Electronic Supplementary Information

Machine learning model-based development of a new PKC ligand, a simplified analogue of alotaketals

Jumpei Maki,[a] Asami Oshimura,[a] Chihiro Tsukano,[*a] Ryo C. Yanagita,[b] Yutaka Saito,[c,d] Yasubumi Sakakibara[e] and Kazuhiro Irie[*a]

[a.] *Division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University, Kyoto 606-8501, Japan, E-mail: tsukano.chihiro.2w@kyoto-u.ac.jp, irie.kazuhiro.2z@kyoto-u.ac.jp*

[b.] *Department of Applied Biological Science, Faculty of Agriculture, Kagawa University, Kagawa 761-0795, Japan*

[c.] *Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Tokyo 135-0064, Japan*

[d.] *AIST-Waseda University Computational Bio Big-Data Open Innovation Laboratory (CBBD-OIL), 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan.*

[e.] *Department of Biosciences and Informatics, Keio University, Kanagawa 223-8522, Japan*

Contents

1. Experimental procedures

General remarks

The following spectroscopic and analytical instruments were used: digital polarimeter, P-2200 (Jasco, Tokyo, Japan); Fourier-transform infrared spectrometer, FT/IR-470 Plus (JASCO, Tokyo, Japan); ^{1}H NMR and ^{13}C NMR, AVANCE III 400 (Bruker, Germany) at 400 and 101 MHz respectively or AVANCE III 500 (Bruker, Germany) at 500 and 126 MHz respectively; HPLC, Model 600E with a Model 2487 UV detector (Waters, Milford, Ma, USA); mass spectrometer for FAB-MS (matrix, 3-nitrobenzyl alcohol), JMS700 (JEOL, Tokyo, Japan). Chemical shifts of ^{1}H NMR are reported relative to TMS (δ 0.00) in $CDCl_3$ and residual solvents of CD_3OD (δ 3.31) and pyridine-d_5 (δ 8.74). Multiplicity is indicated by one or more of the following: s (singlet); d (doublet); t (triplet); q (quartet); m (multiplet), br (broad). Chemical shifts of ^{13}C NMR are reported relative to $CDCl_3$ (δ 77.23), CD_3OD (δ 49.15) and pyridine-d_5 (δ 150.35). Analytical thin-layer chromatography was performed with Silica gel 60 F_{254} (Merck, Darmstadt, Germany). Wakogel C-200 (silica gel, Wako Pure Chemical Industries, Ltd., Osaka, Japan) and CHROMATOREX BW-300 (silica gel, Fuji Silysia Chemical, Ltd., Aichi, Japan) were used for column chromatography. Anhydrous solvents were purchased from Wako Pure Chemical Industries. All other materials were obtained from Tokyo Chemical Industry Co., Ltd., Aldrich Inc. or other chemical suppliers, and used without further purification.

Screening of PKC ligands from PubChem database

The KNIME software and its nodes including RDKit Fingerprint, Similarity Search and XLogP were used for the selection.

1) Machine learning

First, we conducted the literature survey on studies describing the biological activity of PKC C1 domain ligands from 1979 to 2017, which provided 1,146 compounds, including 21 unpublished compounds or results from our group. Of these compounds, 110 were phorbol-type, 24 were daphnane-type, 79 were ingenol-type, 171 were teleocidin-type, 127 were bryostatin-type, 60 were aplysiatoxin-type, 75 were diacylglycerol (DAG)-type and isobenzofuranones, 401 were DAG-lactones, and 99 were other class of compounds including iridals and vibsanins. Then, we labeled these compounds either as “strong” (339 molecules) or “weak or inactive” (805 molecules) based on the threshold of 10 nM of binding inhibition constant, K_i, for PKCδ and its C1B domain, and the reported other biological activities. We also augmented the latter negative dataset with 131 hypothetical inactive ligands created by deoxygenation of natural PKC ligands’ primary or secondary hydroxy group required for the binding to PKC.

We used the DeepChem library[1] for the binary classification with graph convolutional network (GCN). Because the DeepChem library takes only planer structures, the number of molecules in the dataset decreased to 1,127. End-to-end supervised learning using the above dataset provided the binary classifier with 0.908 of the area under the curve (AUC) value. GCN is a machine-learning model widely used for *in silico* screening of compounds due to its ability to extract effective features from chemical structures in the form of graph data structures.[2] We also evaluated conventional machine-learning models based on fingerprints: random forest and fully connected (non-graphical) deep neural network using the Morgan fingerprints (1,024 bits, radius = 2). These models showed the AUC values of 0.879 and 0.875, respectively, which were less accurate than GCN. Therefore we employed GCN as our machine-learning model for compound screening.

We conducted the first screening from the PubChem database that contained approximately 97 million compounds. The first screening afforded many false-positive results (score > 0.7) judging from our experience. Therefore, we added those apparent false positive compounds, which have molecular weights less than 300 and contain sulfonamide or phosphate moiety, to the dataset as negative data. Then, we conducted the supervised learning and the second screening from the PubChem database, which provided 7,595 molecules with score > 0.7. Removing stereoisomers from these results provided 5,055 molecules with unique planer structures.

2) Selection with Tanimoto coefficient of circular fingerprint with known PKC ligands and XLogP values

We calculated the Tanimoto coefficient between the Morgan fingerprints (1,024 bits, radius = 2; equivalent to ECFP4) of the 5,055 molecules and twelve known PKC ligands (Figure S1) which are TPA, 6,7-epoxy-TPA, daphnetoxin, ingenol 3-angelate, teleocidin B-4, benzolactam-V8-310, 1,2-dioctanoyl-*sn*-gylcerol, a DAG-lactone,[3] vibsanin A, bryostatin 1, aplysiatoxin and 10-Me-aplog-1. Then, we excluded molecules with a Tanimoto coefficient ≥ 0.3 either with 12 reference ligands to afford 2,536 molecules. Then, we selected molecules with XLogP values from 2.5 to 8, which provided 1,506 molecules.

3) Manual selection

We visually inspected the chemical structures of the filtered 1,506 molecules, and selected fifteen compounds as PKC ligand candidates (Figure S2). The selection criteria were the balance and arrangement of hydrophobic and hydrophilic parts, the existence of a primary or secondary hydroxy group, and the existence of other hydrogen bond acceptors.

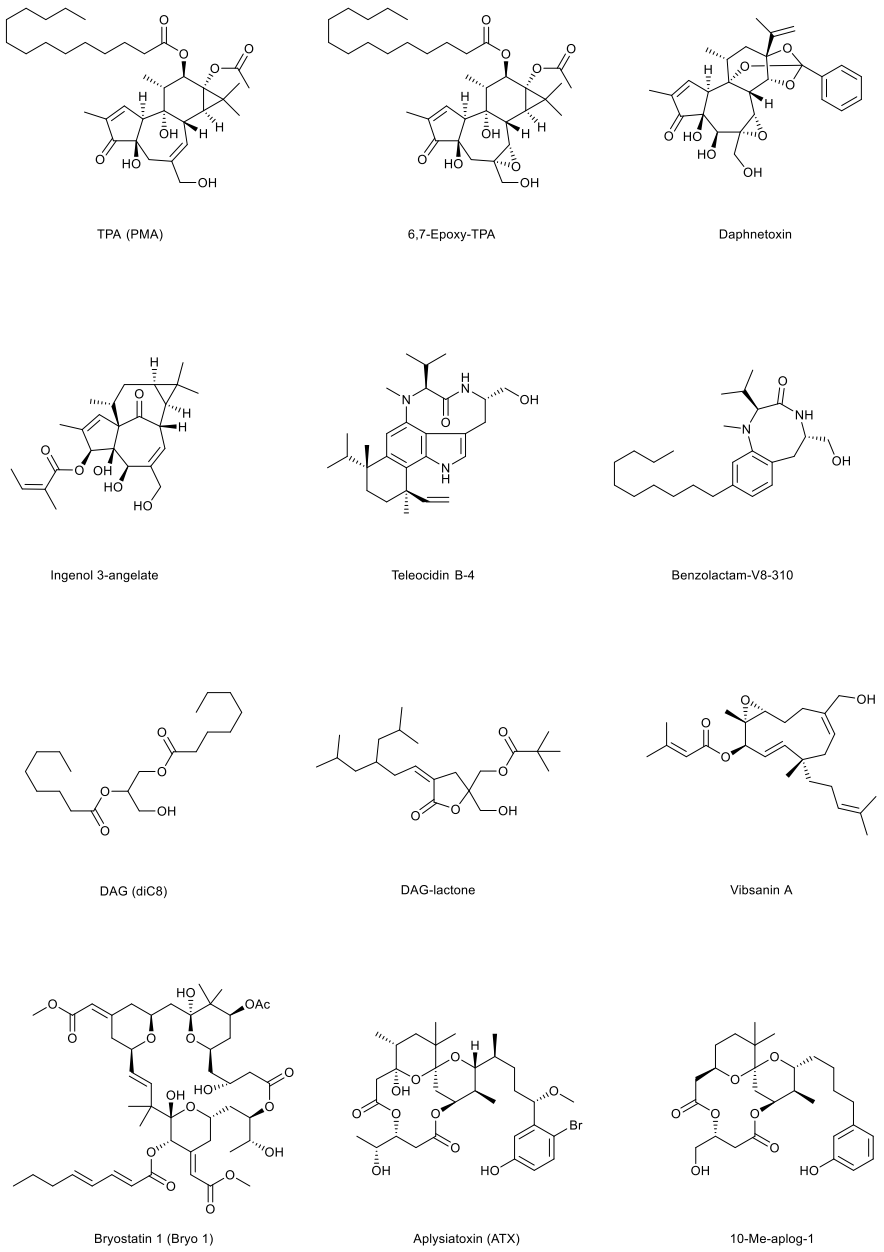


Figure S1 Structures of the known PKC ligands used as reference compounds for the similarity search.

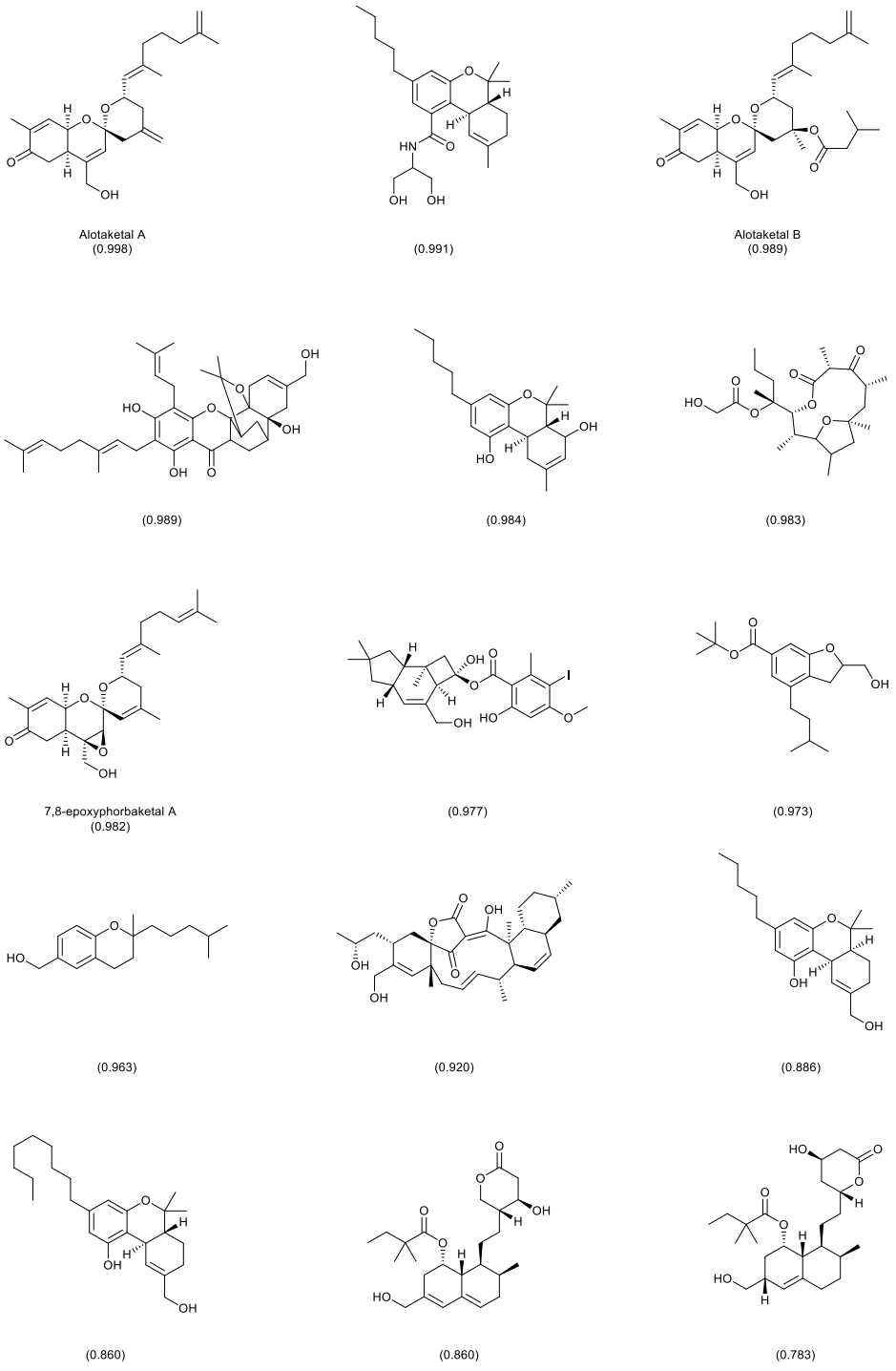


Figure S2 Structures of the finally selected fifteen candidates for PKC ligands. Values in parentheses are scores from the second machine-learning screening.

Molecular modeling study and prediction of the free energy of binding

The following software and servers were used for the calculation: the Avogadro (version 1.2.0) software,[4] the OpenBabel (version 3.1.0) program,[5] the MGLTools (version 1.5.7) software, the AutoDock (version 4.2.6, GPU version) program,[6,7] the CHARMM-GUI server,[8-10] the GROMACS (version 2021.2) software,[11] the AmberTools 21 package,[12] the PPM 3.0 program,[13] and the PyMOL program.[14] Amber ff14SB,[15] ZAFF,[16] LIPID17 and GAFF2 (with AM1-BCC charge) force fields were used for the MD simulation.

Three-dimensional structures of **1**, **2** and **S1**–**S6** (Figure S3) were created using the Avogadro software and energetically minimized with an MMFF94 force field. The docking simulation of these compounds with the PKCδ-C1B domain (PDB id. 1PTR) was performed as described previously.[16] Among 1,000 docking pose obtained for each molecule, the lowest energy result with a primary hydroxy group forming hydrogen bonds with Thr-242 and Leu251 was used as an initial binding pose for the subsequent molecular dynamics (MD) simulation.

The MD simulation of each ligand bound to the PKCδ-C1B domain in the presence of a phophatidylserine bilayer was performed to provide 100 ns of MD trajectory for each ligand as described previously.[17] The free energy of binding was estimated by combining the free energy of binding from the MMPBSA calculation and the free energy of membrane transfer from the PPM program as described previously.[16]

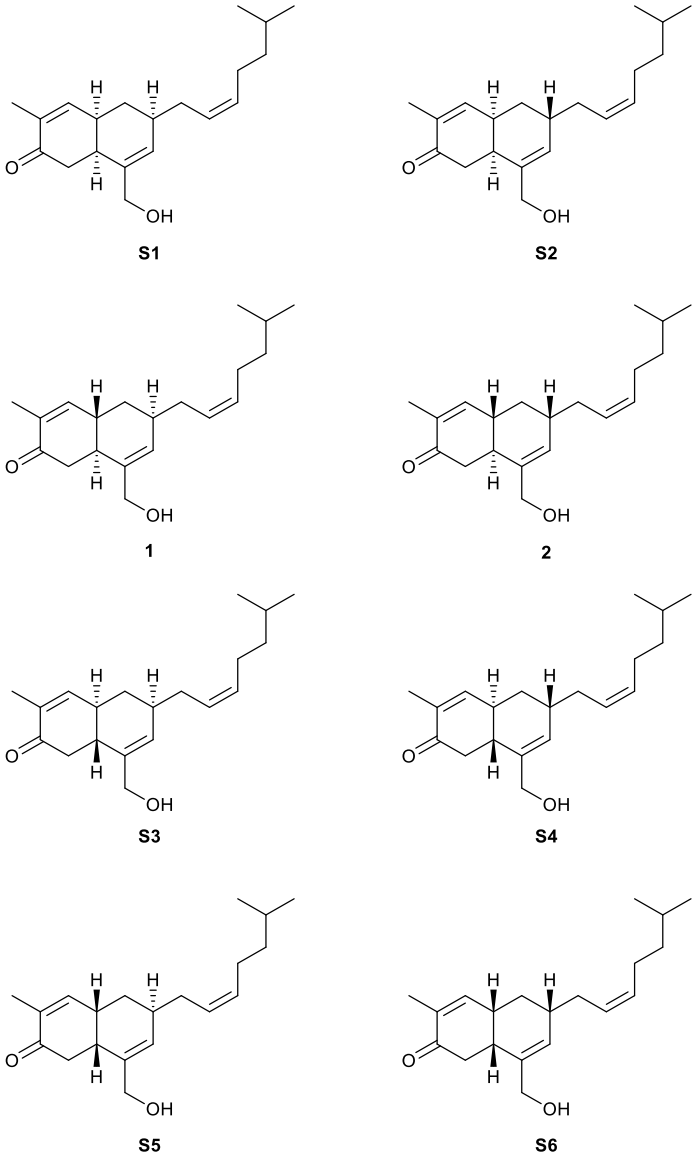


Figure S3 Structures of compounds used for molecular modeling study and prediction of the free energy of binding.

Table S1 Predicted free energy of binding between **1**, **2**, or **S1**–**S6**, and PKCδ-C1B domain.

Compound	$\Delta E_{\text{MM-PBSA}}$ (A)	$\Delta G°_{\text{transfer}}$ [b] (B)	Predicted $\Delta G°_{\text{bind}}$ (A + B)	
S1	−10.2037 (1.0159)[a]	−9.9 (0.2)	−20.1 (1.0)	kcal mol^{-1}
S2	−9.4492 (0.9829)	−10.3 (0.3)	−19.7 (1.0)	kcal mol^{-1}
1	−11.7347 (1.0494)	−10.1 (0.2)	−21.9 (1.1)	kcal mol^{-1}
2	−9.4573 (1.011)	−9.9 (0.2)	−19.3 (1.0)	kcal mol^{-1}
S3	−5.7628 (1.0465)	−11.4 (0.3)	−17.1 (1.1)	kcal mol^{-1}
S4	−5.1981 (0.9775)	−10.7 (0.3)	−15.9 (1.0)	kcal mol^{-1}
S4	−4.4289 (1.0122)	−10.7 (0.2)	−15.1 (1.0)	kcal mol^{-1}
S6	−4.7291 (1.0001)	−11.1 (0.3)	−15.8 (1.0)	kcal mol^{-1}

[a] Values in parentheses are standard errors of the mean: n = 1,000 for A, and n = 25 for B.

[b] Calculated by the PPM 3.0 program built on the local machine.

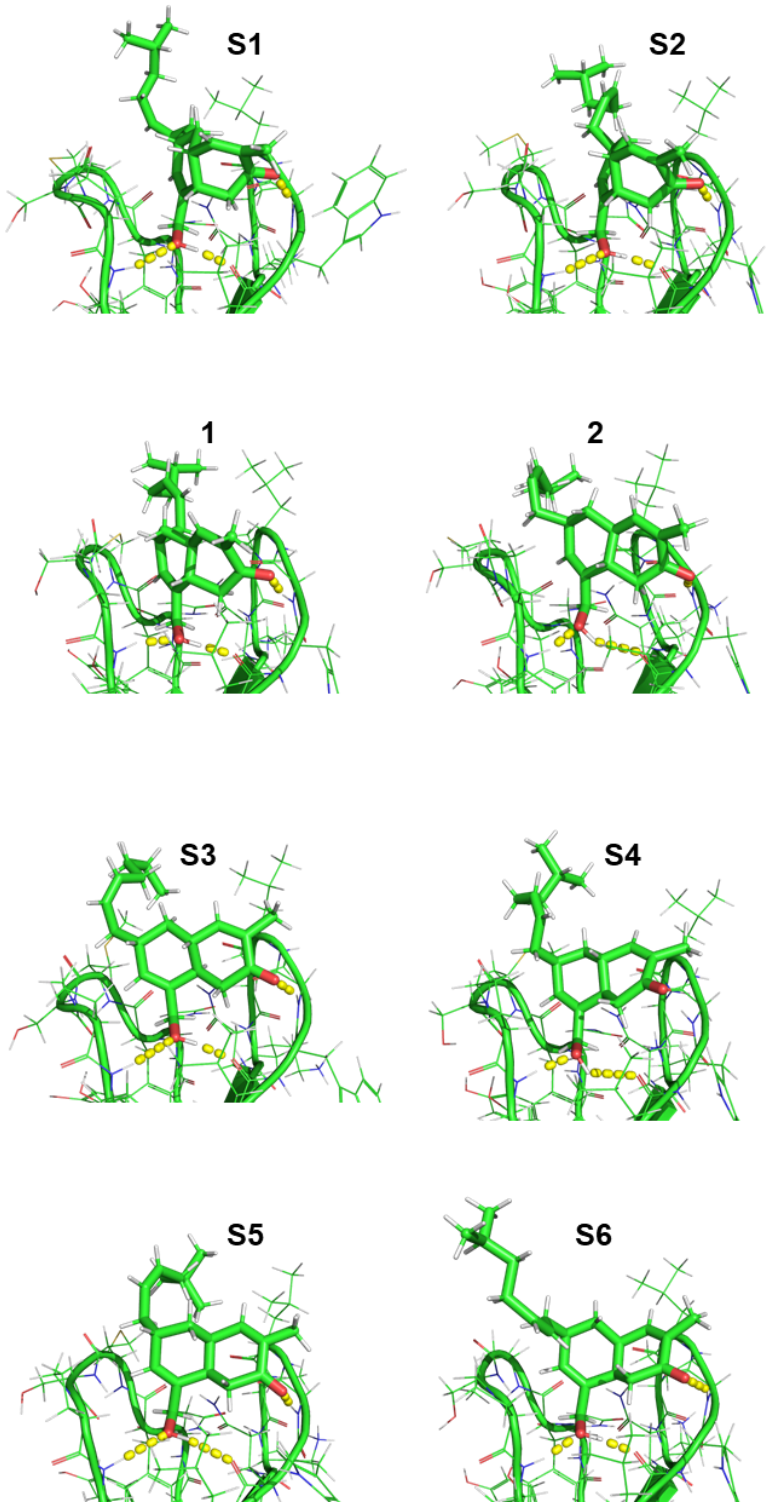
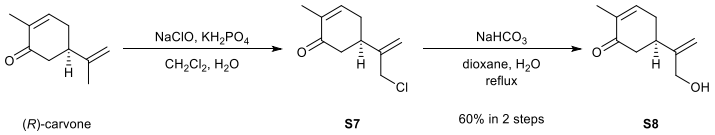


Figure S4 Snapshots from molecular dynamics simulation of membrane-bound PKCδ-C1B domain in complex with **S1**, **S2**, **1**, **2**, **S3**, **S4**, **S5**, or **S6**. The PKCδ-C1B domain is depicted as line and cartoon models. The ligands are depicted as a stick model. Yellow dashed lines represent hydrogen bonds.

*Synthesis of **1** and **2***

Compound S8

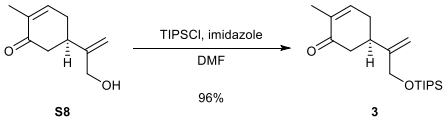


To a solution of (*R*)-carvone (24.0 g, 160 mmol) in CH_2Cl_2 (200 mL) and H_2O (200 mL) were added KH_2PO_4 (44.0 g, 323 mmol) and NaClO aqueous solution (12%, 100 mL) at 0 °C and the resulting mixture was stirred at room temperature for 2 h. After the addition of NaClO aqueous solution (200 mL), the resulting mixture was stirred for 3 h. Additional NaClO aqueous solution (80 mL) was added and the reaction mixture was stirred for 2 h. The reaction was quenched with *satd. aq.* $Na_2S_2O_3$ (250 mL). The aqueous layer was separated and extracted with $CHCl_3$ (3 x 100 mL) The combined organic layers were dried over Na_2SO_4 and then concentrated *in vacuo* to afford crude allyl chloride **S7** (38.3 g).

To a solution of crude **S7** in dioxane (100 mL) and H_2O (300 mL) was added $NaHCO_3$ (33.6 g, 400 mmol). After stirring at reflux for 20 h, the reaction mixture was diluted with $CHCl_3$ (100 mL) and the aqueous layer was separated. The aqueous layer was extracted with $CHCl_3$ (3 x 100 mL) and the combined organic layers were dried over Na_2SO_4 then concentrated *in vacuo*. The residue was purified by column chromatography (silica gel, EtOAc/*n*-hexane = 1/9 to 1/1) to afford silyl ether **S8** (15.7 g, 94.6 mmol, 60% in two steps) as a yellow oil.

The data was reported previously.[18]

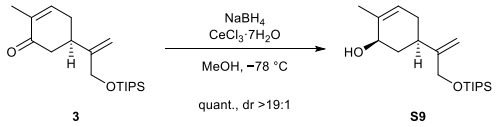
Compound 3



To a solution of **S8** (15.7 g, 94.6 mmol) in DMF (200 mL) were added imidazole (12.9 g, 189 mmol) and TIPSCl (21.0 mL, 99.1 mmol). After stirring at room temperature for 2 h, the resulting mixture was partitioned between water (300 mL) and EtOAc: *n*-hexane = 1:4 solution (4 x 100 mL). The combined organic layers were washed with brine (100 mL), dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by column chromatography (silica gel, EtOAc/*n*-hexane = 1/19) to afford silyl ether **3** (29.4 g, 91.0 mmol, 96%) as a pale yellow oil.

The data was reported previously.[17]

Compound S9



To a solution of **3** (5.23 g, 16.2 mmol) in MeOH (160 mL) under N_2 atmosphere was added $CeCl_3 \cdot 7H_2O$ (7.24 g, 19.4 mmol) and then the reaction mixture was cooled to −78 °C. $NaBH_4$ (739 mg, 19.5 mmol) were added to the reaction mixture and the solution was stirred at −78 °C for 30 min, before the addition of *satd. aq.* NH_4Cl (100 mL). The resulting mixture was partitioned between water (100 mL) and ether (4 x 100 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by column chromatography (silica gel, EtOAc/*n*-hexane = 1/9 to 1/4) to afford allylic alcohol **S9** (5.36 g, 16.5 mmol, quant., dr > 19:1) as a colorless oil.

R_f: 0.36 (silica gel, EtOAc/*n*-hexane = 2:8).

$[\alpha]_D^{25}$: −17.0 (*c* 1.47, $CHCl_3$).

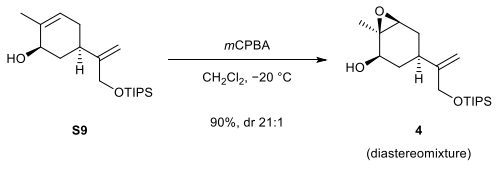
IR (neat on KRS-5): 3328, 2943, 2891, 2866, 1652, 1463, 1386, 1249, 1121, 1102, 1066, 1036, 900, 882, 811, 684, 659 cm^{-1}.

^{1}H NMR (500 MHz, $CDCl_3$): δ 5.50 (dd, *J* = 1.5, 1.5 Hz, 1H), 5.15 (d, *J* = 1.5 Hz, 1H), 4.89 (s, 1H), 4.22 (s, 2H), 4.19 (br-s, 1H), 2.32 (m, 1H), 2.25–2.16 (m, 1H), 2.16–2.07 (m, 1H), 2.05–1.94 (m, 1H), 1.76 (s, 3H), 1.54 (ddd, *J* = 12.2, 12.2, 9.6, 1H), 1.54–1.44 (br-s, 1H), 1.22–0.99 (m, 21H) ppm.

^{13}C NMR (126 MHz, $CDCl_3$): δ 152.2, 136.4, 124.1, 107.7, 71.2, 65.3, 38.5, 36.3, 31.8, 19.2, 18.2 (6C), 12.2 (3C) ppm.

HRMS (FAB+): *m*/*z* calcd for $C_{19}H_{37}O_2Si$ $[M + H]^+$ 325.2563, found 325.2567.

Compound 4



To a solution of **S9** (5.75 g, 17.7 mmol) in CH_2Cl_2 (120 mL) under N_2 atmosphere was added *m*CPBA (65%, 5.22 g, 19.7 mmol) at −20 °C. The reaction mixture was stirred at −20 °C for 3 h, before the addition of *satd. aq.* $Na_2S_2O_3$ (100 mL) and *satd. aq.* $NaHCO_3$ (100 mL). The aqueous layer was separated and extracted with $CHCl_3$ (3 x 50 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by column chromatography (silica gel, EtOAc/*n*-hexane = 1/9 to 2/3) to afford diastereomixture of epoxyalcohol **4** (5.42 g, 15.9 mmol, 90%, dr 21:1) as a colorless oil.

R$_f$: 0.35 (silica gel, EtOAc/*n*-hexane = 4:6).

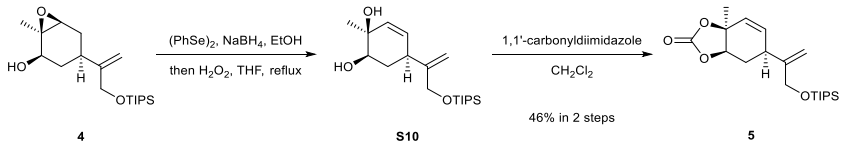
IR (neat on KRS-5): 3423, 2943, 2892, 2866, 1651, 1463, 1383, 1250, 1125, 1096, 1066, 1052, 1014, 900, 883, 848, 813, 687, 661 cm^{-1}.

^{1}H NMR (500 MHz, $CDCl_3$) of major diastereomer: δ 5.10 (ddd, *J* = 1.6, 1.6, 1.6 Hz, 1H), 4.84 (dd, *J* = 1.4, 1.4 Hz, 1H), 4.16 (s, 2H), 3.84 (dd, *J* = 10.6, 5.5 Hz, 1H), 3.16 (d, *J* = 5.0 Hz, 1H), 2.11–2.01 (m, 2H), 1.86–1.80 (m, 1H), 1.70 (m, 1H), 1.61 (br-s, 1H), 1.45 (s, 3H), 1.41–1.29 (m, 1H), 1.17–0.99 (m, 21H) ppm.

^{13}C NMR (126 MHz, $CDCl_3$) of major diastereomer: δ 151.0, 108.2, 72.3, 65.1, 62.4, 60.6, 36.0, 34.3, 30.0, 19.3, 18.1 (6C), 12.1 (3C) ppm.

HRMS (FAB+): *m*/*z* calcd for $C_{19}H_{37}O_3Si$ $[M + H]^+$ 341.2512, found 341.2516.

Compound 5



4 **S10** **5**

To a suspension of diphenyl diselenide (6.60 g, 21.1 mmol) in EtOH (80 mL) at 0 °C under N_2 atmosphere was added $NaBH_4$ (1.62 g, 42.8 mmol) and the resulting mixture was stirred at room temperature for 1 h. After the addition of epoxyalcohol **4** (13.0 g, 38.2 mmol) in EtOH (50 mL) via cannula, the reaction mixture was stirred at room temperature for 3 h. 80 mL of THF was added and the reaction mixture was cooled to 0 °C before the addition of 50 mL of 30% *aq.* H_2O_2. The resulting mixture was warmed up to 90 °C with care. After refluxed for 2 h, the reaction was quenched with brine (200 mL) and partitioned between ether (4 x 70 mL) and H_2O (50 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. Purification by column chromatography (silica gel, EtOAc/*n*-hexane = 1/9 to 2/3) afforded impure diol **S10** (12.7 g).

To a solution of diol **S10** (12.7 g) in CH_2Cl_2 (200 mL) was added 1,1'-carbonyldiimidazole (12.4 g, 76.5 mmol). After stirring at room temperature for 2 h, the resulting mixture was concentrated *in vacuo*. The residue was purified by column chromatography (silica gel, EtOAc/*n*-hexane = 1/9 to 1/4) to afford carbonate **5** (6.48 g, 17.7 mmol, 46% in two steps) as a yellow oil.

R$_f$: 0.27 (silica gel, EtOAc/*n*-hexane = 2:8).

$[\alpha]_D^{26}$: −39.0 (*c* 3.94, $CHCl_3$).

