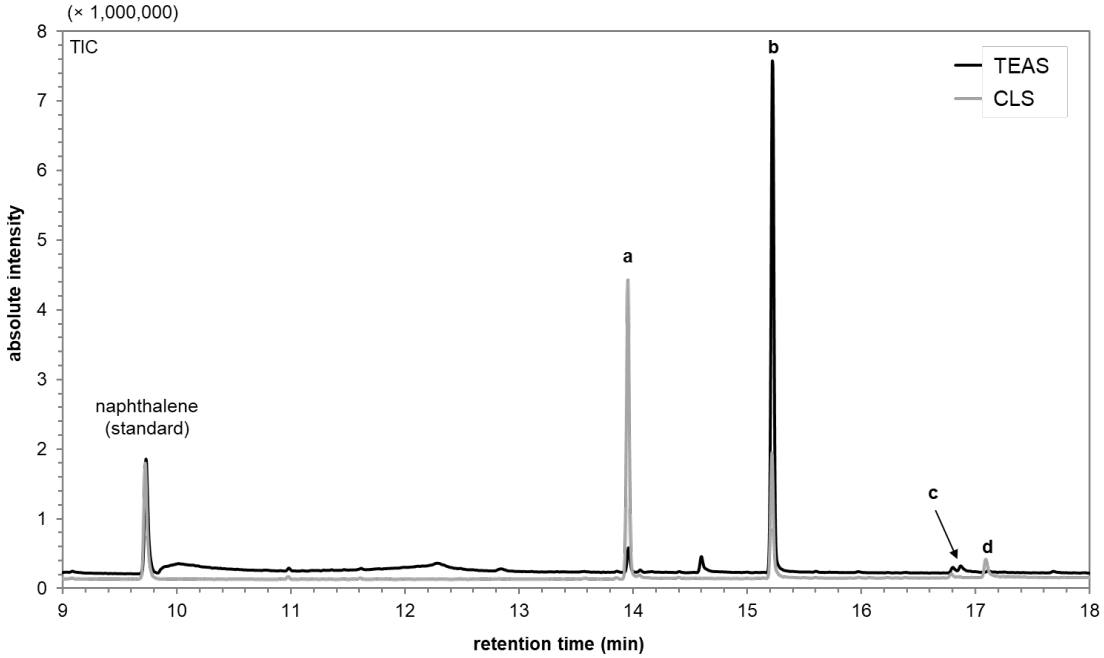


Table S2. GC/MS chromatogram of the single vial assays of CLS and TEAS with corresponding EI mass spectra for each product peak with substrate **2**. RT, retention time; RI, retention index.

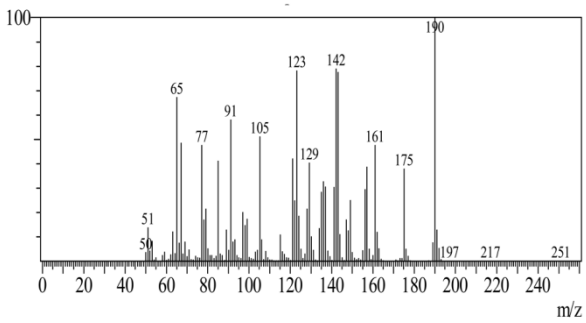


substrate **2**



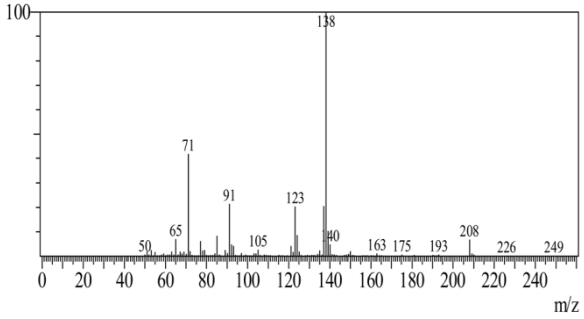
product a

RT: 13.954 min
RI: 1413



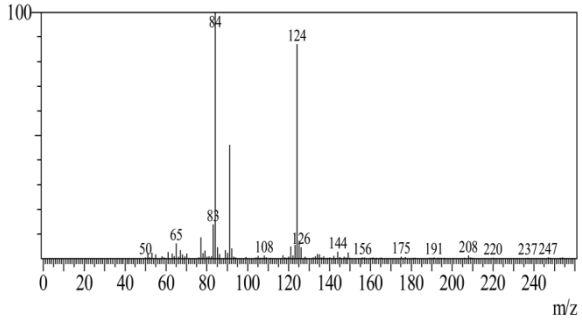
product b

RT: 15.217 min
RI: 1512



product c

RT: 16.802 min
RI: 1645



product d

RT: 17.091 min
RI: 1670

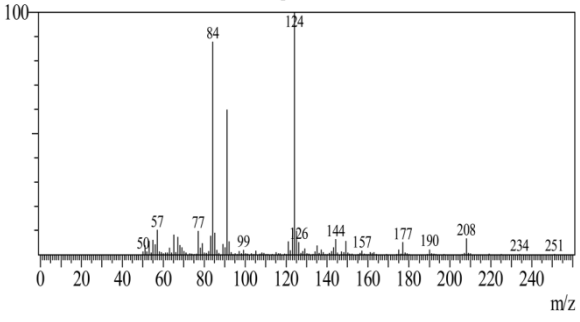
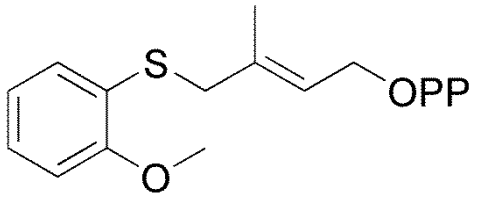
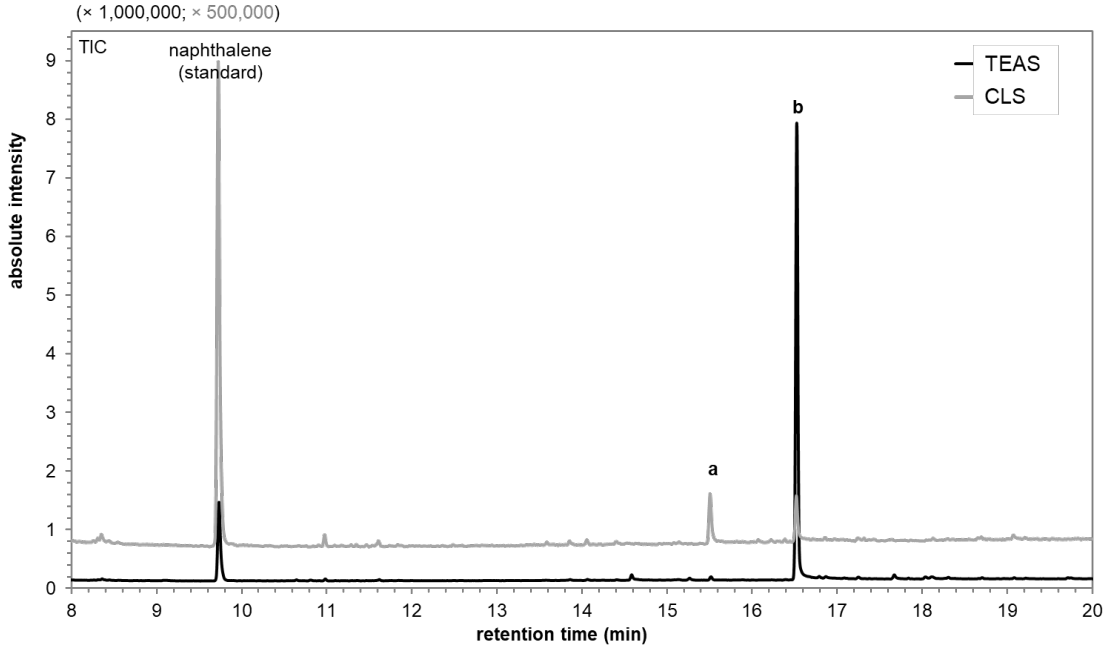


Table S3. GC/MS chromatogram of the single vial assays of CLS and TEAS with corresponding EI mass spectra for each product peak with substrate **3**. RT, retention time; RI, retention index.

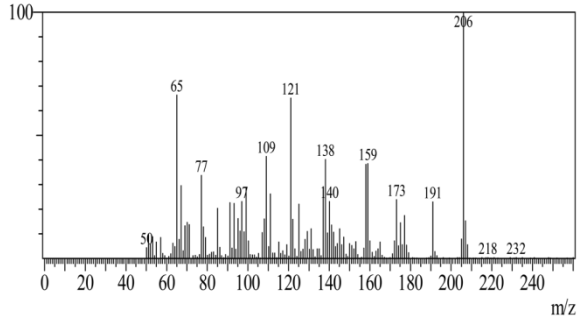


substrate **3**



product a

RT: 15.507 min
RI: 1536



product b

RT: 16.521 min
RI: 1621

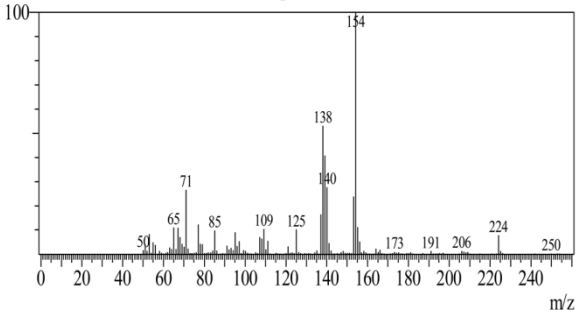
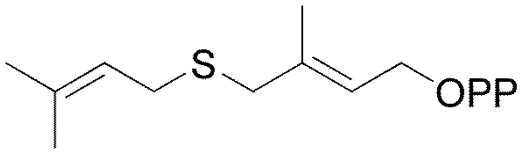
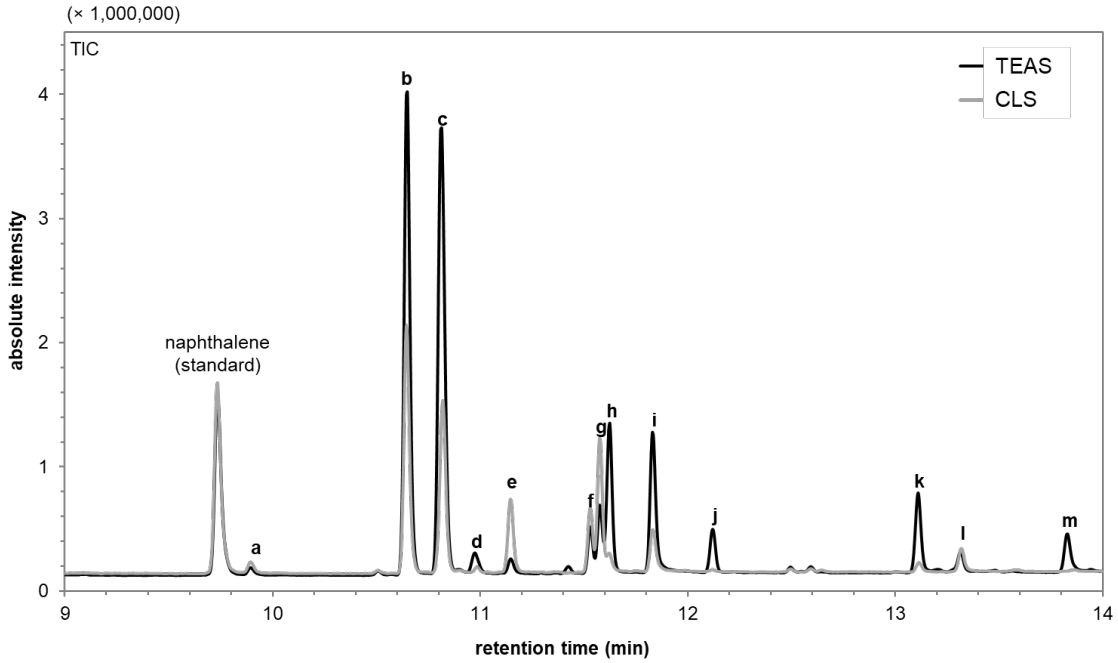


Table S4. GC/MS chromatogram of the single vial assays of CLS and TEAS with corresponding EI mass spectra for each product peak with substrate **4**. RT, retention time; RI, retention index.

