Supplementary Information (SI) for Chemical Science.
This journal is © The Royal Society of Chemistry 2024

Supplementary information

Discovery of polyketide carboxylate phytotoxin from polyketide glycoside hybrid by β-glucosidase mediated ester bond hydrolysis

Xin Wang[a], De-Kun Kong[a], Hua-Ran Zhang[a] and Yi Zou*[a]

[a] College of Pharmaceutical Sciences, Southwest University, Chongqing 400715, P. R. China

* Corresponding author: E-mail: zouyi31@swu.edu.cn

Table of contents

1. Supplementary Experimental Procedures 5

1.1 General methods 5
1.2 Strains and culture conditions 5
1.3 Plasmids construction 5
1.4 Extraction of gDNA and the synthesis of cDNA 7
1.5 Preparation and transformation of *F. proliferatum* or *A. nidulans* protoplasts 7
1.6 Knockout of *pro* cluster in *F. proliferatum* 8
1.7 Heterologous expression of *pro* cluster in *A. nidulans* 8
1.8 Chemical feeding of compounds **1** and **2** into *Fp*-Δ*proA*-*L* strain 8
1.9 Compound **1** in PBS buffer under different pH values 8
1.10 ECD calculations for compound **7** 8
1.11 Protein expression and purification in *E. coli* 9
1.12 Preparation and characterization of *A. nidulans* microsome and crude enzymes 9
1.13 *In vitro* characterization of ProG towards compound **1** 9
1.14 *In vitro* characterization of ProE towards compound **8** 10
1.15 Time-course assays for ProI towards compound **7** 10
1.16 *In vitro* characterization of microsome fractions of ProJ and ProI 10
1.17 Time-course assays for crude enzymes of *AN-proL* towards compound **1** 10
1.18 *In vitro* characterization of crude enzymes of *AN-proK*, *AN-proL* and its mutants 10
1.19 *In vivo* $H_2{}^{18}O$ labeling experiments and LC–MS detection 11
1.20 Acid hydrolysis of compounds **1** and **9** and HPLC detection 11
1.21 The procedure for **3** and **5** configuration determination 11
1.22 Purification and structural characterization of compounds 12
1.23 Root growth inhibition assay of *Arabidopsis thaliana* 13
1.24 Bioinformatics analysis 13

2. Supplementary Tables 14

Table S1. Stains and plasmids used in this study. 14
Table S2. Primers used in this study. 15
Table S3. Plasmids used in this study. 18
Table S4. The protein sequence of *pro* gene cluster. 19
Table S5. HR-MS data of compounds in this study. 22
Table S6. NMR data of compound **1** in DMSO-d_6. 23
Table S7. NMR data of compound **2** in $CDCl_3$ 24
Table S8. NMR data of compound **3** in DMSO-d_6. 25
Table S9. NMR data of compound **4** in DMSO-d_6. 26
Table S10. NMR data of compound **5** in DMSO-d_6. 27
Table S11. NMR data of compound **6** in DMSO-d_6. 28
Table S12. The ^{1}H NMR spectroscopic data of compounds **3*S*** (*R*-MTPA derivative) and **3*R*** (*S*-MTPA derivative) in DMSO-d_6. 29
Table S13. The ^{1}H NMR spectroscopic data of compounds **5*S*** (*R*-MTPA derivative) and **5*R*** (*S*-MTPA derivative) in DMSO-d_6. 30
Table S14. NMR data of compound **7** in DMSO-d_6. 31
Table S15. Energy (298.15 K) analysis for **7*R*** and **7*S***. 32
Table S16. NMR data of compound **8** in DMSO-d_6. 33
Table S17. NMR data of compound **9** in DMSO-d_6. 34
Table S18. NMR data of compound **9** in pyridine-d_5. 35

3. Supplementary Figures 36

Figure S1. Schematic diagram of plasmids used in this study. 36
Figure S2. The well-known fungal phytotoxins produced by *Fusarium* species. 37
Figure S3. The genome-mined homologue gene clusters of PKS16 in public and our lab database. 38
Figure S4. The prediction of transmembrane helices of mGT ProJ. 39
Figure S5. The *pro* genes are all under transcription when *F. proliferatum* is incubated on glycerol medium 40
Figure S6. PCR verification of knockout mutants of *F. proliferatum* in this study. 41
Figure S7. HPLC (refractive index detector) analysis of the glycosyl of compound **1**. 42
Figure S8. LC–MS analysis of the feeding assay of *Fp*-Δ*proA*-*L* towards compound **2**. 43