IR (neat on KRS-5): 2944, 2892, 2866, 1808, 1464, 1381, 1365, 1300, 1247, 1193, 1136, 1117, 1083, 1063, 1040, 1013, 997, 905, 882, 815, 771, 684, 660 cm^{-1}.

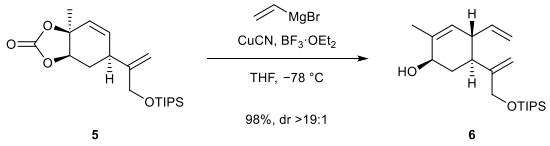
^{1}H NMR (500 MHz, $CDCl_3$): δ 6.06 (dd, *J* = 10.1, 3.0 Hz, 1H), 5.82 (dd, *J* = 10.1, 2.4 Hz, 1H), 5.18 (d, *J* = 0.8 Hz, 1H), 4.88 (s,1H), 4.49 (dd, *J* = 9.0, 4.8 Hz, 1H), 4.28 (d, *J* = 13.5 Hz, 1H),

4.23 (d, *J* = 13.5 Hz, 1H), 2.92 (m, 1H), 2.20 (ddd, *J* = 13.5, 4.9, 4.9 Hz, 1H), 1.86 (ddd, *J* = 13.5, 8.9, 8.9 Hz, 1H), 1.54 (s, 3H), 1.17-1.01 (m, 21H) ppm.

^{13}C NMR (126 MHz, $CDCl_3$): δ 154.0, 149.5, 135.8, 126.2, 111.4, 80.3, 79.2, 65.6, 34.8, 30.9, 25.9, 18.2 (6C), 12.2 (3C) ppm.

HRMS (FAB+): *m*/*z* calcd for $C_{20}H_{35}O_4Si$ $[M + H]^+$ 367.2305, found 367.2300.

Compound 6



To a suspension of CuCN (3.53 g, 39.4 mmol) in dry THF (100 mL) at −78 °C under N_2 atmosphere was added vinylmagnesium bromide solution (1 M in THF, 77 mL, 77 mmol) followed by $BF_3{\cdot}OEt_2$ (1.8 mL, 14 mmol) and then carbonate **5** (4.68 g, 12.8 mmol) in dry THF (30 mL) via cannula. After stirring at −78 °C for 30 min, the reaction was quenched with *satd. aq.* NH_4Cl (100 mL). The resulting mixture was filtered and separated. The aqueous layer was extracted with ether (3 x 50 mL), and the combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by column chromatography (silica gel, EtOAc/*n*-hexane = 1/9 to 1/4) to afforded alcohol **6** (4.18 g, 12.4 mmol, 98%) as a pale yellow oil.

R_f: 0.29 (silica gel, EtOAc/*n*-hexane = 2:8).

$[\alpha]_D^{26}$: 109.2 (*c* 1.22, $CHCl_3$).

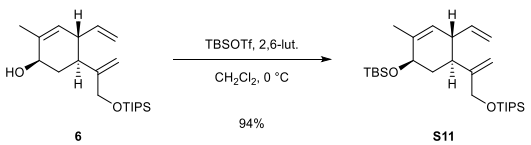
IR (neat on KRS-5): 3318, 2943, 2891, 2866, 1638, 1463, 1385, 1251, 1114, 1066, 1013, 994, 914, 896, 883, 810, 684, 660 cm^{-1}.

^{1}H NMR (500 MHz, $CDCl_3$): δ 5.61 (ddd, *J* = 16.9, 10.5, 8.0 Hz, 1H), 5.34 (s, 1H), 5.22 (dd, *J* = 3.3, 1.6 Hz, 1H), 4.99 (d, *J* = 7.1 Hz, 1H), 4.96 (s, 1H), 4.93 (s, 1H), 4.19-4.12 (m, 3H), 2.87 (m, 1H), 2.13 (ddd, *J* = 12.5, 5.7, 2.7 Hz, 1H), 2.06 (ddd, *J* = 12.0, 9.5, 2.6 Hz, 1H), 1.79 (s, 3H), 1.66 (ddd, *J* = 12.3, 12.3, 9.4, 1H), 1.50 (br, 1H), 1.16-1.02 (m, 21H) ppm.

^{13}C NMR (126 MHz, $CDCl_3$): δ150.7, 141.1, 137.0, 127.3, 115.2, 109.2, 70.7, 65.5, 45.5, 42.7, 39.0, 19.3, 18.3 (6C), 12.2 (3C) ppm.

HRMS (FAB+): *m*/*z* calcd for $C_{21}H_{39}O_2Si$ $[M + H]^+$ 351.2719, found 351.2724.

Compound S11



To a solution of alcohol **6** (4.18 g, 12.4 mmol) in dry CH_2Cl_2 (100 mL) at 0 °C under N_2 atmosphere were added 2,6-lutidine (3.0 mL, 26 mmol) and TBSOTf (3.6 mL, 16 mmol). After stirring at 0 °C for 30 min, the reaction was quenched with *satd. aq.* $NaHCO_3$ (100 mL). The aqueous layer was separated and extracted with $CHCl_3$ (3 x 50 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by column chromatography (silica gel, EtOAc/*n*-hexane = 1/49 to 1/19) to afforded TBS ether **S11** (5.29 g, 11.7 mmol, 94%) as a yellow oil.

R*f*: 0.68 (silica gel, EtOAc/*n*-hexane = 1:9).

$[\alpha]_D^{25}$: 74.5 (*c* 4.88, $CHCl_3$).

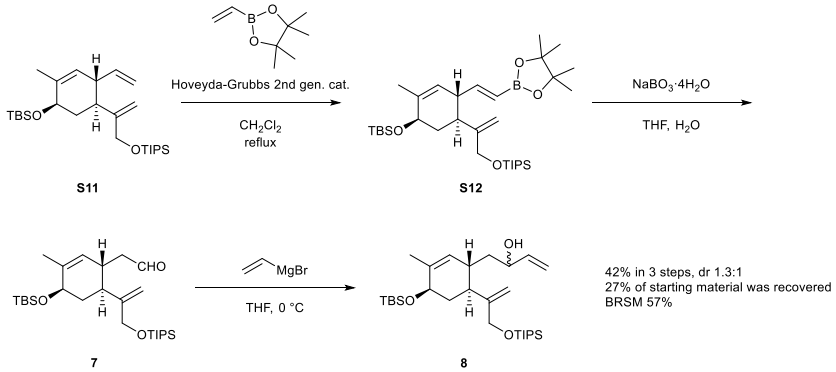
IR (neat on KRS-5): 2944, 2892, 2866, 1639, 1463, 1389, 1362, 1346, 1252, 1089, 1069, 1006, 993, 898, 882, 866, 836, 774, 683 cm^{-1}.

^{1}H NMR: (500 MHz, $CDCl_3$): δ 5.58 (ddd, *J* = 17.2, 10.2, 8.1 Hz, 1H), 5.27 (d, *J* = 1.5 Hz, 1H), 5.20 (dd, *J* = 3.4, 1.6 Hz, 1H), 5.00-4.93 (m, 2H), 4.91 (d, *J* = 1.4 Hz, 1H), 4.21 (m, 1H), 4.17 (ddd, *J* = 14.4, 1.5, 1.5 Hz, 1H), 4.12 (ddd, *J* = 14.4, 1.5, 1.5 Hz, 1H), 2.83 (m, 1H), 2.06-1.97 (m, 2H),1.76–1.63 (m, 4H), 1.16-1.03 (m, 21H), 0.90(s, 9H), 0.09 (s, 3H), 0.07(s, 3H) ppm.

^{13}C NMR: (126 MHz, $CDCl_3$): δ 150.6, 141.3, 138.1, 126.9, 115.0, 109.0, 71.4, 65.5, 45.9, 43.4, 40.1, 26.1, 20.0 (3C), 18.4, 18.3 (6C), 12.2 (3C), −4.0, −4.7 ppm.

HRMS (FAB+): *m*/*z* calcd for $C_{27}H_{52}O_2Si_2Na$ $[M + Na]^+$ 487.3404, found 487.3402.

Compound 8



To a solution of TBS ether **S11** (471 mg, 1.04 mmol) in dry CH_2Cl_2 (10 mL) under N_2 atmosphere were added vinylboronic acid pinacol ester (0.35 mL, 2.1 mmol) and Hoveyda–Grubbs 2nd generation catalyst (31.4 mg, 50.1 mmol). After the reaction mixture was refluxed for 21 h, vinylboronic acid pinacol ester (0.40 mL, 2.36 mmol) was added. The resulting mixture was refluxed for an additional 25 h before cooled and concentrated *in vacuo*. Purification by column chromatography (silica gel, EtOAc/*n*-hexane = 1/49 to 1/9) afforded a mixture of **S12** and **S11** (674 mg).

To a solution of a mixture of **S12** and **S11** in THF (10 mL) and H_2O (10 mL) at 0 °C was

added $NaBO_3 \cdot 4H_2O$ (1.16 g, 7.53 mmol) and stirred vigorously at room temperature for 17 h. The reaction was quenched with H_2O (100 mL). The aqueous layer was separated and extracted with EtOAc (4 x 100 mL). The combined organic layers were washed with brine (50 mL) and dried over Na_2SO_4 then concentrated *in vacuo*. Purification by column chromatography (silica gel, EtOAc/*n*-hexane = 1/49 to 1/9) afforded a mixture of aldehyde **7** and **S11** (361 mg).

To a solution of a mixture of aldehyde **7** and **S11** in THF (10 mL) at 0 °C was added vinylmagnesium bromide solution (1 M in THF, 1.0mL, 1.0 mmol). After stirring at 0 °C for 15 min, the reaction was quenched with *satd. aq.* NH_4Cl (10 mL). The mixture was partitioned between ether (4 x 20 mL) and H_2O (10 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by column chromatography (silica gel, EtOAc/*n*-hexane = 1/49 to 1/9) to afford alcohol **8** (215 mg, 0.434 mmol, 42% in three steps, dr 1.3:1) and TBS ether **S11** (129 mg, 0.285 mmol, 27% recovered in three steps) as a colorless oil.

R_f: 0.47, 0.42 (silica gel, EtOAc/*n*-hexane = 1:9).

IR (neat on KRS-5): 3383, 2943, 2891, 2865, 1648, 1463, 1388, 1362, 1344, 1252, 1079, 1006, 994, 920, 900, 882, 836, 774, 683 cm^{-1}.

^{1}H NMR (500 MHz, $CDCl_3$):

Diastereomer 1: δ 5.88–5.80 (m, 1H), 5.50 (d, *J* = 1.4 Hz,1H), 5.26–5.16 (m, 2H), 5.12–5.04 (m, 1H), 4.96 (d, *J* = 1.4 Hz, 1H), 4.29–4.10 (m, 4H), 2.47 (m, 1H), 2.07–1.94 (m, 2H), 1.71 (m, 3H), 1.67–1.53 (m, 3H), 1.45–1.32 (m, 1H), 1.18–1.01 (m, 21H), 0.90 (s, 9H), 0.08 (s, 3H), 0.06 (s, 3H) ppm.

Diastereomer 2: δ 5.80–5.72 (m, 1H), 5.44 (d, *J* = 1.5 Hz), 2.24 (m, 1H), 5.26–5.16 (m, 2H), 5.12–5.04 (m, 1H), 4.93 (d, *J* = 1.4 Hz,1H), 4.29–4.10 (m, 4H), 2.24 (m, 1H), 2.07–1.94 (m, 2H), 1.71 (m, 3H), 1.67–1.53 (m, 3H), 1.45–1.32 (m, 1H), 1.18–1.01 (m, 21H), 0.89 (s, 9H), 0.08 (s, 3H), 0.06 (s, 3H) ppm.

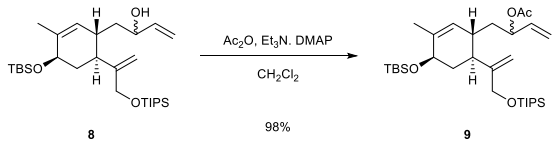
^{13}C NMR (126 MHz, $CDCl_3$):

Diastereomer 1: δ 150.5, 141.8, 137.5, 127.1, 114.2, 109.6, 71.3, 70.5, 65.1, 43.8, 41.0, 40.9, 35.9, 26.1 (3C), 20.1, 18.3, 18.3 (6C), 12.2 (3C), −4.0, −4.7 ppm.

Diastereomer 2: δ 150.8, 141.1, 137.4, 127.6, 114.6, 109.5, 72.2, 71.3, 65.1, 43.9, 41.1, 41.0, 36.8, 26.1 (3C), 20.1, 18.4, 18.3 (6C), 12.2 (3C), −4.0, −4.7 ppm.

HRMS (FAB+): *m*/*z* calcd for $C_{29}H_{56}O_3Si_2Na$ $[M + Na]^+$ 531.3666, found 531.3673.

Compound 9



8 9

To a solution of alcohol **8** (172 mg, 0.347 mmol) in CH_2Cl_2 (10 mL) were added Et_3N (0.30 mL, 2.2 mmol), Ac_2O (0.10 mL, 1.1 mmol) and DMAP (5.9 mg, 0.048 mmol). After stirring at room temperature for 2.5 h, the reaction was quenched with *satd. aq.* $NaHCO_3$ (10 mL). The aqueous layer was separated and extracted with $CHCl_3$ (3 x 10 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by column chromatography (silica gel, EtOAc/*n*-hexane = 1/49 to 1/9) to afford acetate **9** (188 mg, 0.341 mmol, 98%) as a colorless oil.

R_f: 0.60 (silica gel, EtOAc/*n*-hexane = 2:8).

IR (neat on KRS-5): 2944, 2891, 2865, 1742, 1648, 1463, 1433, 1370, 1238, 1081, 1016, 990, 901, 882, 836, 775, 683 cm^{-1}.

^{1}H NMR (500 MHz, $CDCl_3$):

Diastereomer 1: δ 5.73 (m, 1H), 5.49–5.40 (m, 2H), 5.34–5.08 (m, 3H), 4.94 (d, *J* = 1.3 Hz, 1H), 4.23–4.08 (m, 3H), 2.23 (m, 1H), 2.03 (s, 3H), 2.02–1.92 (m, 2H), 1.83 (m, 1H), 1.70 (s, 3H), 1.65–1.54 (m, 1H), 1.41–1.32 (m, 1H), 1.18–1.02 (m, 21H), 0.89 (s, 9H), 0.08 (s, 3H), 0.06 (s, 3H) ppm.

Diastereomer 2: δ 5.67 (m, 1H), 5.49–5.40 (m, 1H), 5.34–5.08 (m, 4H), 4.90 (d, *J* = 1.2 Hz, 1H), 4.23–4.08 (m, 3H), 2.19 (m, 1H), 2.03 (s, 3H), 2.02–1.92 (m, 2H), 1.76–1.68 (m, 1H), 1.70 (s, 3H), 1.65–1.54 (m, 1H), 1.41–1.32 (m, 1H), 1.18–1.02 (m, 21H), 0.89 (s, 9H), 0.08 (s, 3H), 0.06 (s, 3H) ppm.

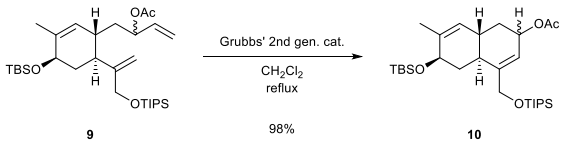
^{13}C NMR (126 MHz, $CDCl_3$):

Diastereomer 1: δ 170.5, 150.1, 137.5, 137.2, 126.4, 116.0, 109.6, 71.8, 71.2, 64.7, 43.8, 40.8, 37.9, 35.6, 26.1 (3C), 21.3, 20.1, 18.3, 18.3 (6C), 12.2 (3C), −4.0, −4.7 ppm.

Diastereomer 2: δ 170.3, 150.2, 137.6, 136.2, 126.7, 118.1, 109.6, 74.2, 71.2, 65.1, 43.9, 40.8, 37.7, 36.4, 26.1 (3C), 21.6, 20.1, 18.3, 18.3 (6C), 12.2 (3C), −4.0, −4.7 ppm.

HRMS (FAB+): *m*/*z* calcd for $C_{31}H_{58}O_4Si_2Na$ $[M + Na]^+$ 573.3771, found 573.3776.

Synthesis of bicyclic compound 10



9 10

To a solution of acetate **9** (219 mg, 0.398 mmol) in dry CH_2Cl_2 (40 mL) under N_2 atmosphere was added Grubbs' 2nd generation catalyst (16.7 mg, 0.0197 mmol). After refluxed for 14 h, the reaction mixture was cooled and concentrated *in vacuo*. The residue was purified by column chromatography (silica gel, EtOAc/*n*-hexane = 1/49 to 1/9) to afford bicyclic compound **10** (204 mg, 0.390 mmol, 98%) as a yellow oil.

R_f: 0.41 (silica gel, EtOAc/*n*-hexane = 1:9).

IR (neat on KRS-5): 2944, 2892, 2865, 1739, 1463, 1371, 1342, 1238, 1200, 1180, 1144, 1084, 1050, 1014, 963, 882, 861, 837, 814, 775, 682 cm^{-1}.

^{1}H NMR (500 MHz, $CDCl_3$):

Diastereomer 1: δ 5.75 (s, 1H), 5.31–5.25 (m, 2H), 4.35–4.22 (m, 2H), 4.16 (d, *J* = 13.7 Hz, 1H), 2.49–2.37 (m, 1H), 2.32–1.84 (m, 3H), 2.02 (s, 3H), 1.69 (m, 3H), 1.53–1.30 (m, 2H), 1.16–1.01 (m, 21H), 0.91 (s, 9H), 0.10 (s, 3H), 0.09 (s, 3H) ppm.

Diastereomer 2: δ 5.54 (s, 1H), 5.46 (m, 1H), 5.31–5.25 (m, 1H), 4.35–4.22 (m, 2H), 4.13 (d, *J* = 13.5 Hz, 1H), 2.49–2.37 (m, 1H), 2.32–2.13 (m, 3H), 2.05 (s, 3H), 1.69 (m, 3H), 1.53–1.30 (m, 2H), 1.16–1.01 (m, 21H), 0.91 (s, 9H), 0.10 (s, 3H), 0.09 (s, 3H) ppm.

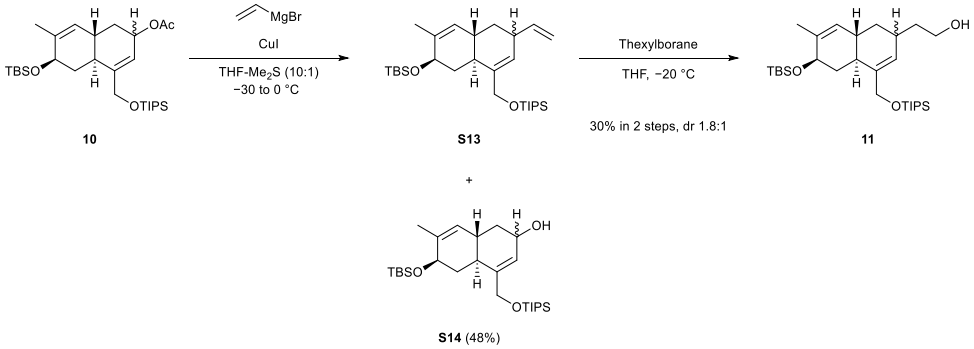
^{13}C NMR (126 MHz, $CDCl_3$):

Diastereomer 1: δ 170.9, 145.8, 137.7, 128.4, 119.9, 72.8, 67.6, 65.1, 39.7, 35.8, 35.2, 34.8, 26.1 (3C), 21.6, 20.0, 18.3, 18.2 (6C), 12.2 (3C), −4.0, −4.6 ppm.

Diastereomer 2: δ 171.1, 143.3, 137.8, 128.1, 120.5, 72.7, 71.5, 64.7, 39.7, 38.7, 35.9, 35.1, 26.1 (3C), 21.6, 20.0, 18.3, 18.2 (6C), 12.2 (3C), −4.0, −4.6 ppm.

HRMS (FAB+): *m*/*z* calcd for $C_{29}H_{54}O_4Si_2Na$ $[M + Na]^+$ 545.3458, found 545.3455.

Compound 11



To a solution of bicyclic compound **10** (66.1 mg, 0.126 mmol) in dry THF (4 mL) under N_2 atmosphere were added CuI (6.0 mg, 0.032 mmol) and Me_2S (400 μL) and the resulting mixture was cooled to −30 °C. After the dropwise addition of vinylmagnesium bromide solution (1 M in THF, 200 μL, 0.20 mmol), the resulting mixture was stirred at −30 °C for 1 h. After another addition of vinylmagnesium bromide solution (1 M in THF, 500 μL, 0.50 mmol), the mixture was

stirred at −30 °C for 20 min and at 0 °C for an additional 1 h. The reaction was quenched with *satd. aq.* NH_4Cl (10 mL). The resultant mixture was partitioned between ether (4 x 10 mL) and H_2O (5 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. Purification by column chromatography (silica gel, EtOAc/*n*-hexane = 1/19 to 1/4) afforded impure alkene **S13** (27.5 mg) and alcohol **S14** (28.9 mg, 0.0601 mmol, 48%).

To a solution of alkene **S13** (27.5 mg) in dry THF (5 mL) at −20 °C under N_2 atmosphere was added *in situ* prepared thexylborane solution (1 M, 0.11 mL, 0.11 mmol). After stirring at −20 °C for 1 h, the reaction was quenched with 30% *aq.* H_2O_2 (300 μl) and 3M *aq,* NaOH (300 μL). The resulting mixture was stirred at room temperature for 1 h and then partitioned between ether (4 x 10 mL) and brine (10 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by column chromatography (silica gel, EtOAc/*n*-hexane = 1/9 to 1/4) to afforded alcohol **11** (19.3 mg, 0.0379 mmol, 30% in two steps, dr 1.8:1) as a colorless oil.

Data of compound **11**

R_f: 0.24 (silica gel, EtOAc/*n*-hexane = 2:8).

IR (neat on KRS-5): 3357, 2943, 2893, 2864, 1717, 1463, 1388, 1362, 1342, 1251, 1082, 1007, 916, 882, 861, 837, 815, 775, 681 cm^{-1}.

^{1}H NMR (500 MHz, $CDCl_3$):

Diastereomer 1: δ 5.57 (d, *J* = 3.5 Hz, 1H), 5.29 (d, *J* = 1.2 Hz, 1H), 4.27 (d, *J* = 11.6 Hz, 2H), 4.08 (d, *J* = 12.4 Hz, 1H), 3.71 (t, *J* = 6.6 Hz, 2H), 2.51–2.31 (m, 2H), 2.17–1.79 (m, 3H), 1.69 (s, 3H), 1.51–1.42 (m, 1H), 1.39–1.29 (m, 2H), 1.29–1.16 (m, 2H), 1.16–1.00 (m, 21H), 0.91 (s, 9H), 0.10 (s, 3H), 0.09 (s, 3H) ppm.

Diastereomer 2: δ 5.47 (s, 1H), 5.33 (d, *J* = 1.3 Hz, 1H), 4.27 (d, *J* = 11.6 Hz, 2H), 4.08 (d, *J* = 12.4 Hz, 1H), 3.71 (t, *J* = 6.6 Hz, 2H), 2.51–2.31 (m, 2H), 2.17–1.79 (m, 3H), 1.69 (s, 3H), 1.51–1.42 (m, 1H), 1.39–1.29 (m, 2H), 1.29–1.16 (m, 2H), 1.16–1.00 (m, 21H), 0.91 (s, 9H), 0.10 (s, 3H), 0.09 (s, 3H) ppm.

^{13}C NMR (126 MHz, $CDCl_3$):

Diastereomer 1: δ 139.0, 137.4, 129.7, 126.7, 73.1, 65.7, 61.4, 39.9, 39.1, 36.4, 35.5, 33.8, 31.2, 26.1 (3C), 20.1, 18.4, 18.3 (6C), 12.2 (3C), −4.0, −4.6 ppm.

Diastereomer 2: δ 139.0, 137.2, 129.7, 127.1, 73.1, 65.5, 60.9, 39.9, 39.7, 39.4, 36.4, 36.2, 33.8, 26.1 (3C), 20.1, 18.4, 18.3 (6C), 12.2 (3C), −4.0, −4.6 ppm.

HRMS (FAB+): *m*/*z* calcd for $C_{29}H_{56}O_3Si_2Na$ $[M + Na]^+$ 531.3666, found 531.3661.

Data of Compound **S14**

R_f: 0.30 (silica gel, EtOAc/*n*-hexane = 2:8).

IR (neat on KRS-5): 3330, 2944, 2893, 2865. 1463, 1251, 1083, 1047, 1011, 882, 860, 837, 775, 682 cm^{-1}.

^{1}H NMR (500 MHz, $CDCl_3$):

Diastereomer 1: δ 5.81 (s, 1H), 5.29 (s, 1H), 4.44–4.20 (m, 3H), 4.14 (dd, *J* = 12.9, 12.9 Hz, 1H), 2.40 (m, 1H), 2.29–2.08 (m, 2H), 2.02–1.84 (m, 1H), 1.69 (br-s, 4H), 1.51–1.20 (m, 2H), 1.19–0.98 (m, 21H), 0.91 (s, 9H), 0.09 (m, 6H) ppm.

Diastereomer 2: δ 5.62 (s, 1H), 5.30 (s, 1H). 4.44–4.20 (m, 3H), 4.14 (dd, *J* = 12.9, 12.9 Hz, 1H), 2.40 (m, 1H), 2.29–2.08 (m, 2H), 2.02–1.84 (m, 1H), 1.69 (br-s, 4H), 1.51–1.20 (m, 2H), 1.19–0.98 (m, 21H), 0.91 (s, 9H), 0.09 (m, 6H) ppm.

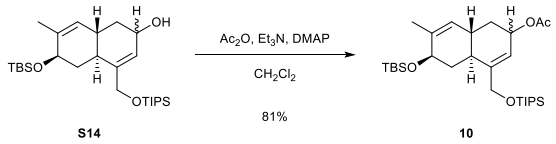
^{13}C NMR (126 MHz, $CDCl_3$):

Diastereomer 1: δ 143.7, 137.6, 128.9, 126.1, 72.8, 64.9, 64.7, 40.0, 38.9, 36.1, 34.1, 26.1 (3C), 20.0, 18.4, 18.3 (6C), 12.2 (3C), −4.0, −4.6 ppm.

Diastereomer 2: δ 141.5, 137.6, 128.6, 123.4, 72.8, 68.9, 64.7, 39.9, 39.6, 38.2, 35.8, 26.1 (3C), 20.0, 18.4, 18.3 (6C), 12.2 (3C), −4.0, −4.6 ppm.

HRMS (FAB+): *m*/*z* calcd for $C_{27}H_{52}O_3Si_2Na$ $[M + Na]^+$, 503.3353, found. 503.3356

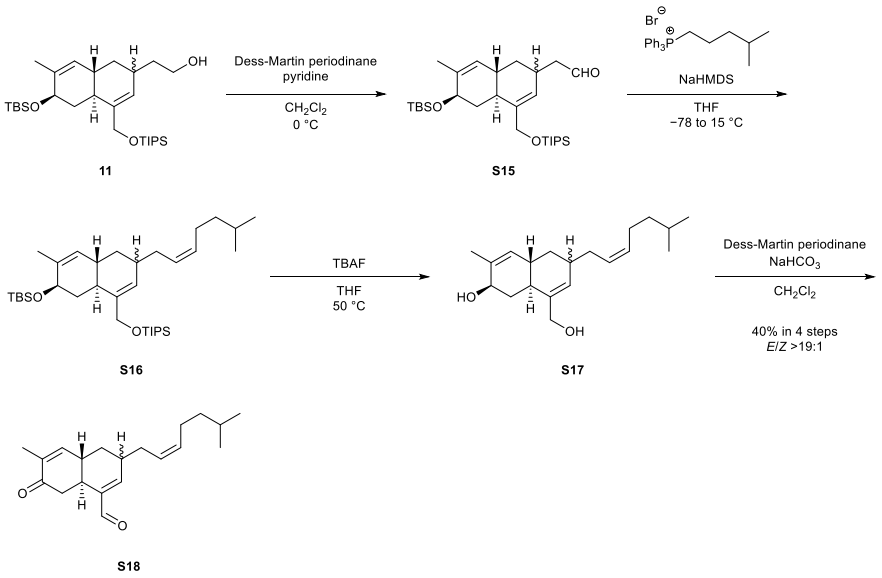
Compound 10 from S14



S14 **10**

To a solution of **S14** (150 mg, 0.312 mmol) in CH_2Cl_2 (5 mL) were added Et_3N (130 μL, 0.94 mmol) and Ac_2O (60 μL, 0.63 mmol) and DMAP (4.7 mg, 0.038 mmol). The reaction mixture was stirred at room temperature for 19 h. The reaction was quenched with *satd. aq.* $NaHCO_3$ (5 mL). The aqueous layer was separated and extracted with $CHCl_3$ (3 x 5 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by column chromatography (silica gel, EtOAc/*n*-hexane = 1/49 to 1/9) to afford **10** (133 mg, 0.254 mmol, 81%) as a colorless oil.

Compound S18



11 **S15** **S16** **S17** **S18**

To a solution of alcohol **11** (71.3 mg, 0.140 mmol) in CH_2Cl_2 (10 mL) at 0 °C were added pyridine (0.10 mL) and Dess-Martin periodinane (122 mg, 0.288 mmol). After stirring at 0 °C for 20 min, the reaction mixture was stirred at room temperature for 3.5 h. The reaction was quenched with *satd. aq.* $Na_2S_2O_3$ (10 mL) and *satd. aq.* $NaHCO_3$ (10 mL) and stirred for 15 min. The aqueous layer was separated and extracted with $CHCl_3$ (3 x 10 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. Purification by column chromatography (silica gel, EtOAc/*n*-hexane = 1/19 to 1/9) afforded impure aldehyde **S15** (49.9 mg).

To a suspension of (4-methylpentyl)(triphenyl)phosphonium bromide (236 mg, 0.552 mmol) in dry THF (7 mL) at −78 °C under N_2 atmosphere was added NaHMDS solution (1 M in THF, 0.42 mL, 0.42 mmol). After the resulting mixture was stirred at −78 °C for 30 min, aldehyde **S15** (49.9 mg) in dry THF (3 mL) was added via cannula. After stirring at −78 °C for 2 h, the reaction mixture was gradually warmed up to 15 °C over 17 h. The reaction was quenched with *satd. aq.* NH_4Cl (10 mL). The resultant mixture was partitioned between ether (4 x 10 mL) and H_2O (5 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. Purification by column chromatography (silica gel, EtOAc/*n*-hexane = 1/99 to 1/19) afforded impure alkene **S16** (88.3 mg).

To a solution of alkene **S16** (88.3 mg) in dry THF (10 mL) under N_2 atmosphere was added TBAF solution (1M in THF, 0.70 mL, 0.70 mmol). After stirring at room temperature for 1 h, the reaction mixture was stirred at 50 °C for 16 h. After cooling to room temperature, the reaction was quenched with *satd. aq.* NH_4Cl (10 mL). The resultant mixture was partitioned between EtOAc (4 x 10 mL) and H_2O (5 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. Purification by column chromatography (silica gel, EtOAc/*n*-hexane = 1/1 to 9/1) afforded impure diol **S17** (24.3 mg).

To a solution of **S17** (24.3 mg) in CH_2Cl_2 (10 mL) were added $NaHCO_3$ (122 mg, 1.45 mmol) and Dess-Martin periodinane (240 mg, 0.566 mmol). After stirring at room temperature, the reaction was quenched with *satd. aq.* $Na_2S_2O_3$ (5 mL) and *satd. aq.* $NaHCO_3$ (5 mL). The resultant mixture was stirred for 20 min. The aqueous layer was separated and extracted with $CHCl_3$ (3 x 5 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by column chromatography twice (silica gel, EtOAc/*n*-hexane = 3/7 and then 1/19 to 2/3) to afford **S18** (16.9 mg, 0.0563 mmol, 40% in four steps, *E*/*Z* >19:1) as a pale yellow oil.

R_f: 0.45 (silica gel, EtOAc/*n*-hexane = 3:7).

IR (neat on KRS-5): 3005, 2953, 2924, 2867, 2713, 2360, 2342, 1675, 1625, 1450, 1381, 1245, 1189, 1147, 1116, 924, 881, 712 cm^{-1}.

^{1}H NMR (500 MHz, $CDCl_3$):

Diastereomer 1: δ 9.42 (s, 1H), 6.77 (m, 1H), 6.58 (s, 1H), 5.61–5.50 (m, 1H), 5.46–5.33 (m, 1H), 3.68 (m, 1H), 2.72–2.57 (m, 2H), 2.43–2.15 (m, 3H), 2.12–1.75 (m, 4H), 1.80 (s, 3H), 1.65–1.49 (m, 1H), 1.30–1.14 (m, 3H), 0.88 (m, 6H) ppm.