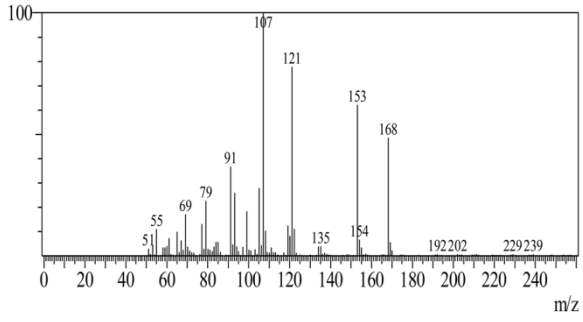


substrate **4**



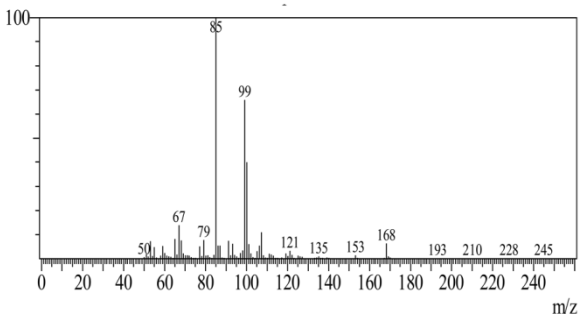
product a

RT: 9.895 min
RI: 1129



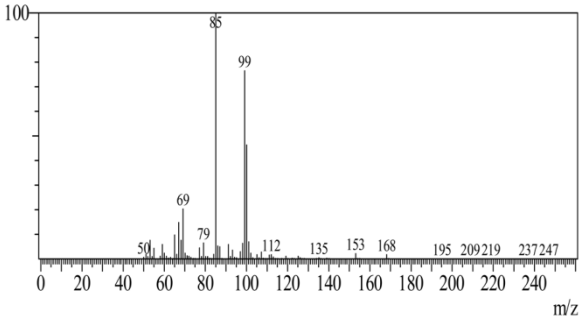
product b

RT: 10.646 min
RI: 1178



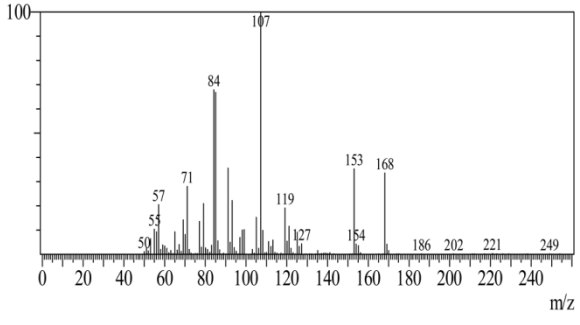
product c

RT: 10.820 min
RI: 1189



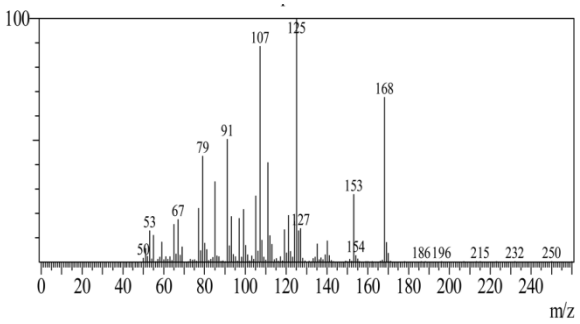
product d

RT: 10.975 min
RI: 1200



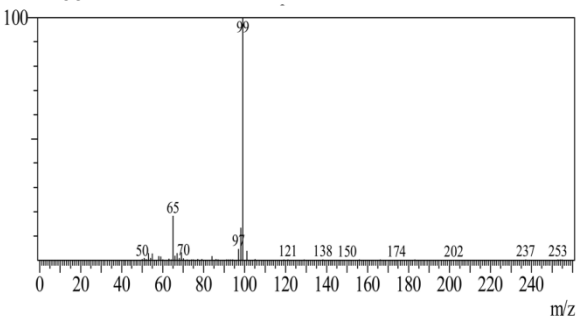
product e

RT: 11.147 min
RI: 1211



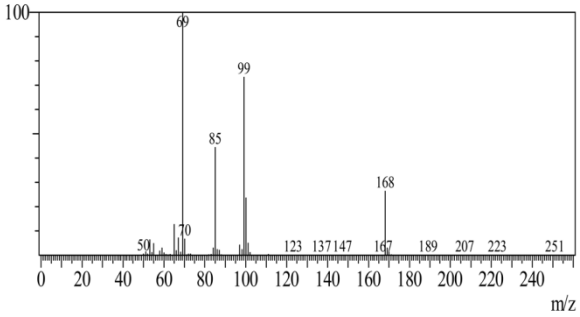
product f

RT: 11.531 min
RI: 1238



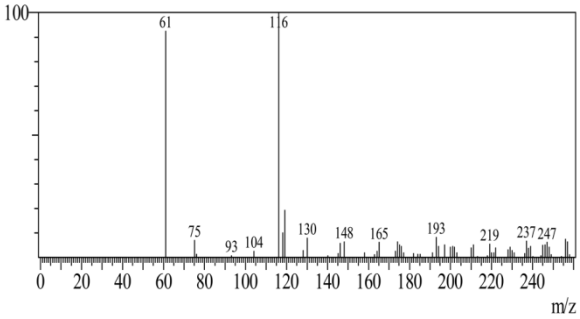
product g

RT: 11.577 min
RI: 1241



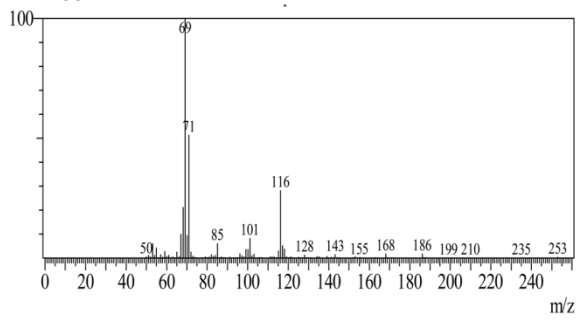
product h

RT: 11.621 min
RI: 1244



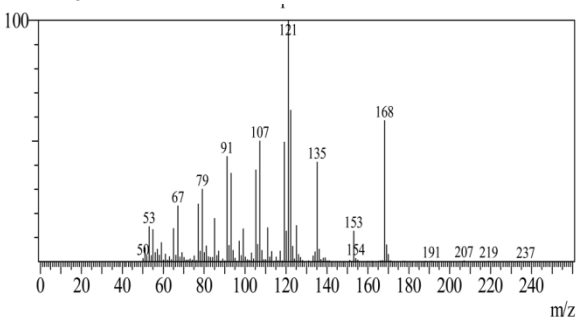
product i

RT: 11.832 min
RI: 1259



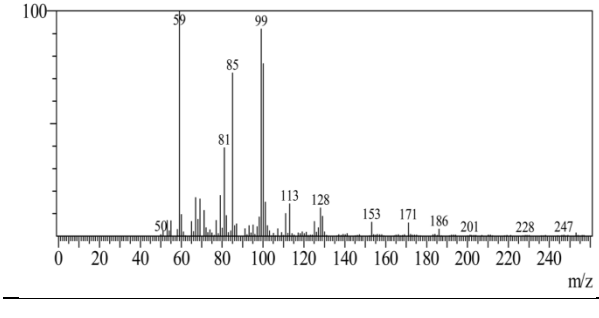
product j

RT: 12.121 min
RI: 1279



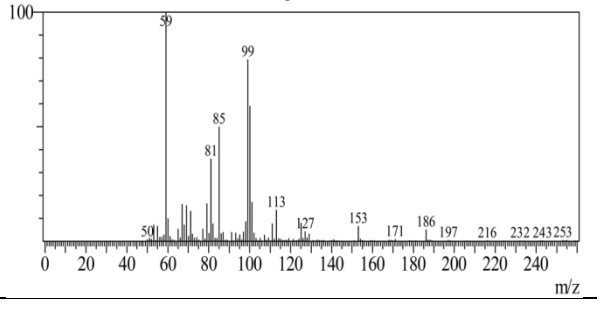
product k

RT: 13.114 min
RI: 1351



product l

RT: 13.318 min
RI: 1366



product m

RT: 13.828 min

RI: 1404

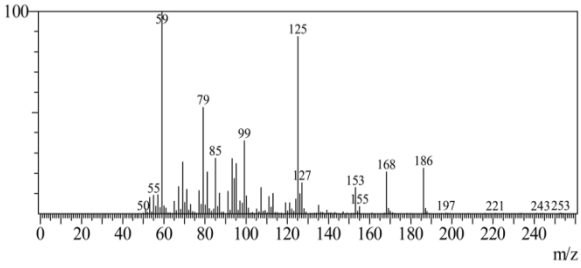
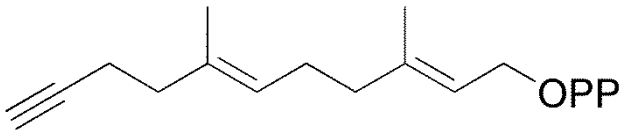
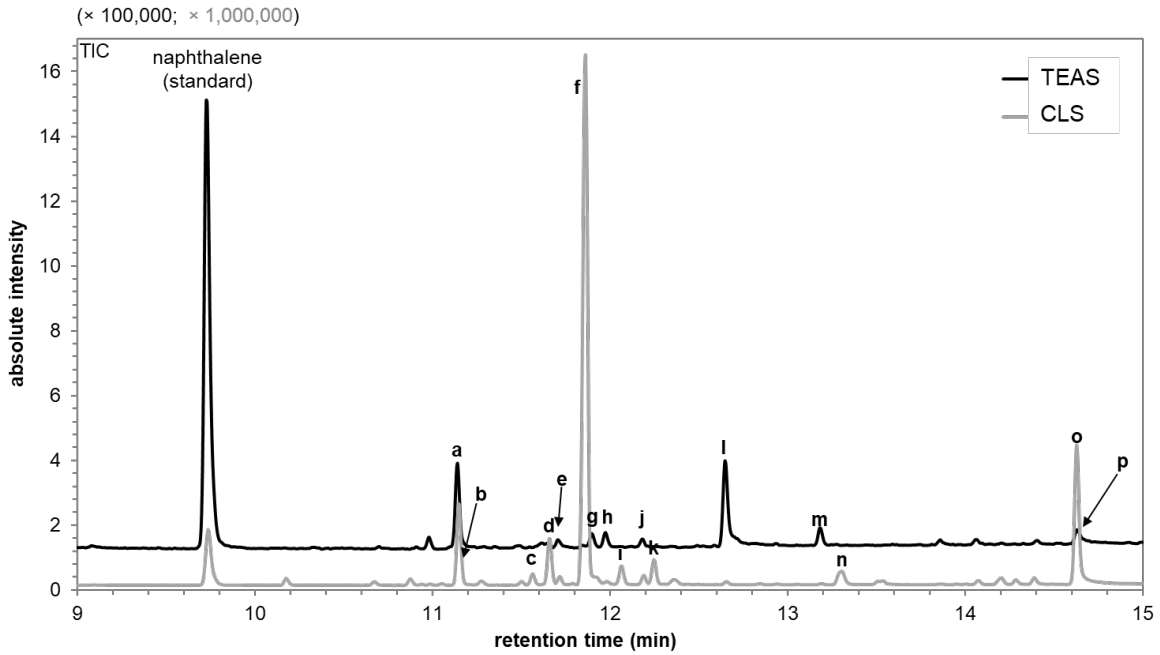


Table S5. GC/MS chromatogram of the single vial assays of CLS and TEAS with corresponding EI mass spectra for each product peak with substrate **5**. RT, retention time; RI, retention index.

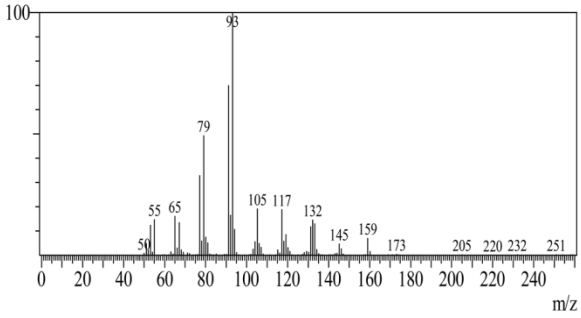


substrate 5



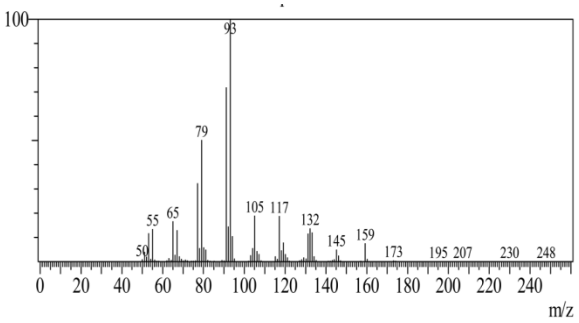
product a

RT: 11.139 min
RI: 1211



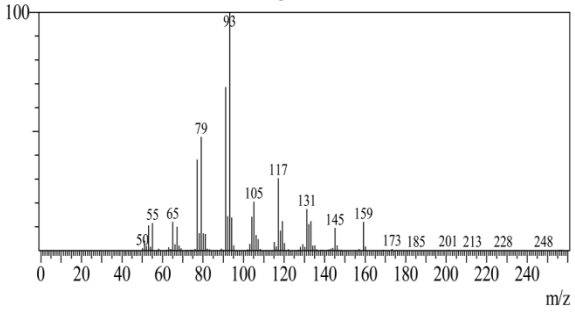
product b

RT: 11.147 min
RI: 1211



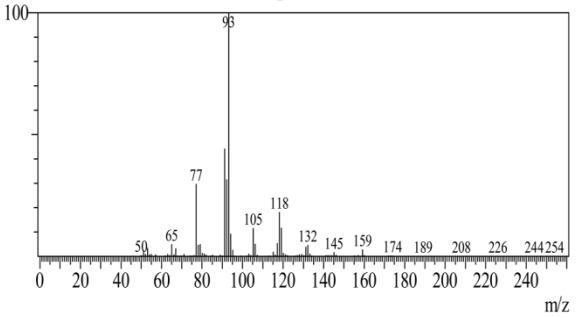
product c

RT: 11.563 min
RI: 1124



product d

RT: 11.658 min
RI:1247

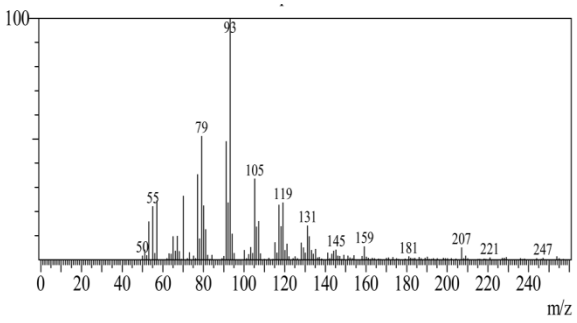


product e

RT: 11.705 min
RI: 1250

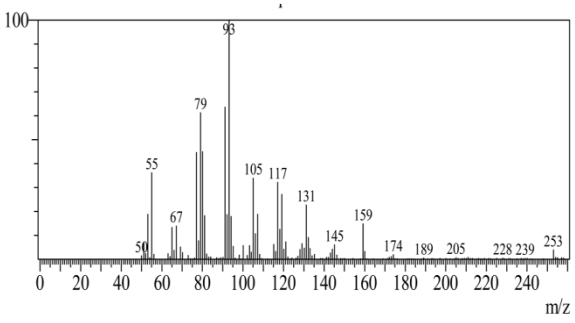
product f

RT: 11.860 min
RI:1261



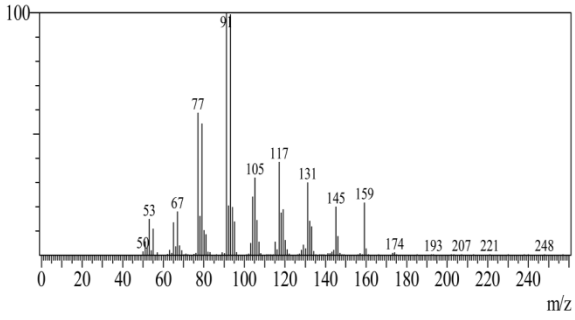
product g

RT: 11.896 min
RI: 1263



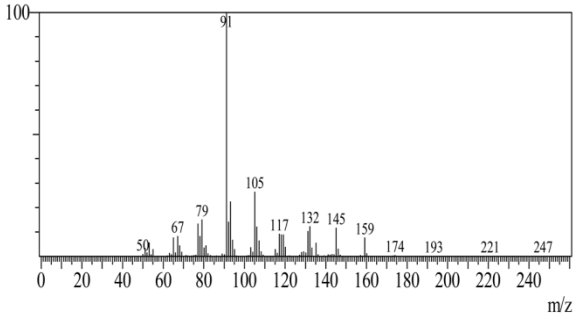
product i

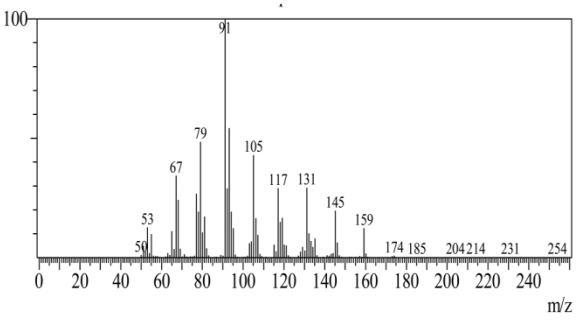
RT: 12.063 min
RI: 1275



product k

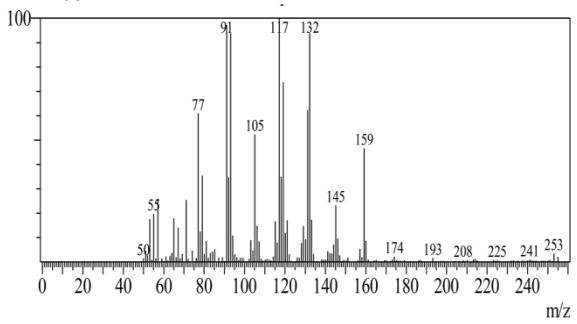
RT: 12.246 min
RI: 1288





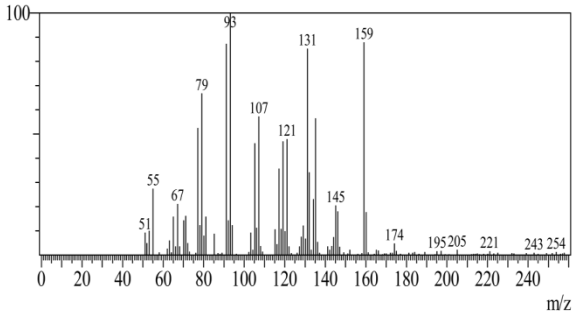
product h

RT: 11.972 min
RI: 1269



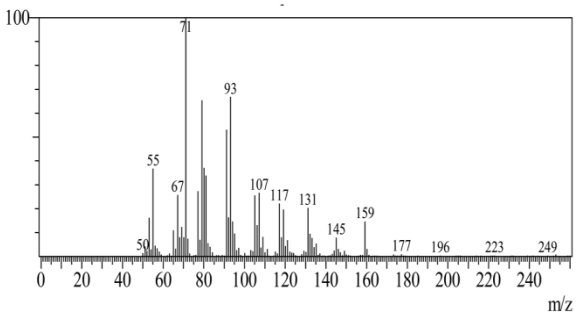
product j

RT: 12.181min
RI: 1283



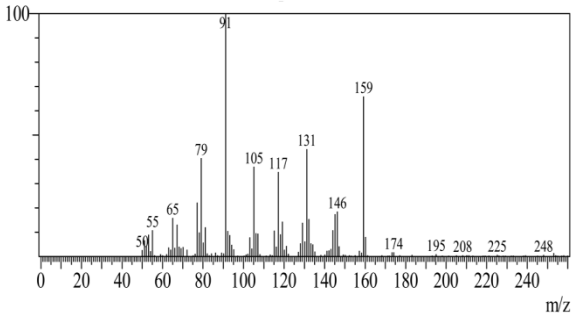
product l

RT: 12.647 min
RI: 1316



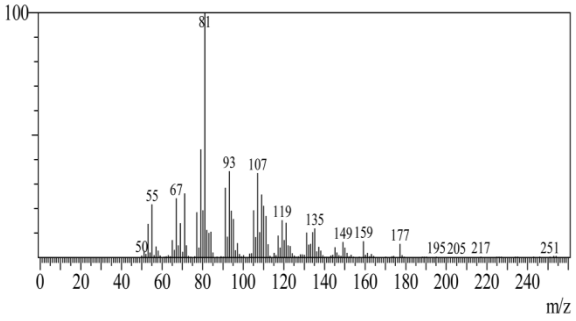
product m

RT: 13.181 min
RI: 1356



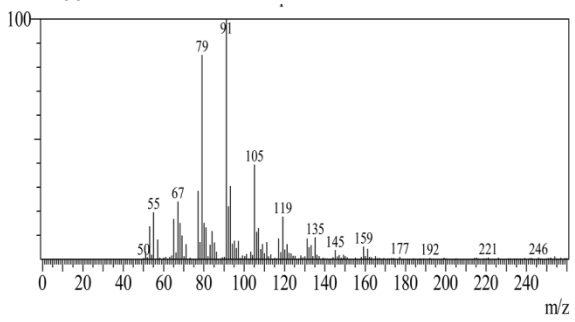
product n

RT: 13.303 min
RI: 1365



product o

RT: 14.626 min
RI: 1466



product p

RT: 14.628 min
RI: 1466

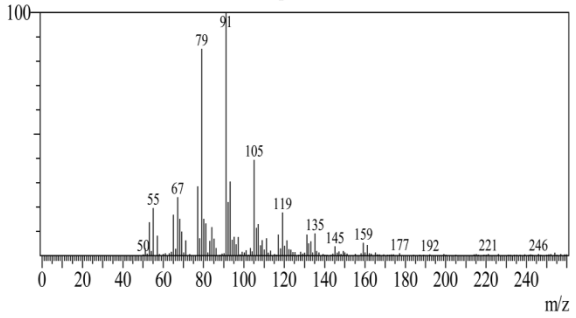
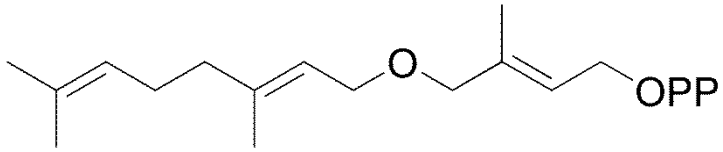
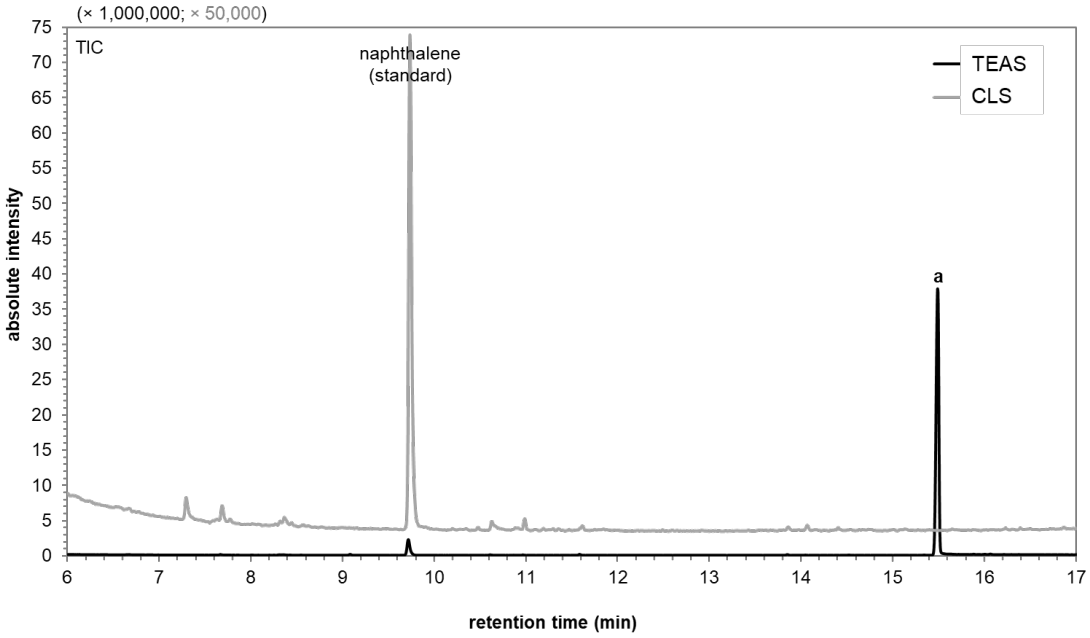


Table S6. GC/MS chromatogram of the single vial assays of CLS and TEAS with corresponding EI mass spectra for each product peak with substrate **6**. RT, retention time; RI, retention index.