Figure S9. SDS-PAGE of proteins purified from *E. coli* BL21 in this study. 44
Figure S10. LC–MS analysis of the *in vitro* biochemical assays of α,β-hydrolase ProG towards compound **1**. 45
Figure S11. LC–MS analysis of the feeding assay of *Fp*-Δ*proA*-*L* towards compound **1**. 46
Figure S12. LC–MS analysis of compound **1** in PBS buffer under different pH values. 47
Figure S13. Comparison of calculated ECD spectra of (7*S*)- and (7*R*)-isomers with the experimental spectra of **7**. 48
Figure S14. Time-course assays for FMO ProI towards compound **7**. 49
Figure S15. Bioinformatic analysis of β-glucosidase ProL. 50
Figure S16. Traditional catalytic mechanism of β-glucosidase that hydrolyze glycosidic bonds. 51
Figure S17. HPLC (evaporative light scattering detector) analysis of the glycosyl of compound **9**. 52
Figure S18. LC–MS analysis of the *in vitro* biochemical assays of mGT ProJ towards compound **1**. 53
Figure S19. Sequence similarity network (SSN) analysis of unknown function protein ProK. 54
Figure S20. The β-glucosidase ProL with His-tag in *E. coli* BL21 is insoluble. 55
Figure S21. LC–MS analysis of *in vitro* biochemical assays of crude enzymes of *AN-proK* towards compound **1**. 56
Figure S22. LC–MS analysis of mutagenesis experiments on the conventional active sites D266 and E493 of β-glucosidases in ProL towards compound **1**. 57
Figure S23. UV absorption and HRMS spectrum (positive ionization) analysis of compound **1**. 58
Figure S24. UV absorption and HRMS spectrum (positive ionization) analysis of compound **2**. 58
Figure S25. UV absorption and HRMS spectrum (positive ionization) analysis of compound **3**. 59
Figure S26. UV absorption and HRMS spectrum (positive ionization) analysis of compound **4**. 59
Figure S27. UV absorption and HRMS spectrum (positive ionization) analysis of compound **5**. 60
Figure S28 UV absorption and HRMS spectrum (positive ionization) analysis of compound **6**. 60
Figure S29. UV absorption and HRMS spectrum (negative ionization) analysis of compound **7**. 61
Figure S30. UV absorption and HRMS spectrum (positive ionization) analysis of compound **8**. 61
Figure S31. UV absorption and HRMS spectrum (negative ionization) analysis of compound **9**. 62
Figure S32. UV absorption and HRMS spectrum (negative ionization) analysis of compound **3*S***. 62
Figure S33. UV absorption and HRMS spectrum (negative ionization) analysis of compound **3*R***. 63
Figure S34. UV absorption and HRMS spectrum (negative ionization) analysis of compound **5*S***. 63
Figure S35. UV absorption and HRMS spectrum (negative ionization) analysis of compound **5*R***. 64
Figure S36. ^{1}H NMR spectrum of compound **1** in DMSO-d_6 (400 MHz). 65
Figure S37. ^{13}C NMR spectrum of compound **1** in DMSO-d_6 (100 MHz). 66
Figure S38. DEPT-135° spectrum of compound **1** in DMSO-d_6. 67
Figure S39. HSQC spectrum of compound **1** in DMSO-d_6. 68
Figure S40. HMBC spectrum of compound **1** in DMSO-d_6. 69
Figure S41. ^{1}H-^{1}H COSY spectrum of compound **1** in DMSO-d_6. 70
Figure S42. ^{1}H-^{1}H NOESY spectrum of compound **1** in DMSO-d_6. 71
Figure S43. ^{1}H NMR spectrum of compound **2** in $CDCl_3$ (400 MHz). 72
Figure S44. ^{13}C NMR spectrum of compound **2** in $CDCl_3$ (100 MHz). 73
Figure S45. DEPT-135° spectrum of compound **2** in $CDCl_3$. 74
Figure S46. HSQC spectrum of compound **2** in $CDCl_3$. 75
Figure S47. HMBC spectrum of compound **2** in $CDCl_3$. 76
Figure S48. ^{1}H-^{1}H COSY spectrum of compound **2** in $CDCl_3$. 77
Figure S49. ^{1}H-^{1}H NOESY spectrum of compound **2** in $CDCl_3$. 78
Figure S50. ^{1}H NMR spectrum of compound **3** in DMSO-d_6 (400 MHz). 79
Figure S51. ^{13}C NMR spectrum of compound **3** in DMSO-d_6 (100 MHz). 80
Figure S52. DEPT-135° spectrum of compound **3** in DMSO-d_6. 81
Figure S53. HSQC spectrum of compound **3** in DMSO-d_6. 82
Figure S54. HMBC spectrum of compound **3** in DMSO-d_6. 83
Figure S55. ^{1}H-^{1}H COSY spectrum of compound **3** in DMSO-d_6. 84
Figure S56. ^{1}H-^{1}H NOESY spectrum of compound **3** in DMSO-d_6. 85
Figure S57. ^{1}H NMR spectrum of compound **4** in DMSO-d_6 (400 MHz). 86
Figure S58. ^{13}C NMR spectrum of compound **4** in DMSO-d_6 (100 MHz). 87
Figure S59. DEPT-135° spectrum of compound **4** in DMSO-d_6. 88
Figure S60. HSQC spectrum of compound **4** in DMSO-d_6. 89
Figure S61. HMBC spectrum of compound **4** in DMSO-d_6. 90
Figure S62. ^{1}H-^{1}H COSY spectrum of compound **4** in DMSO-d_6. 91
Figure S63. ^{1}H-^{1}H NOESY spectrum of compound **4** in DMSO-d_6. 92
Figure S64. ^{1}H NMR spectrum of compound **5** in DMSO-d_6 (400 MHz). 93