Diastereomer 2: δ 9.42 (s, 1H), 6.66 (s, 1H), 6.63 (s, 1H) 5.61–5.50 (m, 1H), 5.46–5.33 (m, 1H), 3.68 (m, 1H), 2.72–2.57 (m, 2H), 2.43–2.15 (m, 3H), 2.12–1.92 (m, 4H), 1.92–1.75 (m, 1H), 1.80 (s, 3H), 1.65–1.49 (m, 1H), 1.30–1.14 (m, 3H), 0.88 (m, 6H) ppm.

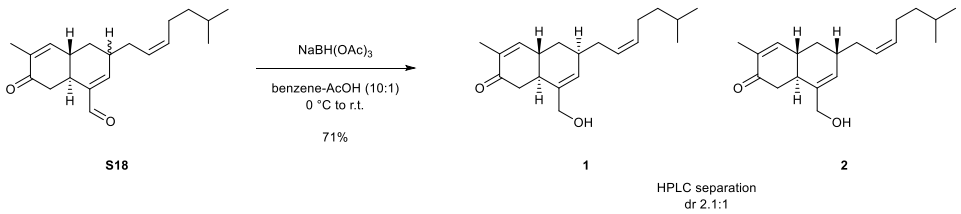
^{13}C NMR (126 MHz, $CDCl_3$):

Diastereomer 1: δ 199.5, 193.5, 156.0, 148.9, 141.8, 136.6, 133.3, 126.3, 41.6, 39.5, 39.0, 36.5, 35.4, 32.1, 31.2, 27.8, 25.5, 22.7, 22.7, 15.9 ppm.

Diastereomer 2: δ 199.4, 193.6, 156.3, 148.7, 141.9, 136.4, 133.4, 125.6, 41.8, 39.6, 39.4, 39.0, 38.9, 33.7, 32.5, 27.8, 25.5, 22.7, 22.7, 15.9 ppm.

HRMS (FAB+): *m*/*z* calcd for $C_{20}H_{29}O_2$ $[M + H]^+$ 201.2168, found 301.2170.

Compound 1 and 2



To a solution of **S18** (14.4 mg, 0.0479 mmol) in benzene (10 mL) was added acetic acid (1.0 mL) and the resulting solution was cooled to 0 °C. After the addition of $NaBH(OAc)_3$ (38.8 mg, 0.183 mmol), the reaction mixture was stirred at 0 °C for 1.5 h. Additional $NaBH(OAc)_3$ (56.4 mg, 0.266 mmol) was added and stirred at 0 °C for 30 min and then at room temperature for 3 h. The reaction was quenched with *satd. aq.* $NaHCO_3$ (10 mL). The resultant solution was stirred

for 10 min. The aqueous layer was separated and extracted with EtOAc (3 x 10 mL). The combined organic layers were dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by column chromatography (silica gel, EtOAc/*n*-hexane = 1/9 to 2/3) to afford a mixture of **1** and **2** (10.3 mg, 0.0341 mmol, 71%) as a colorless oil. Compounds **1** and **2** were purified by HPLC (column, YMC-Pack ODS-A AA12S05-2510WT; solvent, CH_3CN:H_2O = 7:3; flow rate, 3.0 mL/min; UV detector, 254 nm; retention time of **1**, 25.8 min; retention time of **2**, 29.7 min). $\mathbf{R_f}$: 0.18 (silica gel, EtOAc/*n*-hexane = 3:7).

Diastereomer 1

$\mathbf{[\alpha]_D^{18}}$: 101.1 (*c* 0.28, $CHCl_3$).

IR (neat on KRS-5): 3414, 3005, 2954, 2924, 2868, 1673, 1450, 1365, 1250, 1021 cm^{-1}.

^{1}H NMR (500 MHz, CD_3OD): δ 6.71 (s, 1H). 5.71 (s, 1H), 5.53–5.45 (m, 1H), 5.45–5.38 (m, 1H), 4.11 (d, *J* = 12.5 Hz, 1H), 3.94 (d, *J* = 12.5 Hz, 1H), 2.92 (dd, *J* = 16.4, 3.5 Hz,1H), 2.50 (m, 1H), 2.39 (m, 1H), 2.32 (m, 1H), 2.21–2.12 (m, 3H), 2.06 (dd, *J* = 7.5, 7.5 Hz, 2H), 1.88 (d, *J* = 14.4 Hz, 1H), 1.76 (m,1H), 1.62–1.51 (m, 2H), 1.29–1.20 (m, 2H), 0.90 (d, *J* = 1.7 Hz, 3H), 0.89 (d, *J* = 1.7 Hz, 3H) ppm.

^{13}C NMR (126 MHz, CD_3OD): δ 202.5, 153.1, 139.0, 136.8, 132.8, 130.1, 129.0, 64.9, 42.8, 42.4, 40.3, 36.9, 36.5, 34.2, 32.8, 29.0, 26.5, 23.1, 23.1, 15.9 ppm.

HRMS (FAB+): *m*/*z* calcd for $C_{20}H_{31}O_2$ $[M + H]^+$ 303.2324, found 303.2321.

Diastereomer 2

$\mathbf{[\alpha]_D^{20}}$: 23.6 (*c* 0.17, $CHCl_3$).

IR (neat on KRS-5): 3430, 2954, 2925, 2852, 1672, 1450, 1365, 1240, 1018 cm^{-1}.

^{1}H NMR (500 MHz, pyridine-d_5): δ 6.55 (dd, *J* = 1.5, 1.5 Hz, 1H), 6.35 (dd, *J* = 6.0, 4.8 Hz, 1H), 5.83 (s, 1H), 5.53 (m, 2H), 4.41 (dd, *J* = 13.0, 5.0 Hz, 1H), 4.31 (dd, *J* =13.0, 3.7 Hz, 1H), 3.35 (dd, *J* = 16.2, 3.5 Hz,1H), 2.72 (m, 1H), 2.41 (br-s, 1H), 2.35 (m, 1H), 2.29 (dd, *J* = 16.2, 14.3 Hz, 1H), 2.26–2.18 (m, 1H), 2.16–2.05 (m, 3H), 1.96 (br-dd, *J* = 11.4, 5.4 Hz, 1H), 1.85 (dd, *J* = 2.3, 1.4 Hz, 3H), 1.54 (m, 1H), 1.26 (m, 2H), 1.15 (dd, *J* = 24.0, 12.7 Hz, 1H), 0.87 (d, *J* = 6.7 Hz, 6H) ppm.

^{13}C NMR (126 MHz, pyridine-d_5): δ 199.9, 150.9, 140.0, 135.8, 132.2, 128.1, 128.0, 64.4, 42.8, 42.1, 40.8, 39.5, 38.0, 35.5, 34.5, 28.3, 26.0, 23.1, 23.1, 16.3 ppm.

HRMS (FAB+): *m*/*z* calcd for $C_{20}H_{31}O_2$ $[M + H]^+$ 303.2324, found 303.2318.

PKC binding assay (competitive inhibition test using [^{3}H]PDBu)

The binding of [^{3}H]PDBu to PKC C1 domain peptides was evaluated by the procedure of Sharkey and Blumberg[19] with modifications as reported previously[20] with 50 mM Tris-maleate buffer (pH 7.4 at 4 °C), 20 nM (for α-, β-, and γ-C1A) or 13.8 nM (for δ-, ε-, η-, and θ-C1B) of

peptides, 20 nM [^{3}H]PDBu (17.16 Ci/mmol, Perkin-Elmer Life Science), 50 μg/ml, 1,2-dioleoyl-*sn*-glycero-3-phospho-L-serine (Sigma-Aldrich), 3 mg/mL bovine γ-globulin (Sigma-Aldrich), and various concentrations of ligands. Binding affinity was evaluated on the basis of the concentration required to cause 50% inhibition of the specific binding of [^{3}H]PDBu, IC_{50}, which was calculated by logit analysis using Microsoft Excel. The inhibition constant, K_i, was calculated by the equation of Goldstein and Barrett, $K_i = IC_{50}/(2[L_{50}]/[L_0] - 1 + [L_{50}]/K_d)$, where $[L_{50}]$ and $[L_0]$ are the free concentration of [^{3}H]PDBu at 50% and 0% inhibition, respectively.[21]

Notes and references

1 B. Ramsundar, P. Eastman, P. Walters and V. Pande, *Deep learning for the life sciences: Applying deep learning to genomics, microscopy, drug discovery, and more,* O'Reilly Media, inc., Sebastopol, California, 2019.

2. S. Kearnes, K. McCloskey, M. Berndl, V. Pande, P. Riley Molecular graph convolutions: moving beyond fingerprints. *J. Comput. Aided Mol. Des.* 2016, **30**, 595. doi: 10.1007/s10822-016-9938-8. Epub 2016 Aug 24. PMID: 27558503; PMCID: PMC5028207.

3 J. Lee, K. C. Han, J. H. Kang, L. L. Pearce, N. E. Lewin, S. Yan, S. Benzaria, M. C. Nicklaus, P. M. Blumberg and V. E. Marquez, *J. Med. Chem.*, 2001, **44**, 4309.

4 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminf.*, 2012, **4**, 17.

5 N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *J. Cheminf.*, 2011, **3**, 33.

6 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson., *J. Comput. Chem.*, 2009, **30**, 2785.

7 D. Santos-Martins, L. Solis-Vasquez, A. F. Tillack, M. F. Sanner, A. Koch and S. Forli, *J. Chem. Theory Comput.*, 2021, **17**, 1060.

8 S. Jo, T. Kim, V. G. Iyer and W. Im, *J. Comput. Chem.*, 2008, **29**, 1859.

9 J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. Brooks, A. D. MacKerell, J. B. Klauda and W. Im, *J. Chem. Theory Comput.*, 2016, **12**, 405.

10 S. Jo, T. Kim and W. Im., *PLoS ONE*, 2007, **2**, e880.

11 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindah, *SoftwareX*. 2015, **1–2**, 19.

12 D. A. Case, H. M. Aktulga, K. Belfon, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham, III, G. A. Cisneros, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, G. Giambasu, M. K. Gilson, H. Gohlke, A. W. Goetz, R. Harris, S. Izadi, S. A. Izmailov, C. Jin, K. Kasavajhala, M. C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K. M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K. A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, N. R. Skrynnikov, J. Smith, J. Swails, R. C. Walker, J. Wang, H. Wei, R. M. Wolf, X. Wu, Y. Xue, D. M. York, S. Zhao, and P. A. Kollman, *Amber 2021*, University of California, San Francisco, 2021.

13 M. A. Lomize, I. D. Pogozheva, H. Joo, H. I. Mosberg and A. L. Lomize, *Nucleic Acids Res.*, 2012, **40**, D370.

14 The PyMOL Molecular Graphics System, Schrödinger, LLC.
15 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696.
16 M. B. Peters, Y. Yang, B. Wang, L. Füsti-Molnár, M. N. Weaver, K. M. Merz, *J. Chem. Theory Comput.*, 2010, **6**, 2935.
17 A. Gonda, K. Takada, R. C. Yanagita, S. Dan and K. Irie, *Biosci. Biotechnol. Biochem.*, 2021, **85**, 168.
18 M. Xuan, I. Paterson and S. M. Dalby, *Org. Lett.*, 2012, **14**, 5492.
19 N. A. Sharkey and P. M. Blumberg, *Cancer Res.*, 1985, **45**, 19.
20 K. Irie, K. Oie, A. Nakahara, Y. Yanai, H. Ohigashi, P. A. Wender, H. Fukuda, H. Konishi and U. Kikkawa, *J. Am. Chem. Soc.*, 1998, **120**, 9159.
21 A. Goldstein and R. W. Barrett, *Mol. Pharmacol.*, 1987, **31**, 603.

2. Comparison of K_i values of PKC ligands

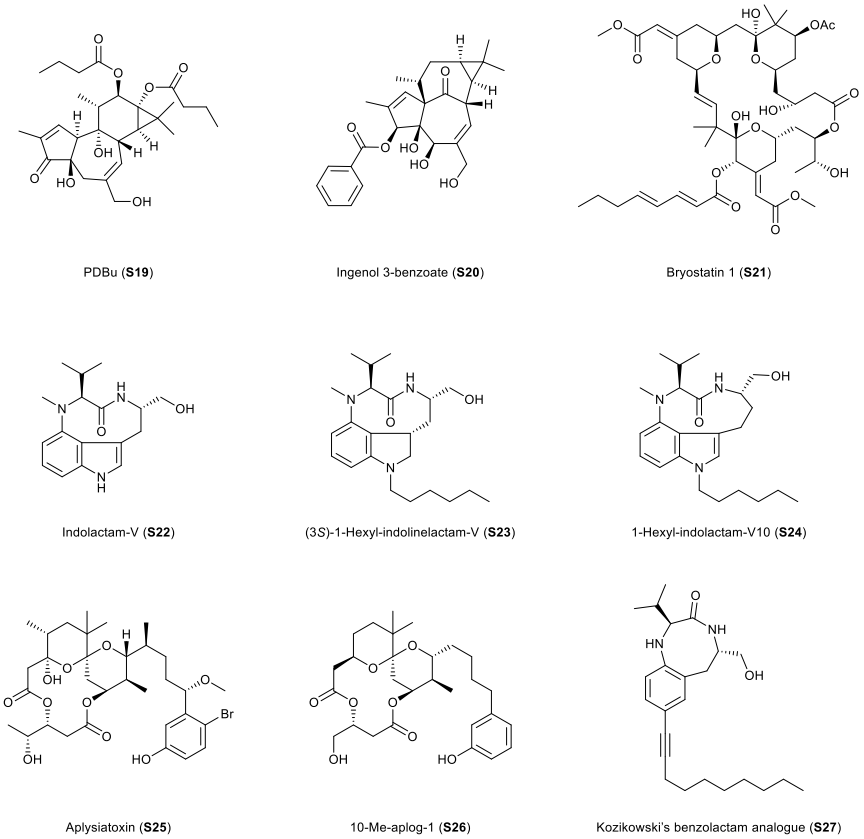


Figure S5 Structures of various compounds whose K_i values of for inhibition of the specific [^{3}H]PDBu binding to PKC C1 peptides and whole-PKC.

Table S2 K_i values of **1**, **2** and various reported compounds for inhibition of the specific [^{3}H]PDBu binding to PKC C1 peptides and whole-PKC.

PKC subfamilies	C1 peptides	K_i values (nM)[1]										
		1	**2**	S19 (K_d)[2–5]	S20[5]	S21[5,6]	S22[7]	S23[8]	S24[9]	S25[10]	S26[10]	S27[11]
Conventional	α-C1A	62 (11)	1800 (290)	1.1	0.78	7.4	20.8	550	6700	0.26	4.7	–
	α-C1B	5200 (40)	NT[12]	5.3	3.7	7.3	4000	>10000	>10000	–	–	–
	whole-PKCα	NT	NT	0.15, 0.46	–	0.81	20	–	–	–	–	14.7
	β-C1A	77 (5.6)	1800 (90)	1.3	1.1	12.7	18.9	1200	9200	0.17	12	–
	β-C1B	NT	NT	1.3	0.48	0.94	137	250	440	–	–	–
	whole-PKCβ	NT	NT	0.14, 0.54	–	2.2	31	–	–	–	–	17.4
	γ-C1A	150 (3)	4100 (110)	1.5	0.34	3.4	138	1800	11000	0.38	5.5	–
	γ-C1B	NT	NT	1.2	0.40	2.1	213	1600	420	–	–	–
	whole-PKCγ	NT	NT	0.37, 1.80	–	2.2	91	–	–	–	–	40.7
Novel	δ-C1A	1200 (50)	NT	51.9	10.4	5.3	1900	>10000	–	–	–	–
	δ-C1B	460 (23)	>4000	0.53	0.17	0.60	8.3	16	15	0.20	0.46	–
	whole-PKCδ	NT	NT	0.71, 0.76	–	1.1	12	–	–	–	–	122
	ε-C1A	NT	NT	5.6	2.7	2.1	4110	>10000	–	–	–	–
	ε-C1B	590 (24)	>7000	0.81	0.14	15.6	7.7	14	29	0.63	2.0	–
	whole-PKCε	NT	NT	0.63, 0.56	–	3.0	6.6	–	–	–	–	142
	η-C1A	NT	NT	4.3	1.6	3.8	3770	>10000	–	–	–	–
	η-C1B	280 (21)	>4000	0.45	0.11	2.1	5.5	12	25	0.11	0.45	–
	whole-PKCη	NT	NT	0.58, 0.95	-	2.2	4.8	–	–	–	–	–
	θ-C1B	520 (45)	>6000	0.72	0.24	1.4	8.7	12	12	0.11	0.54	–
	whole-PKCθ	NT	NT	–	–	1.5	–	–	–	–	–	–
Ratio (K_i for PKCα-C1A/K_i for PKCδ-C1B)		0.13	–	2.1	4.6	12	2.5	34	450	1.3	10	0.12[13]

Notes and references

1 Values in parentheses represent the standard deviation from at least two separate experiments.

2 K_i values for inhibition of the specific [^{3}H]PDBu binding to PKC C1 peptides, see: M. Shindo, K. Irie, A. Nakahara, H. Ohigashi, H. Konishi, U. Kikkawa, H. Fukuda and P. A. Wender, *Bioorg. Med. Chem.*, 2001, **9**, 2073.

3 K_i values for inhibition of the specific [^{3}H]PDBu binding to whole-PKC by Kazanietz *et al*., see: M. G. Kazanietz, L. B. Areces, A. Bahador, H. Mischak, J. Goodnight, J. F. Mushinski and P. M. Blumberg, *Mol. Pharmacol.* 1993, **44**, 298.

4 K_i values for inhibition of the specific [^{3}H]PDBu binding to whole-PKC by Irie *et al*., see: (a) (in Japanese) K. Irie and H. Ohigashi, *J. Syn. Org. Chem. Japan*, 2002, **60**, 563; (b) (in Japanese) K. Irie and H. Ohigashi, *Kagaku to Seibutsu*, 2004, **42**, 54; the reference in English, see reference 5.

5 K. Irie, Y. Nakagawa and H. Ohigashi, *Curr. Pharm. Design*, 2004, **10**, 1371.

6 (a) M. D. Marsden, B. A. Loy, X. Wu, C. M. Ramirez, A. J. Schrier, D. Murray, A. Shimizu, S. M. Ryckbosch, K. E. Near, T. W. Chun, P. A. Wender and J. A. Zack, *PLoS Pathog.*, 2017, **13**, e1006575; (b) M. D. Marsden, X. Wu, S. M. Navab, B. A. Loy, A. J. Schrier, B. A. DeChristopher, A. J. Shimizu, C. T. Hardman, S. Ho, C. M. Ramirez, P. A. Wender and J. A. Zack, *Virology*, 2018, **520**, 83.

7 (a) A. Masuda, K. Irie, Y. Nakagawa and H. Ohigashi, *Biosci. Biotechnol. Biochem*, 2002, **66**, 1615; (b) K. Irie, Y. Nakagawa and H. Ohigashi, *Chem. Rec.*, 2005, **5**, 185.

8 (a) Y. Nakagawa, K. Irie, Y. Komiya, H. Ohigashi and K. Tsuda, *Tetrahedron*, 2004, **60**, 7077; (b) Y. Nakagawa, K. Irie, R. C. Yanagita, H. Ohigashi, K. Tsuda, K. Kashiwagi and N. Saito, *J. Med. Chem.*, 2006, **49**, 2681.

9 R. C. Yanagita, Y. Nakagawa, N. Yamanaka, K. Kashiwagi, N. Saito and K. Irie, *J. Med. Chem.*, 2008, **51**, 46.

10 M. Kikumori, R. C. Yanagita, H. Tokuda, N. Suzuki, H. Nagai, K. Suenaga and K. Irie, *J. Med. Chem.*, 2012, **55**, 5614.

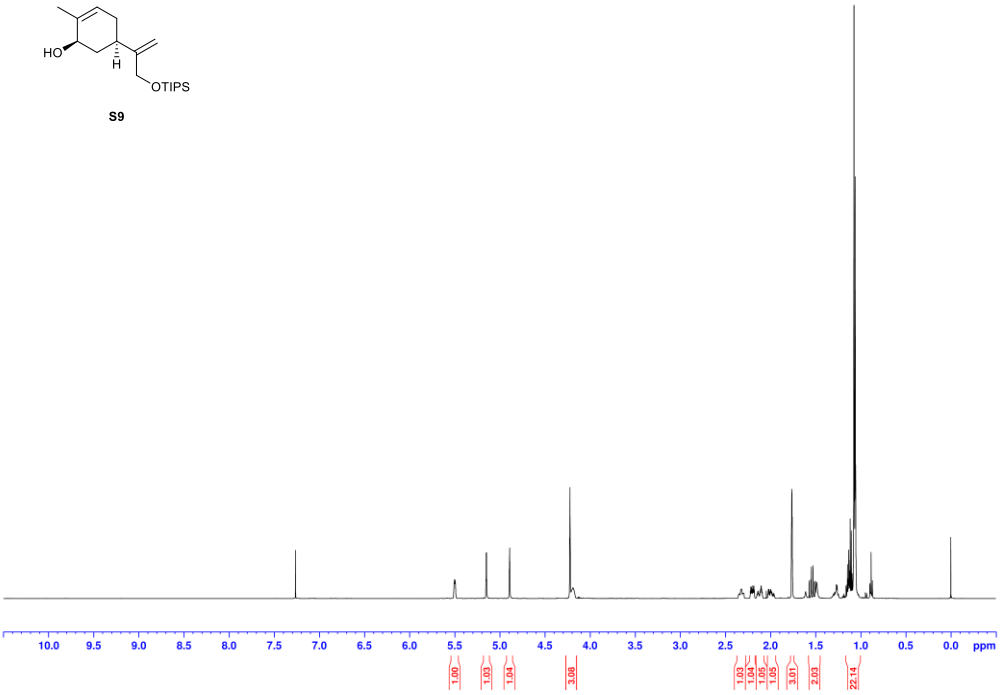
11 A. P. Kozikowski, S. Wang, D. Ma, J. Yao, S. Ahmad, R. I. Glazer, K. Bogi, P. Acs, S. Modarres, N. E. Lewin and P. M. Blumberg, *J. Med. Chem.*, 1997, **40**, 1316.

12 NT: not tested.

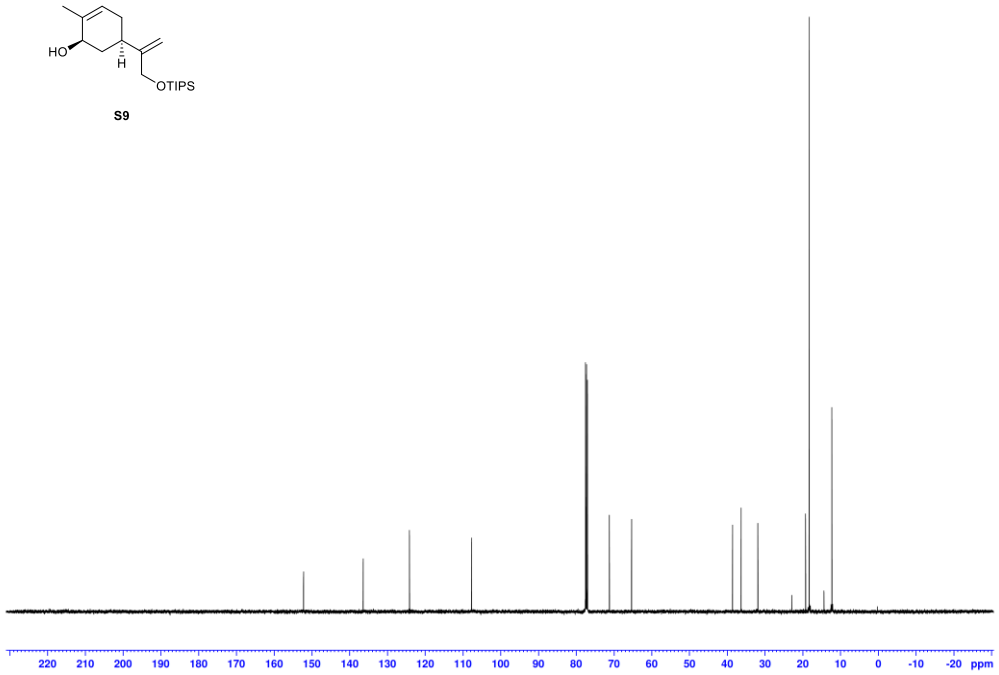
13 Ratio is calculated by "K_i for whole-PKCα/K_i for whole-PKCδ".

3. ^{1}H, ^{13}C and 2D NMR spectra

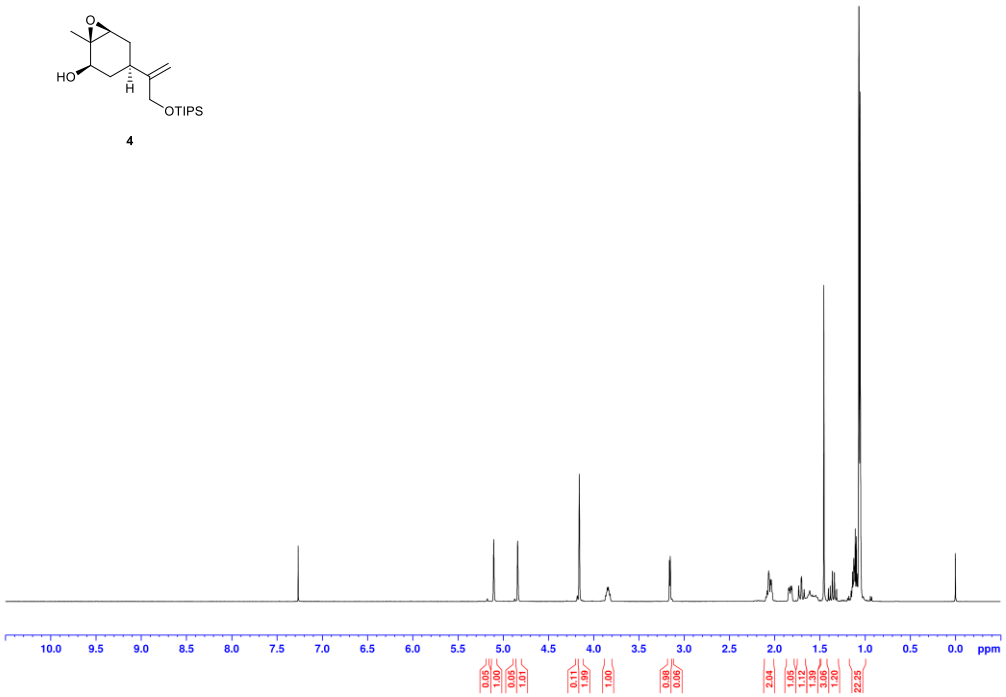
^{1}H NMR spectrum of S9 (500 MHz, 297.0 K, $CDCl_3$, 0.26 M)



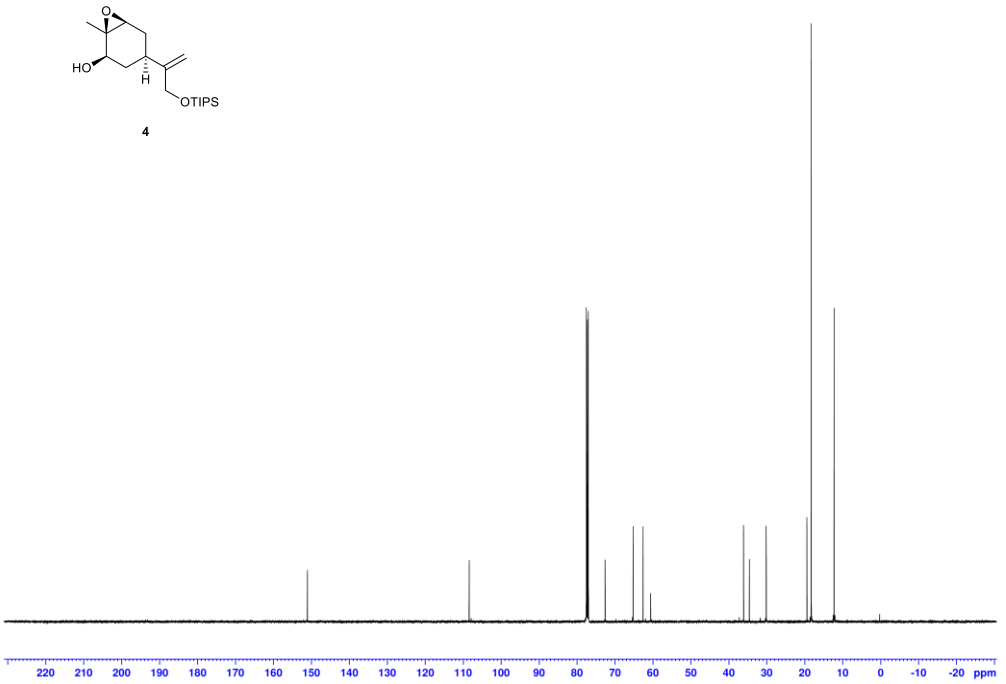
^{13}C NMR spectrum of S9 (126 MHz, 298.0 K, $CDCl_3$, 0.26 M)



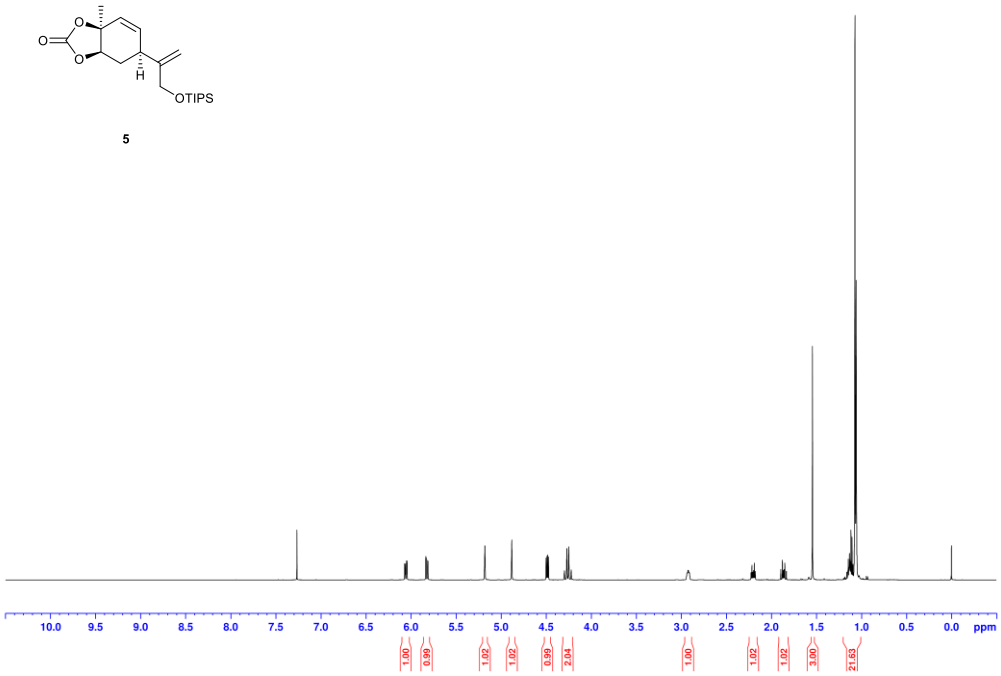
^{1}H NMR spectrum of 4 (500 MHz, 297.0 K, $CDCl_3$, 0.13 M)