substrate **6**



product a

RT: 15.492 min

RI: 1535

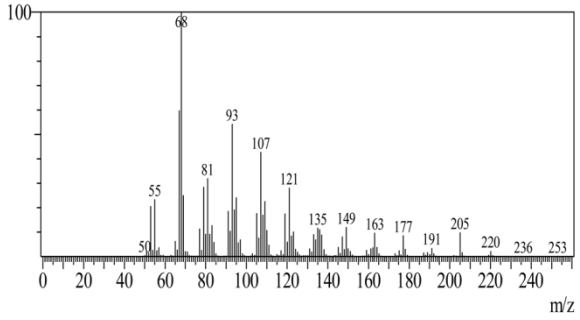
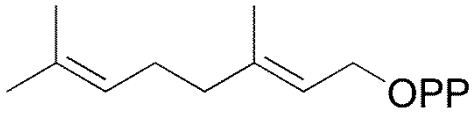
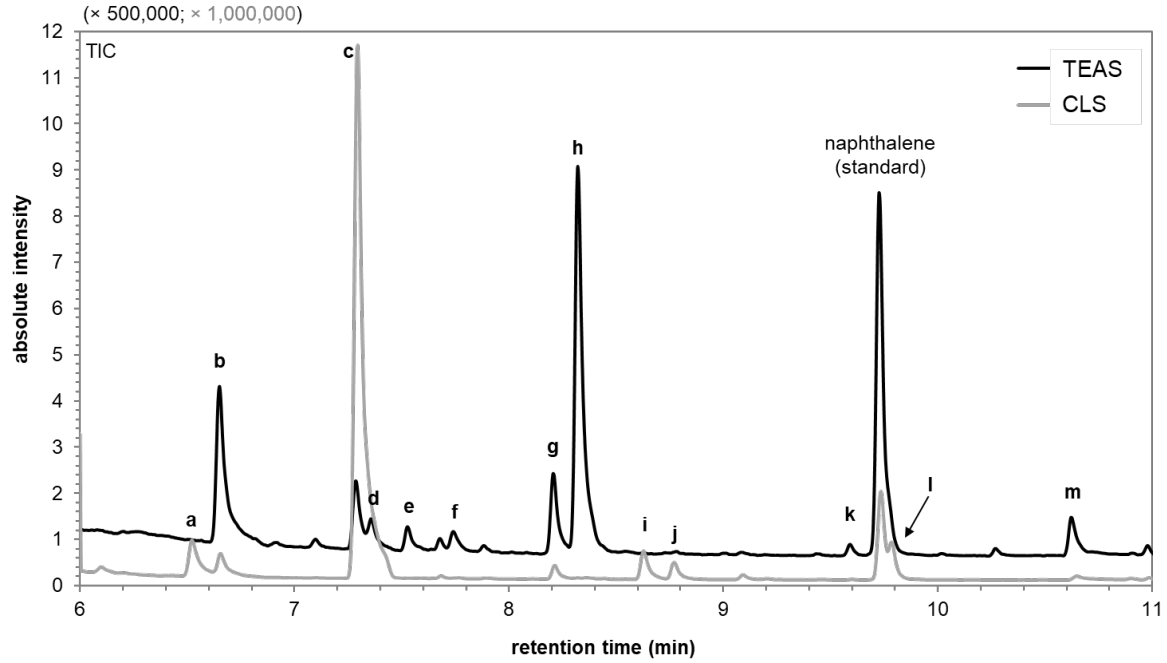


Table S7. GC/MS chromatogram of the single vial assays of CLS and TEAS with corresponding EI mass spectra for each product peak with substrate **GPP**. RT, retention time; RI, retention index.



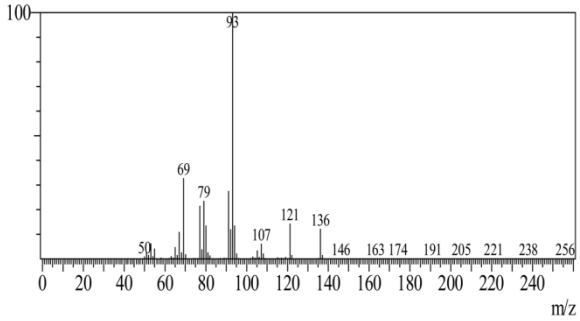
substrate **GPP**



product a (*β*-pinene)

RT: 6.529 min

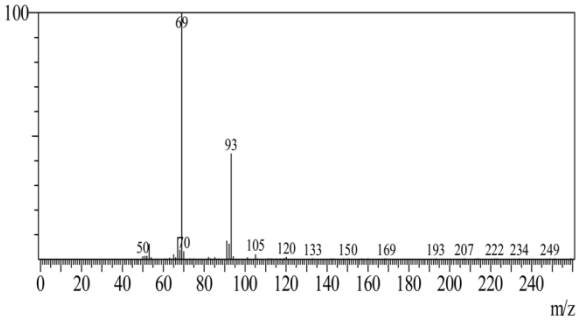
RI: 921



product b (*β*-myrcene)

RT: 6.655 min

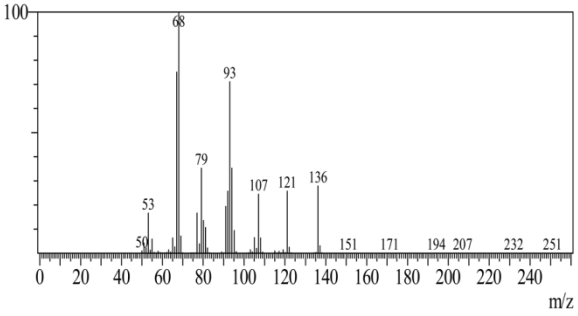
RI: 929



product c ((-)-limonene)

RT: 7.294 min

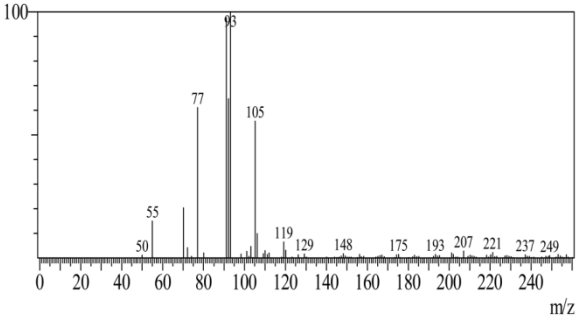
RI: 967



product d

RT: 7.357 min

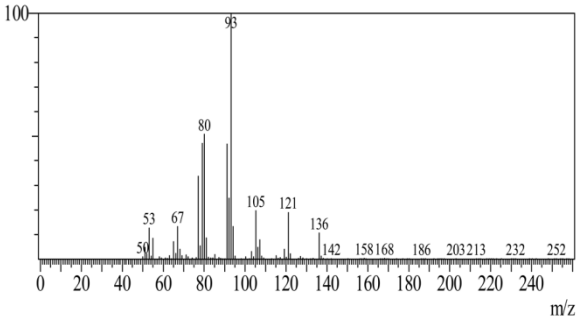
RI: 970



product e (*β*-ocimene)

RT: 7.527 min

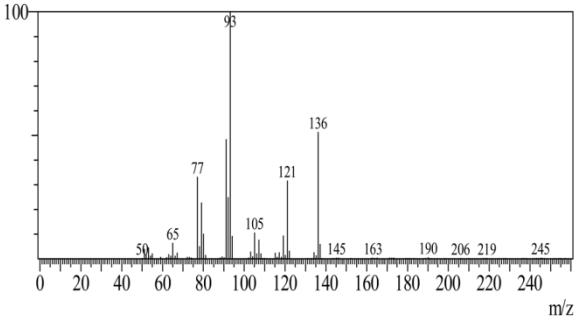
RI: 981



product f (*γ*-terpinene)

RT: 7.740 min

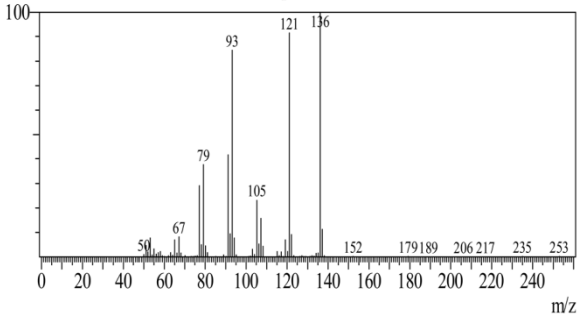
RI: 993



product g ((+)-4-carene)

RT: 8.207 min

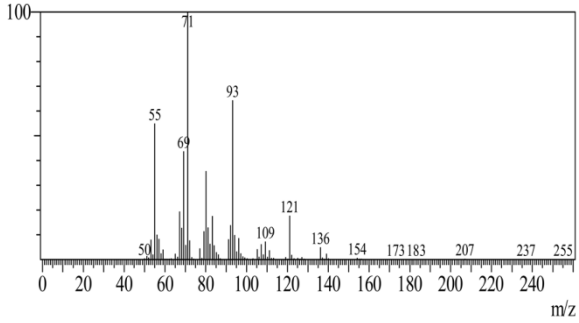
RI: 1022



product h (linalool)

RT: 8.321 min

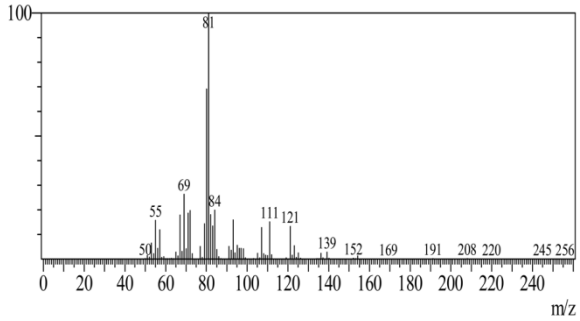
RI: 1029



product i (fenchol)

RT: 8.627 min

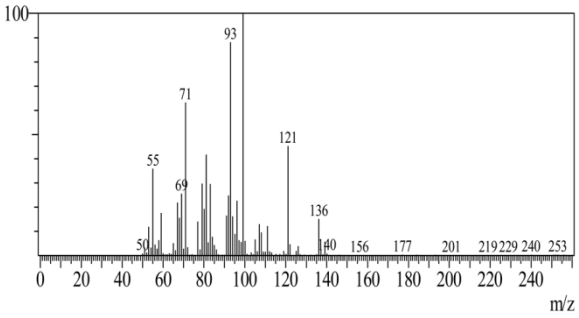
RI: 1049



product j ((E)-2-pinanol)

RT: 8.770 min

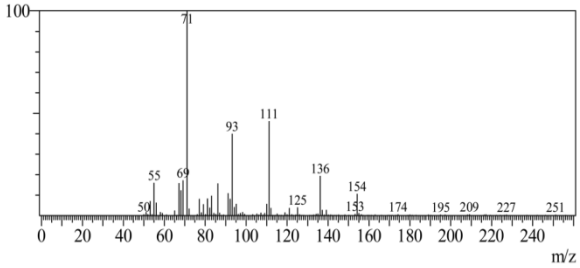
RI: 1058



product k (4-terpineol)

RT: 9.590 min

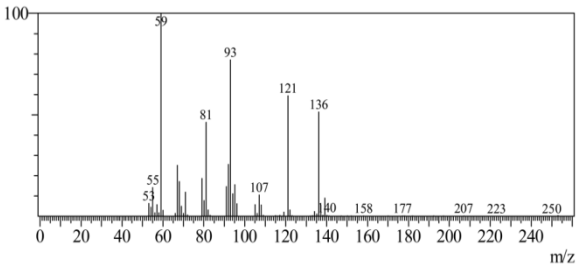
RI: 1109



product l (*α*-terpineol)

RT: 9.783 min

RI: 1122



product m (geraniol)

RT: 10.622 min

RI: 1176

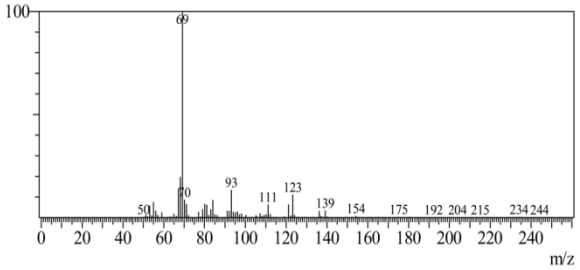
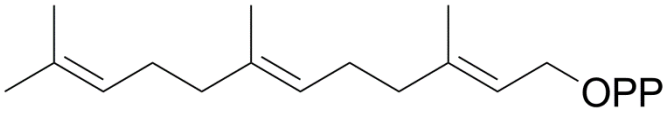
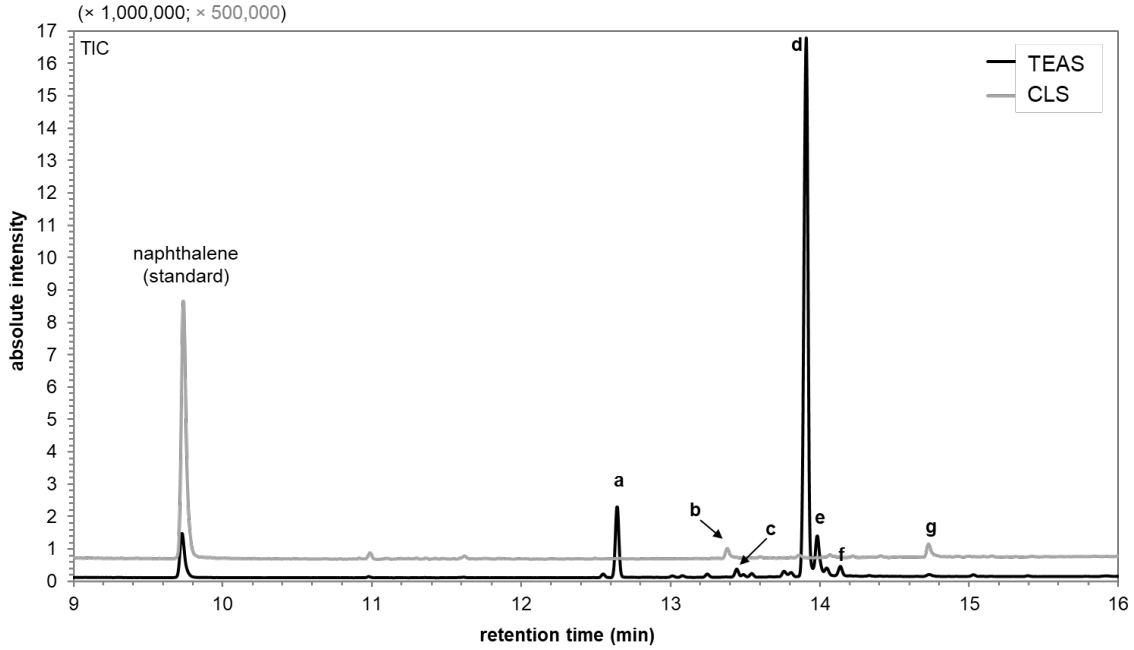


Table S8. GC/MS chromatogram of the single vial assays of CLS and TEAS with corresponding EI mass spectra for each product peak with substrate **FPP**. RT, retention time; RI, retention index.



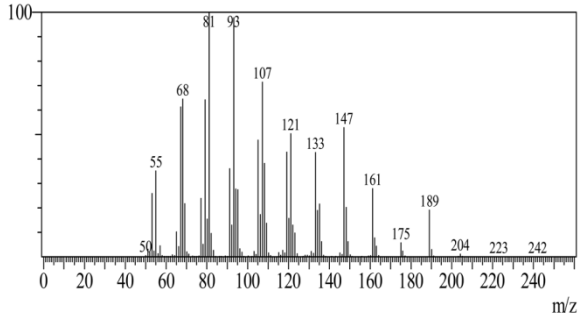
substrate **FPP**



product a

RT: 12.642 min

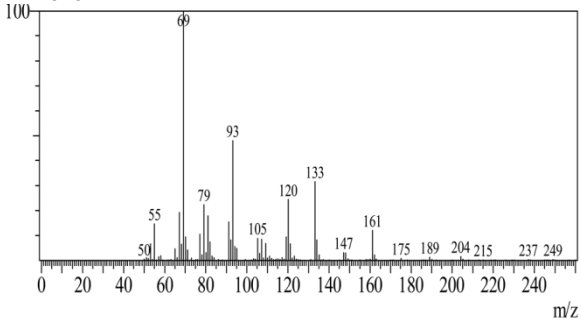
RI: 1316



product b (farnesene)

RT: 13.380 min

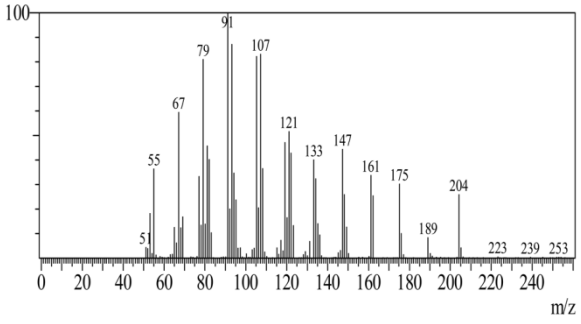
RI: 1370



product c (*β*-selinene)

RT: 13.444 min

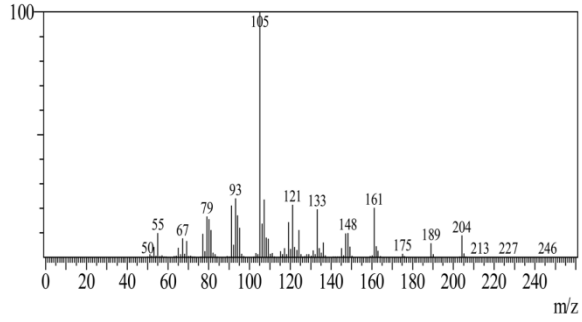
RI: 1375



product d (5-*epi*-aristolochene)

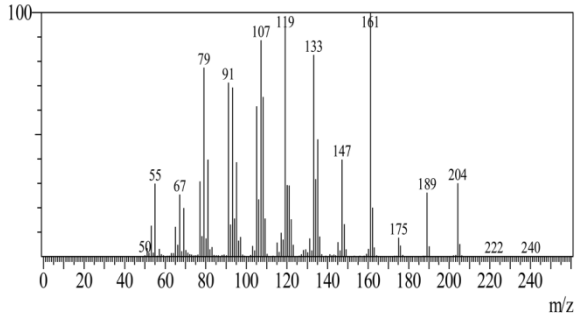
RT: 13.909 min

RI: 1410



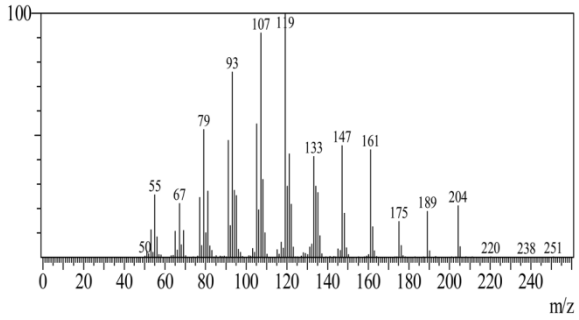
product e

RT: 13.983 min
RI: 1416



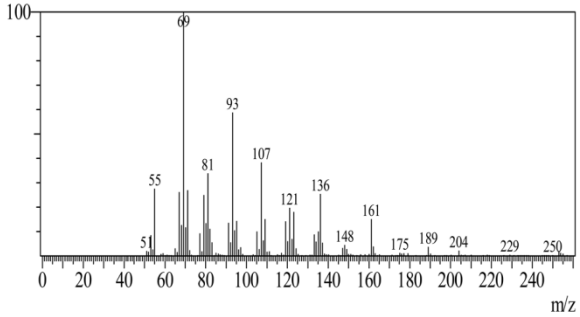
product f

RT: 14.138 min
RI: 1428



product g ((*E*)-farnesol)