Figure S65. ^{13}C NMR spectrum of compound **5** in DMSO-d_6 (100 MHz) 94
Figure S66. DEPT-135° spectrum of compound **5** in DMSO-d_6. 95
Figure S67. HSQC spectrum of compound **5** in DMSO-d_6. 96
Figure S68. HMBC spectrum of compound **5** in DMSO-d_6. 97
Figure S69. ^{1}H-^{1}H COSY spectrum of compound **5** in DMSO-d_6. 98
Figure S70. ^{1}H-^{1}H NOESY spectrum of compound **5** in DMSO-d_6. 99
Figure S71. ^{1}H NMR spectrum of compound **6** in DMSO-d_6 (400 MHz) 100
Figure S72. ^{13}C NMR spectrum of compound **6** in DMSO-d_6 (100 MHz) 101
Figure S73. DEPT-135° spectrum of compound **6** in DMSO-d_6. 102
Figure S74. HSQC spectrum of compound **6** in DMSO-d_6. 103
Figure S75. HMBC spectrum of compound **6** in DMSO-d_6. 104
Figure S76. ^{1}H-^{1}H COSY spectrum of compound **6** in DMSO-d_6. 105
Figure S77. ^{1}H-^{1}H NOESY spectrum of compound **6** in DMSO-d_6. 106
Figure S78. ^{1}H NMR spectrum of compound **3*S*** in DMSO-d_6 (400 MHz). 107
Figure S79. ^{1}H NMR spectrum of compound **3*R*** in DMSO-d_6 (400 MHz). 108
Figure S80. ^{1}H NMR spectrum of compound **5*S*** in DMSO-d_6 (400 MHz). 109
Figure S81. ^{1}H NMR spectrum of compound **5*R*** in DMSO-d_6 (400 MHz). 110
Figure S82. ^{1}H NMR spectrum of compound **7** in DMSO-d_6 (400 MHz) 111
Figure S83. ^{13}C NMR spectrum of compound **7** in DMSO-d_6 (100 MHz) 112
Figure S84. DEPT-135° spectrum of compound **7** in DMSO-d_6. 113
Figure S85. HSQC spectrum of compound **7** in DMSO-d_6. 114
Figure S86. HMBC spectrum of compound **7** in DMSO-d_6. 115
Figure S87. ^{1}H-^{1}H COSY spectrum of compound **7** in DMSO-d_6. 116
Figure S88. ^{1}H-^{1}H NOESY spectrum of compound **7** in DMSO-d_6. 117
Figure S89. ^{1}H NMR spectrum of compound **8** in DMSO-d_6 (400 MHz) 118
Figure S90. ^{13}C NMR spectrum of compound **8** in DMSO-d_6 (100 MHz) 119
Figure S91. DEPT-135° spectrum of compound **8** in DMSO-d_6. 120
Figure S92. HSQC spectrum of compound **8** in DMSO-d_6. 121
Figure S93. HMBC spectrum of compound **8** in DMSO-d_6. 122
Figure S94. ^{1}H-^{1}H COSY spectrum of compound **8** in DMSO-d_6. 123
Figure S95. ^{1}H-^{1}H NOESY spectrum of compound **8** in DMSO-d_6. 124
Figure S96. ^{1}H NMR spectrum of compound **9** in DMSO-d_6 (400 MHz) 125
Figure S97. ^{13}C NMR spectrum of compound **9** in DMSO-d_6 (100 MHz) 126
Figure S98. DEPT-135° spectrum of compound **9** in DMSO-d_6. 127
Figure S99. HSQC spectrum of compound **9** in DMSO-d_6. 128
Figure S100. HMBC spectrum of compound **9** in DMSO-d_6. 129
Figure S101. ^{1}H-^{1}H COSY spectrum of compound **9** in DMSO-d_6. 130
Figure S102. ^{1}H-^{1}H NOESY spectrum of compound **9** in DMSO-d_6. 131
Figure S103. ^{1}H NMR spectrum of compound **9** in pyridine-d_5 (400 MHz). 132
Figure S104. ^{13}C NMR spectrum of compound **9** in pyridine-d_5 (100 MHz). 133
Figure S105. DEPT-135° spectrum of compound **9** in pyridine-d_5. 134
Figure S106. HSQC spectrum of compound **9** in pyridine-d_5. 135
Figure S107. HMBC spectrum of compound **9** in pyridine-d_5. 136
Figure S108. ^{1}H-^{1}H COSY spectrum of compound **9** in pyridine-d_5. 137
Figure S109. ^{1}H-^{1}H NOESY spectrum of compound **9** in pyridine-d_5 138