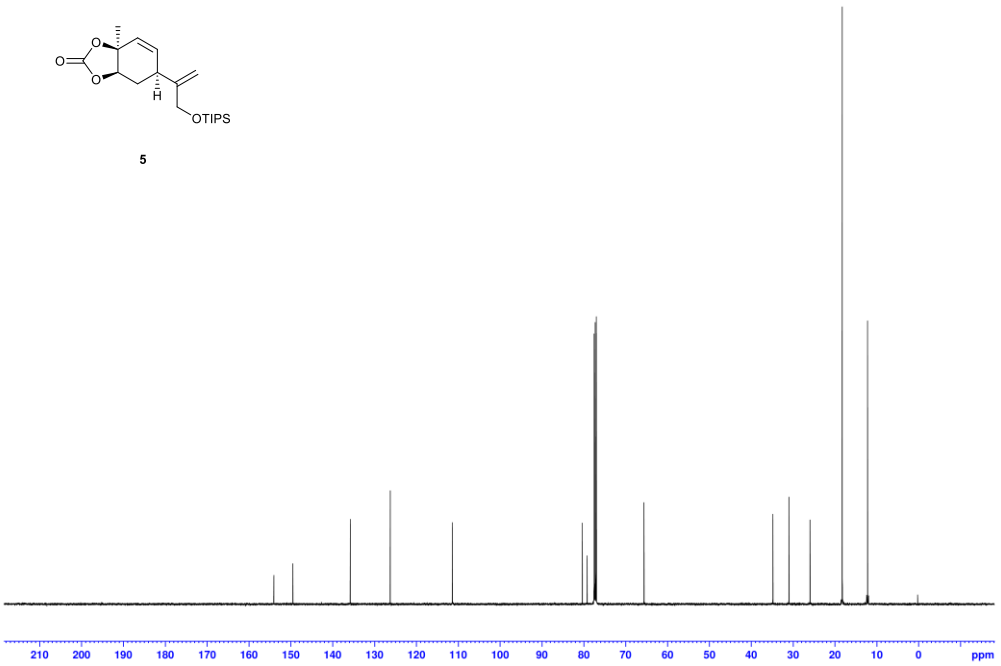
^{13}C NMR spectrum of 4 (126 MHz, 298.0 K, $CDCl_3$, 0.13 M)



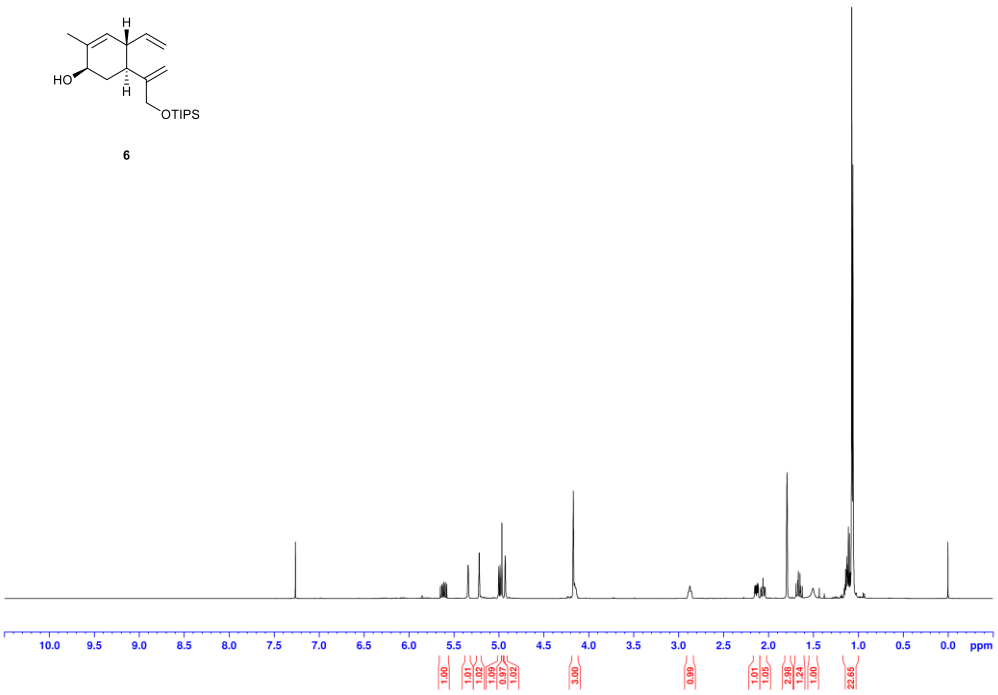
^{1}H NMR spectrum of 5 (500 MHz, 297.0 K, $CDCl_3$, 0.16 M)



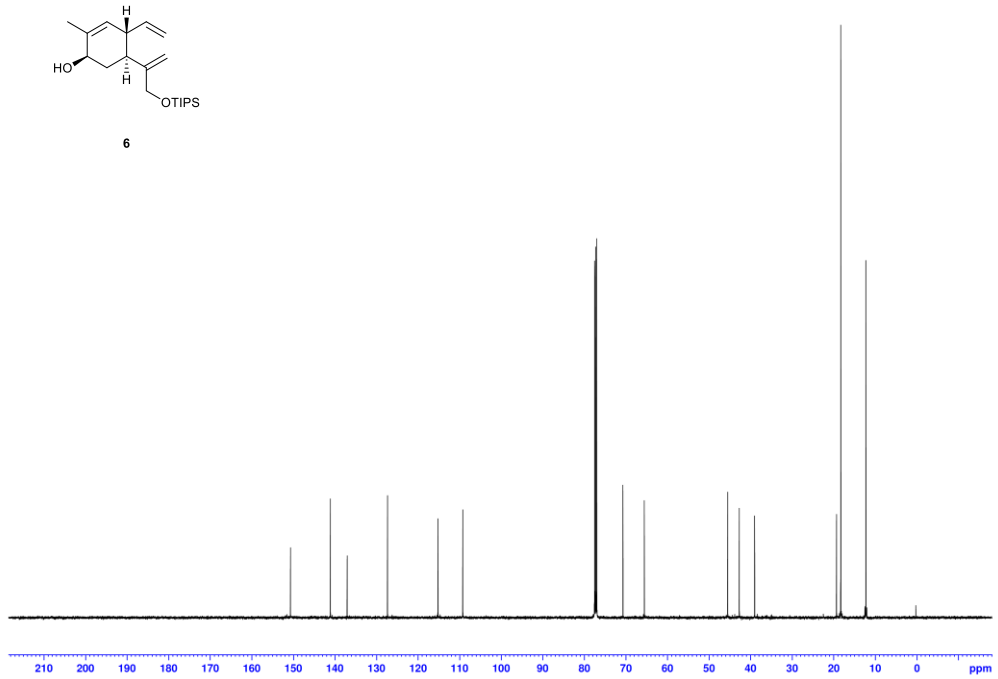
^{13}C NMR spectrum of 5 (126 MHz, 298.0 K, $CDCl_3$, 0.16 M)



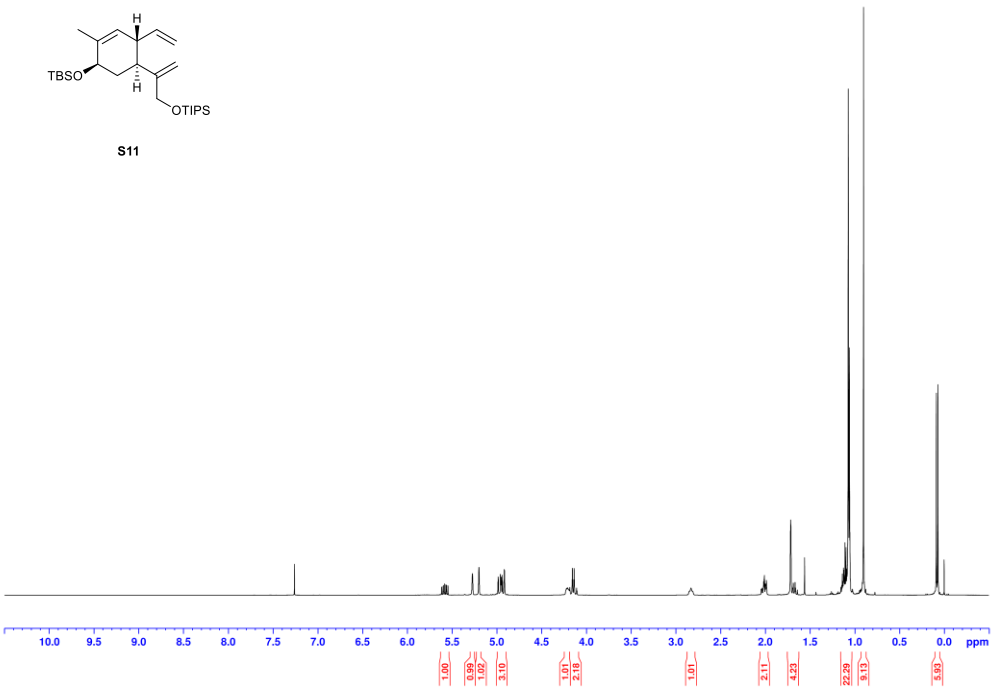
^{1}H NMR spectrum of 6 (500 MHz, 297.0 K, $CDCl_3$, 0.12 M)



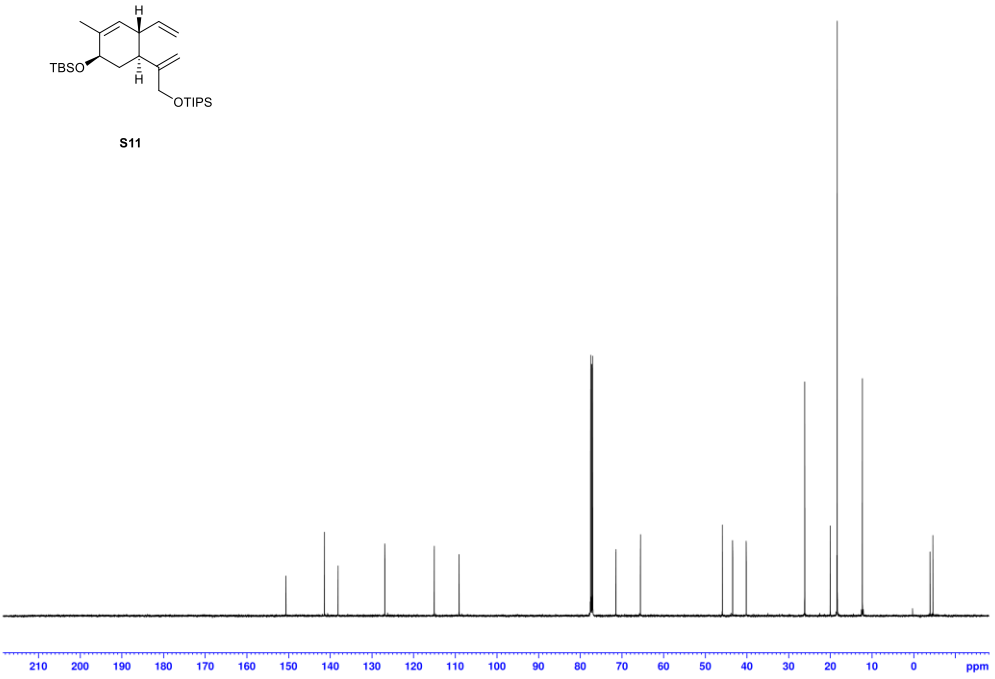
^{13}C NMR spectrum of 6 (126 MHz, 298.0 K, $CDCl_3$, 0.12 M)



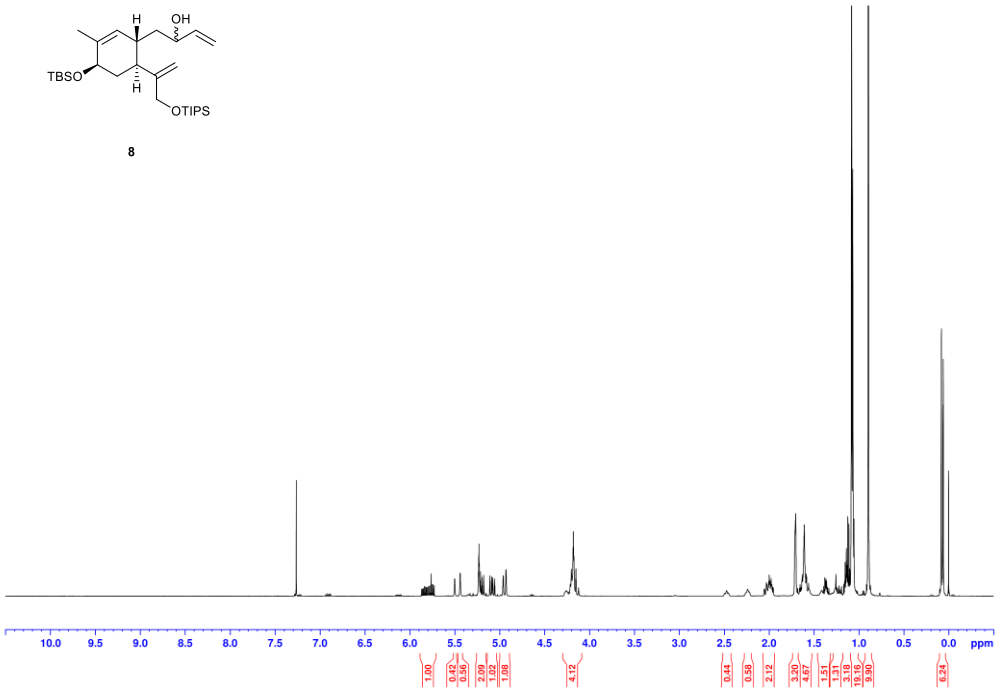
^{1}H NMR spectrum of S11 (500 MHz, 297.0 K, $CDCl_3$, 0.080 M)



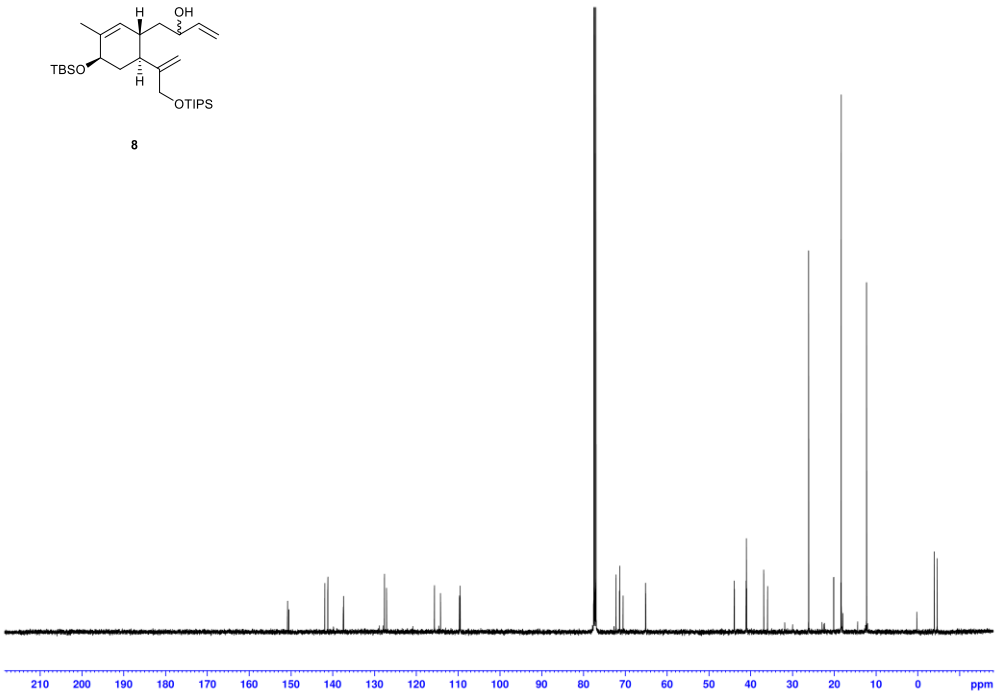
^{13}C NMR spectrum of S11 (126 MHz, 298.0 K, $CDCl_3$, 0.080 M)



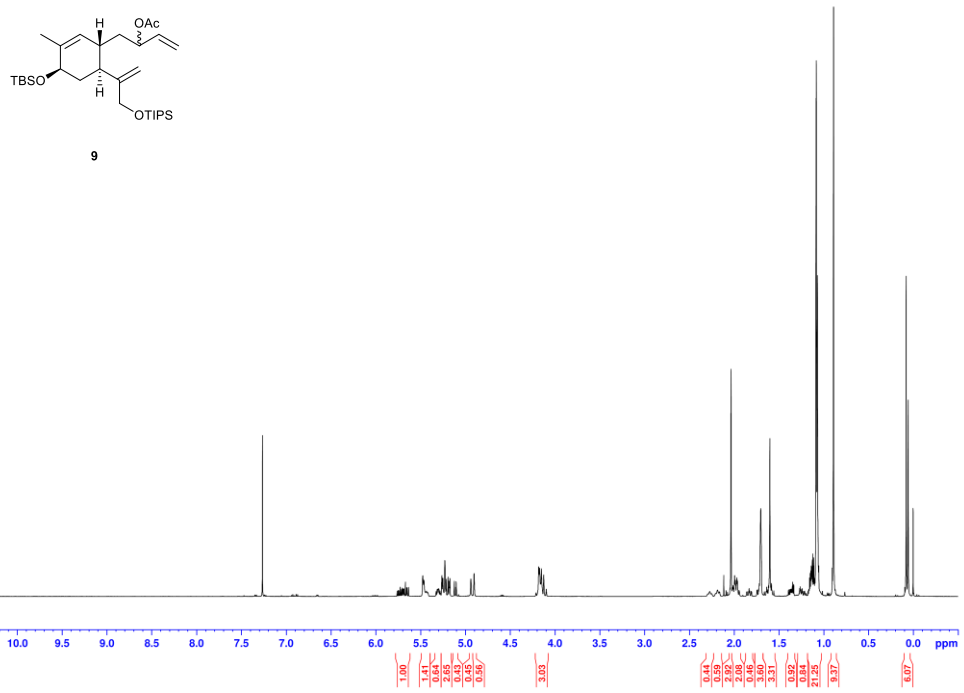
^{1}H NMR spectrum of 8 (500 MHz, 297.0 K, $CDCl_3$, 0.053 M)



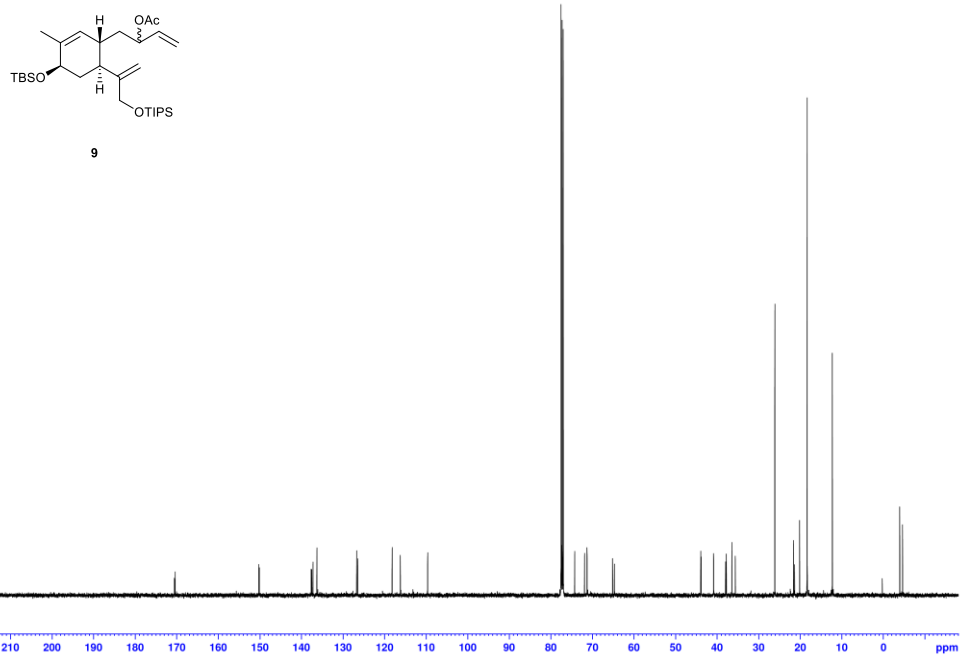
^{13}C NMR spectrum of 8 (126 MHz, 298.0 K, $CDCl_3$, 0.053 M)



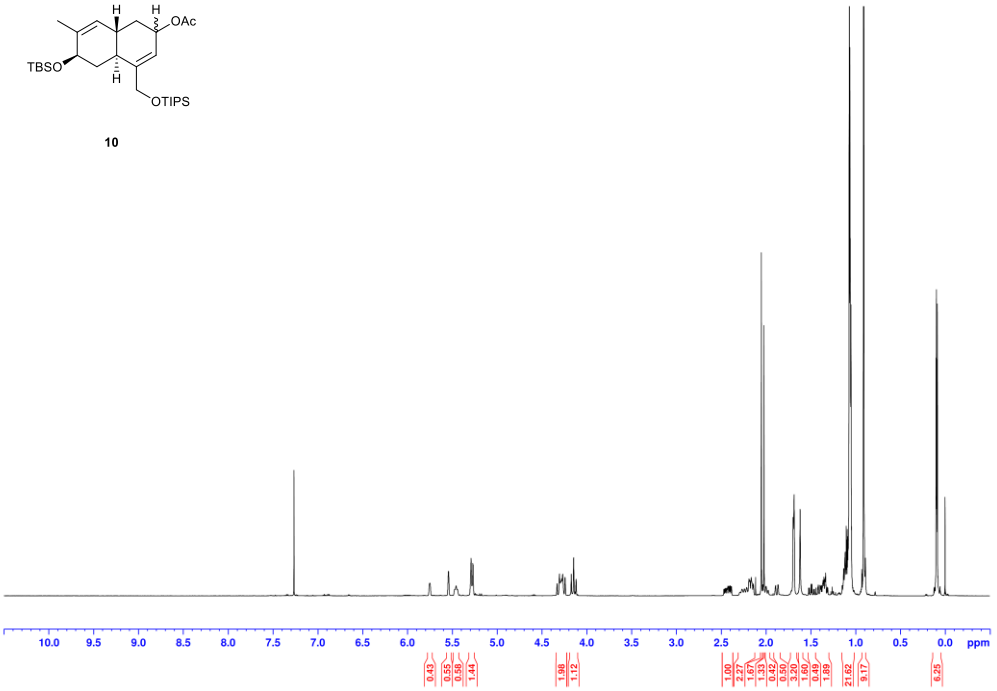
^{1}H NMR spectrum of 9 (500 MHz, 297.0 K, $CDCl_3$, 0.065 M)



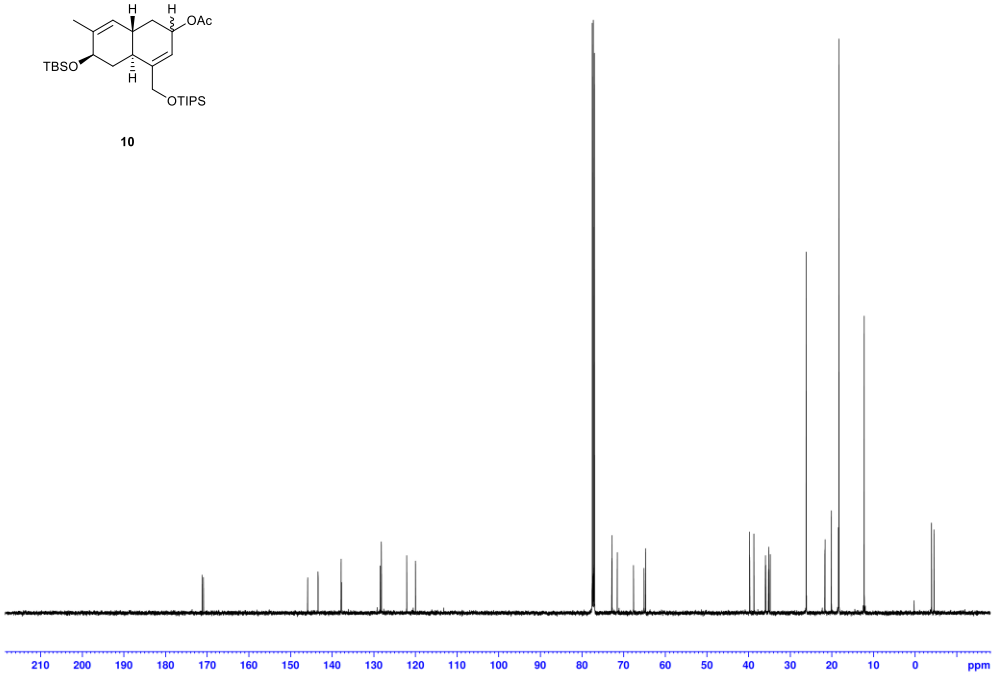
^{13}C NMR spectrum of 9 (126 MHz, 298.0 K, $CDCl_3$, 0.065 M)



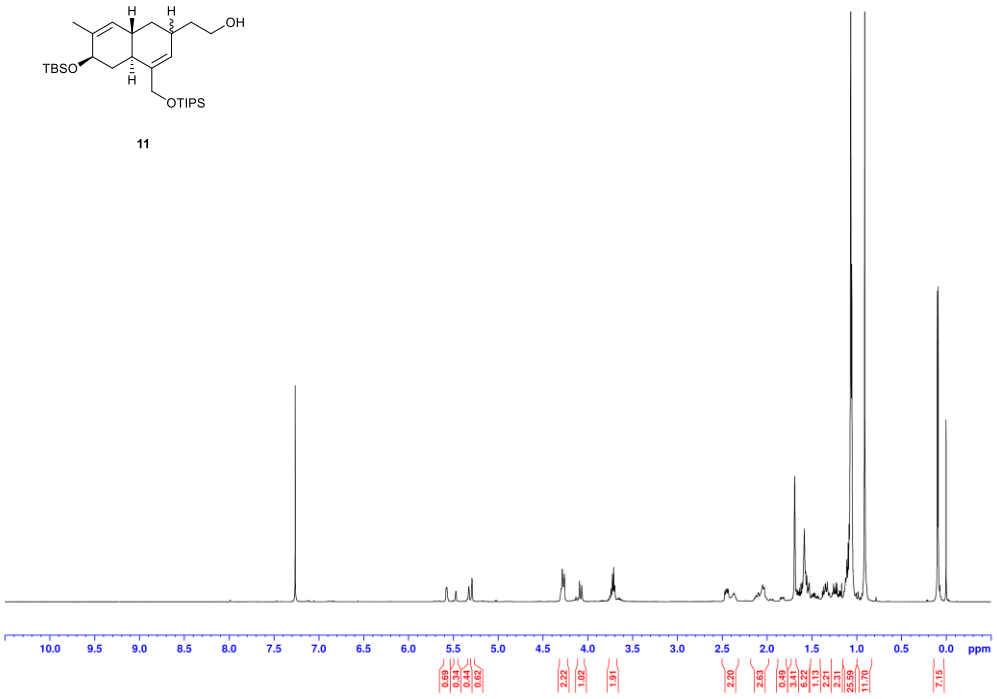
^{1}H NMR spectrum of 10 (500 MHz, 297.0 K, $CDCl_3$, 0.069 M)



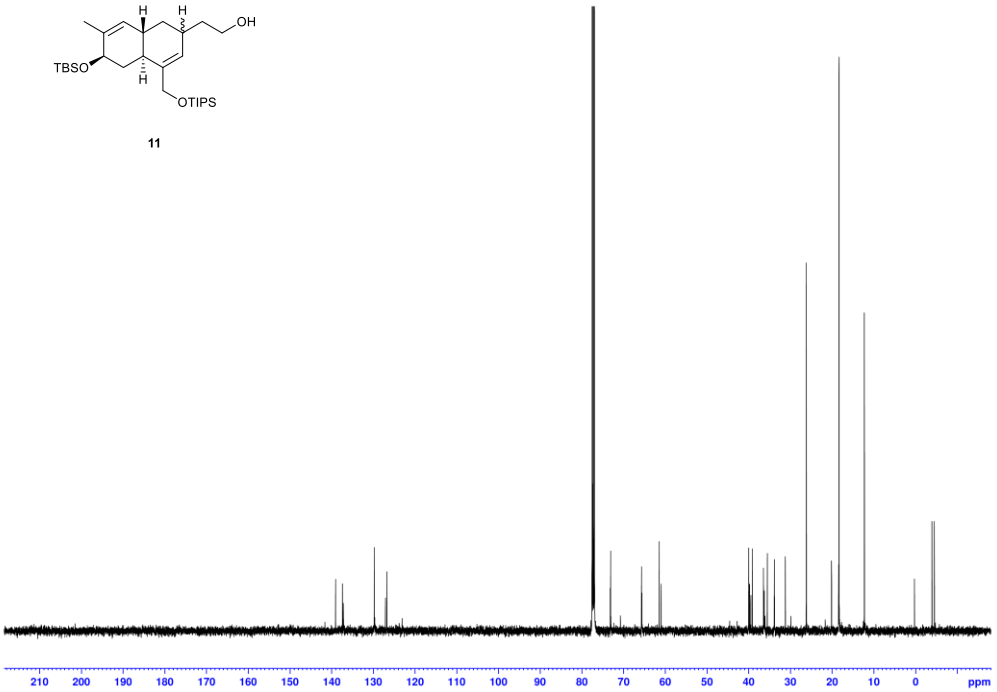
^{13}C NMR spectrum of 10 (126 MHz, 298.0 K, $CDCl_3$, 0.069 M)



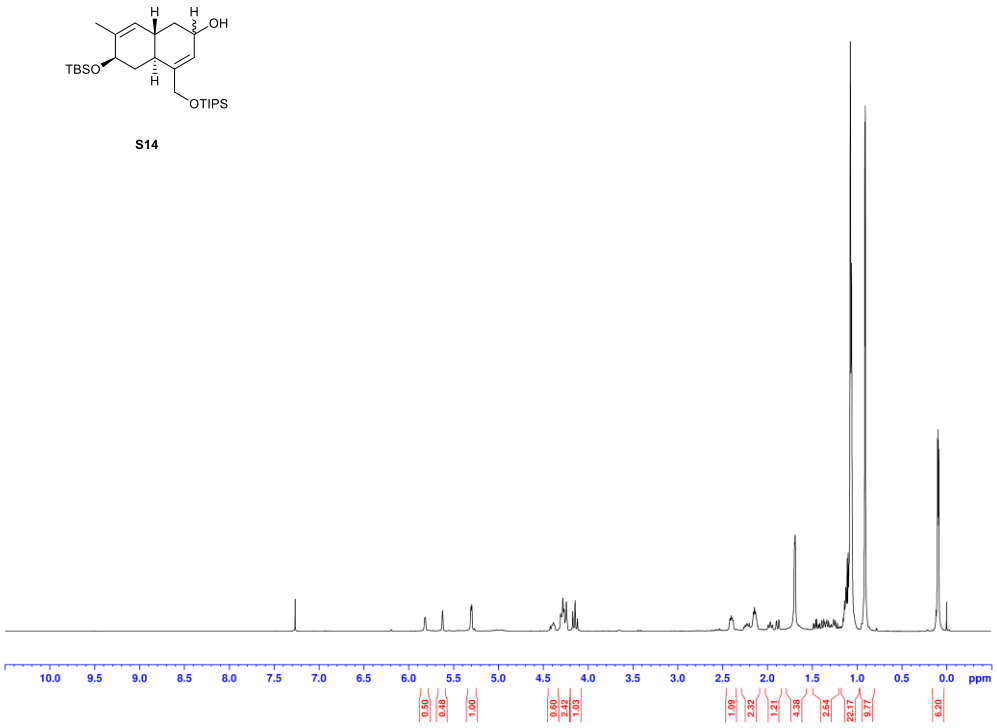
^{1}H NMR spectrum of 11 (500 MHz, 297.0 K, $CDCl_3$, 0.076 M)



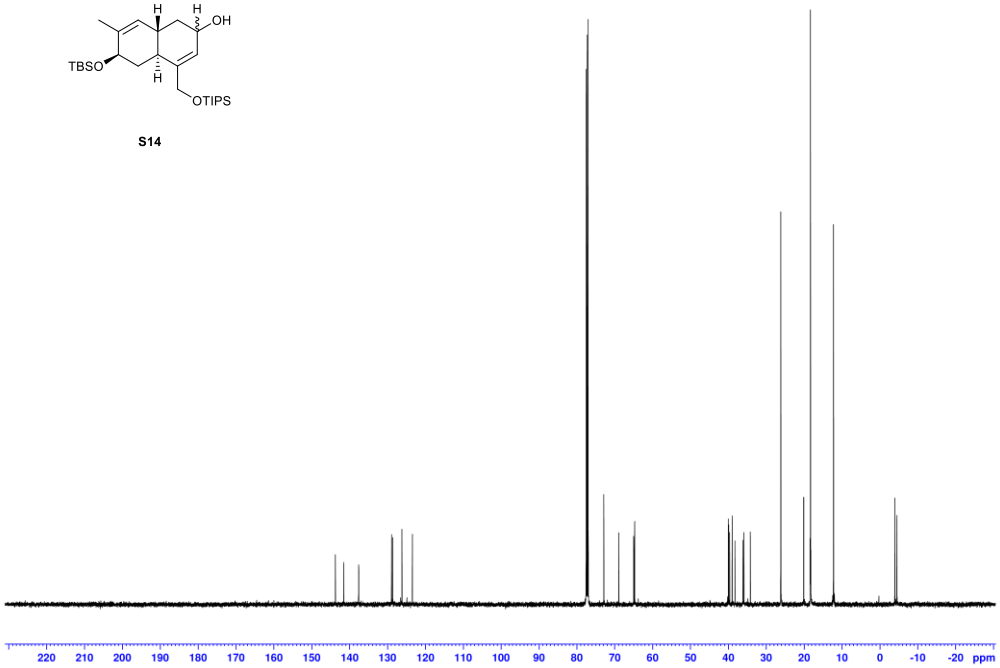
^{13}C NMR spectrum of 11 (126 MHz, 298.0 K, $CDCl_3$, 0.076 M)