RT: 14.728 min
RI: 1474



Determination of the volume of active site pocket

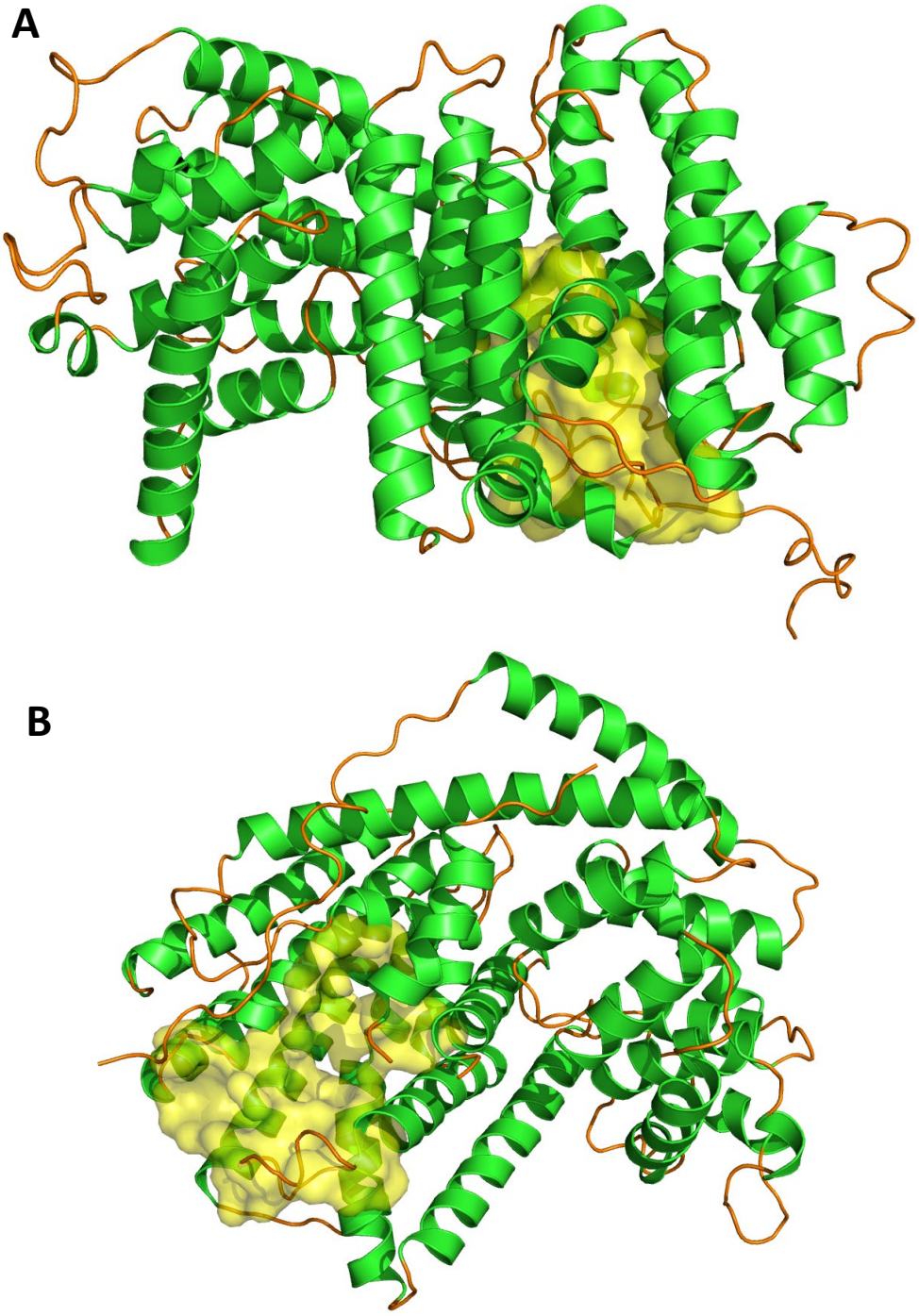


Figure S5: Active site pocket in CLS (**A**) and TEAS (**B**); Fpocket tool of HotSpot Wizard 3.0[1] was used for the calculation of the volume of the active site pocket. Calculated binding pocket (in yellow) in CLS (1076 Å^3) is smaller than TEAS (1940 Å^3).

Determination of the dimensions of the access tunnel

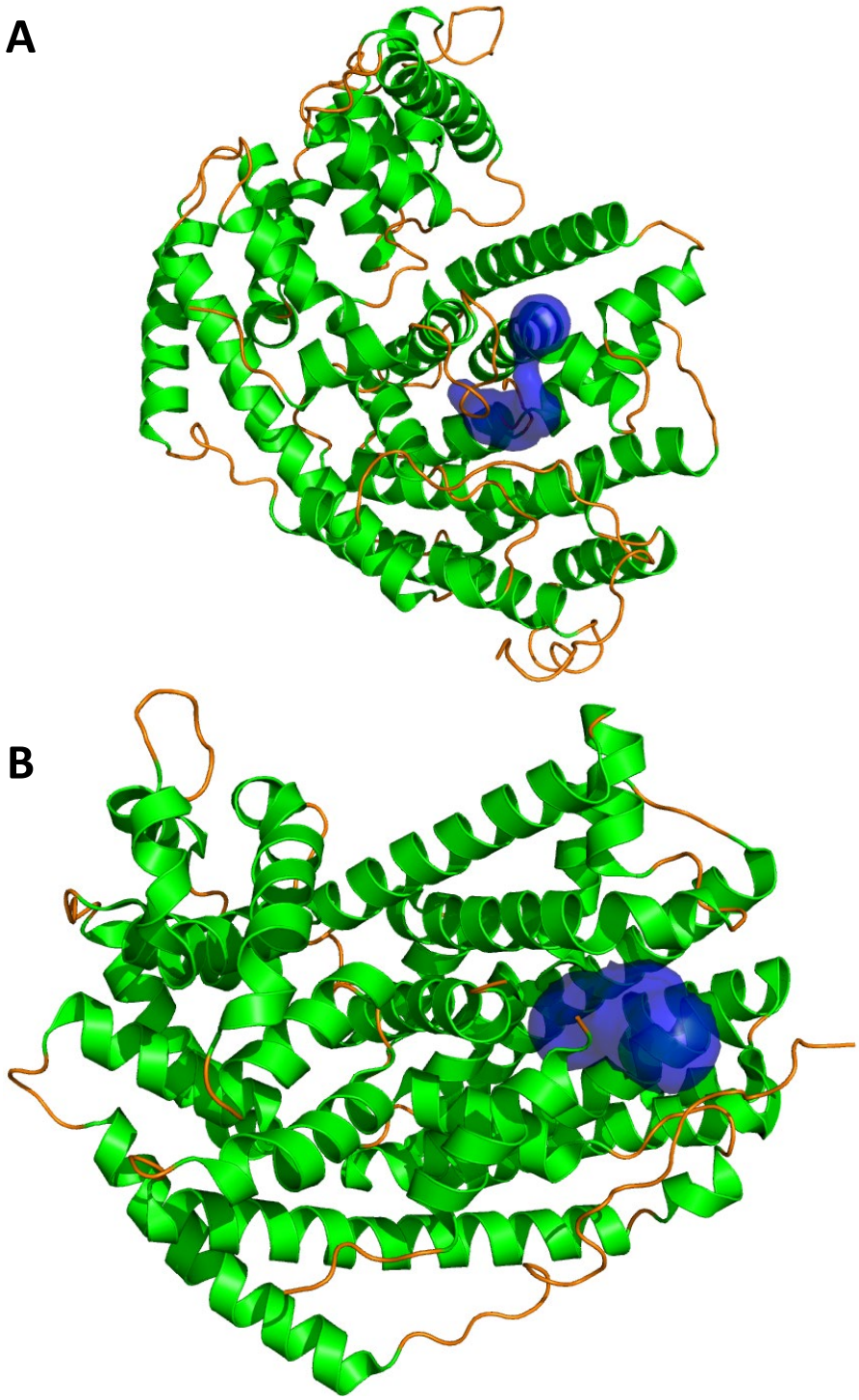


Figure S6: Substrate access tunnel in CLS (**A**) and TEAS (**B**); MoleOnline[2] was used to identify and visualize the substrate access tunnels (in blue). Comparison shows that CLS has a longer but narrower substrate-binding pocket (bottleneck radius = 1.3 Å, length = 33.5 Å) compared with TEAS (bottleneck radius = 3.2 Å, length = 18.8 Å

Flexibility analyses of CLS and TEAS protein structures

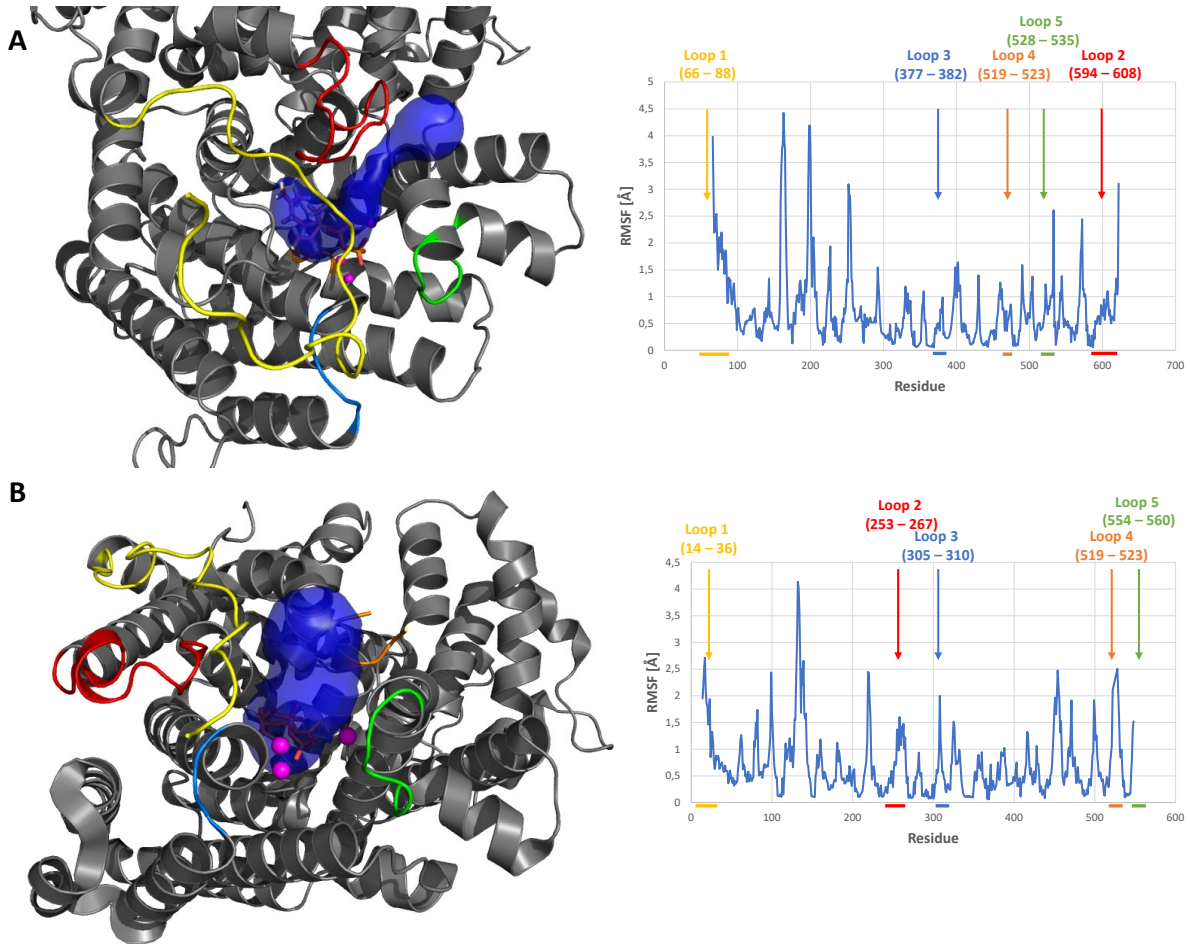


Figure S7: Substrate-binding loops (1-5) and the residue fluctuation profile for amino acid residues (RMSF) in CLS (**A**) and TEAS (**B**). Flexibility analysis of protein structures was performed by CABS-flex 2.0[3] web server. Loops (1-5) forming the substrate entrance tunnels are marked with the same color in the RMSF plot as in the protein structure.

Comparison between Carbocation Docking of GPP and substrate 6 in CLS

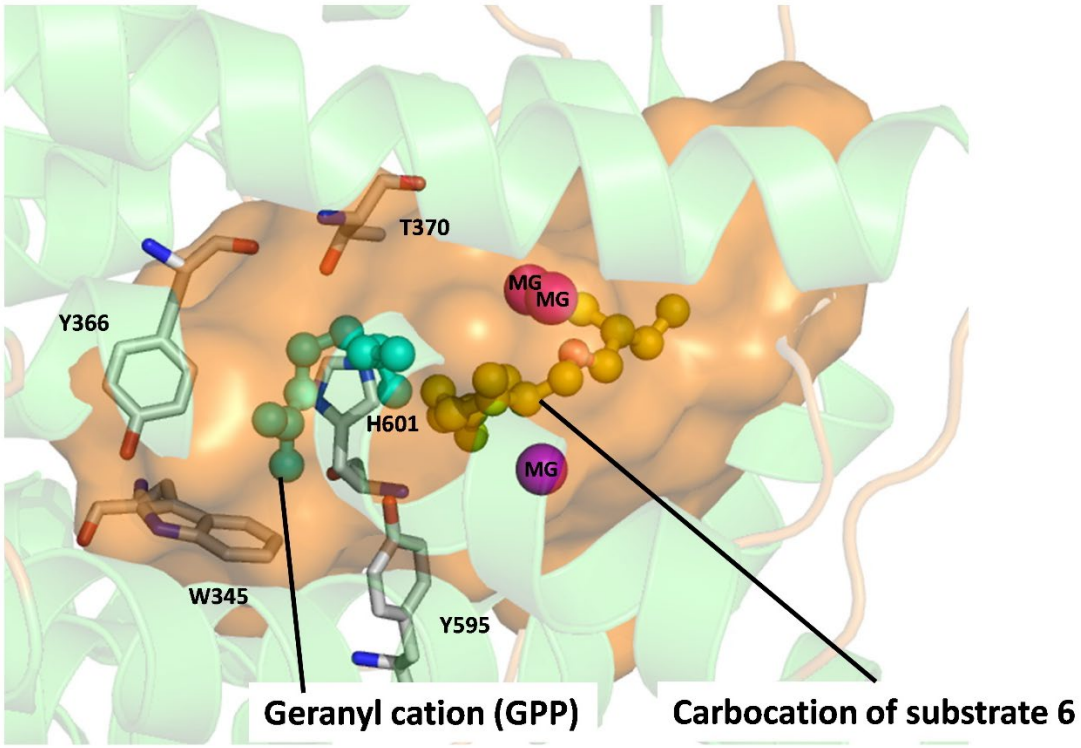


Figure S8: Representation of docking positions of geranyl cation and carbocation of substrate **6** in the active site pocket of limonene synthase from *Cannabis sativa* (CLS). The active site pocket was visualized as a surface in orange, complexed Mg^{2+} ions as magenta spheres. Residues involved in the migration and stabilization of the carbocation intermediate are highlighted as grey stick representations and labeled accordingly. Docked conformations of the geranyl cation and the carbocation of substrate **6** are shown as stick and ball in cyan and yellow, respectively (with the oxygen of substrate **6** bound to Mg^{2+} complex highlighted in red)

Simplified GPP, FPP, and substrates 5, and 6 binding in CLS and TEAS

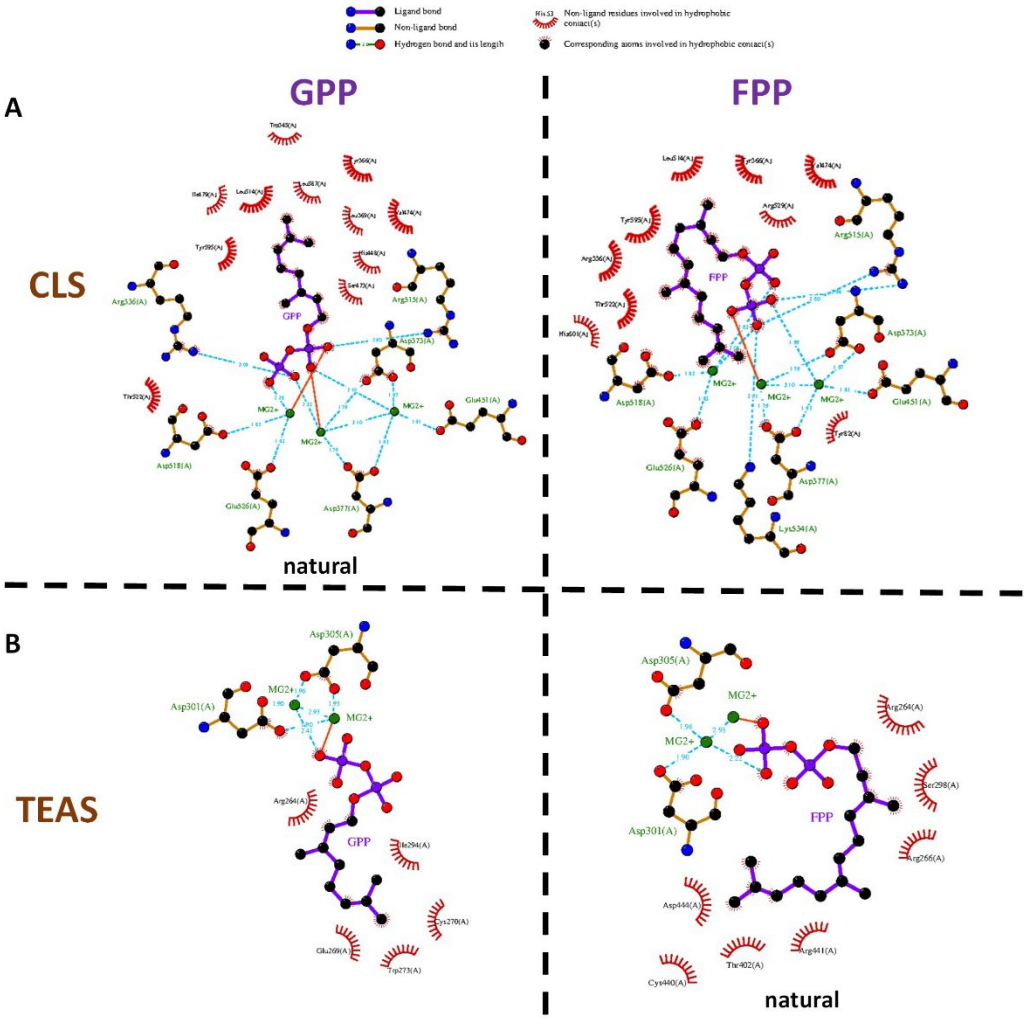


Figure S9: A 2D view of GPP and FPP substrate binding in the CLS (**A**) and TEAS (**B**) active site pockets. Molecular docking of substrates was used to identify productive binding poses. LigPlot+[4] was used for the visualization of 2D enzyme-substrate interactions. Key interactions between substrates and Mg^{2+} and amino acids in the binding pocket are shown.