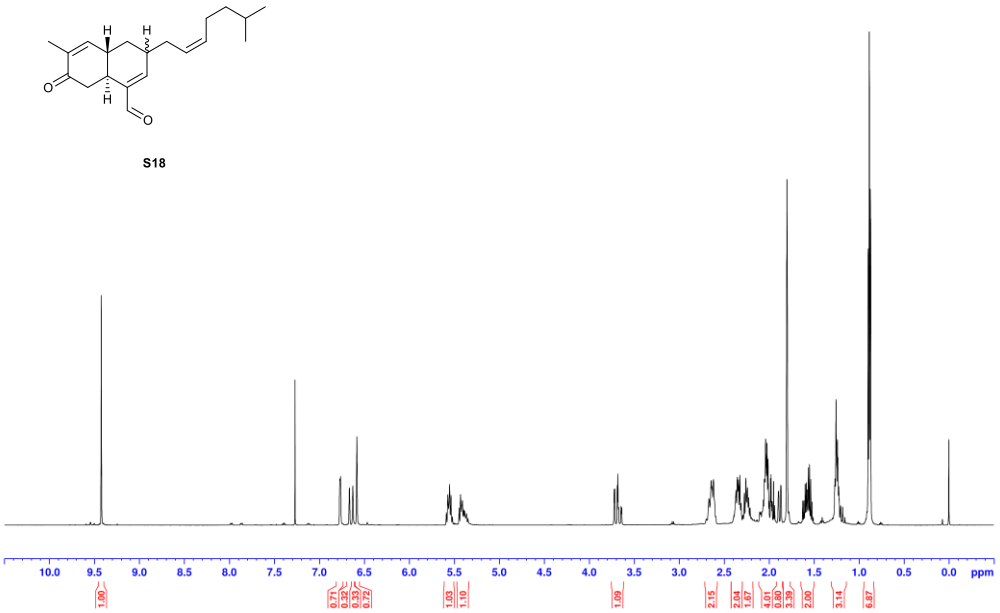
^{1}H NMR spectrum of S14 (500 MHz, 297.0 K, $CDCl_3$, 0.10 M)



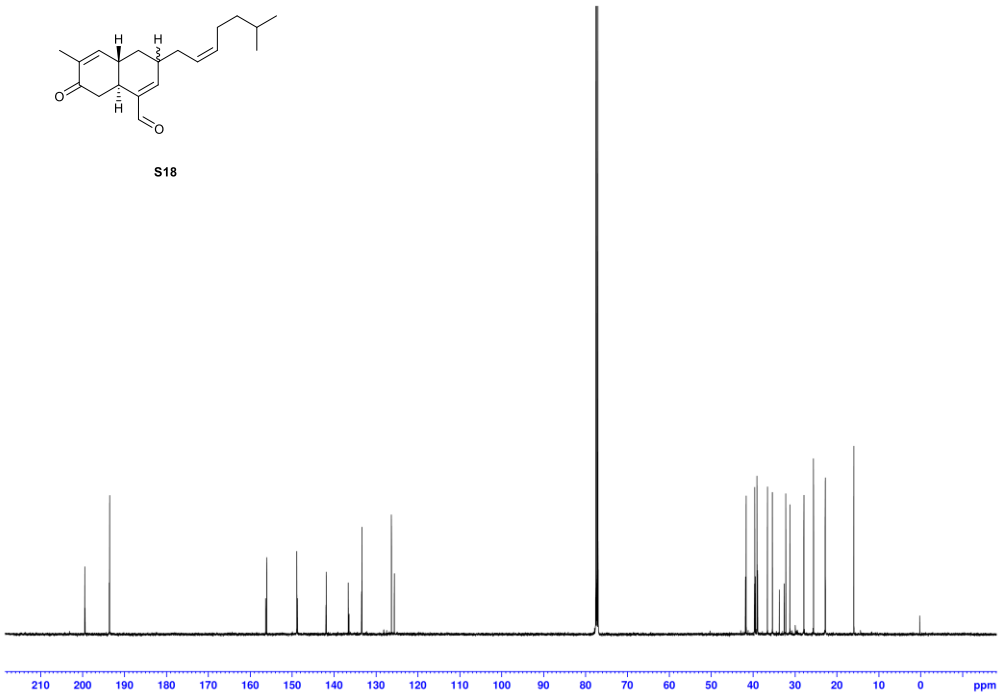
^{13}C NMR spectrum of S14 (126 MHz, 298.0 K, $CDCl_3$, 0.10 M)



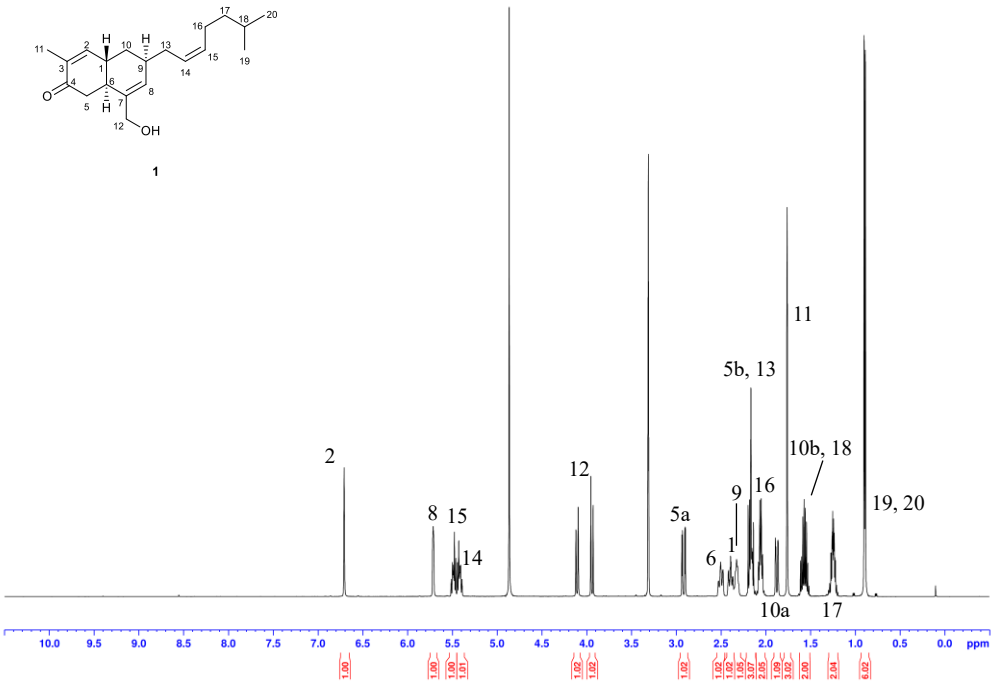
^{1}H NMR spectrum of S18 (500 MHz, 297.0 K, $CDCl_3$, 0.11 M)



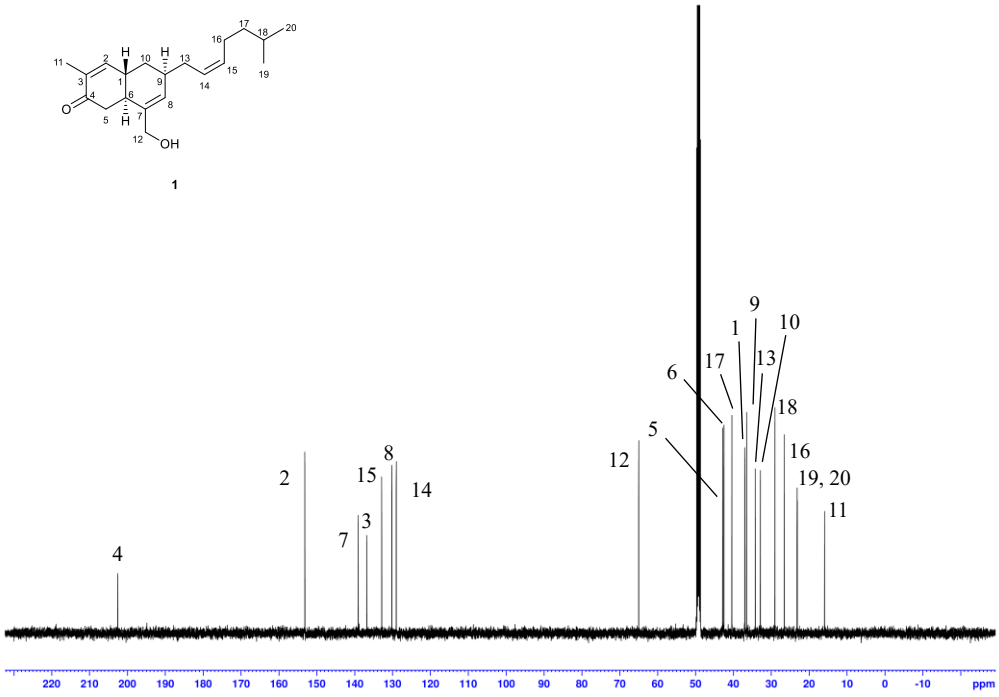
^{13}C NMR spectrum of S18 (126 MHz, 298.0 K, $CDCl_3$, 0.11 M)



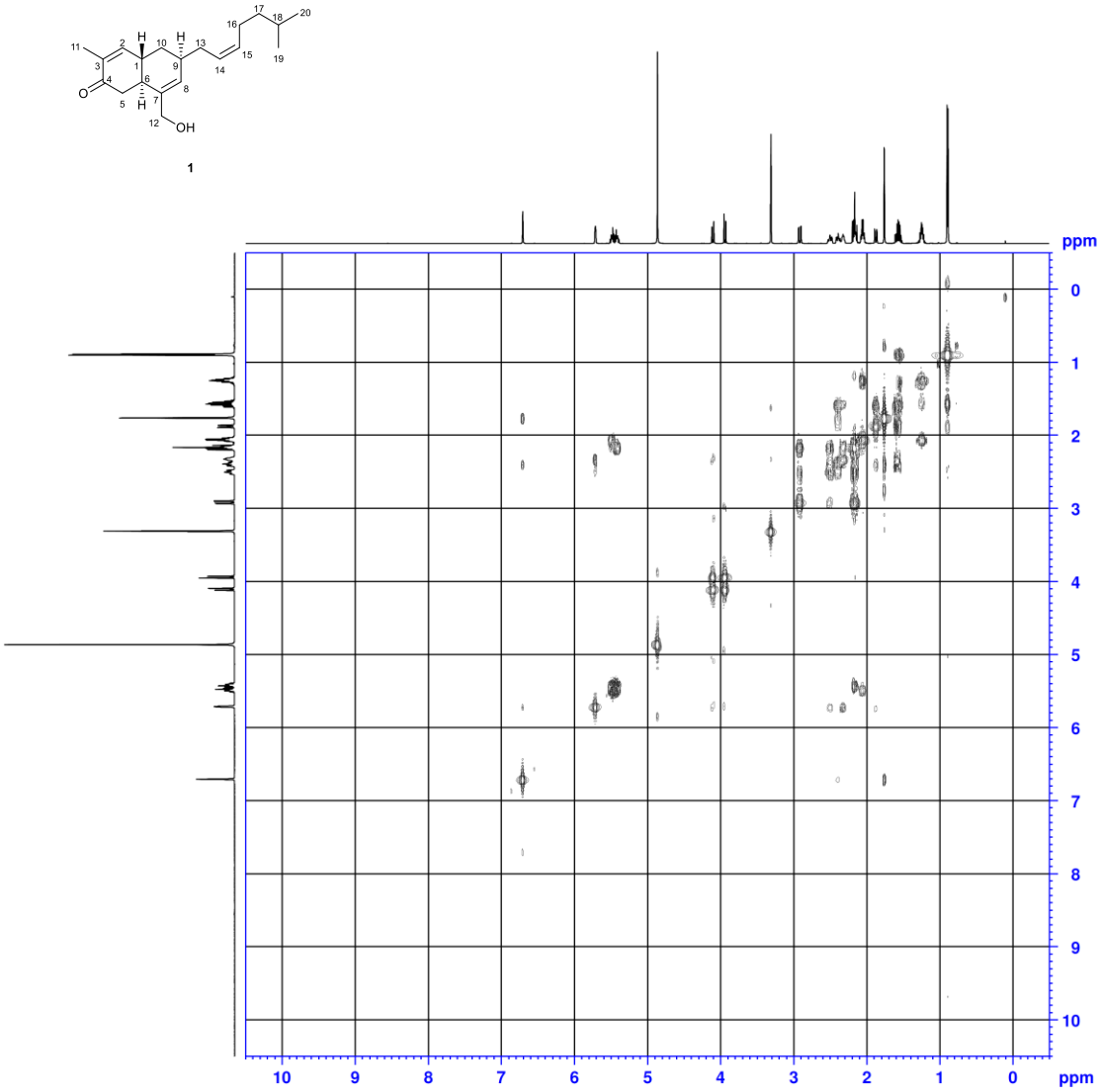
^{1}H NMR spectrum of 1 (500 MHz, 297.0 K, CD_3OD, 0.042 M)



^{13}C NMR spectrum of 1 (126 MHz, 298.0 K, CD_3OD, 0.042 M)

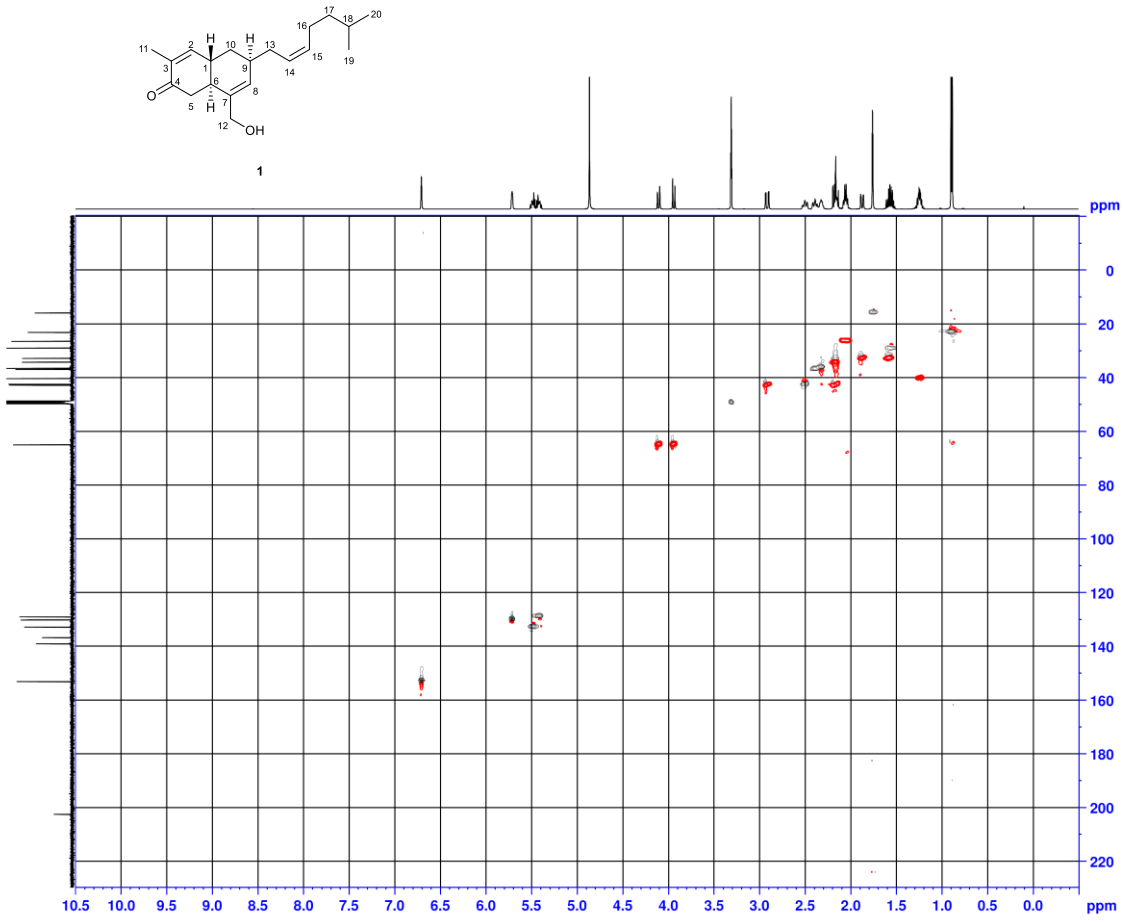


COSY spectrum of 1 (500 MHz, 297.0 K, CD_3OD, 0.042 M)

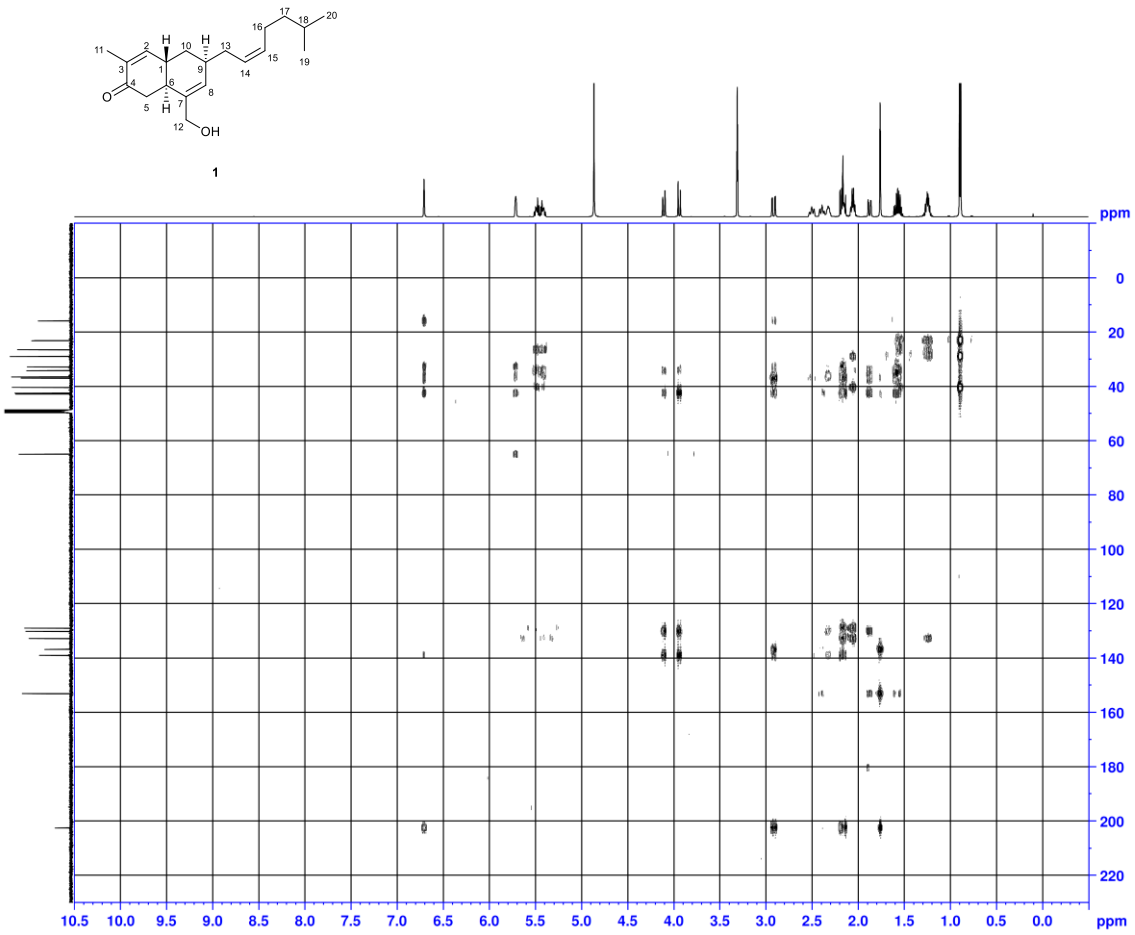


1

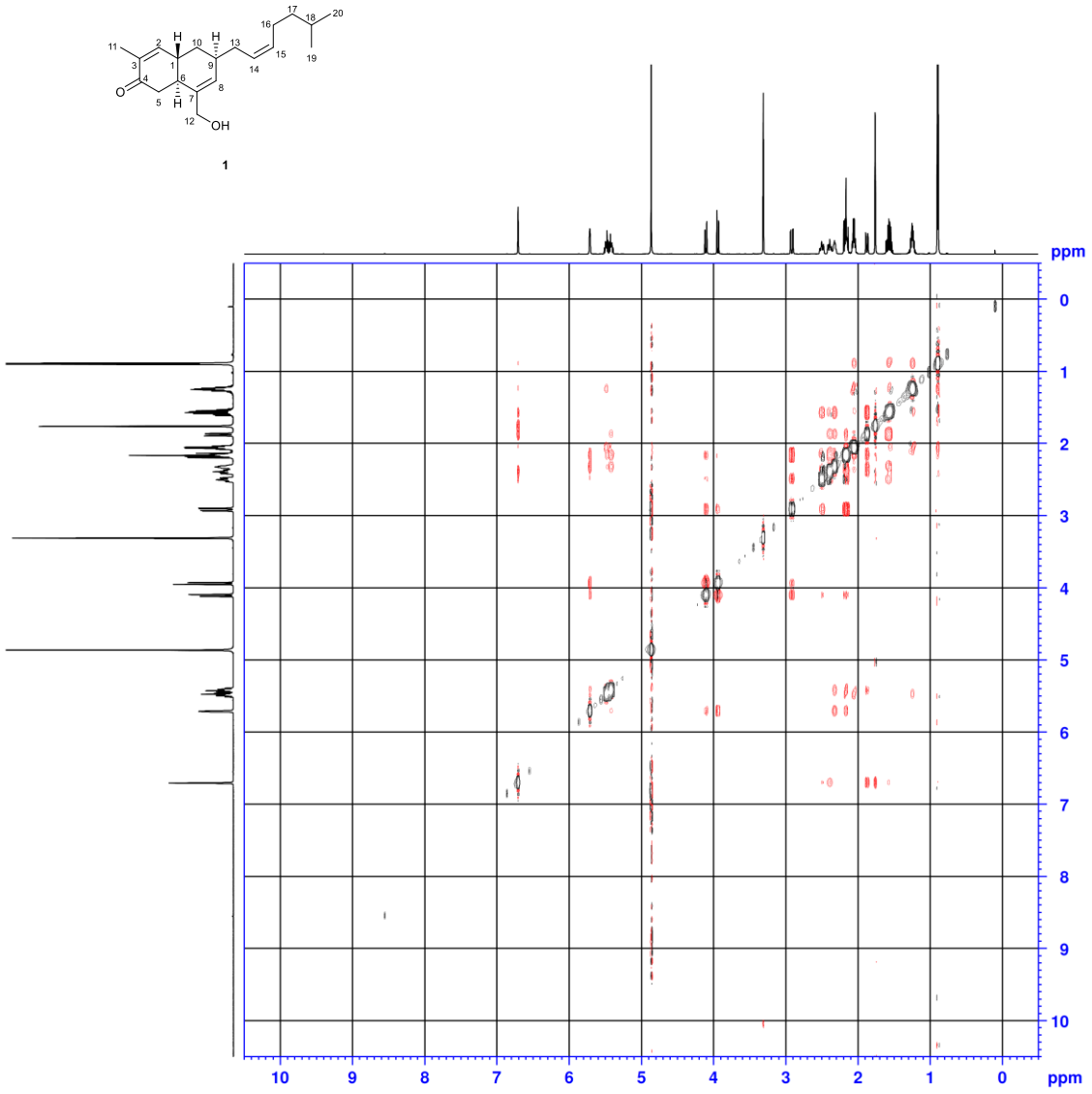
Edited-HSQC spectrum of 1 (500 MHz, 297.0 K, CD_3OD, 0.042 M)



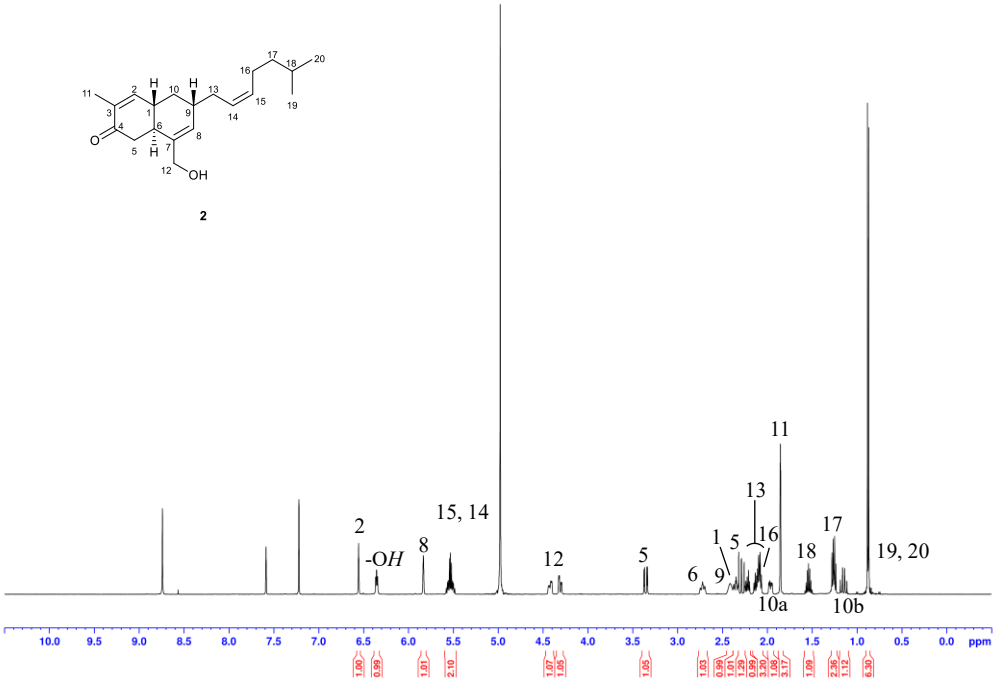
HMBC spectrum of 1 (500 MHz, 297.0 K, CD_3OD, 0.042 M)



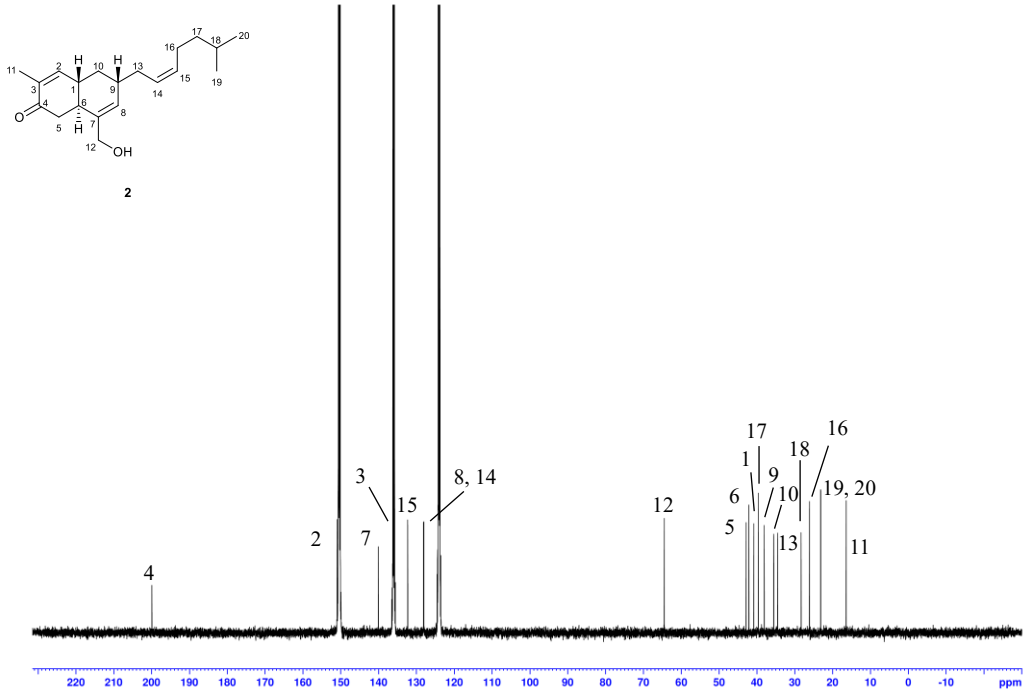
NOESY spectrum of 1 (500 MHz, 297.0 K, CD_3OD, 0.042 M)



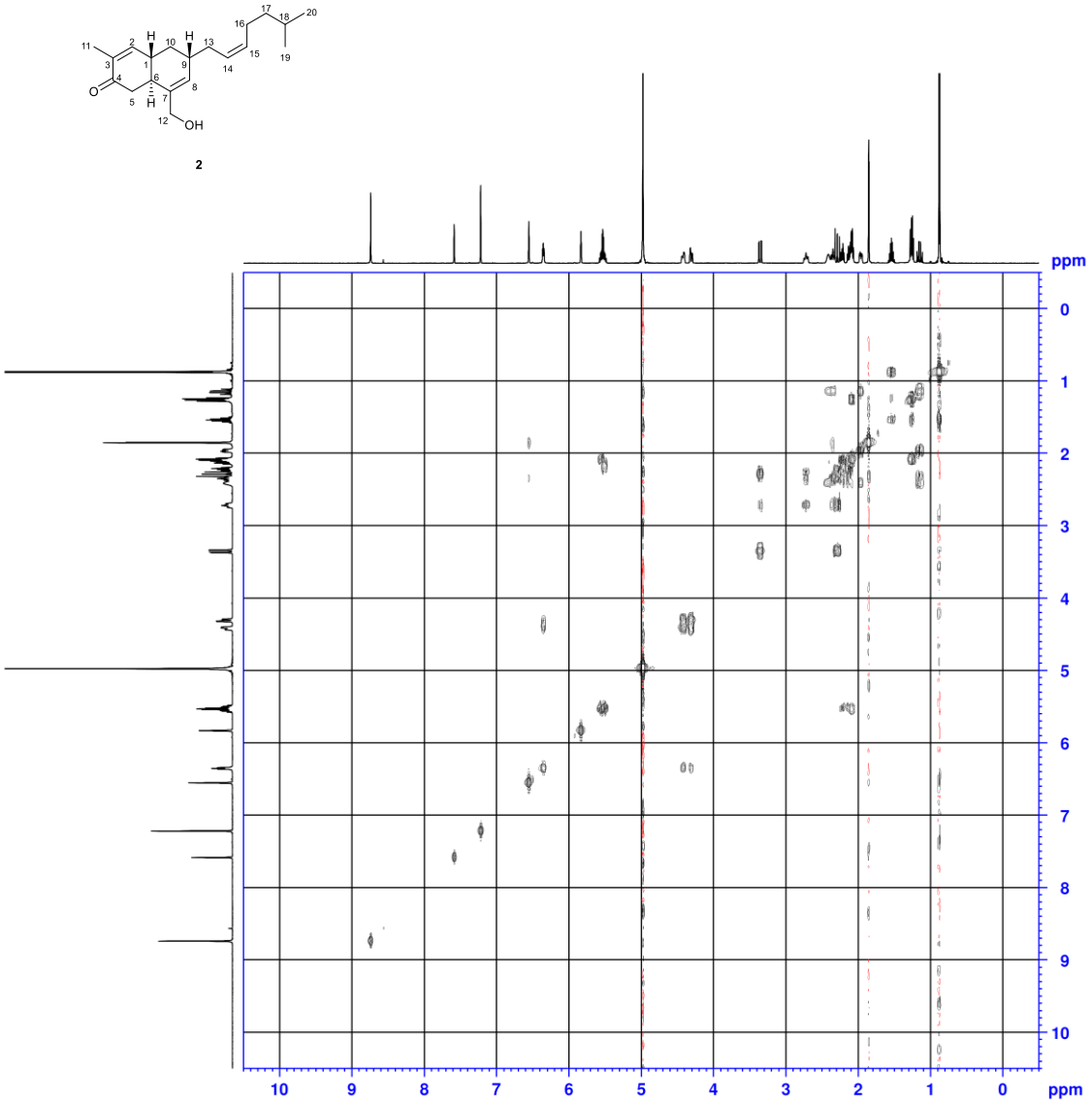
^{1}H NMR spectrum of 2 (500 MHz, 297.0 K, pyridine-d_5, 0.020 M)



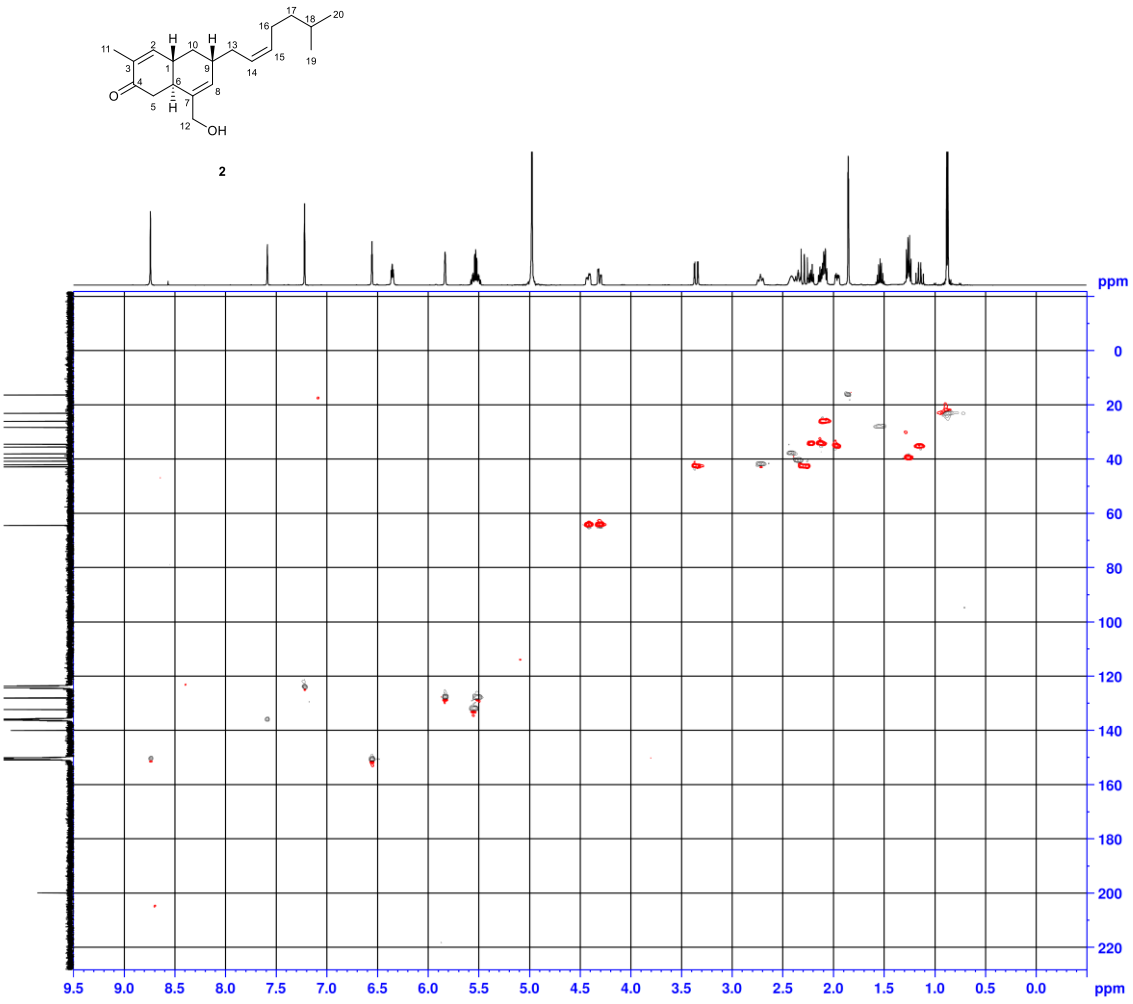
^{13}C NMR spectrum of 2 (126 MHz, 298.0 K, pyridine-d_5, 0.020 M)



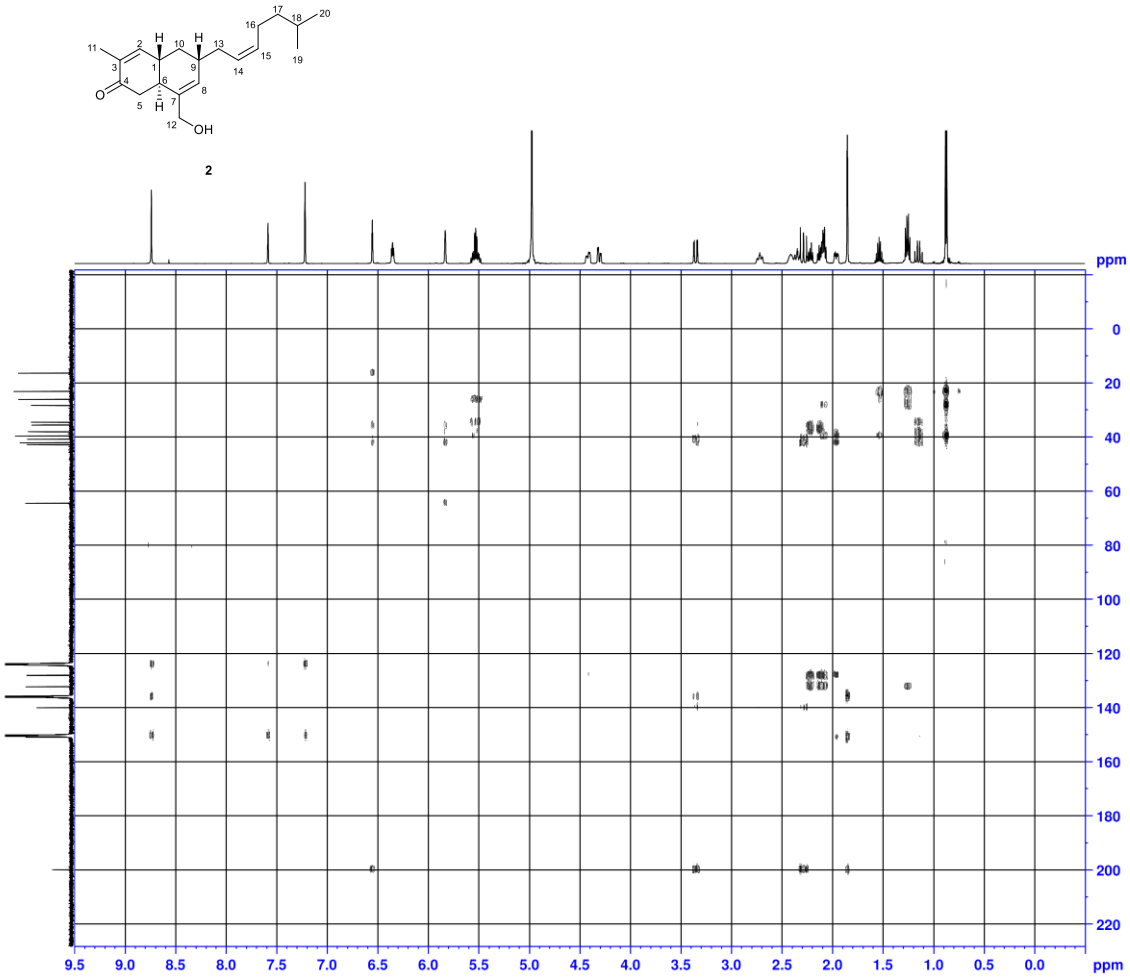
COSY spectrum of 2 (500 MHz, 297.0 K, pyridine-d_5, 0.020 M)



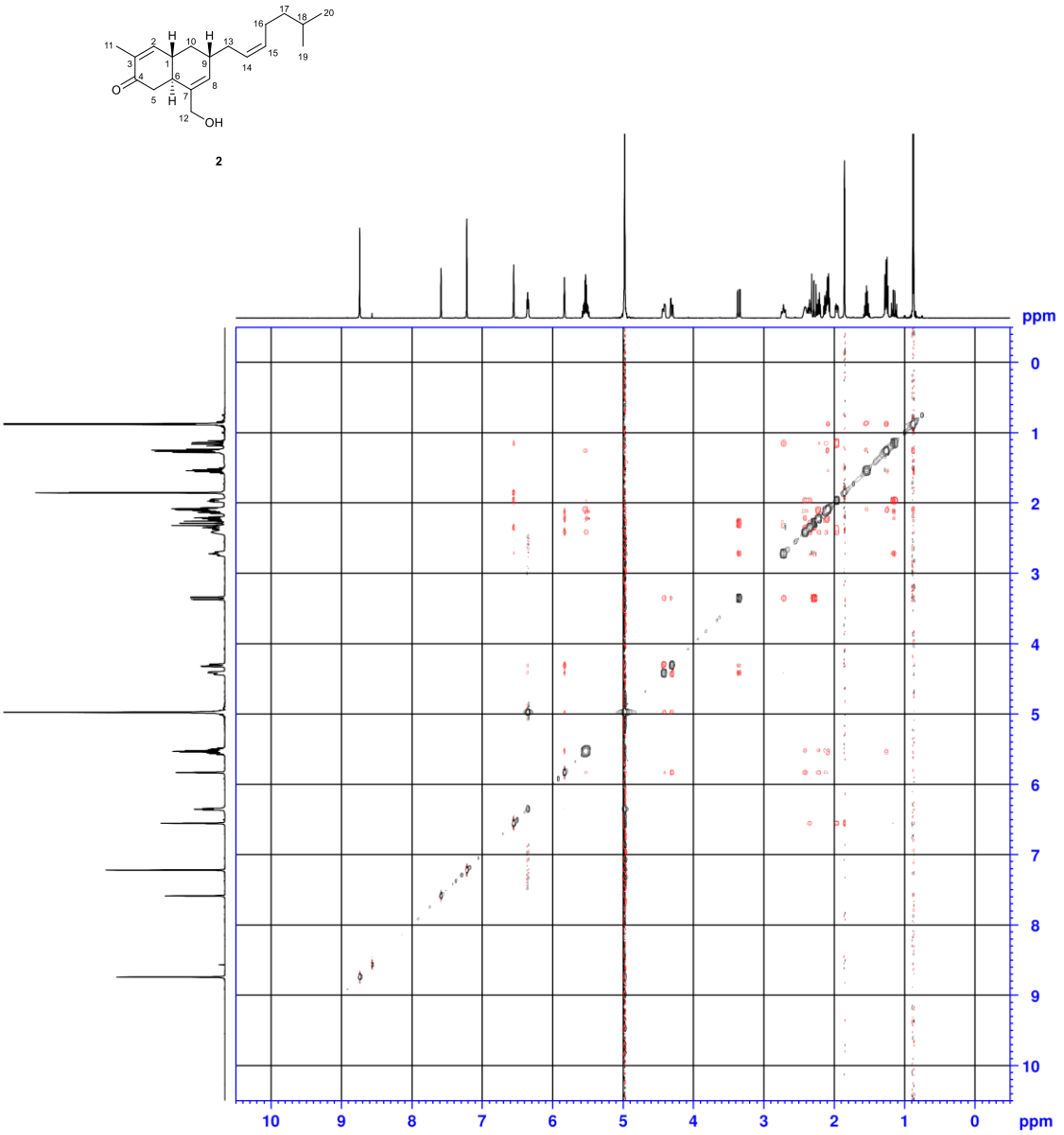
Edited-HSQC spectrum of 2 (500 MHz, 297.0 K, pyridine-d_5, 0.020 M)



HMBC spectrum of 2 (500 MHz, 297.0 K, pyridine-d_5, 0.020 M)



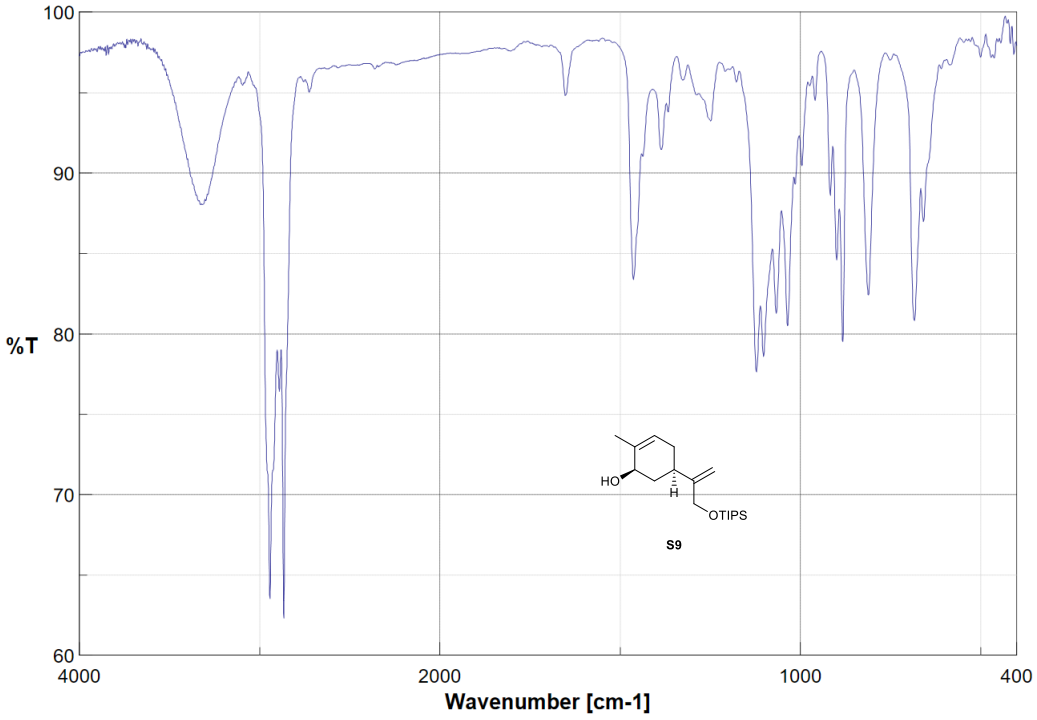
NOESY spectrum of 2 (500 MHz, 297.0 K, pyridine-d_5, 0.020 M)



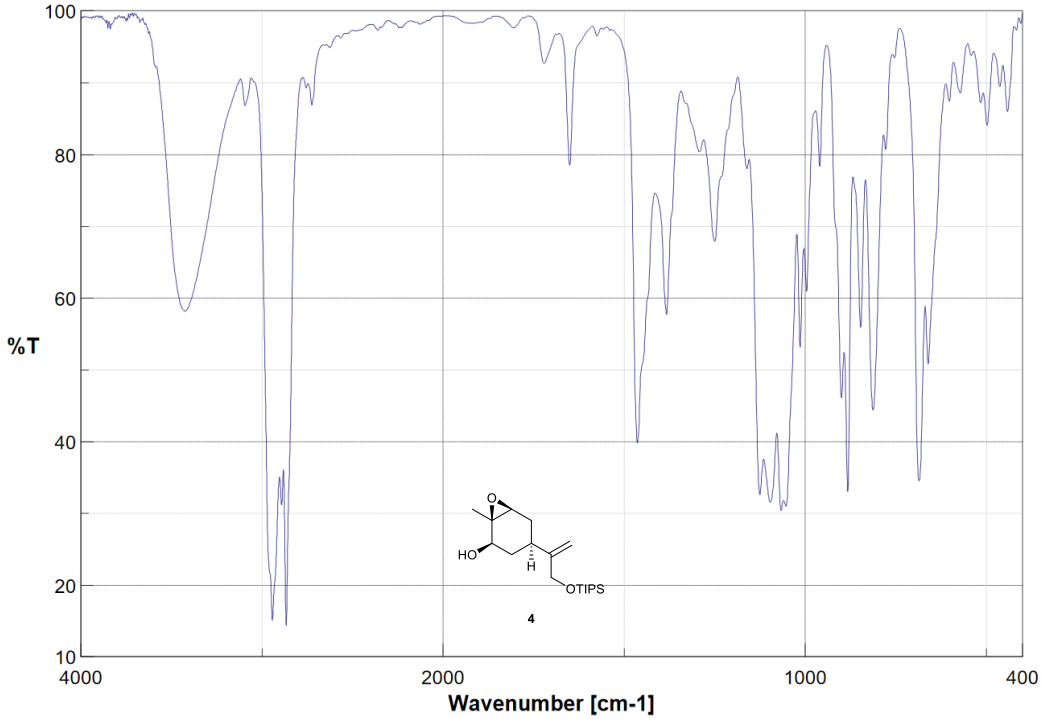
2

4. IR spectra

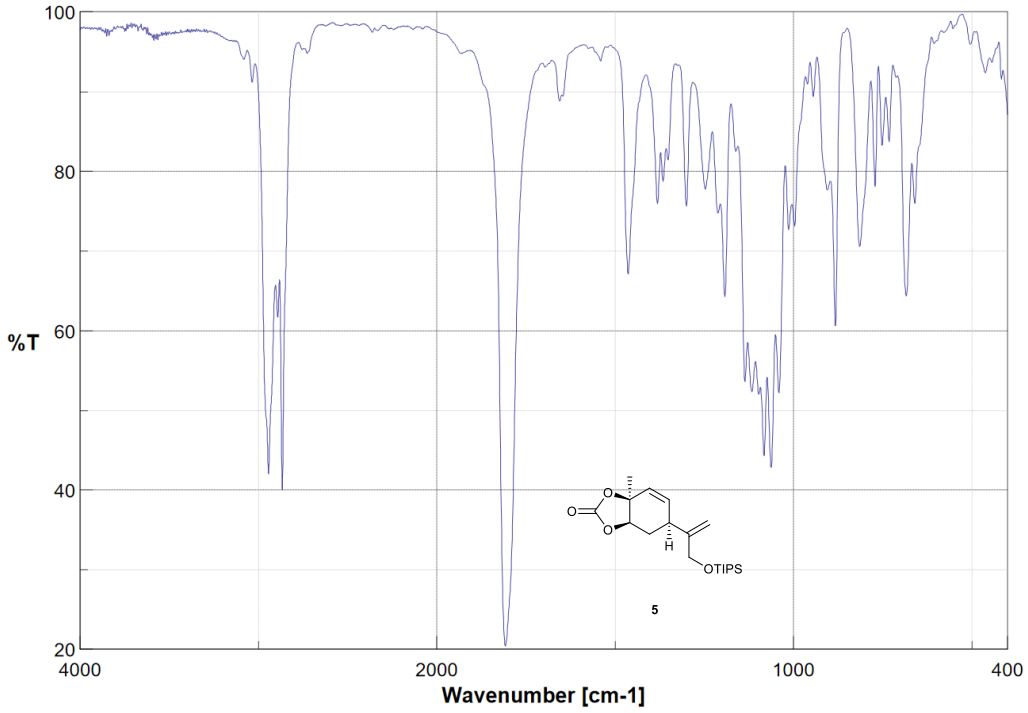
IR spectrum of S9 (neat on KRS-5)



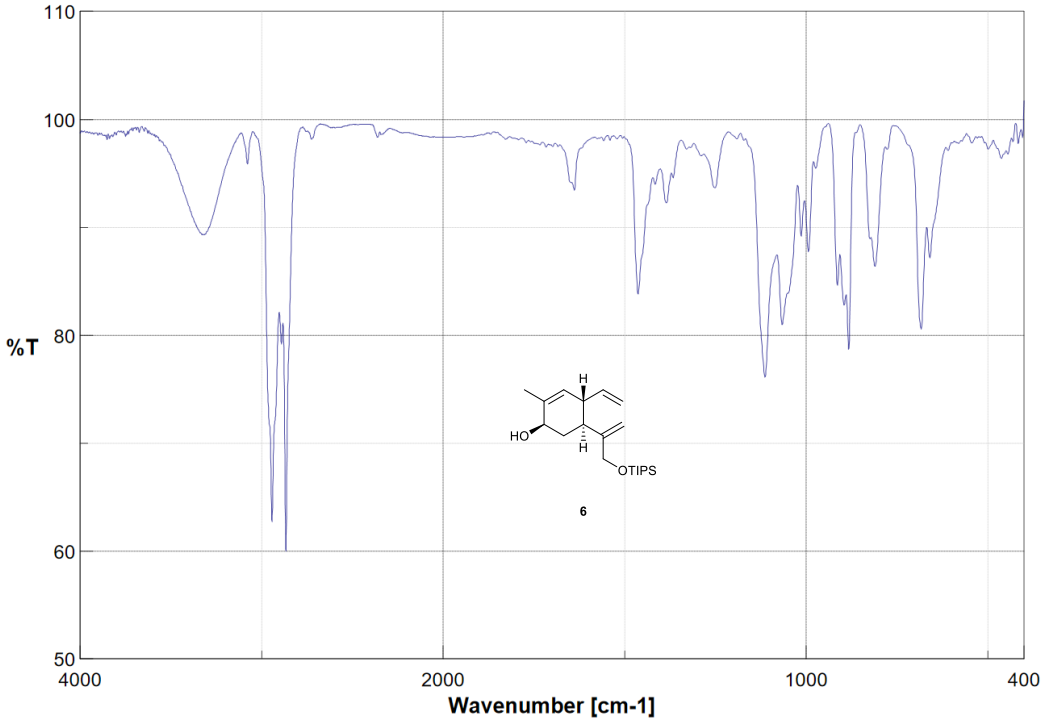
IR spectrum of 4 (neat on KRS-5)



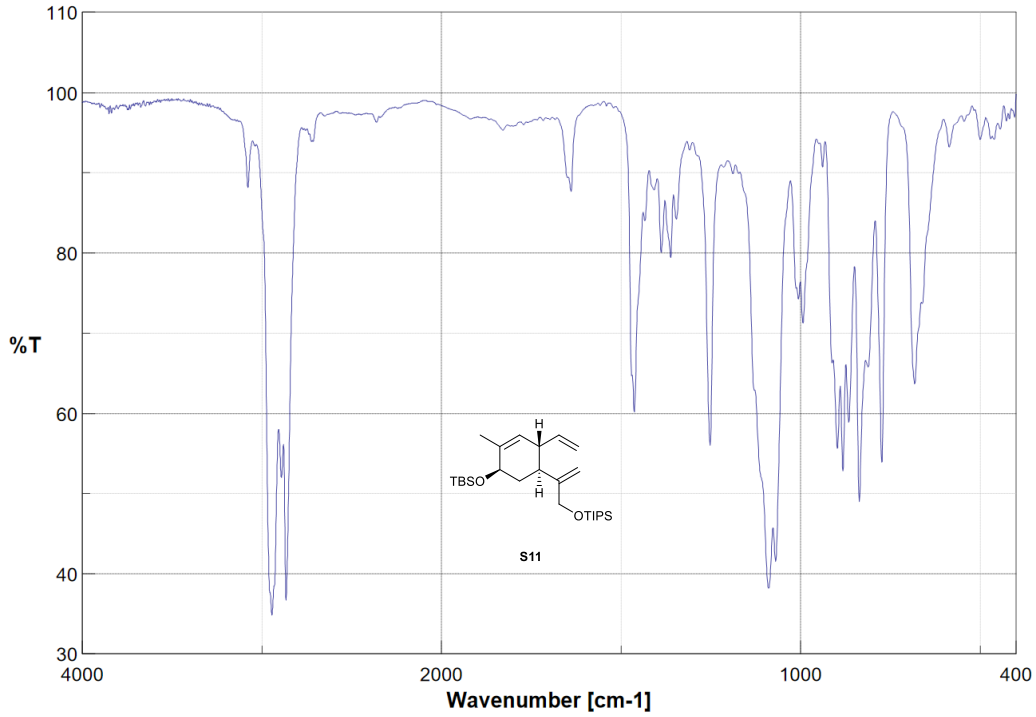
IR spectrum of 5 (neat on KRS-5)



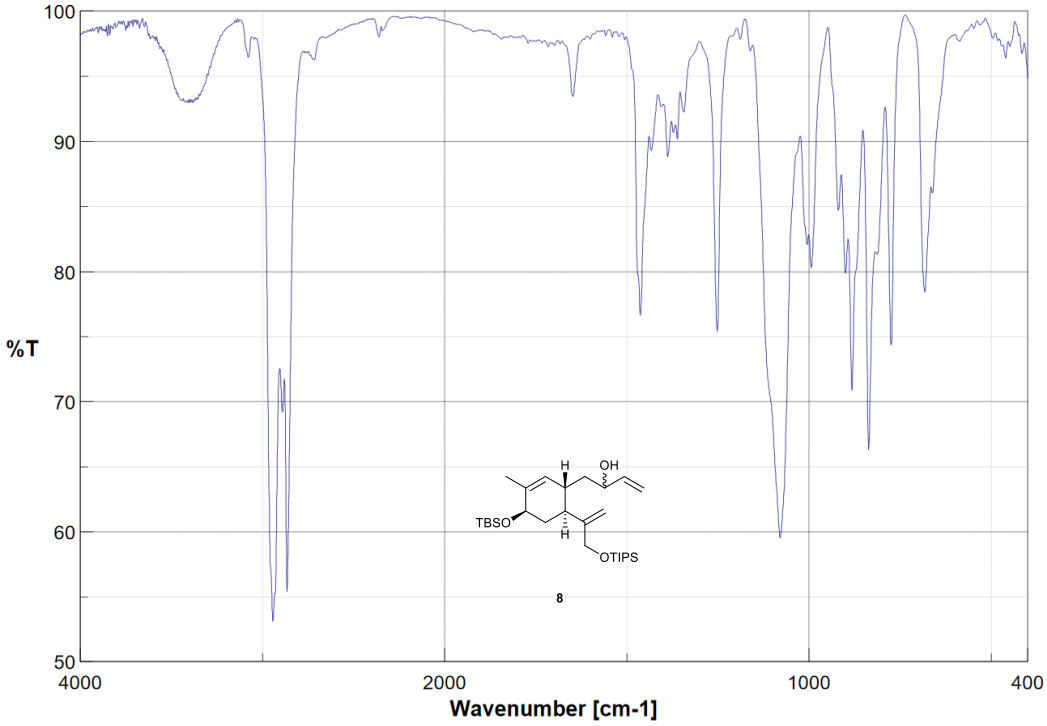
IR spectrum of 6 (neat on KRS-5)



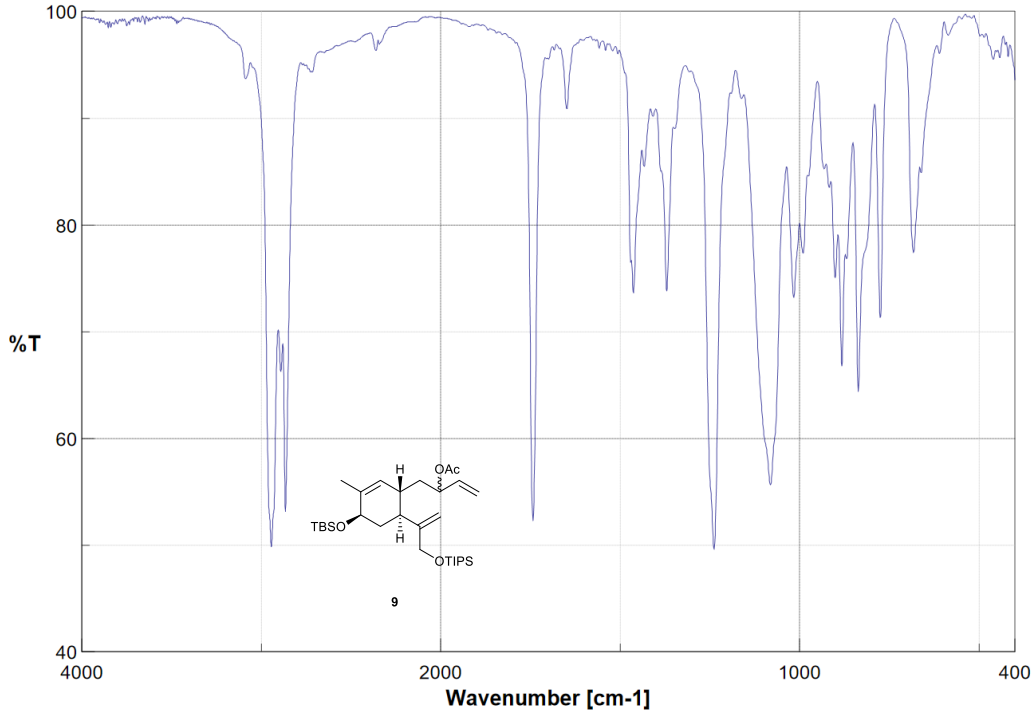
IR spectrum of S11 (neat on KRS-5)



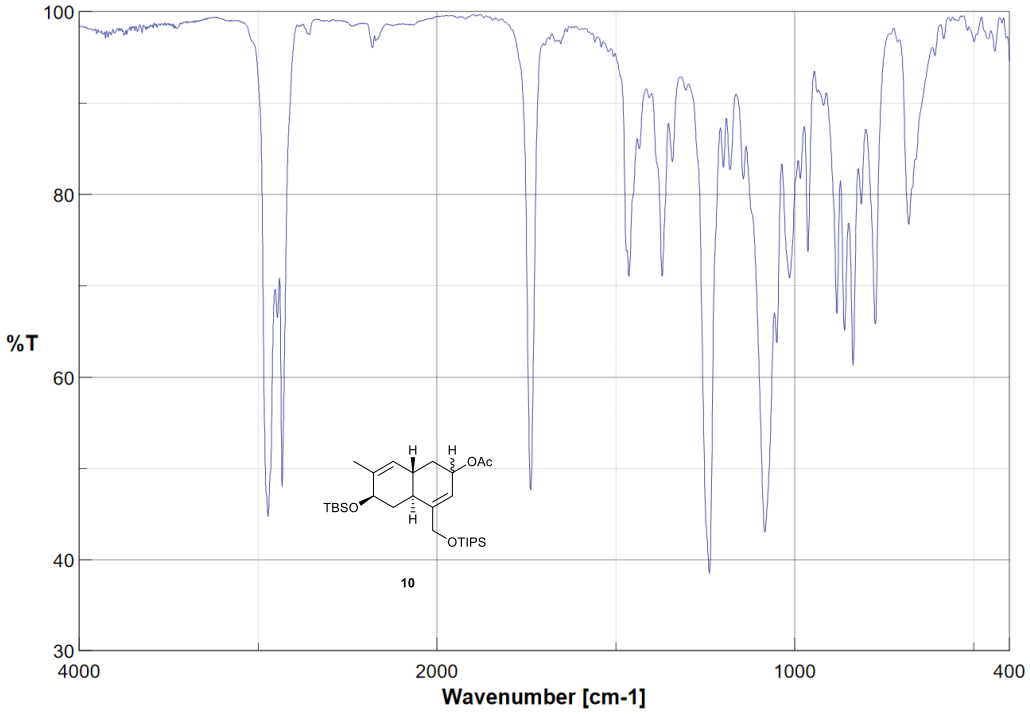
IR spectrum of 8 (neat on KRS-5)



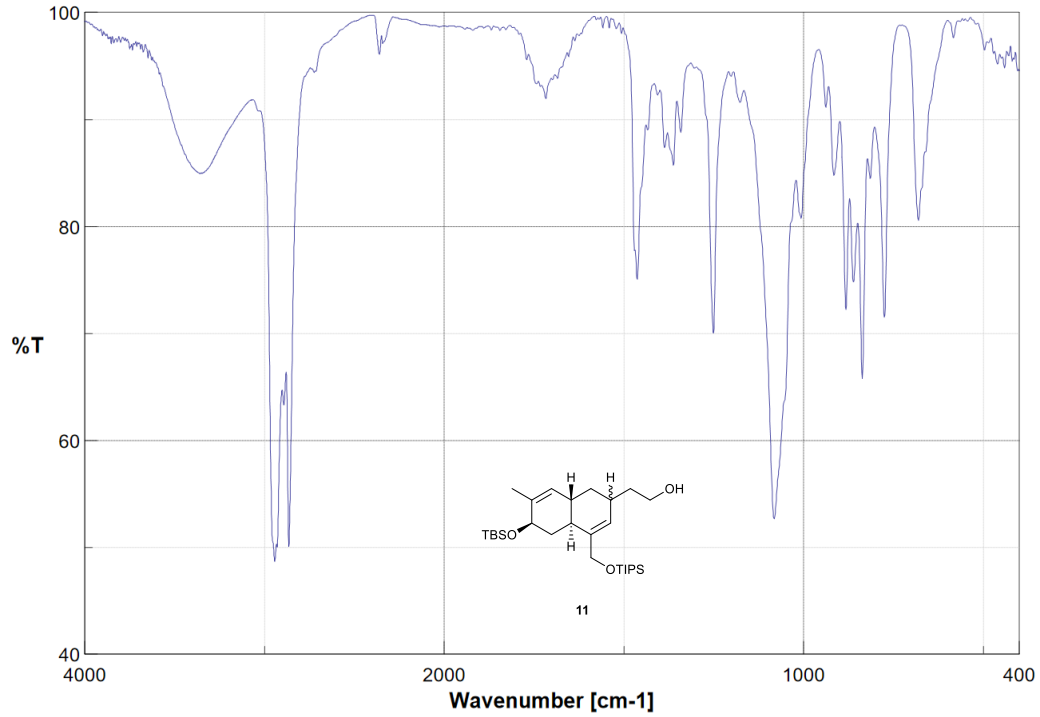
IR spectrum of 9 (neat on KRS-5)