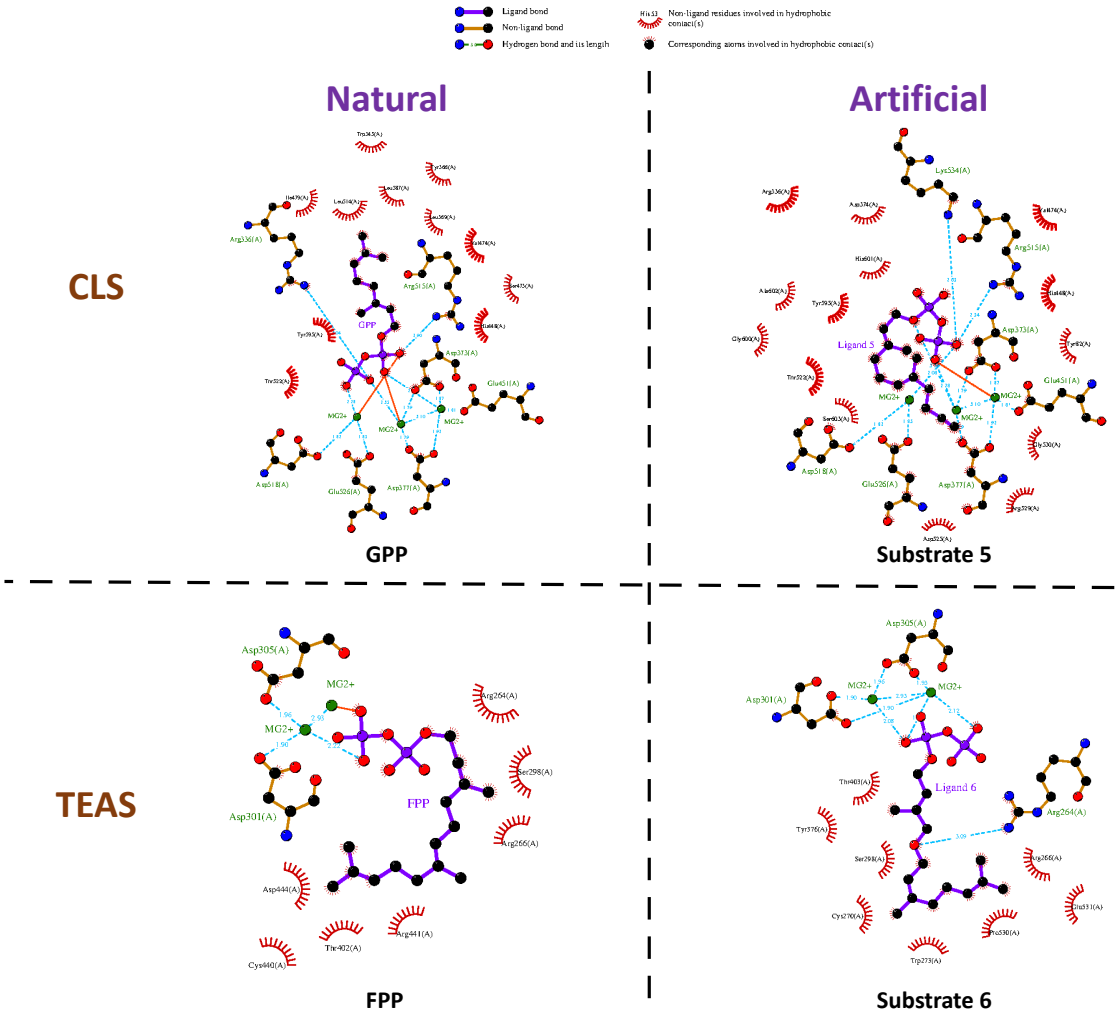


Figure S10: A 2D view of substrates **5** and **6** binding in the active site pocket of CLS and TEAS, respectively. Molecular docking of the substrate was used to identify productive binding poses. LigPlot+[4] was used for the visualization of 2D enzyme-substrate interactions. Key interactions between substrates and Mg^{2+} and amino acids in the binding pocket are shown.

Carbocations form a helical conformation inside the active pocket of TEAS

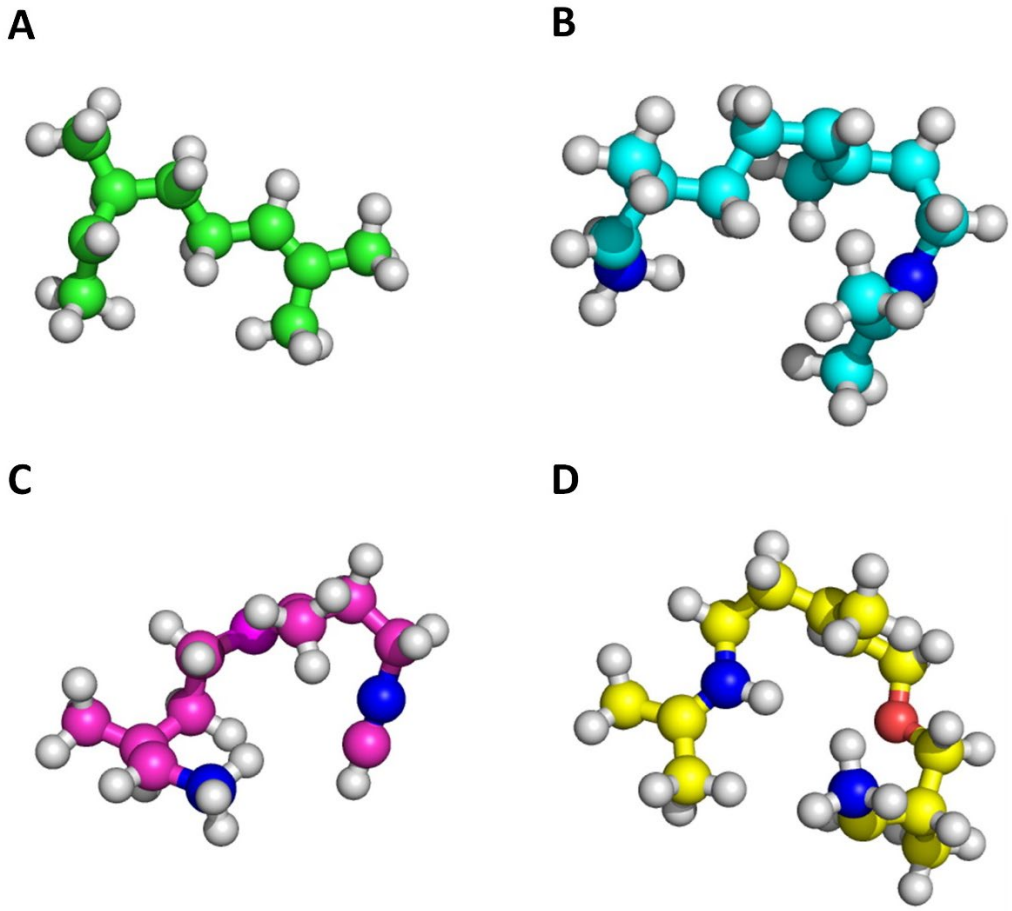


Figure S11: Resulting conformations of carbocations of GPP (**A**, green), FPP (**B**, cyan), substrate **5** (**C**, magenta), and substrate **6** (**D**, yellow) after docking in TEAS. C1 and C10 (if existing in the molecules) are highlighted in deep blue. Substrates FPP (B) and **6** (D) show a helical conformation but GPP (A) and substrate **5** (C) were (more) linear.

Docking of linalyl diphosphate and its analogues from FPP, and substrates 5 and 6

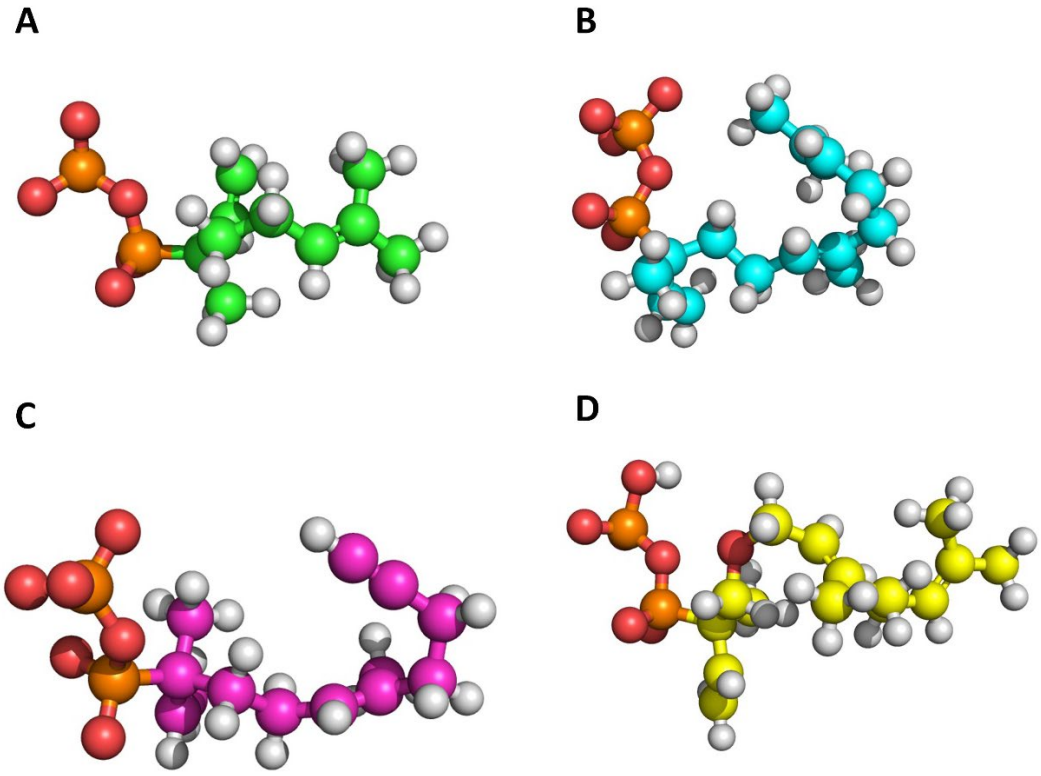


Figure S12: Resulting conformations of linalyl diphosphate or its analogs of GPP (**A**, green), FPP (**B**, cyan), substrate **5** (**C**, magenta), and substrate **6** (**D**, yellow) after docking in CLS.

Docking of the carbocations of FPP and substrate 6 in TEAS

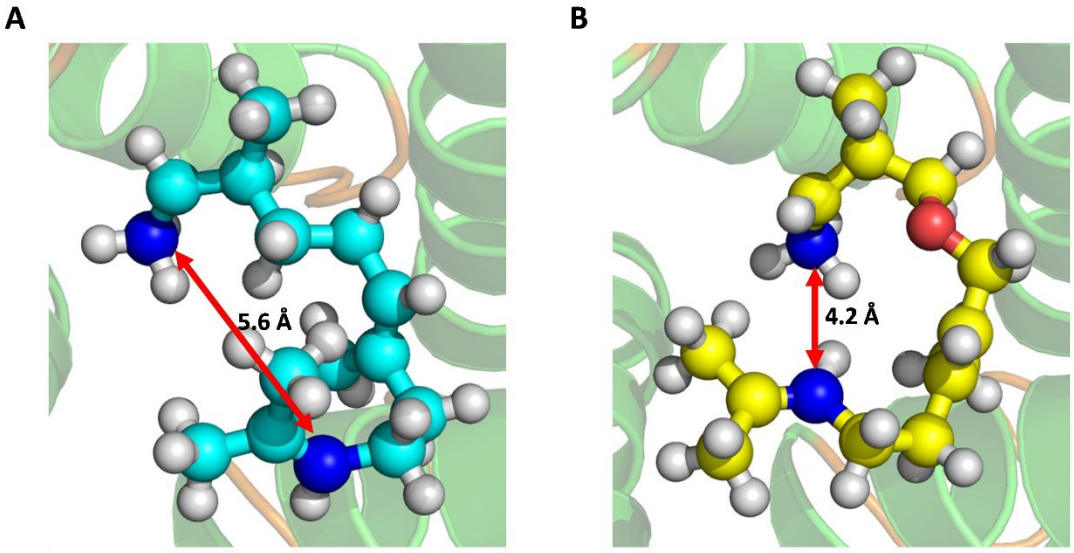


Figure S13: Comparison of the dockings of carbocations of FPP and substrate **6** in TEAS. The distance between C1 and C10 of resulting helical conformations of the carbocations from FPP (**A**; cyan) and substrate **6** (**B**; yellow) after docking in TEAS. The representative distance between C1 and C10 (carbons are highlighted in deep blue) is shown as a red arrow

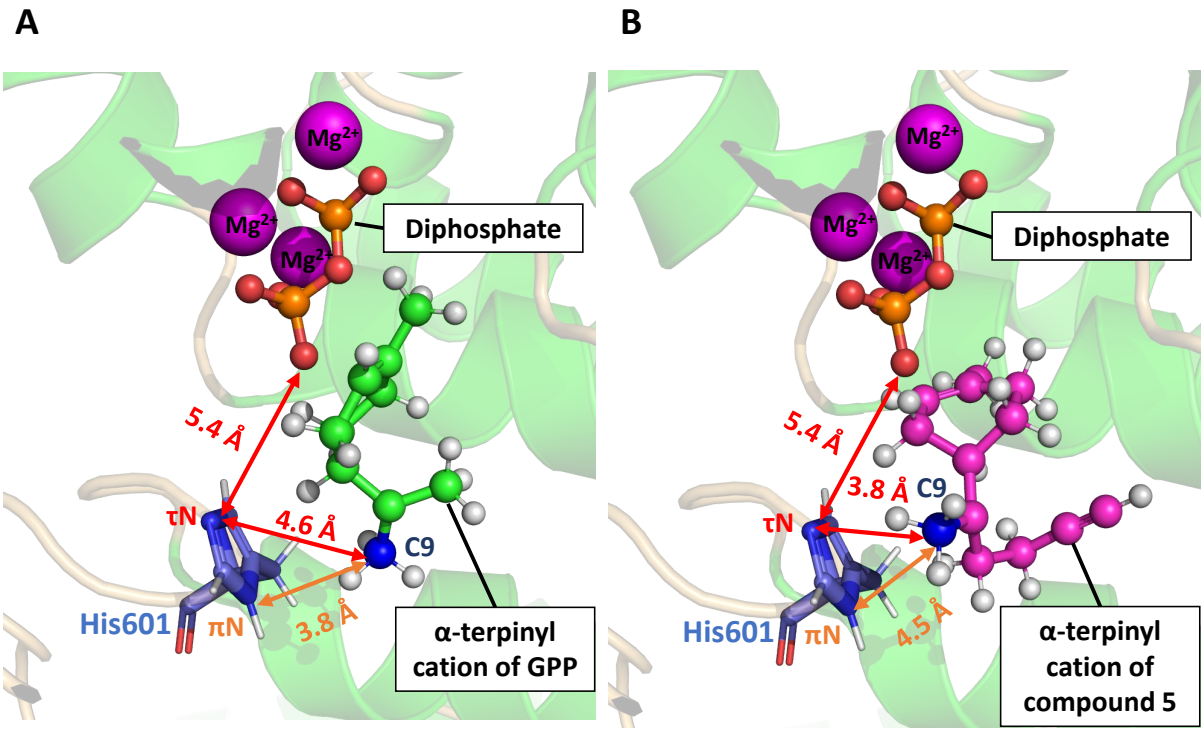
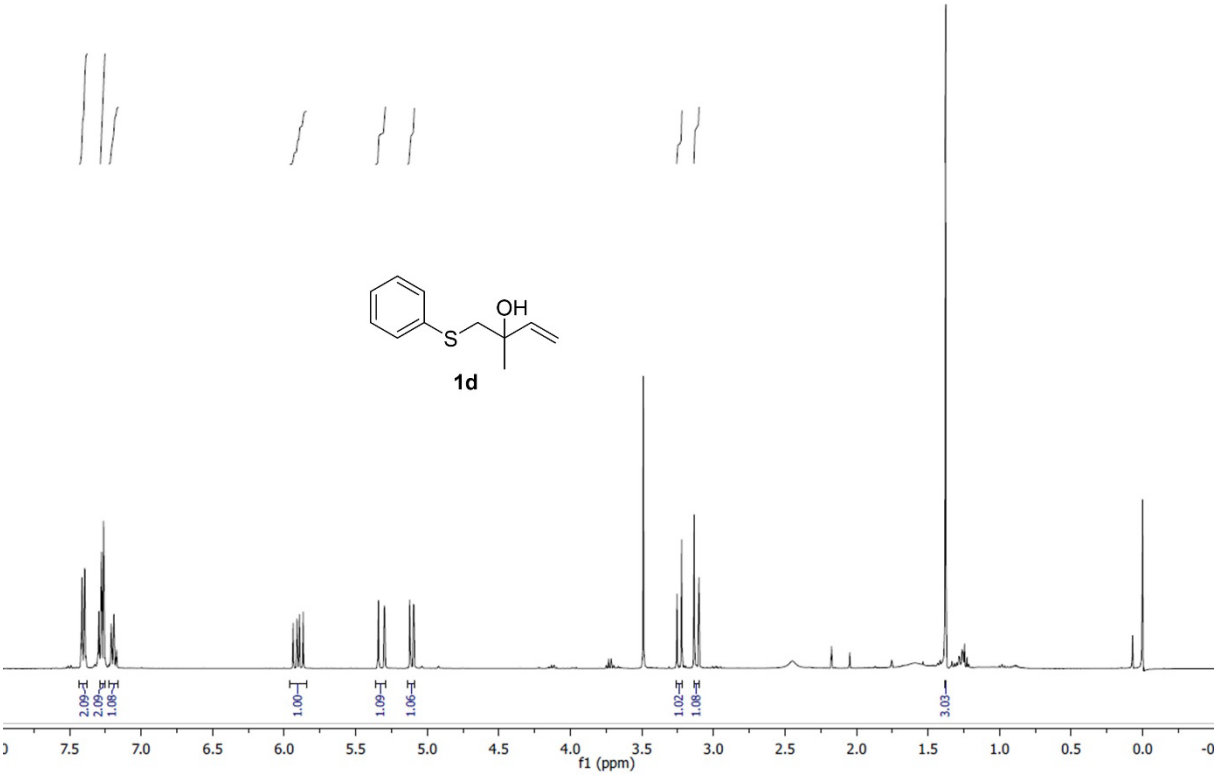
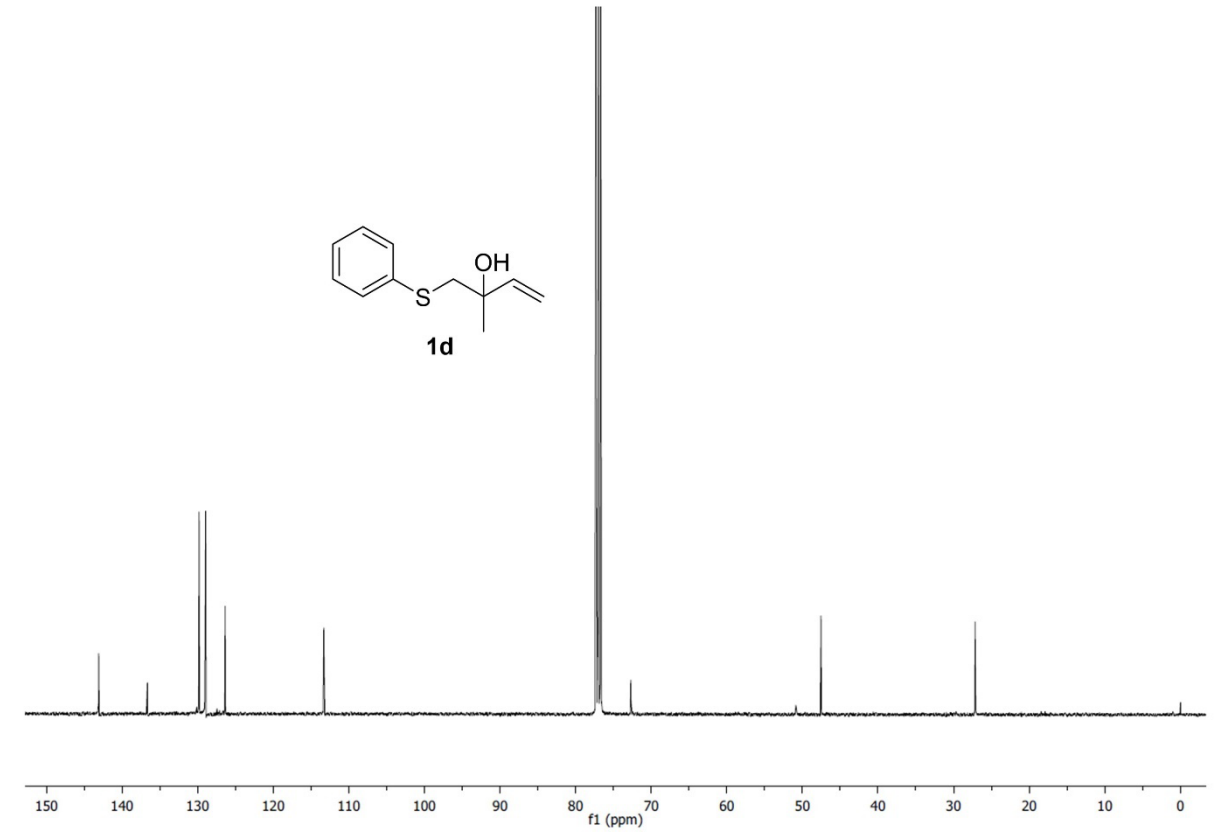


Figure S14: Docking pose of α-terpinyl cations of GPP (A; green) and compound **5** (B; magenta) into the active site of CLS. The distances between C9 of terpinyl cation of GPP and πN of His601 and τN of His601 are marked in red and orange, respectively. The distances between diphosphate and τN are shown in red.

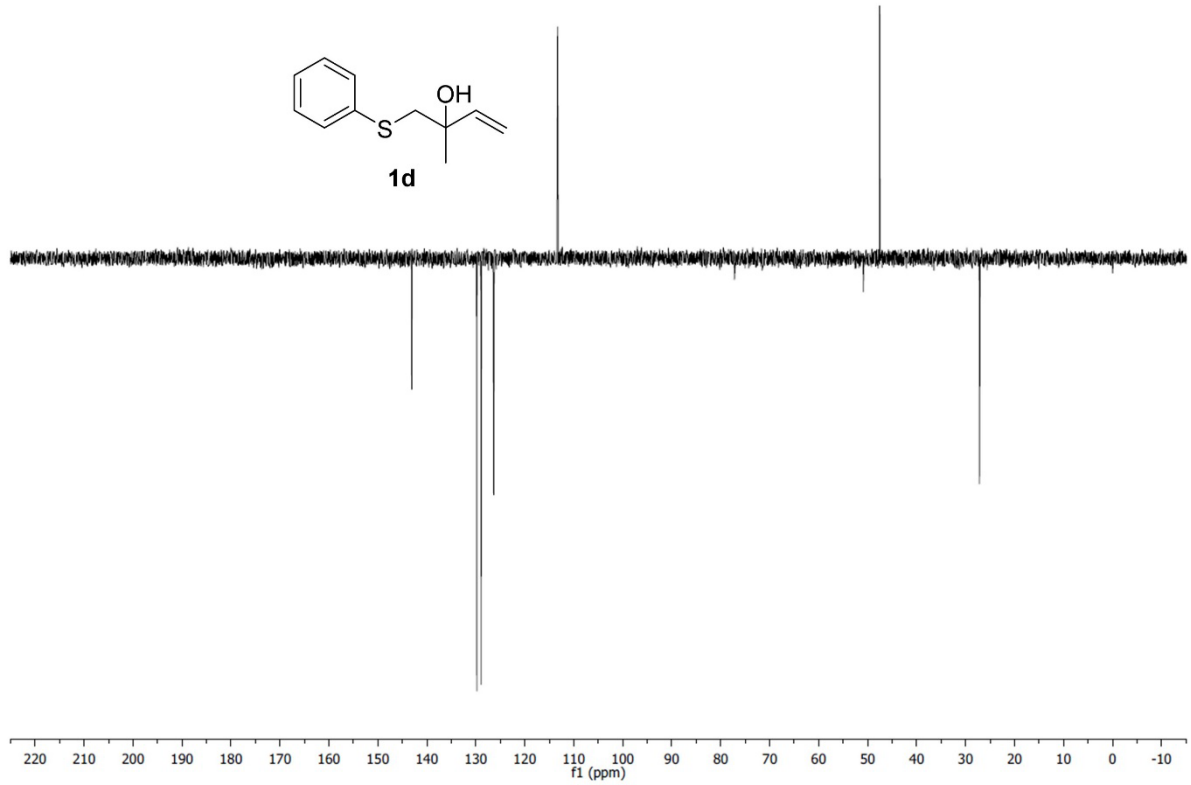
Spectral data of the enzymatic products



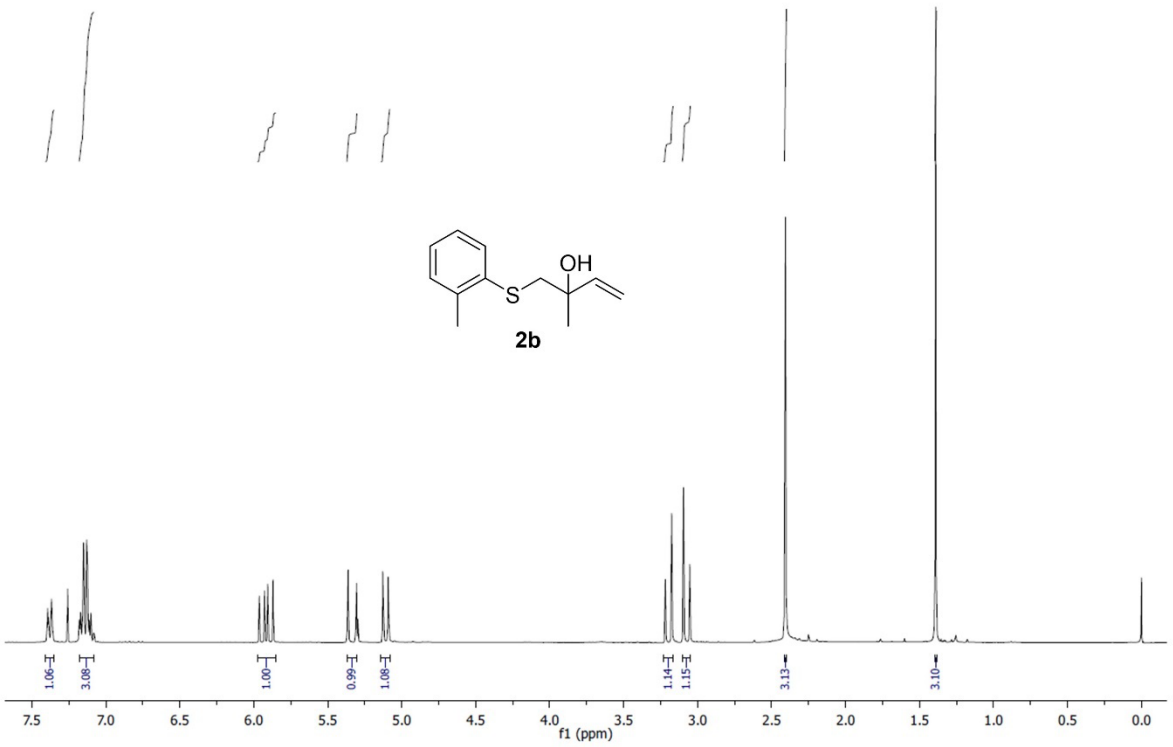
^{1}H NMR spectrum of product **1d**



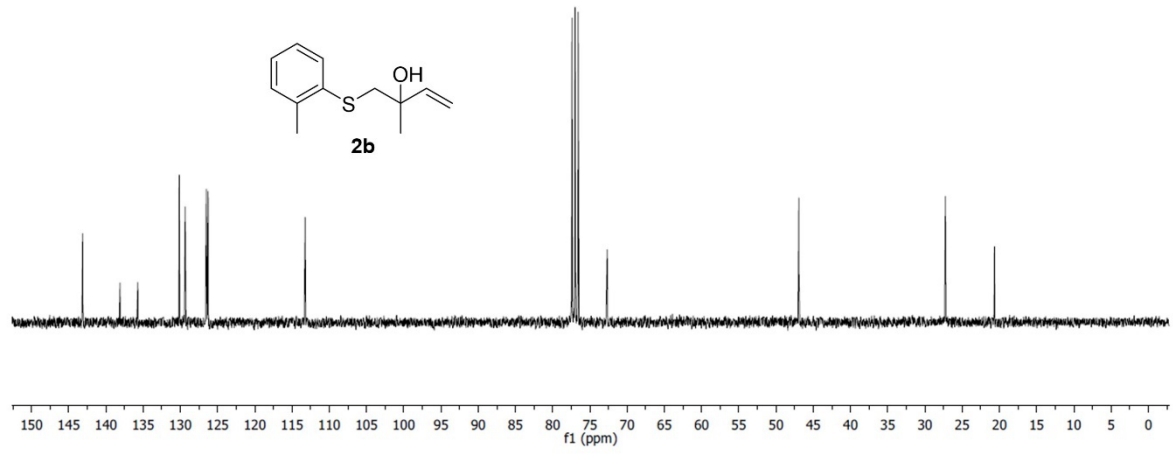
^{13}C NMR spectrum of product **1d**



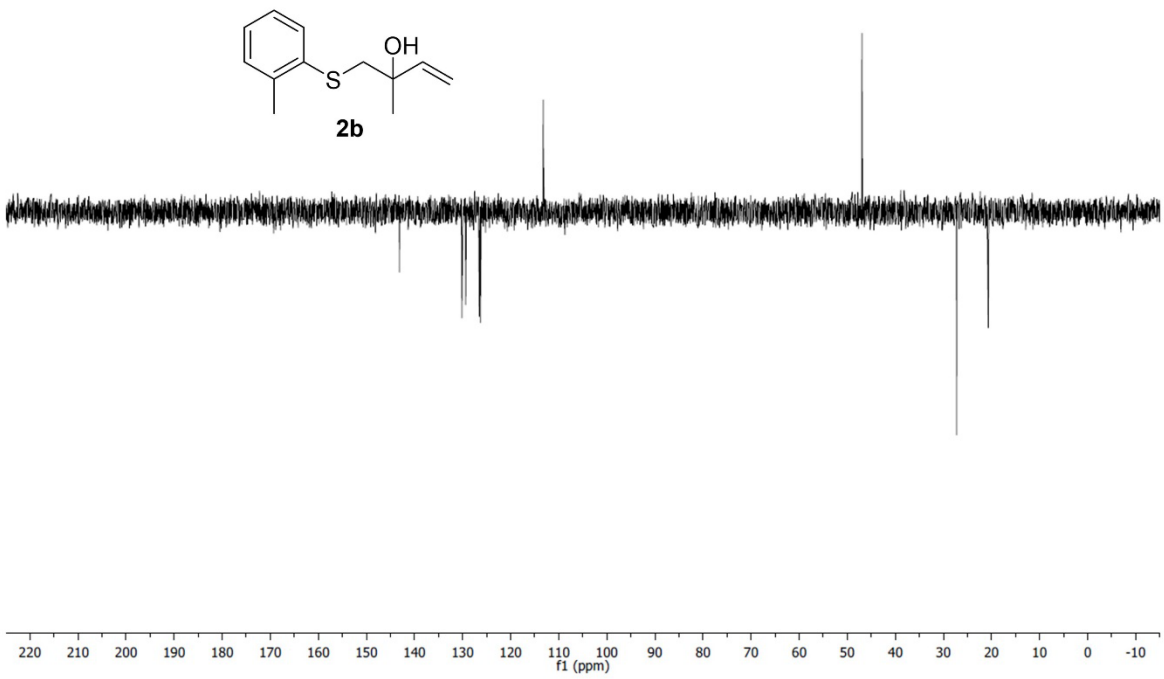
^{13}C NMR Dept spectrum of product **1d**



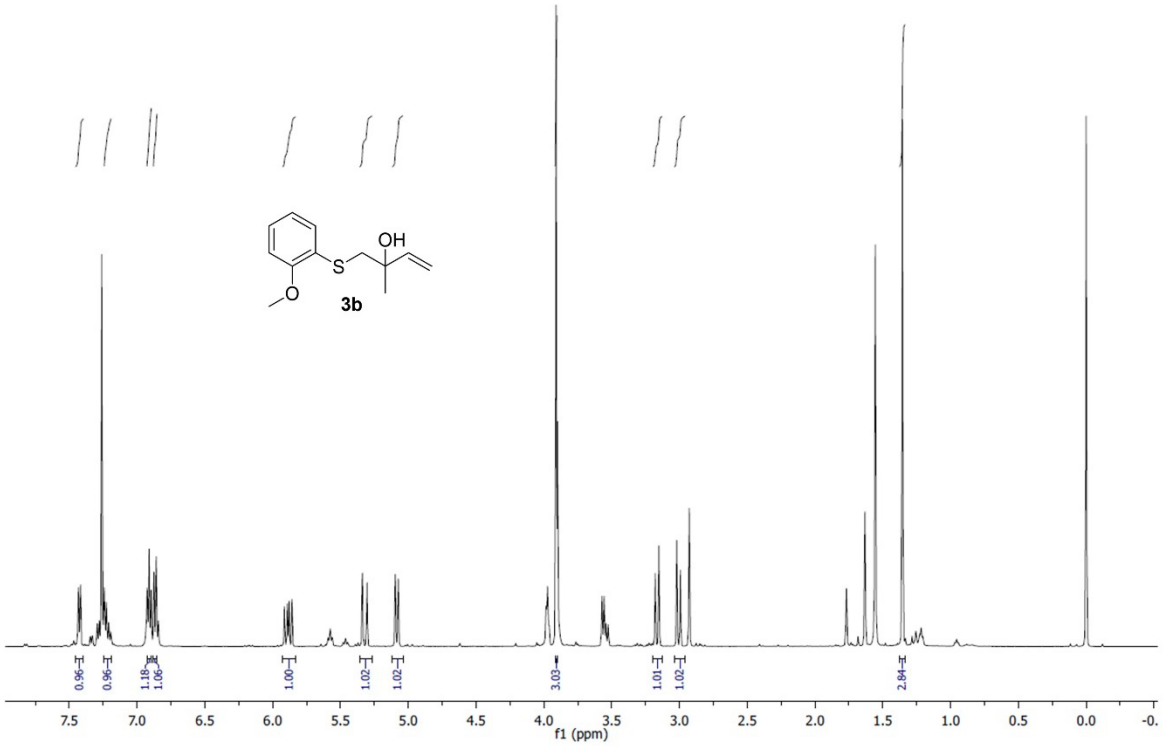
^{1}H NMR of the synthetic product **2b**