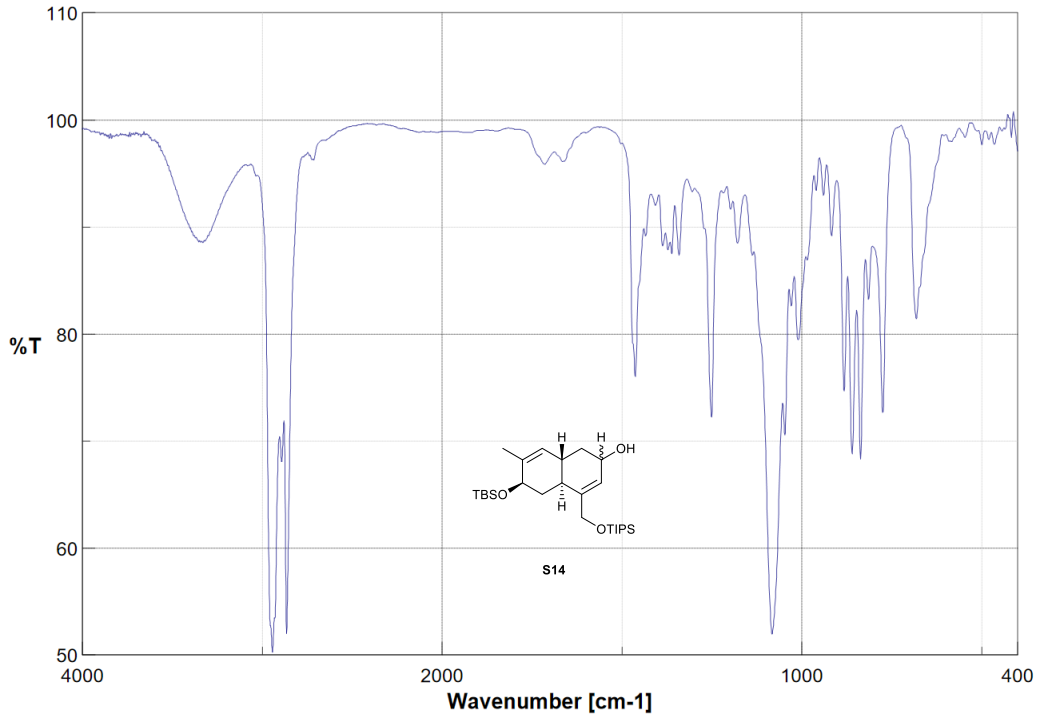
IR spectrum of 10 (neat on KRS-5)



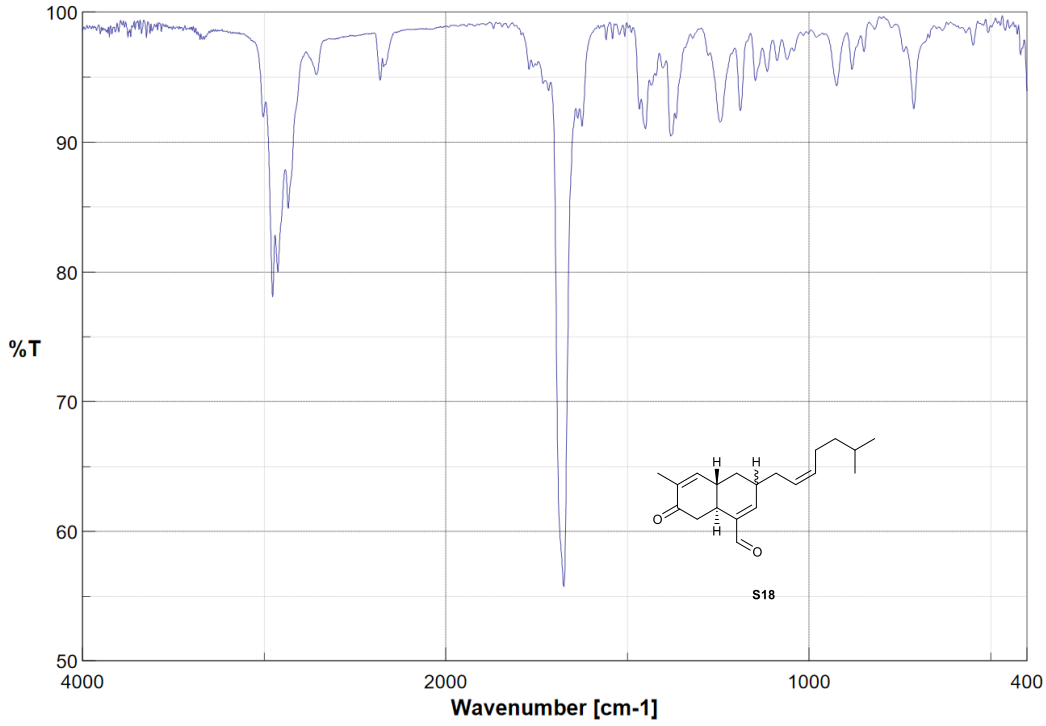
IR spectrum of 11 (neat on KRS-5)



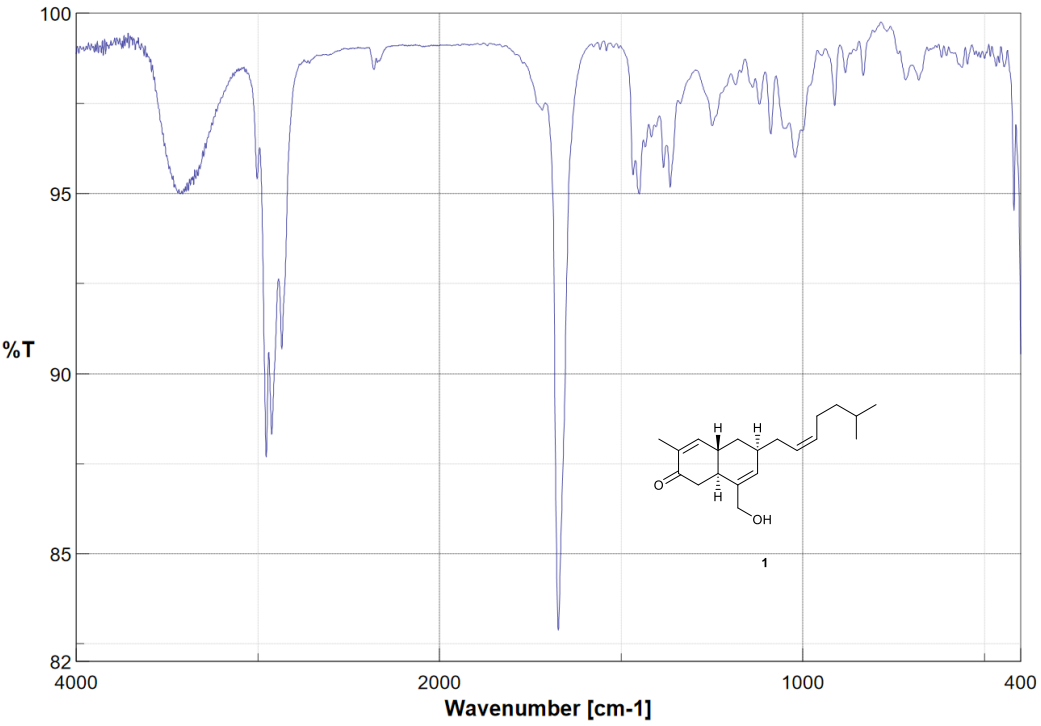
IR spectrum of S14 (neat on KRS-5)



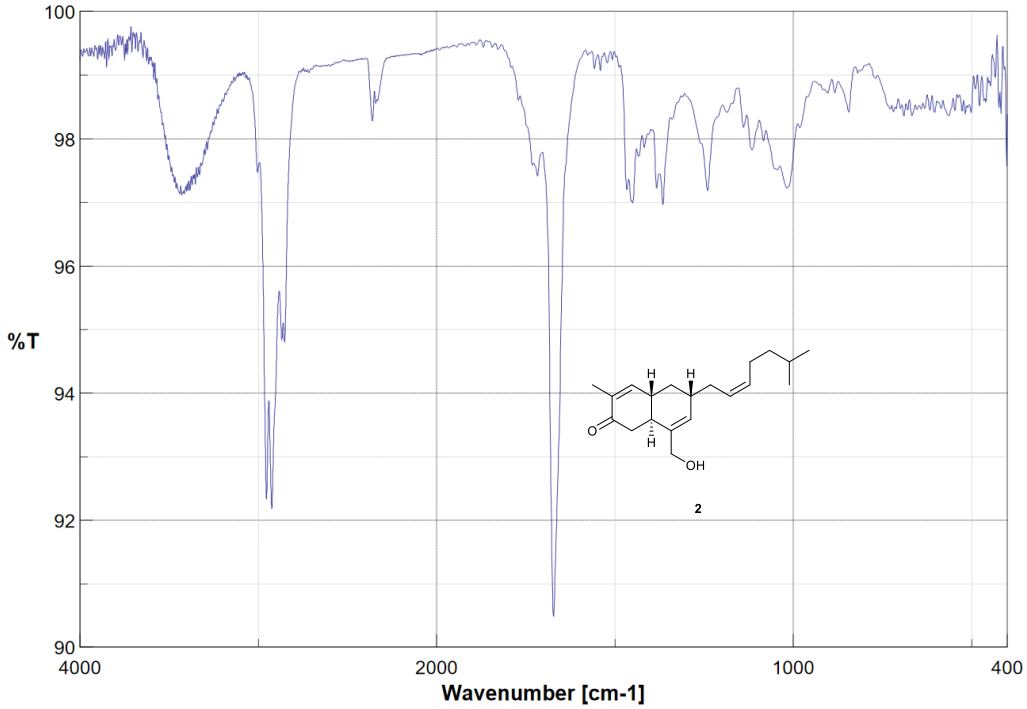
IR spectrum of S18 (neat on KRS-5)



IR spectrum of 1 (neat on KRS-5)



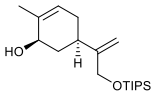
IR spectrum of 2 (neat on KRS-5)



5. MS data

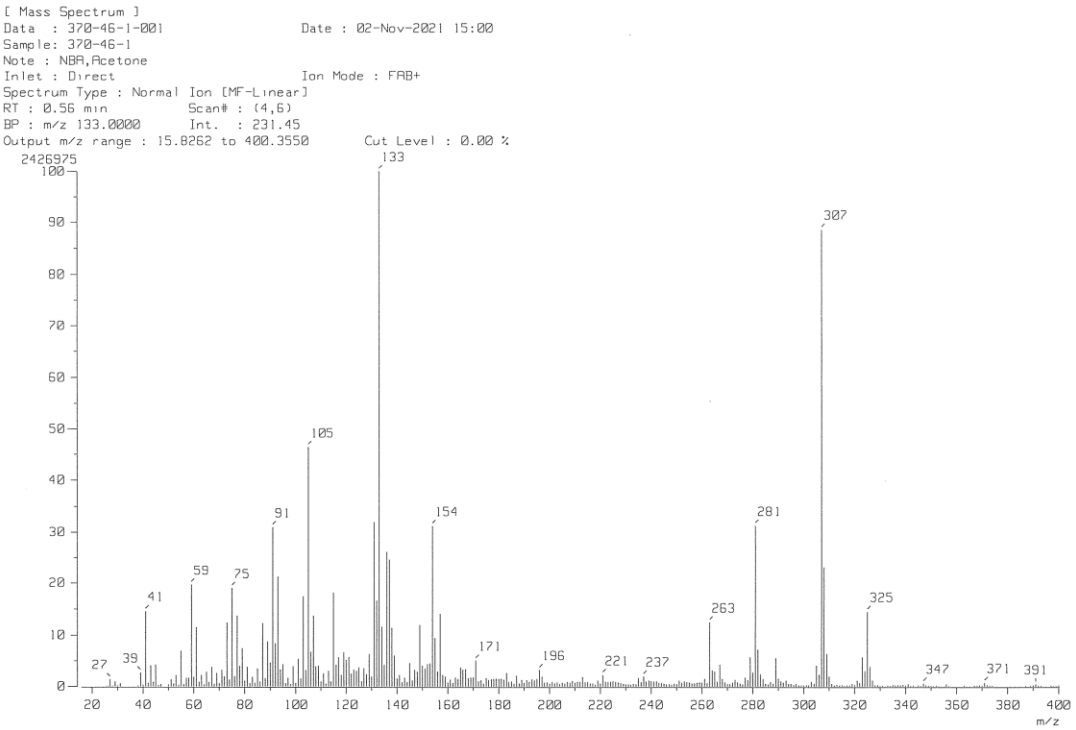
MS data of S9 (FAB+, NBA, acetone)

```
 [ Elemental Composition ]                                          Page: 1
 Data  : 370-46-1-001                 Date : 02-Nov-2021 15:05
 Sample: 370-46-1
 Note : NBA,Acetone
 Inlet : Direct                       Ion Mode : FAB+
 RT : 1.00 min                        Scan#: 6
 Elements : C 19/0, H 40/0, O 2/0, Si 1/0
 Mass Tolerance      : 1000ppm, 3mmu if m/z < 3, 5mmu if m/z > 5
 Unsaturation (U.S.) : -0.5 - 30.0

 Observed m/z Int%   Err[ppm / mmu]    U.S.  Composition
   325.2567    19.6     +1.1 /  +0.4    2.5  C 19 H 37 O 2 Si
```



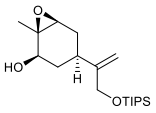
S9

```
[ Theoretical Ion Distribution ]                                    Page: 1
Molecular Formula : C19 H37 O2 Si
                    (m/z   325.2563, MW   325.5871, U.S.  2.5)
Base Peak :   325.2563, Averaged MW :   325.5833(a),   325.5845(w)

    m/z       INT.
  325.2563 100.0000 ************************************************************
  326.2589  26.2721 ***************
  327.2575   6.9675 ****
  328.2590   1.0609 *
  329.2612   0.1112
  330.2635   0.0090
  331.2660   0.0006
```



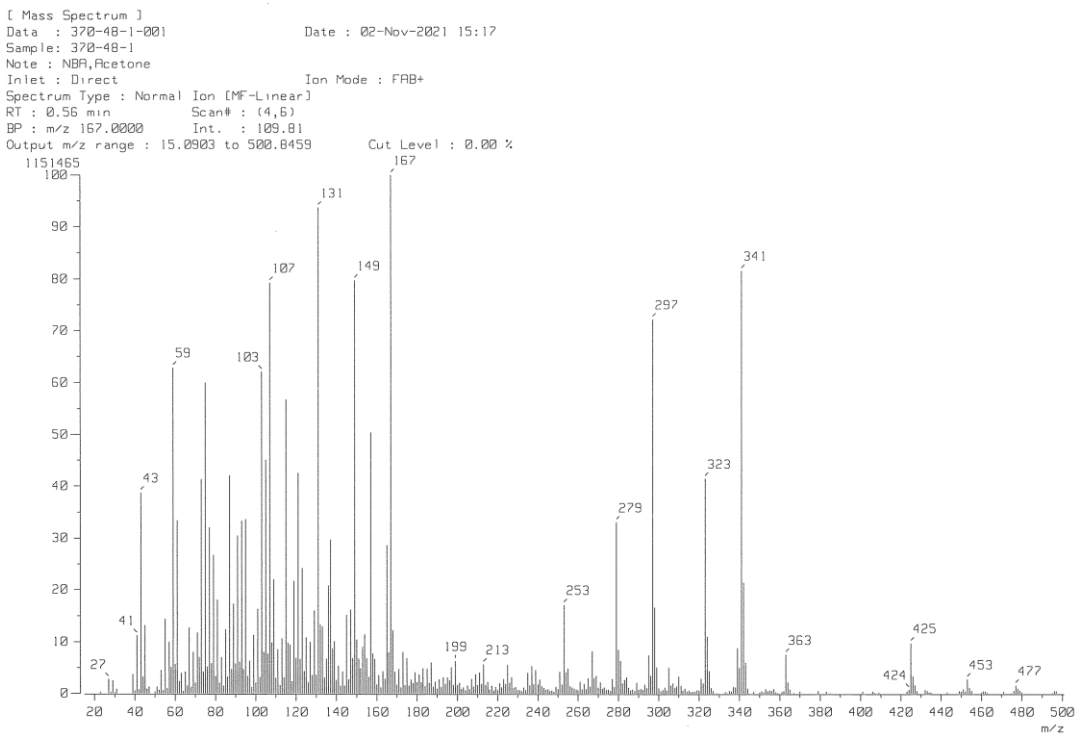
MS data of 4 (FAB+, NBA, acetone)

```
[ Elemental Composition ]                                          Page: 1
Data  : 370-48-1-001                  Date : 02-Nov-2021 15:44
Sample: 370-48-1
Note : NBA,Acetone
Inlet : Direct                        Ion Mode : FAB+
RT : 0.00 min                         Scan#: 1
Elements : C 19/0, H 40/0, O 3/0, Si 1/0
Mass Tolerance      : 1000ppm, 3mmu if m/z < 3, 5mmu if m/z > 5
Unsaturation (U.S.) : -0.5 - 30.0

Observed m/z Int%   Err[ppm / mmu]    U.S.  Composition
  341.2516   100.0     +1.1 /  +0.4    2.5  C 19 H 37 O 3 Si
```



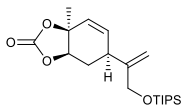
4

```
[ Theoretical Ion Distribution ]                                   Page: 1
Molecular Formula : C19 H37 O3 Si
                    (m/z   341.2512, MW   341.5865, U.S.  2.5)
Base Peak :   341.2512, Averaged MW :   341.5827(a),   341.5838(w)

    m/z       INT.
  341.2512 100.0000 ************************************************************
  342.2538  26.3102 ***************
  343.2525   7.1780 ****
  344.2541   1.1162 *
  345.2562   0.1256
  346.2584   0.0111
  347.2607   0.0008
```



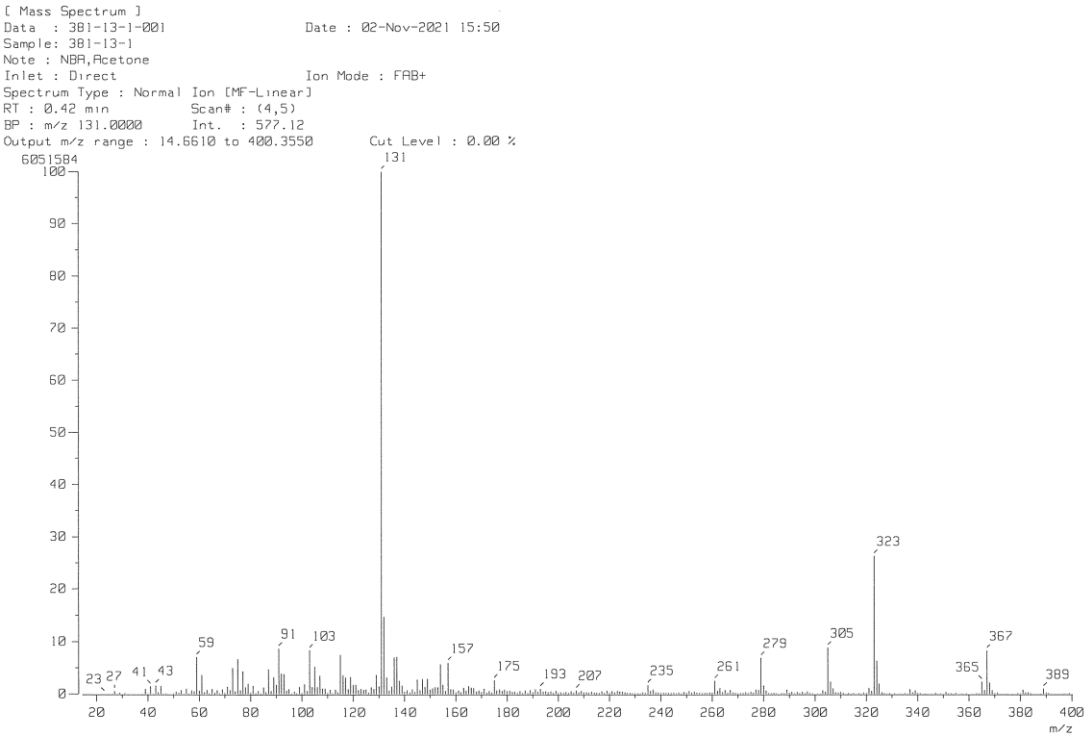
MS data of 5 (FAB+, NBA, acetone)

```
[ Elemental Composition ]                                          Page: 1
Data  : 381-13-1-001                  Date : 02-Nov-2021 16:06
Sample: 381-13-1
Note : NBA,Acetone
Inlet : Direct                        Ion Mode : FAB+
RT : 1.00 min                         Scan#: 6
Elements : C 20/0, H 40/0, O 4/0, Si 1/0
Mass Tolerance      : 1000ppm, 3mmu if m/z < 3, 5mmu if m/z > 5
Unsaturation (U.S.) : -0.5 - 30.0

Observed m/z Int%   Err[ppm / mmu]    U.S.  Composition
  367.2300    36.1     -1.3 /  -0.5    4.5  C 20 H 35 O 4 Si
```



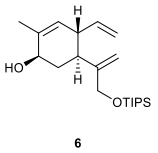
5

```
[ Theoretical Ion Distribution ]                                   Page: 1
Molecular Formula : C20 H35 O4 Si
                    (m/z   367.2305, MW   367.5810, U.S.  4.5)
Base Peak :   367.2305, Averaged MW :   367.5773(a),   367.5785(w)

    m/z       INT.
  367.2305 100.0000 ************************************************************
  368.2331  27.4605 ****************
  369.2320   7.6815 ****
  370.2337   1.2539 *
  371.2356   0.1534
  372.2378   0.0150
  373.2401   0.0012
```



MS data of 6 (FAB+, NBA, acetone)

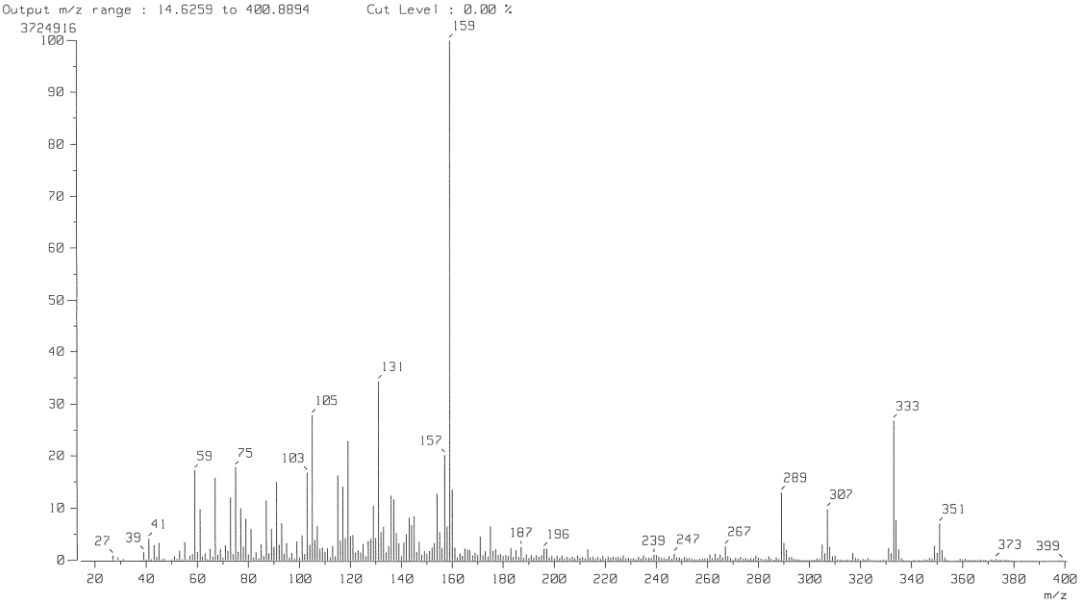
```
[ Elemental Composition ]                                          Page: 1
Data  : 370-88-1-001                Date : 02-Nov-2021 16:54
Sample: 370-88-1
Note : NBA,Acetone
Inlet : Direct                      Ion Mode : FAB+
RT : 0.20 min                       Scan#: 2
Elements : C 21/0, H 40/0, O 2/0, Si 1/0
Mass Tolerance      : 1000ppm, 3mmu if m/z < 3, 5mmu if m/z > 5
Unsaturation (U.S.) : -0.5 - 30.0

Observed m/z Int%   Err[ppm / mmu]    U.S.  Composition
  351.2724    24.2     +1.3 /  +0.5    3.5  C 21 H 39 O 2 Si
```



6

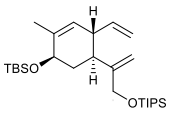
```
[ Theoretical Ion Distribution ]                                   Page: 1
Molecular Formula : C21 H39 O2 Si
                    (m/z   351.2719, MW   351.6250, U.S.  3.5)
Base Peak :   351.2719, Averaged MW :   351.6211(a),   351.6222(w)

    m/z       INT.
  351.2719 100.0000 ************************************************************
  352.2746  28.4965 ******************
  353.2735   7.5643 ****
  354.2749   1.2191 *
  355.2771   0.1357
  356.2794   0.0116
  357.2819   0.0008
```

```
[ Mass Spectrum ]
Data  : 370-88-1-001                Date : 02-Nov-2021 16:16
Sample: 370-88-1
Note : NBA,Acetone
Inlet : Direct                      Ion Mode : FAB+
Spectrum Type : Normal Ion [MF-Linear]
RT : 0.42 min                       Scan# : (3,5)
BP : m/z 159.0000                   Int.  : 355.24
Output m/z range : 14.6259 to 400.8894       Cut Level : 0.00 %
```



MS data of S11 (FAB+, NBA, acetone + *aq.* NaI)

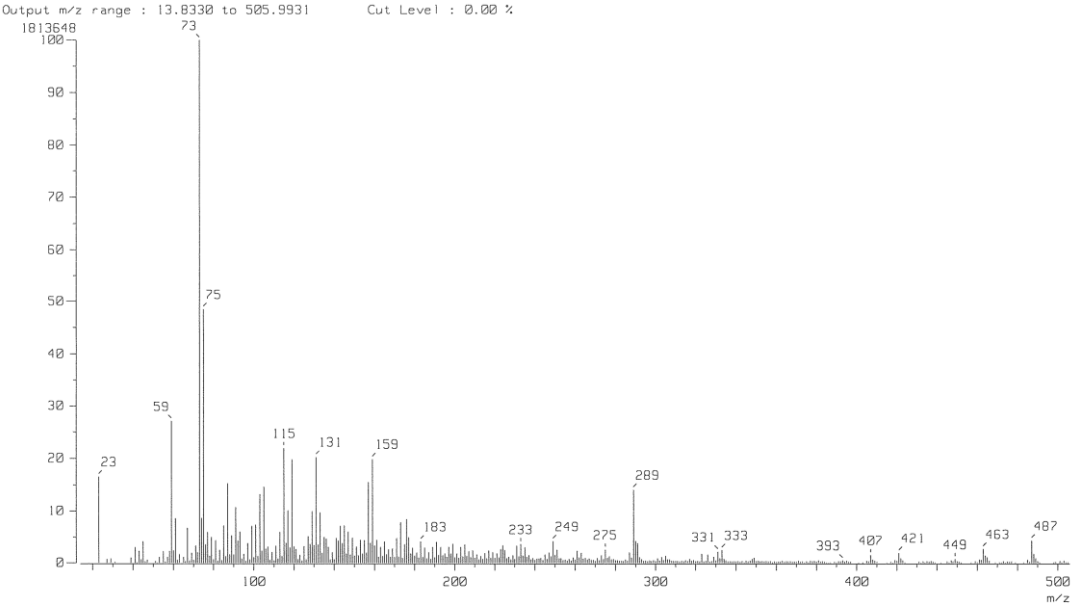
```
[ Elemental Composition ]                                          Page: 1
Data  : 381-37-1-002                 Date : 05-Nov-2021 14:52
Sample: 381-37-1
Note : NBA,Acetone+NaIaq.
Inlet : Direct                       Ion Mode : FAB+
RT : 0.00 min                        Scan#: 1
Elements : C 27/0, H 60/0, O 2/0, Si 2/0, Na 1/0
Mass Tolerance      : 1000ppm, 3mmu if m/z < 3, 5mmu if m/z > 5
Unsaturation (U.S.) : -0.5 - 30.0
```



S11

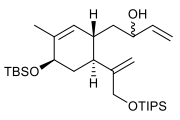
```
Observed m/z Int%   Err[ppm / mmu]    U.S.  Composition
  487.3402    6.9     -0.3 /  -0.1    3.5  C 27 H 52 O 2 Si 2 Na


[ Theoretical Ion Distribution ]                                   Page: 1
Molecular Formula : C27 H52 O2 Si2 Na
                    (m/z   487.3404, MW   487.8694, U.S.  3.5)
Base Peak :   487.3404, Averaged MW :   487.8643(a),   487.8656(w)

    m/z       INT.
  487.3404 100.0000 ************************************************************
  488.3428  40.2334 ************************
  489.3416  14.7935 *********
  490.3425   3.4504 **
  491.3426   0.6464
  492.3434   0.0947
  493.3449   0.0108
  494.3468   0.0010
```

```
[ Mass Spectrum ]
Data  : 381-37-1-Na-001              Date : 02-Nov-2021 17:13
Sample: 381-37-1
Note : NBA,Acetone+NaIaq.
Inlet : Direct                       Ion Mode : FAB+
Spectrum Type : Normal Ion [MF-Linear]
RT : 0.42 min                        Scan# : (4,5)
BP : m/z 73.0000                     Int.  : 172.96
Output m/z range : 13.8330 to 505.9931       Cut Level : 0.00 %
```



MS data of 8 (FAB+, NBA, acetone + *aq.* NaI)

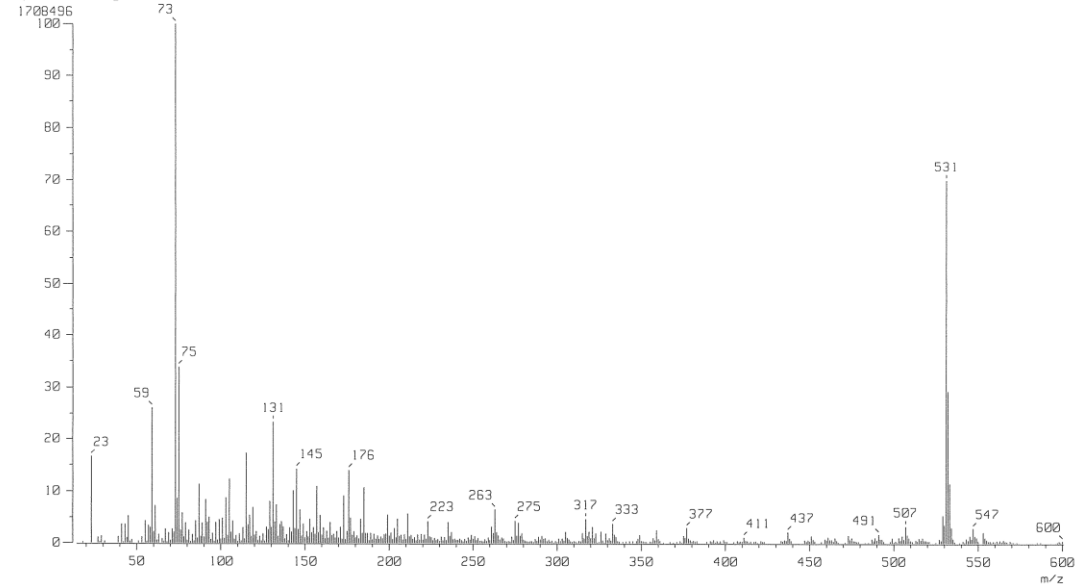
```
[ Elemental Composition ]                                          Page: 1
Data  : 381-47-1-001                  Date : 07-Sep-2021 15:09
Sample: 381-47-1
Note : NBA,Acetone+NaIaq.
Inlet : Direct                        Ion Mode : FAB+
RT : 0.40 min                         Scan#: 3
Elements : C 29/0, H 60/0, O 3/0, Si 2/0, Na 1/0
Mass Tolerance      : 1000ppm, 3mmu if m/z < 3, 5mmu if m/z > 5
Unsaturation (U.S.) : -0.5 - 30.0

Observed m/z Int%   Err[ppm / mmu]    U.S.  Composition
  531.3673   100.0     +1.4 /  +0.7    3.5  C 29 H 56 O 3 Si 2 Na
```