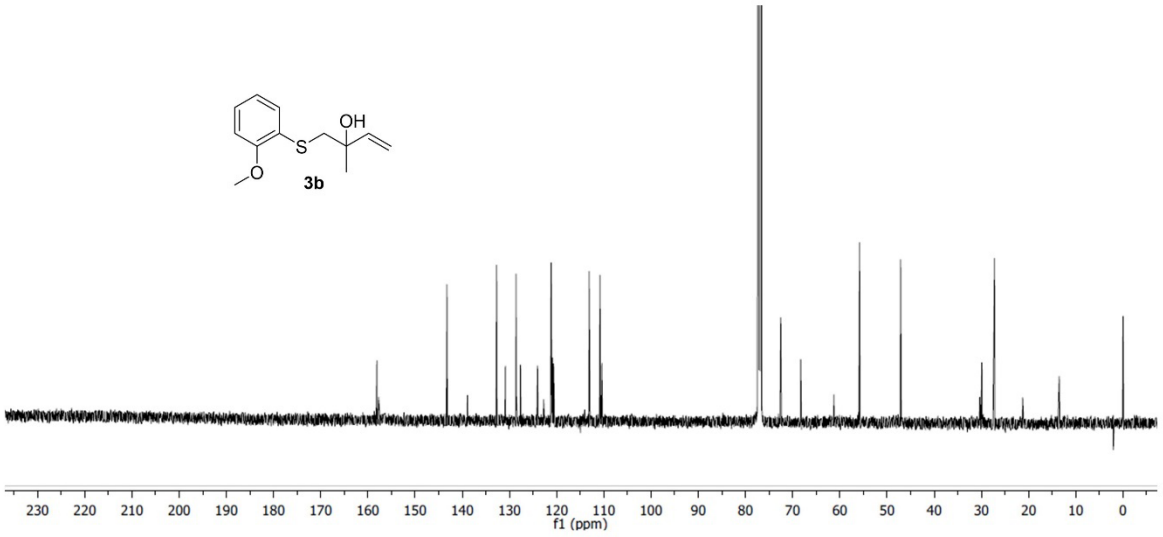
^{13}C NMR of the product **2b**



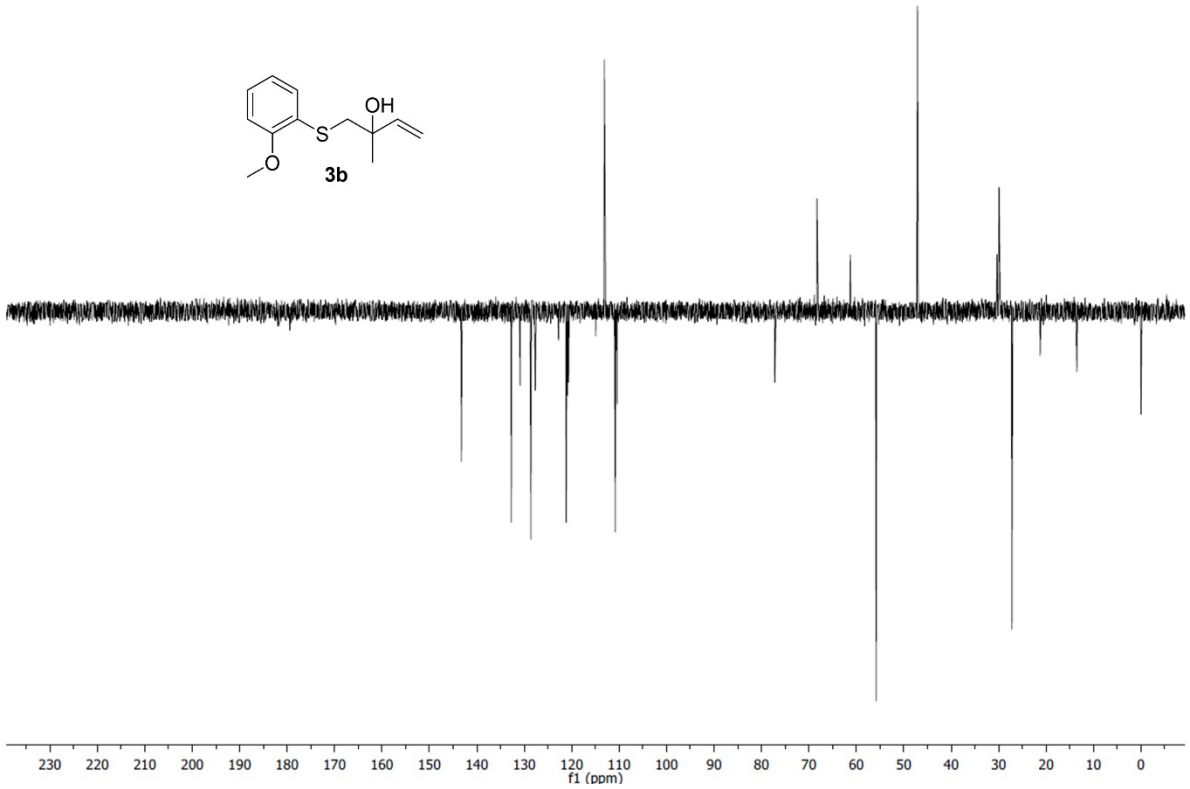
^{13}C NMR Dept spectrum of product **2b**



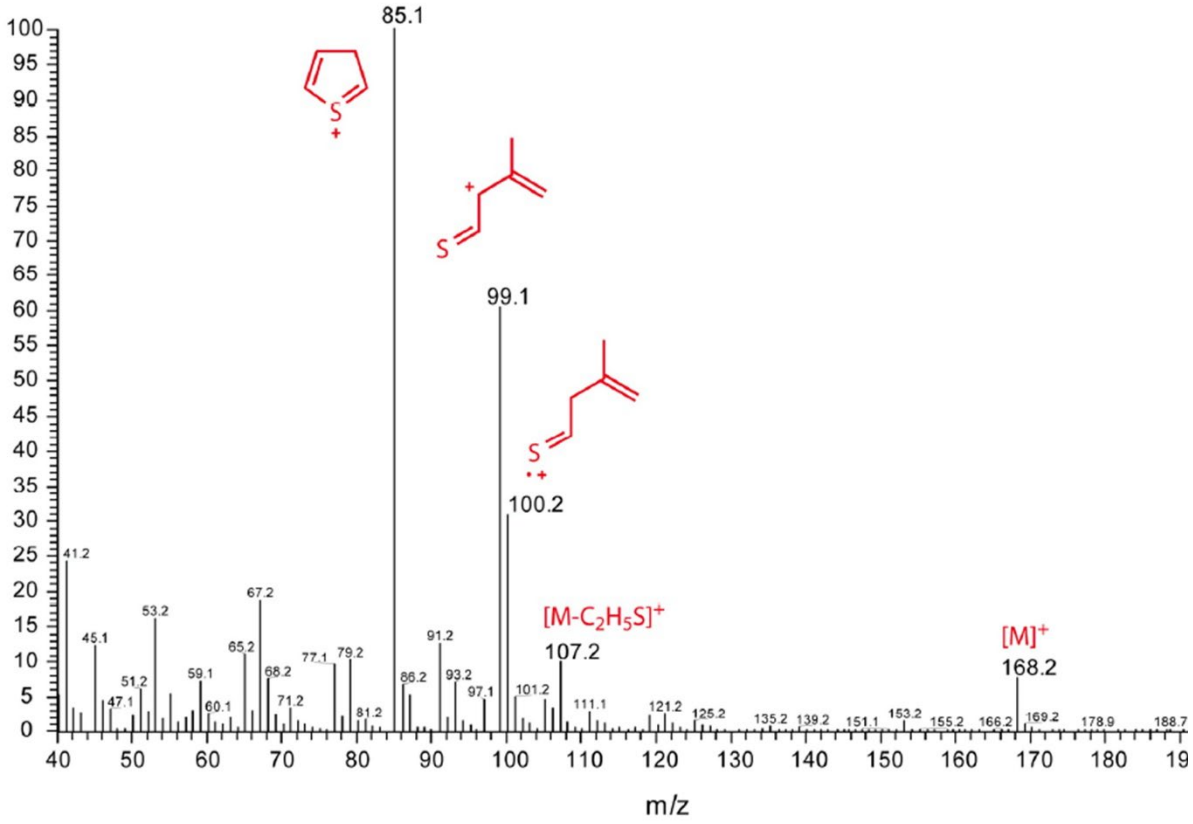
^{1}H NMR of product **3b**



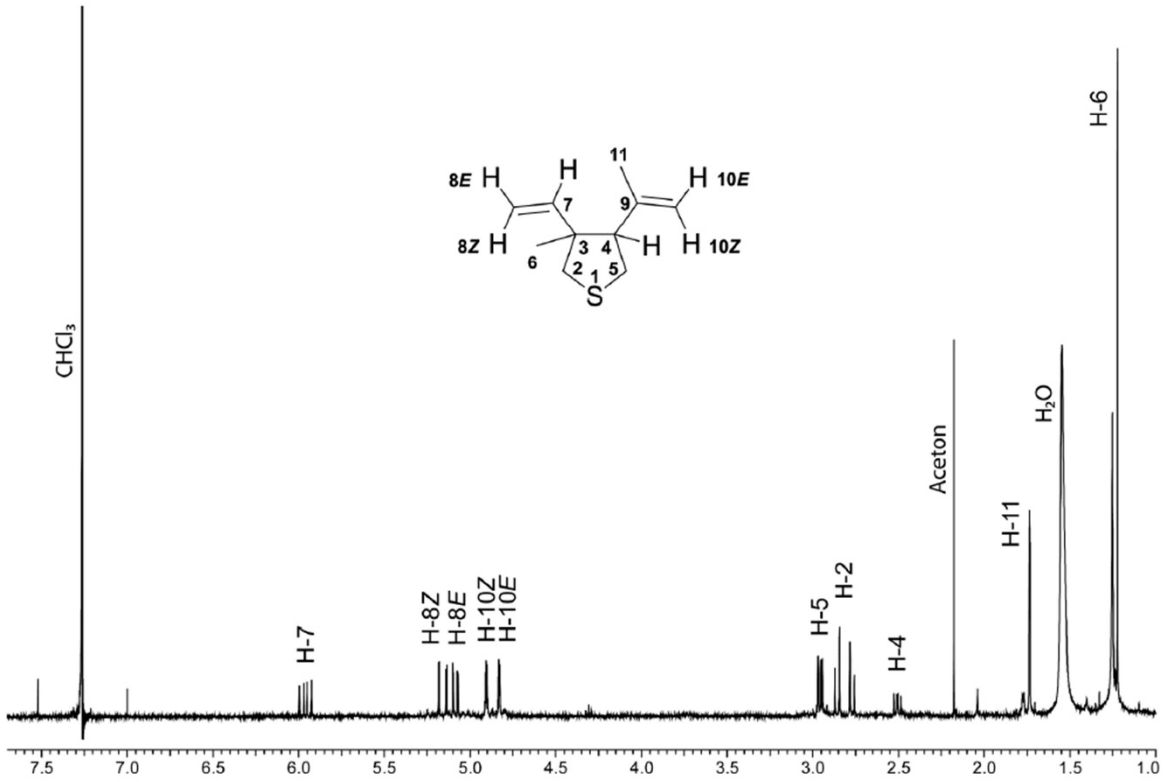
^{13}C NMR of the product **3b**



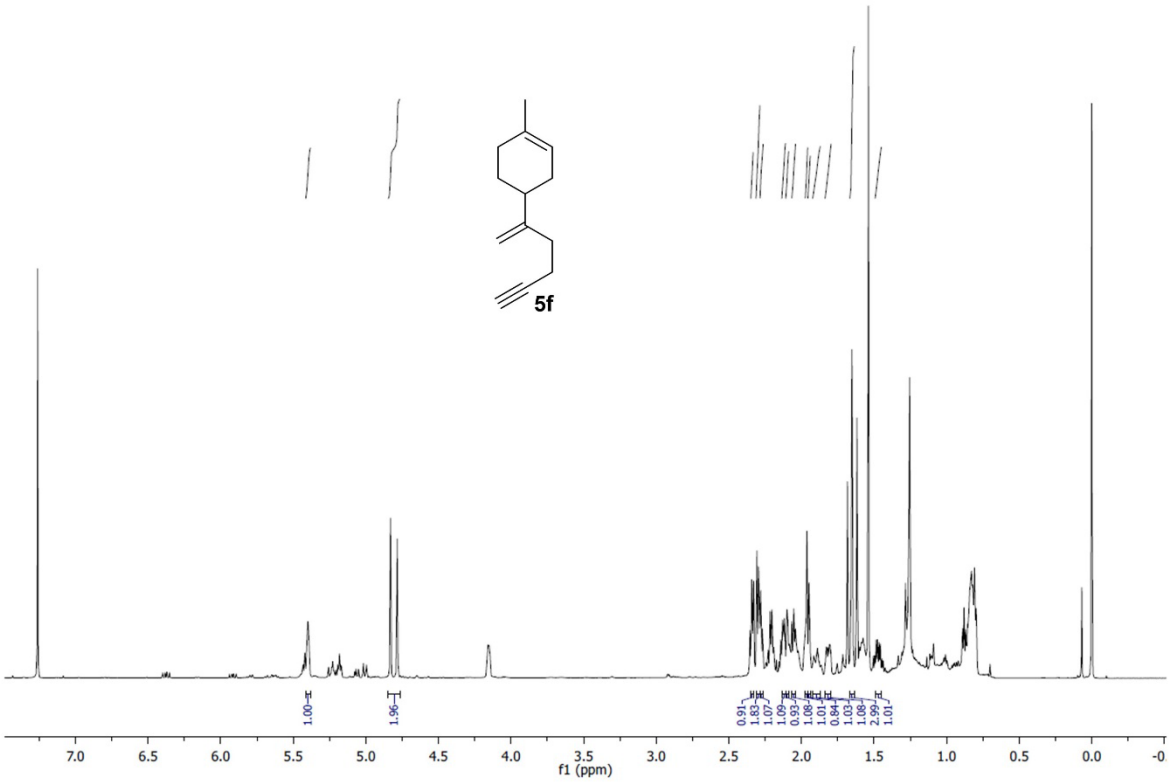
^{13}C NMR Dept spectrum of product **3b**



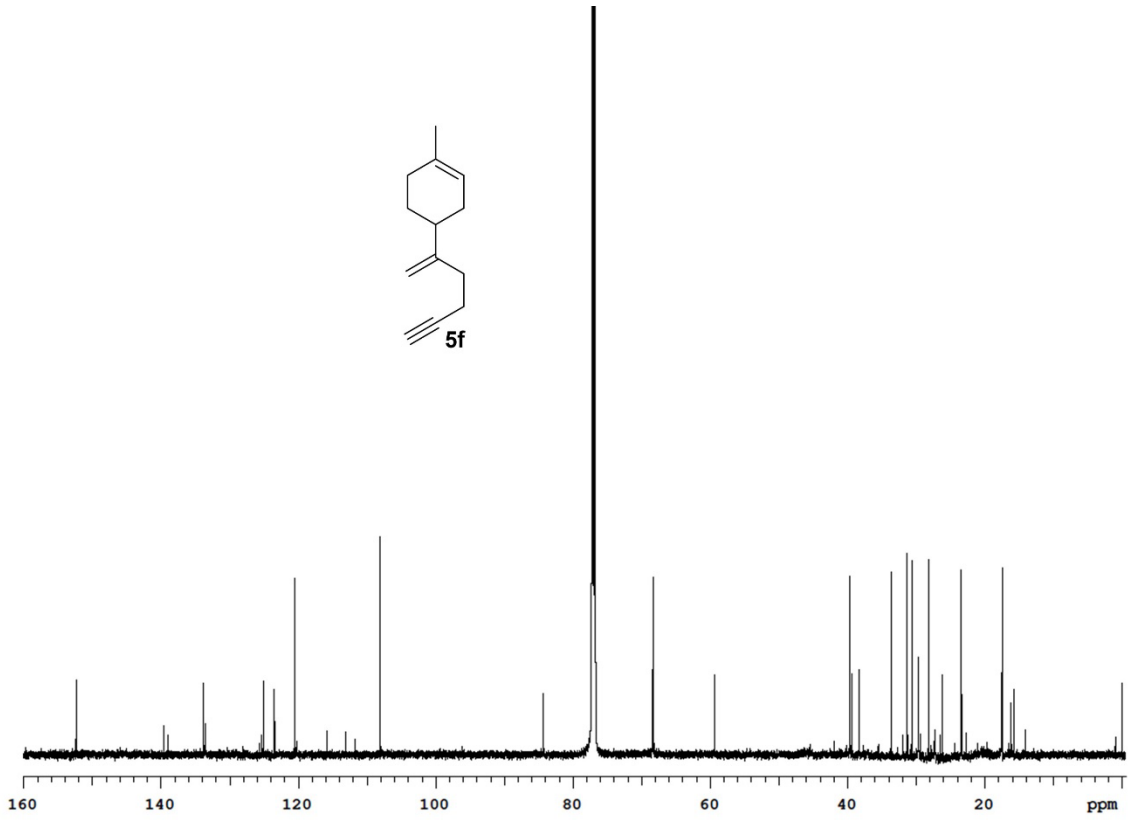
EI-MS of enzymatic product **4b**



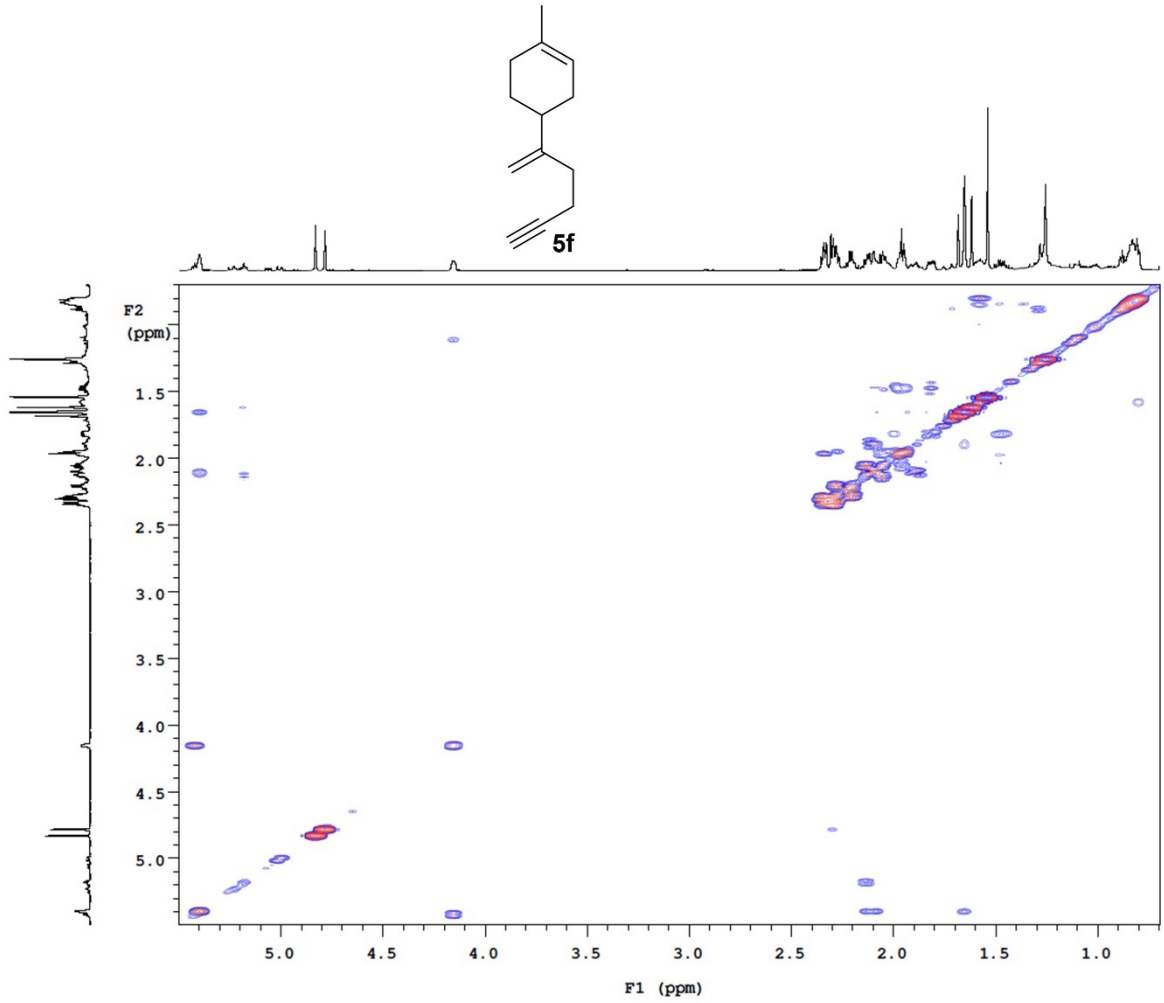
[1]H NMR of the enzymatic product **4b**



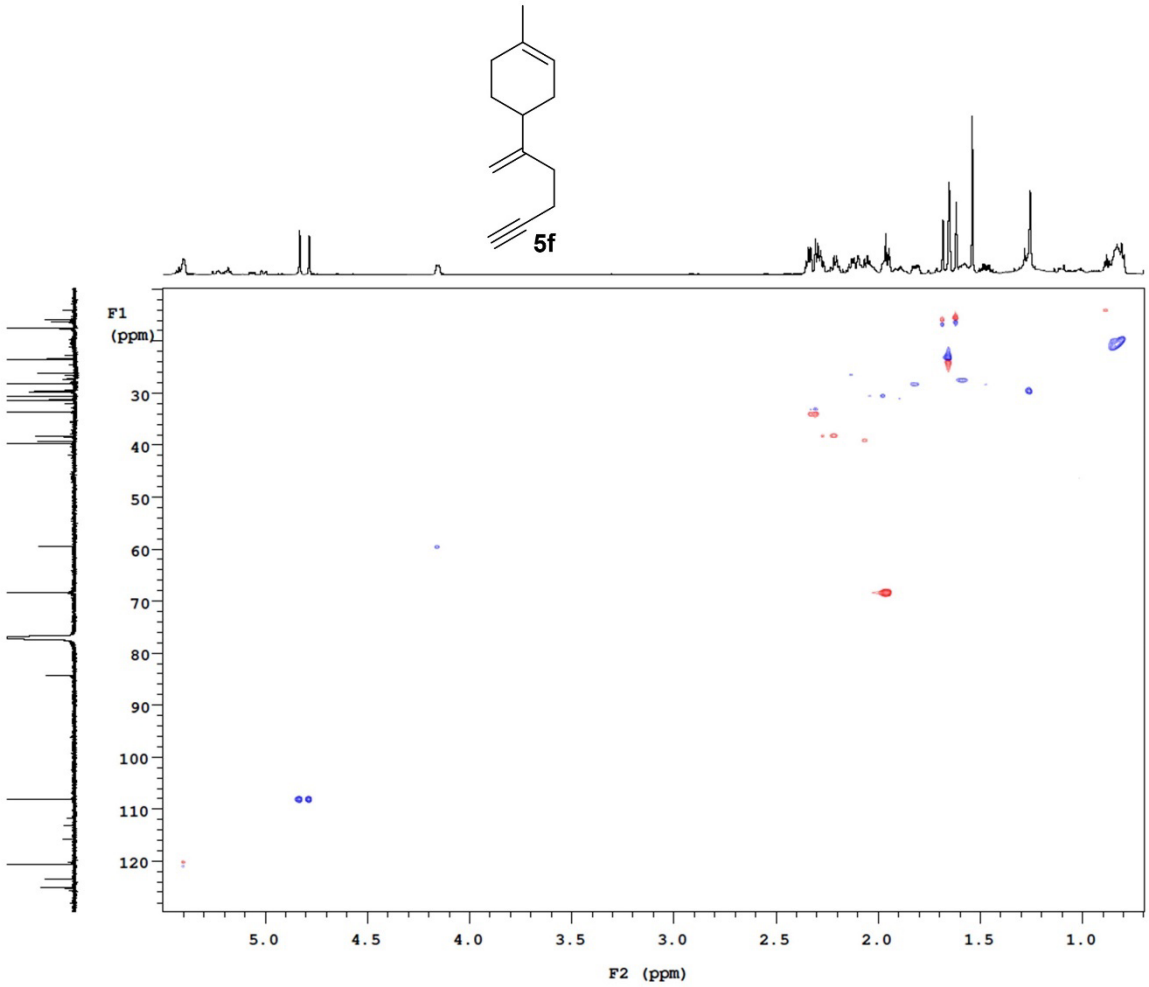
^{1}H NMR of the enzymatic product **5f**



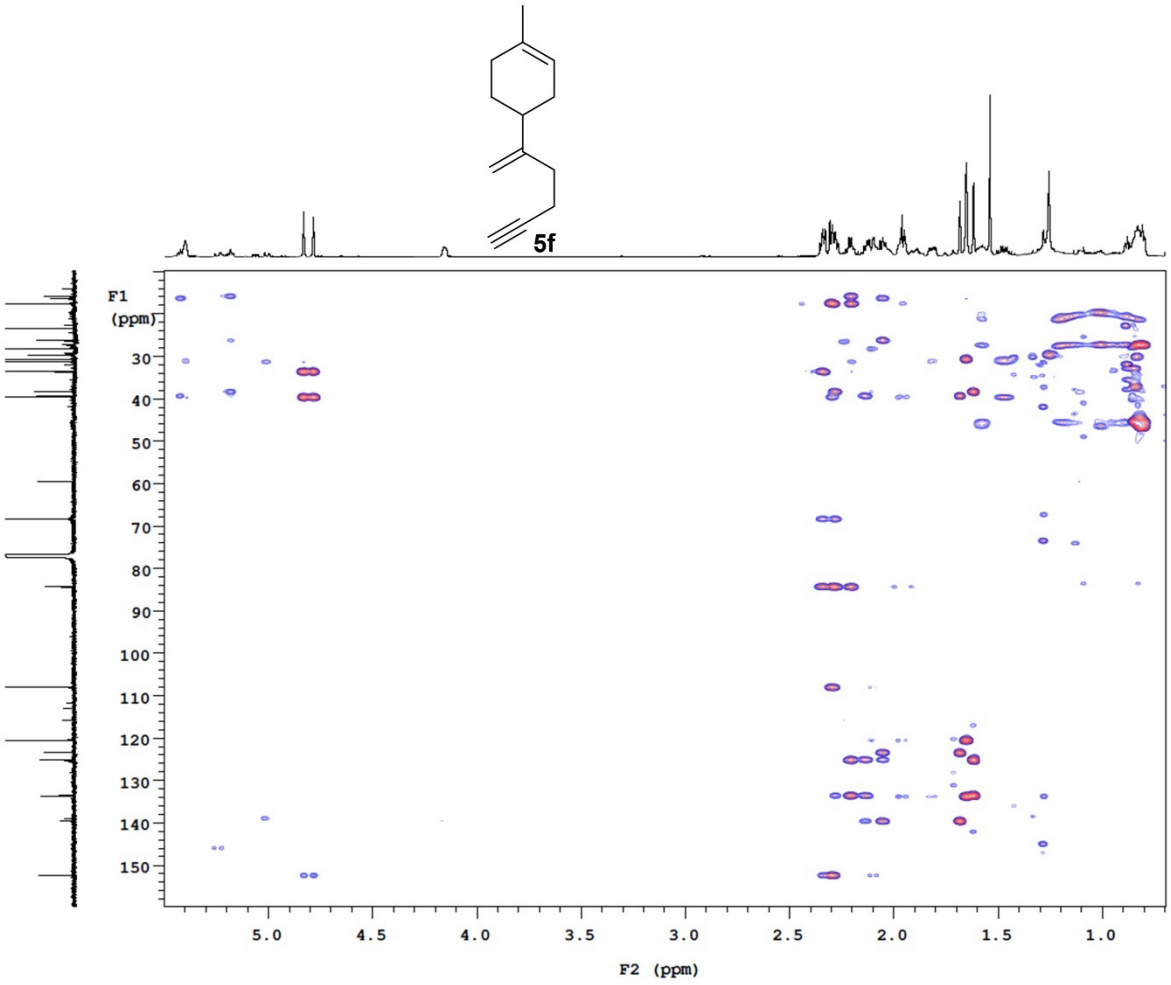
^{13}C NMR of the enzymatic product **5f**



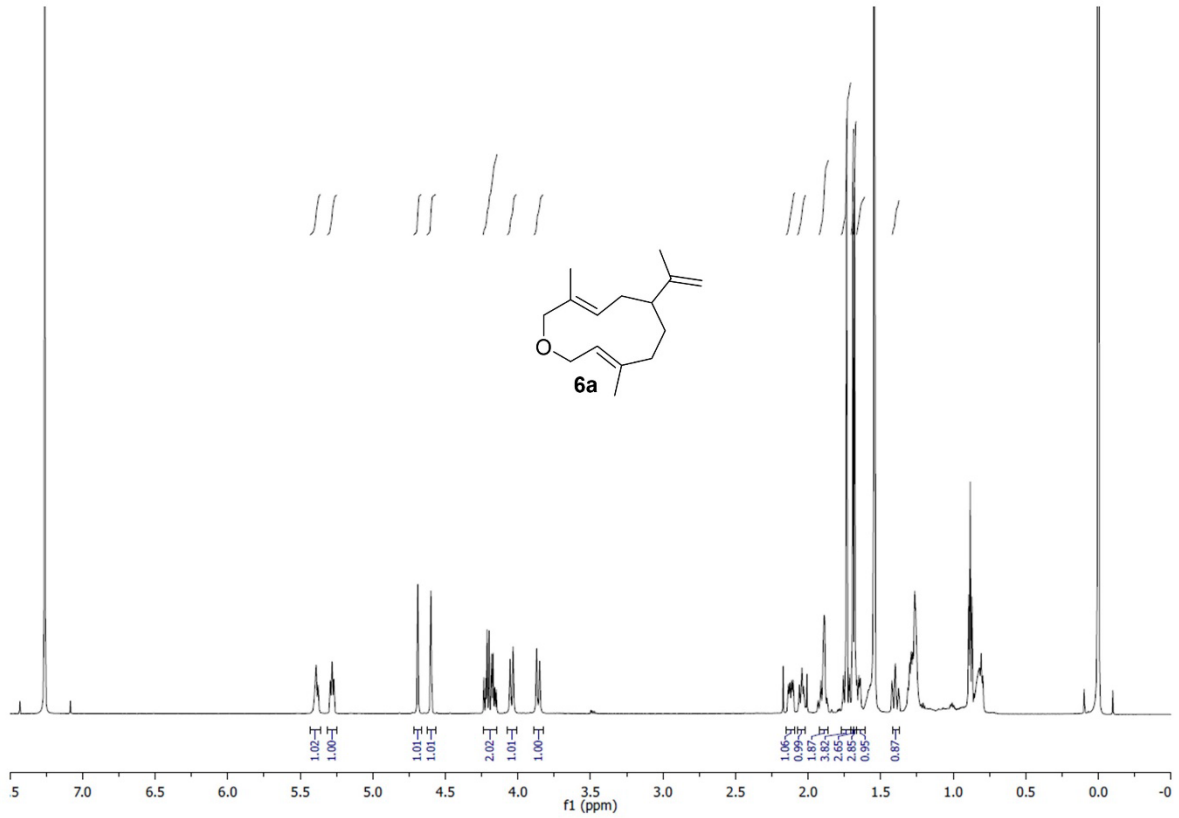
COSY spectrum of the enzymatic product **5f**



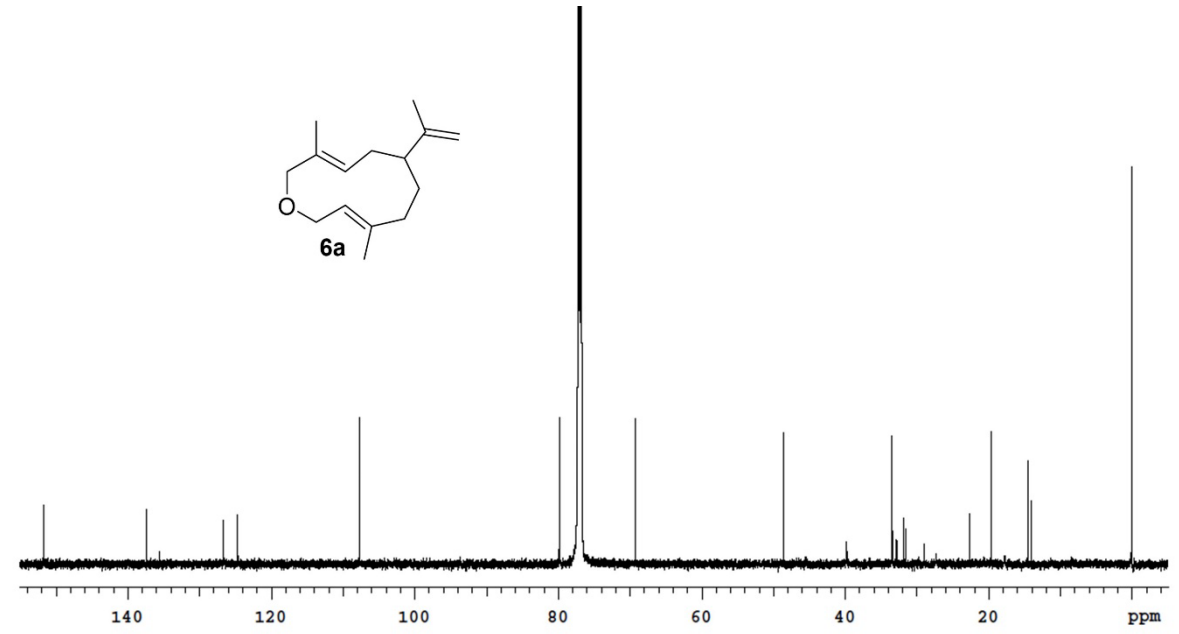
HSQC spectrum of the enzymatic product **5f**



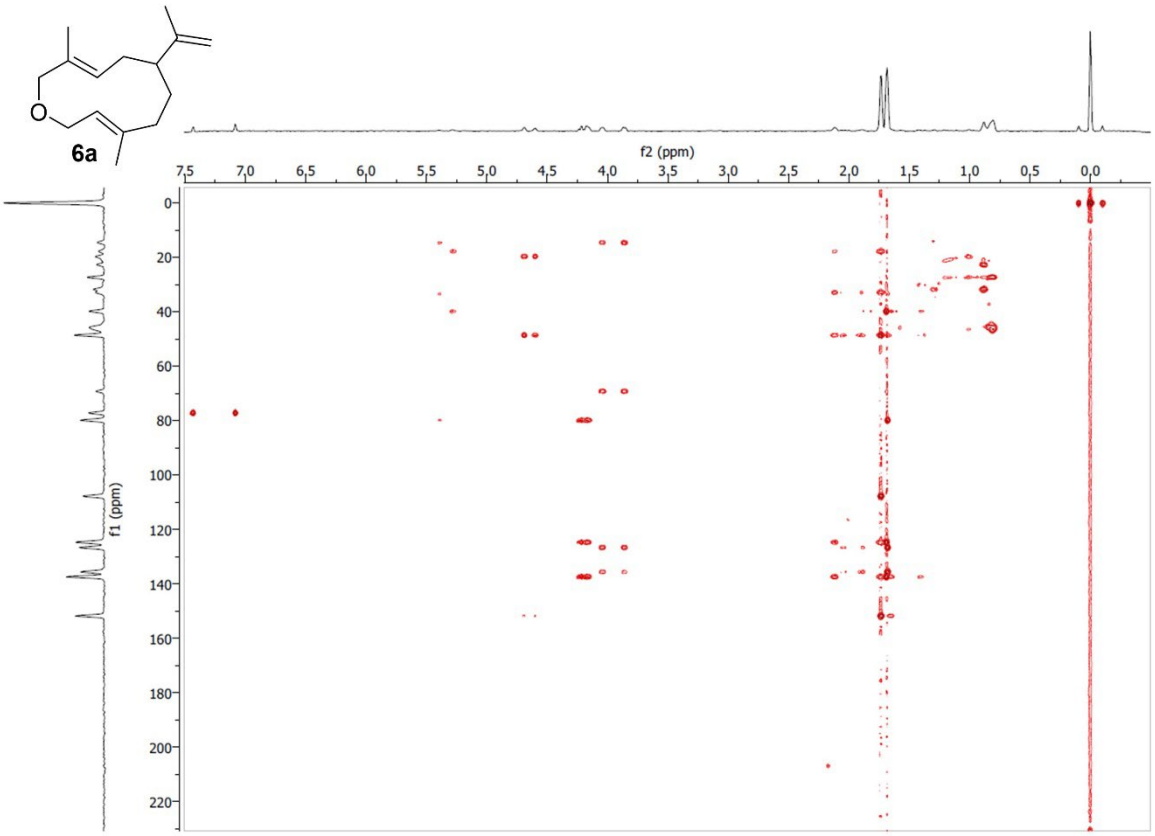
HMBC spectrum of the enzymatic product **5f**



^{1}H NMR of the enzymatic product **6a**



^{13}C NMR of the enzymatic product **6a**



HMBC spectrum of the enzymatic product **6a**

Supplementary References

[1] L. Sumbalova, J. Stourac, T. Martinek, D. Bednar, J. Damborsky, *Nucleic Acids Research* **2018**, *46*, W356-W362.

[2] K. Berka, O. Hanák, D. Sehnal, P. Banáš, V. Navratilova, D. Jaiswal, C.-M. Ionescu, R. Svobodová Vařeková, J. Koča, M. Otyepka, *Nucleic Acids Research* **2012**, *40*, W222-W227.

[3] A. Kuriata, A. M. Gierut, T. Oleniecki, M. P. Ciemny, A. Kolinski, M. Kurcinski, S. Kmiecik, *Nucleic acids research* **2018**, *46*, W338-W343.

[4] R. A. Laskowski, M. B. Swindells, *J. Chem. Inf. Model.* 2011, *51*, 2778-2786.