8

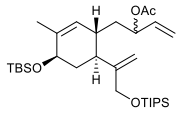
```
[ Theoretical Ion Distribution ]                                   Page: 1
Molecular Formula : C29 H56 O3 Si2 Na
                    (m/z   531.3666, MW   531.9226, U.S.  3.5)
Base Peak :   531.3666, Averaged MW :   531.9170(a),   531.9182(w)

    m/z       INT.
  531.3666 100.0000 ************************************************************
  532.3690  42.4959 *************************
  533.3681  15.9175 *********
  534.3690   3.8755 **
  535.3693   0.7579
  536.3702   0.1173
  537.3716   0.0145
  538.3735   0.0015
  539.3755   0.0001
```

```
[ Mass Spectrum ]
Data  : 381-47-1-Na-001               Date : 07-Sep-2021 14:53
Sample: 381-47-1
Note : NBA,Acetone+NaIaq.
Inlet : Direct                        Ion Mode : FAB+
Spectrum Type : Normal Ion [MF-Linear]
RT : 0.14 min                         Scan# : (2,3)
BP : m/z 73.0000                      Int.  : 162.94
Output m/z range : 14.4025 to 600.8155        Cut Level : 0.00 %
```



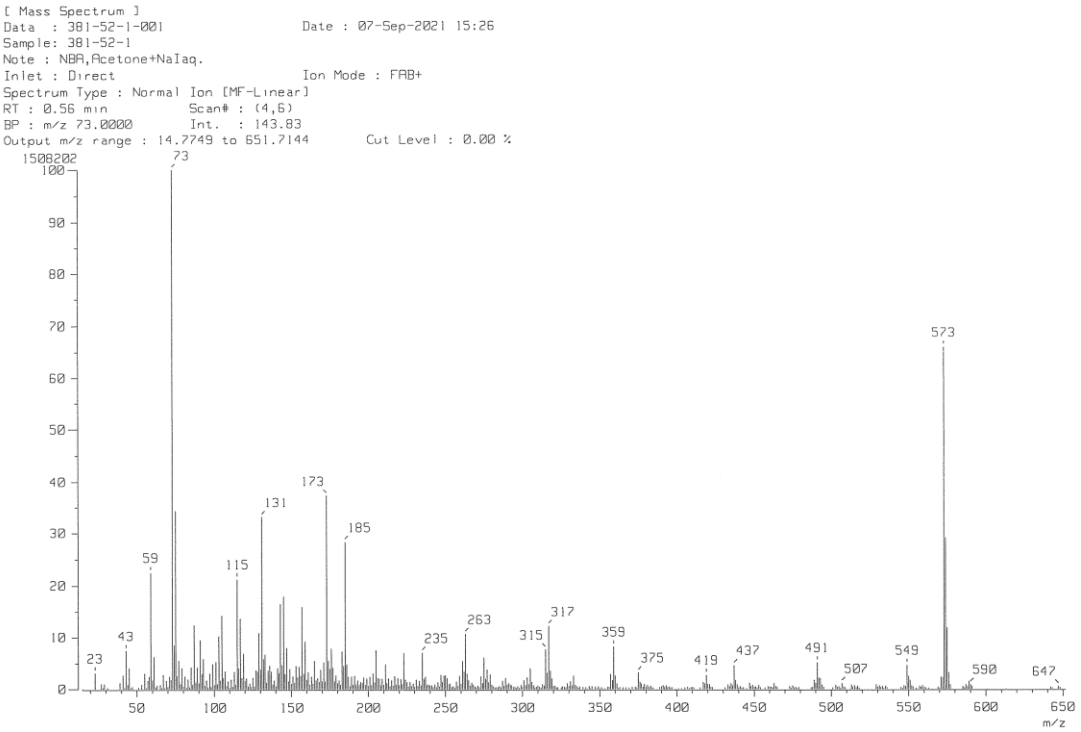
MS data of 9 (FAB+, NBA, acetone + *aq.* NaI)

```
[ Elemental Composition ]                                          Page: 1
Data  : 381-52-1-001                    Date : 07-Sep-2021 15:57
Sample: 381-52-1
Note : NBA,Acetone+NaIaq.
Inlet : Direct                          Ion Mode : FAB+
RT : 1.20 min                           Scan#: 7
Elements : C 31/0, H 60/0, O 4/0, Si 2/0, Na 1/0
Mass Tolerance       : 1000ppm, 3mmu if m/z < 3, 5mmu if m/z > 5
Unsaturation (U.S.) : -0.5 - 30.0
```



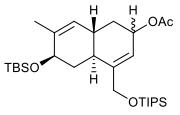
9

```
Observed m/z Int%   Err[ppm / mmu]    U.S.  Composition
  573.3776   100.0     +0.8 /  +0.5     4.5  C 31 H 58 O 4 Si 2 Na

[ Theoretical Ion Distribution ]                                   Page: 1
Molecular Formula : C31 H58 O4 Si2 Na
                    (m/z   573.3771, MW   573.9599, U.S.  4.5)
Base Peak :   573.3771, Averaged MW :   573.9540(a),   573.9552(w)

    m/z       INT.
  573.3771 100.0000 ************************************************************
  574.3796  44.7585 ************************
  575.3790  17.0927 **********
  576.3799   4.3310 ***
  577.3803   0.8815 *
  578.3812   0.1435
  579.3827   0.0190
  580.3845   0.0021
  581.3865   0.0002
```



MS data of 10 (FAB+, NBA, acetone + *aq.* NaI)

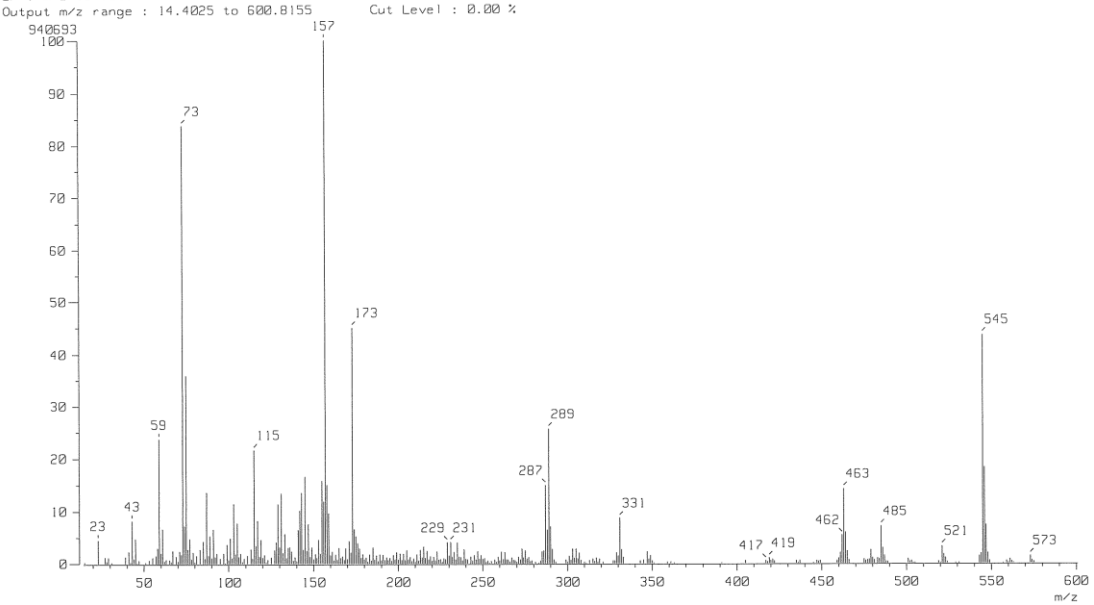
```
[ Elemental Composition ]                                          Page: 1
Data  : 381-56-1-001                  Date : 07-Sep-2021 16:14
Sample: 381-56-1
Note : NBA,Acetone+NaIaq.
Inlet : Direct                        Ion Mode : FAB+
RT : 0.60 min                         Scan#: 4
Elements : C 29/0, H 60/0, O 4/0, Si 2/0, Na 1/0
Mass Tolerance       : 1000ppm, 3mmu if m/z < 3, 5mmu if m/z > 5
Unsaturation (U.S.) : -0.5 - 30.0

Observed m/z Int%   Err[ppm / mmu]    U.S.  Composition
  545.3455   100.0     -0.7 /  -0.4    4.5  C 29 H 54 O 4 Si 2 Na
```



10

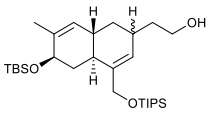
```
[ Theoretical Ion Distribution ]                                   Page: 1
Molecular Formula : C29 H54 O4 Si2 Na
                    (m/z   545.3458, MW   545.9061, U.S.  4.5)
Base Peak :   545.3458, Averaged MW :   545.9006(a),   545.9019(w)

    m/z       INT.
  545.3458 100.0000 ************************************************************
  546.3483  42.5340 ************************
  547.3474  16.1342 *********
  548.3484   3.9668 **
  549.3487   0.7913
  550.3496   0.1254
  551.3511   0.0161
  552.3529   0.0017
  553.3549   0.0002
```

```
[ Mass Spectrum ]
Data  : 381-56-1-001                  Date : 07-Sep-2021 16:08
Sample: 381-56-1
Note : NBA,Acetone+NaIaq.
Inlet : Direct                        Ion Mode : FAB+
Spectrum Type : Normal Ion [MF-Linear]
RT : 0.28 min                         Scan# : (2,4)
BP : m/z 157.0000                     Int.  : 89.71
Output m/z range : 14.4025 to 600.8155        Cut Level : 0.00 %
```



MS data of 11 (FAB+, NBA, acetone + *aq.* NaI)

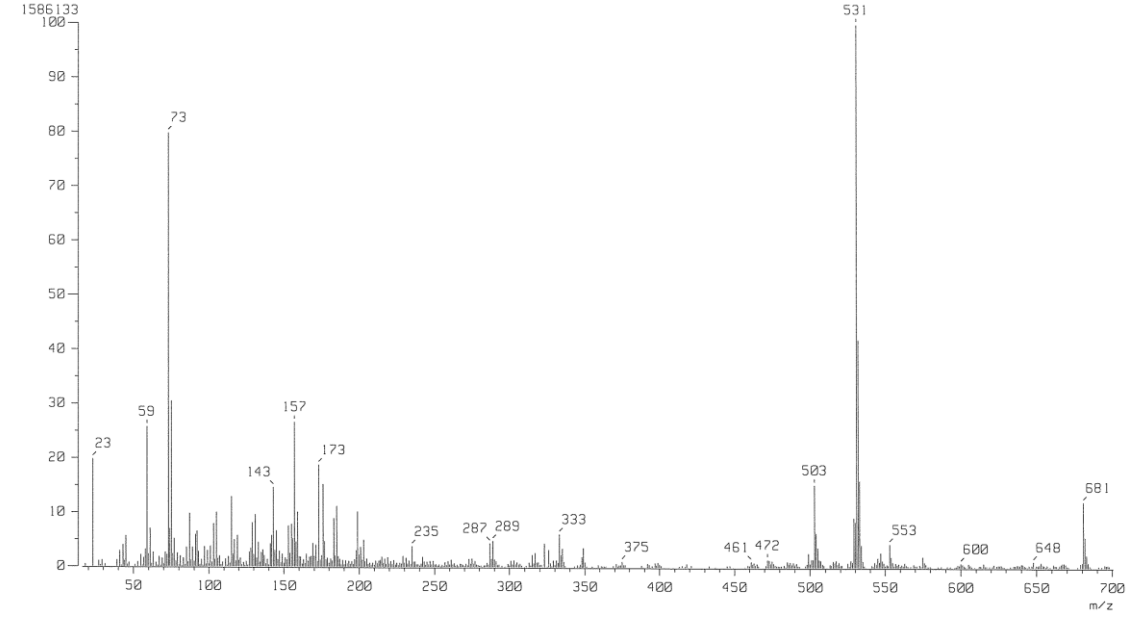
```
[ Elemental Composition ]                                          Page: 1
Data  : 381-69-1-001                  Date : 07-Sep-2021 16:52
Sample: 381-69-1
Note : NBA,Acetone+NaIaq.
Inlet : Direct                        Ion Mode : FAB+
RT : 0.40 min                         Scan#: 3
Elements : C 29/0, H 60/0, O 3/0, Si 2/0, Na 1/0
Mass Tolerance       : 1000ppm, 3mmu if m/z < 3, 5mmu if m/z > 5
Unsaturation (U.S.) : -0.5 - 30.0

Observed m/z Int%   Err[ppm / mmu]    U.S.  Composition
  531.3661   100.0     -0.8 /  -0.4    3.5  C 29 H 56 O 3 Si 2 Na
```



11

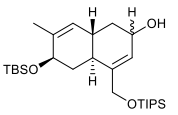
```
[ Theoretical Ion Distribution ]                                   Page: 1
Molecular Formula : C29 H56 O3 Si2 Na
                    (m/z   531.3666, MW   531.9226, U.S.  3.5)
Base Peak :   531.3666, Averaged MW :   531.9170(a),   531.9182(w)

    m/z       INT.
  531.3666 100.0000 ************************************************************
  532.3690  42.4959 *************************
  533.3681  15.9175 *********
  534.3690   3.8755 **
  535.3693   0.7579
  536.3702   0.1173
  537.3716   0.0145
  538.3735   0.0015
  539.3755   0.0001
```

```
[ Mass Spectrum ]
Data  : 381-69-1-001                  Date : 07-Sep-2021 16:20
Sample: 381-69-1
Note : NBA,Acetone+NaIaq.
Inlet : Direct                        Ion Mode : FAB+
Spectrum Type : Normal Ion [MF-Linear]
RT : 0.70 min          Scan# : (5,7)
BP : m/z 531.0000      Int.  : 151.27
Output m/z range : 16.1766 to 700.7544       Cut Level : 0.00 %
```



MS data of S14 (FAB+, NBA, acetone + *aq.* NaI)

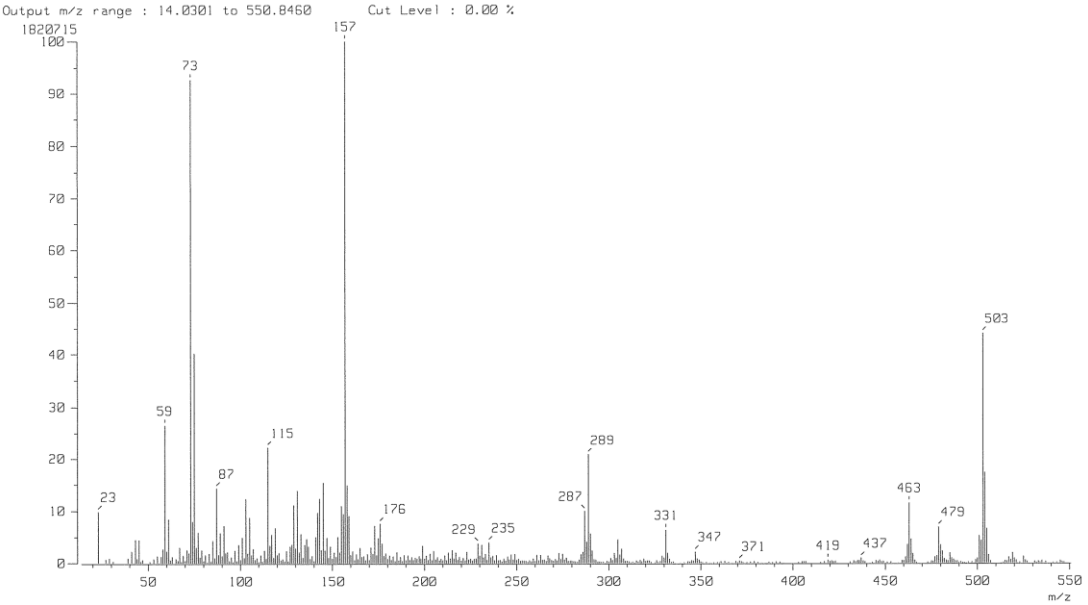
```
 [ Elemental Composition ]                                          Page: 1
 Data  : 381-89-2-001                 Date : 08-Mar-2022 16:18
 Sample: 381-89-2
 Note : NBA,Acetone+NaIaq.
 Inlet : Direct                       Ion Mode : FAB+
 RT : 1.00 min                        Scan#: 6
 Elements : C 27/0, H 60/0, O 3/0, Si 2/0, Na 1/0
 Mass Tolerance      : 1000ppm, 3mmu if m/z < 3, 5mmu if m/z > 5
 Unsaturation (U.S.) : -0.5 - 30.0

 Observed m/z Int%   Err[ppm / mmu]    U.S.  Composition
   503.3356   100.0     +0.7 /  +0.3    3.5  C 27 H 52 O 3 Si 2 Na
```



S14

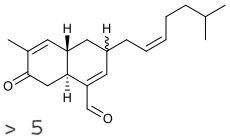
```
[ Theoretical Ion Distribution ]                                    Page: 1
Molecular Formula : C27 H52 O3 Si2 Na
                    (m/z   503.3353, MW   503.8688, U.S.  3.5)
Base Peak :   503.3353, Averaged MW :   503.8636(a),   503.8649(w)

    m/z       INT.
  503.3353 100.0000 ************************************************************
  504.3377  40.2715 ************************
  505.3366  15.0093 *********
  506.3375   3.5367 **
  507.3377   0.6773
  508.3385   0.1018
  509.3400   0.0121
  510.3419   0.0012
```

```
[ Mass Spectrum ]
Data  : 381-89-2-Na-001              Date : 08-Mar-2022 15:42
Sample: 381-89-2
Note : NBA,Acetone+NaIaq.
Inlet : Direct                       Ion Mode : FAB+
Spectrum Type : Normal Ion [MF-Linear]
RT : 0.42 min          Scan# : (3,6)
BP : m/z 157.0000      Int.  : 173.64
Output m/z range : 14.0301 to 550.8460       Cut Level : 0.00 %
```



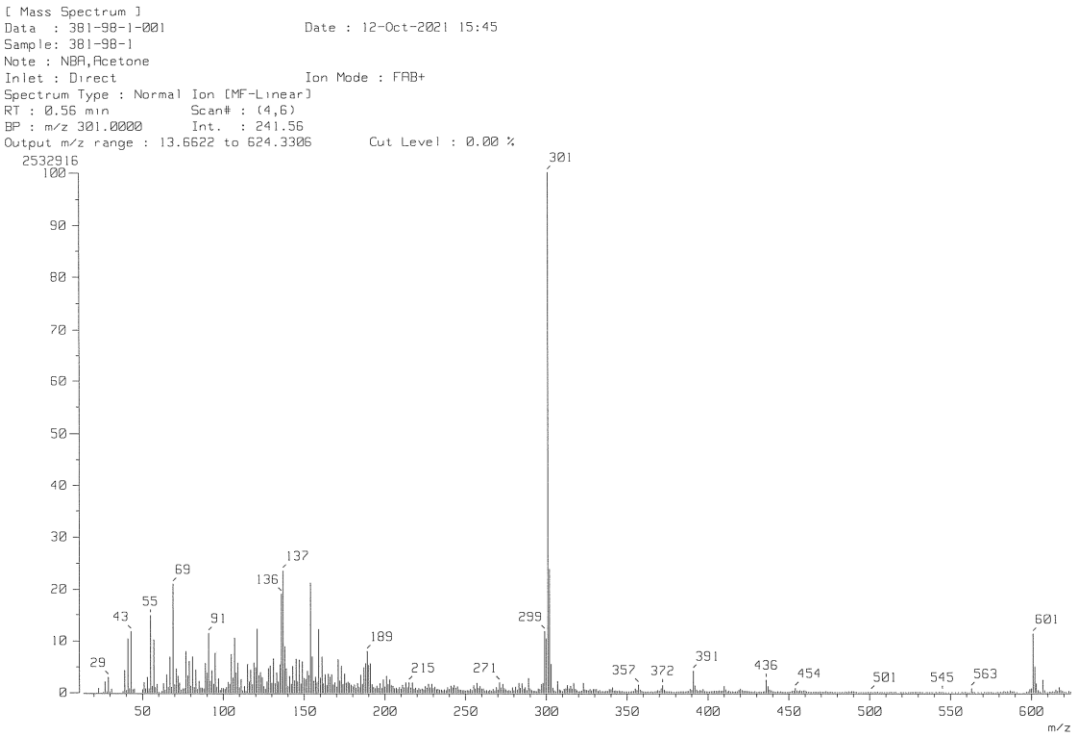
MS data of S18 (FAB+, NBA, acetone)

```
 [ Elemental Composition ]                                          Page: 1
 Data  : 381-98-1-001                 Date : 12-Oct-2021 15:58
 Sample: 381-98-1
 Note : NBA,Acetone
 Inlet : Direct                       Ion Mode : FAB+
 RT : 2.20 min                        Scan#: 12
 Elements : C 20/0, H 40/0, O 2/0
 Mass Tolerance      : 1000ppm, 3mmu if m/z < 3, 5mmu if m/z > 5
 Unsaturation (U.S.) : -0.5 - 30.0

 Observed m/z Int%   Err[ppm / mmu]    U.S.  Composition
   301.2170   100.0     +0.7 /  +0.2    6.5  C 20 H 29 O 2
```



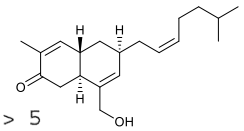
S18

```
 [ Theoretical Ion Distribution ]                                   Page: 1
Molecular Formula : C20 H29 O2
                    (m/z   301.2168, MW   301.4491, U.S.  6.5)
Base Peak :   301.2168, Averaged MW :   301.4463(a),   301.4471(w)

    m/z       INT.
  301.2168 100.0000 ************************************************************
  302.2201  22.3209 *************
  303.2231   2.7683 **
  304.2259   0.2480
  305.2287   0.0174
  306.2315   0.0010
```



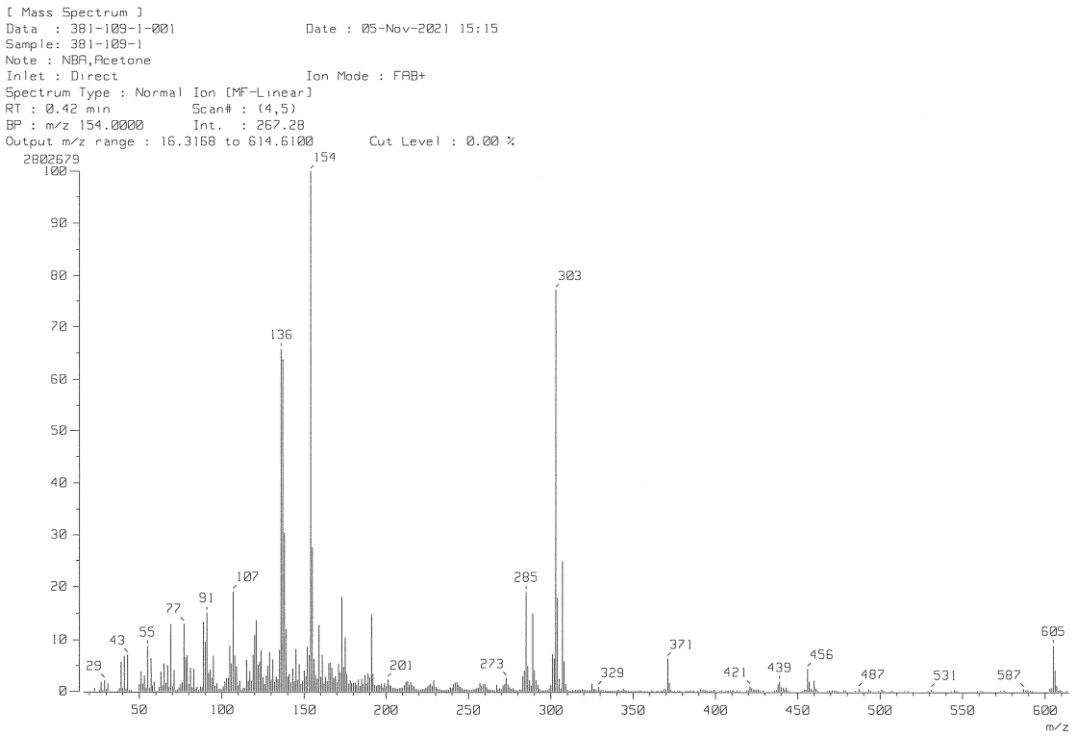
MS data of 1 (FAB+, NBA, acetone)

```
 [ Elemental Composition ]                                          Page: 1
 Data  : 381-109-1-001                Date : 05-Nov-2021 15:31
 Sample: 381-109-1
 Note : NBA,Acetone
 Inlet : Direct                       Ion Mode : FAB+
 RT : 0.20 min                        Scan#: 2
 Elements : C 20/0, H 40/0, O 2/0
 Mass Tolerance      : 1000ppm, 3mmu if m/z < 3, 5mmu if m/z > 5
 Unsaturation (U.S.) : -0.5 - 30.0

 Observed m/z Int%   Err[ppm / mmu]    U.S.  Composition
   303.2321    44.1     -1.0 /  -0.3    5.5  C 20 H 31 O 2
```



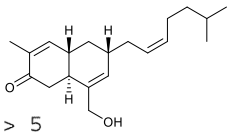
1

```
 [ Theoretical Ion Distribution ]                                   Page: 1
Molecular Formula : C20 H31 O2
                    (m/z   303.2324, MW   303.4649, U.S.  5.5)
Base Peak :   303.2324, Averaged MW :   303.4619(a),   303.4627(w)

    m/z       INT.
  303.2324 100.0000 ************************************************************
  304.2358  22.3209 *************
  305.2388   2.7683 **
  306.2416   0.2480
  307.2444   0.0174
  308.2472   0.0010
```



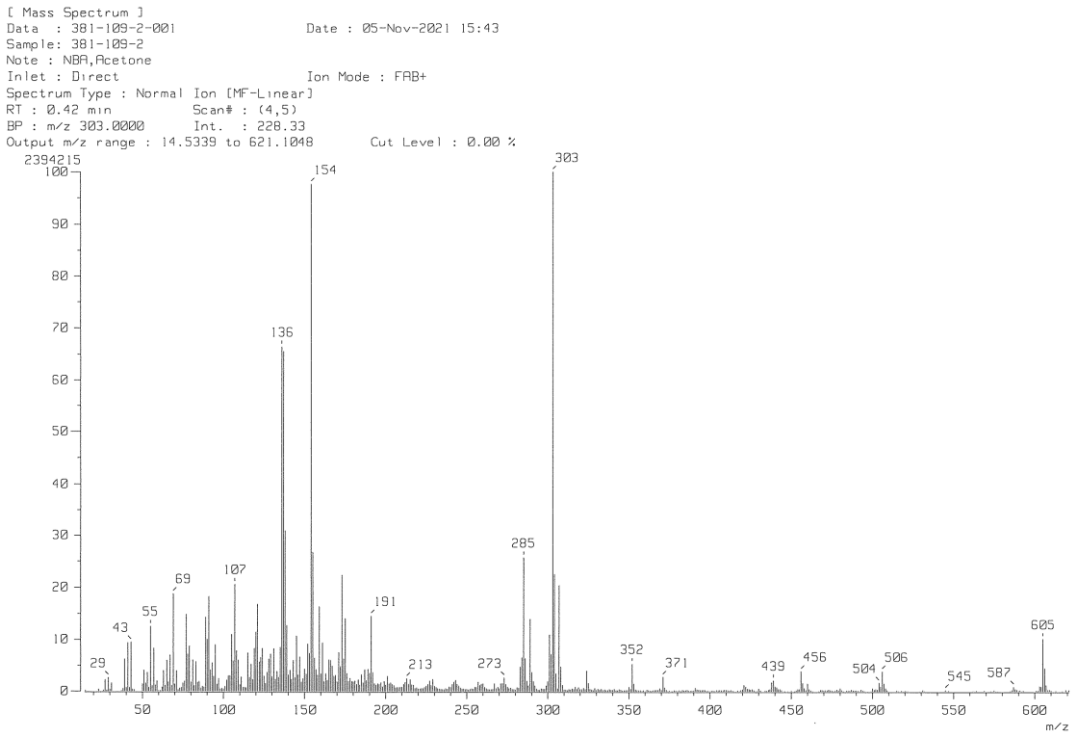
MS data of 2 (FAB+, NBA, acetone)

```
[ Elemental Composition ]                                          Page: 1
Data  : 381-109-2-001                Date : 05-Nov-2021 15:49
Sample: 381-109-2
Note : NBA,Acetone
Inlet : Direct                       Ion Mode : FAB+
RT : 0.00 min                        Scan#: 1
Elements : C 20/0, H 40/0, O 2/0
Mass Tolerance       : 1000ppm, 3mmu if m/z < 3, 5mmu if m/z > 5
Unsaturation (U.S.) : -0.5 - 30.0

Observed m/z Int%   Err[ppm / mmu]    U.S.  Composition
  303.2318   100.0     -1.9 /  -0.6    5.5  C 20 H 31 O 2
```



2

```
[ Theoretical Ion Distribution ]                                   Page: 1
Molecular Formula : C20 H31 O2
                    (m/z   303.2324, MW   303.4649, U.S.  5.5)
Base Peak :   303.2324, Averaged MW :   303.4619(a),   303.4627(w)

    m/z       INT.
  303.2324 100.0000 ************************************************************
  304.2358  22.3209 *************
  305.2388   2.7683 **
  306.2416   0.2480
  307.2444   0.0174
  308.2472   0.0010
```



6. Green chemistry metrics

Reference: Q. H. Pho, M. Escriba-Gelonch, D. Losic, E. V. Rebrov, N. N. Tran and V. Hessel, *ACS Sustainable Chem. Eng.* 2021, **9**, 4755.

Raw material: a substrate (1 eq.) of the reaction.
Product: a product of the reaction.
Auxiliary reagents: All reagents used for the reaction, except for solvents.
Solvents: All organic solvents used for the reaction, extraction and purification, except for water.
Water: All water used for the reaction, as a quencher and for washing the organic extracts.

CY: Chemical yield.
RME: Reaction mass efficiency, which is the mass of product divided by the mass of raw material.
CE: Carbon economy, which is the mass of carbon in the product divided by the total mass of carbon in the raw material.
AE: Atom economy, which is the molecular weight of the product divided by the molecular weight of raw material and auxiliary reagents.
E-factor: E factor is the mass of wastes produced considering the reaction, extraction, purification, including raw material (subtracted the mass of product), auxiliary reagents and solvent, yet excluding water, divided by the mass of product.
MP: Mass productivity, which is the mass of product divided by the mass of all input materials including raw material, auxiliary reagents, solvent, and even water.
MI: Mass intensity, which is an inversely defined metrics as to mass productivity (MP).
SI: Solvent intensity, which is the mass intensity (MI) only considering solvents as input materials.
EMY: Effective mass yield, which is the mass of product divided by the mass of raw material and auxiliary reagents.
RMI: Reaction mass intensity, which is an inversely defined metrics as to effective mass yield (EMY).

Table S3 Green metrics calculation.

Metrics	Unit	Ideal value	Reactions												
			Carvone –>**S8**	**S8–>3**	**3–>S9**	**S9–>4**	**4–>5**	**5–>6**	**6–>S11**	**S11–>8**	**8–>9**	**9–>10**	**10–>11**	**S14–>10**	**11–>S18**	**S18–>1&2**
CY	%	100	60	96	quant.	90	46	98	94	42	98	98	30	81	40	71
RME	%	100	65	>100	>100	94	50	89	>100	46	>100	93	29	89	24	72
CE	%	100	59	92	>100	90	49	>100	>100	45	>100	92	30	88	28	71
AE	%	100	37	76	44	69	41	46	64	33	67	37	49	65	13	59
E-factor	kg/kg	0	32	7.0	25	30	75	42	26	160	72	260	480	51	2800	960
MP	%	100	0.26	1.0	0.41	0.20	0.14	0.39	0.40	0.023	0.11	0.13	0.013	0.10	0.0038	0.010
MI	kg/kg	0	380	99	250	500	700	260	250	4400	930	790	7800	1000	27000	10000
SI	kg/kg	0	300	84	210	470	640	230	230	3600	880	780	6500	970	23000	9100
EMY	%	100	10	62	41	49	13	21	48	8.7	37	87	10	42	1.5	9.4
RMI	kg/kg	0	10	1.6	2.5	2.0	7.8	4.9	2.1	12	2.7	1.2	9.6	2.4	68	11