4. Sequence information of *pro* gene cluster 139

5. Supplementary references 148

1. Supplementary Experimental Procedures

1.1 General methods

Reagents were purchased from Sigma-Aldrich, Thermo Fisher Scientific, or New England BioLabs. Sangon Biotech Co., Ltd. (Shanghai, China) carried out the synthesis of primers and DNA sequencing. LC–MS analysis were performed on Waters ACQUITY H-Class UPLC–MS system coupled to a PDA detector and a SQD2 mass spectrometer (MS) detector with an ESI source. Chromatographic separation was executed at 35 °C utilizing a C18 column (ACQUITY UPLC® BEH, 1.7 μm, 2.1 mm × 100 mm, Waters). LC–MS metabolite profiles were generated on a Waters UPLC–MS system with the method outlined below: Chromatographic separation was achieved with a linear gradient of 5–99% MeCN-H_2O (both containing 0.02% *v/v* formic acid) in 10 minutes followed by 99% MeCN for 3 minutes and then 5% MeCN-H_2O for 3 minutes, at a flow rate of 0.4 mL/min. The MS data were collected in the *m/z* range 50–1500 in positive modes, simultaneously. MPLC was performed on BUCHI Reveleris® X2 Flash Chromatography System, with UV detectors and a BUCHI Reveleris® C18 column (40 μm, 80 g). Semi-preparative HPLC was performed on Shimadzu Prominence HPLC system using a YMC-Pack ODS-A column (5 μm, 10 × 250 mm). MCI Column chromatography (CC) was performed using MCI gel CHP 20P/P120 (37–75 μm, Mitsubishi Chemical Corporation, Japan). NMR spectra were recorded on Bruker AVANCE III NMR (400 MHz) with a 5 mm broadband probe and TMS as an internal standard. HRMS data were generated using Fourier-transform ion cyclotron resonance-mass spectrometry (FT-ICR-MS) (Bruker SolariII, Bremen, Germany) or quadrupole time-of-flight (QTOF) mass spectrometer (Bruker IMPACT II, Bremen, Germany). ECD spectra were recorded on a JASCO J-810 spectrometer.

1.2 Strains and culture conditions

Fusarium proliferatum CGMCC 3.4710 was used as the host for knockout of the *pro* gene cluster and the isolation of compounds **1**–**6** and **9**.The extraction of the genome DNA (gDNA) and complementary DNA (cDNA) was performed by culturing *F. proliferatum* on solid glycerol medium (20 mL/L glycerol, 2 g/L tryptone, 2 g/L yeast extract, 20 g/L agar) at 25 °C for 4 days. For sporulation, *F. proliferatum* was maintained in a 4% mung bean water medium at 25 °C for 2 days. *Aspergillus nidulans* LO8030 was employed as the host for the heterologous expression of the *pro* gene cluster, ProJ microsome and ProL crude enzymes. *A. nidulans* was cultured at 37 °C for 3 days on solid CD medium (10 g/L glucose, 50 mL/L 20 × nitrate salts, 1 mL/L trace elements, 20 g/L agar) for sporulation or at 25 °C, 3.5 days on solid CD-ST medium (20 g/L starch, 10 g/L casein hydrolysate (acid), 50 mL/L 20 ×nitrate salts, 1 mL/L trace elements, 20 g/L agar) for heterologous expression and compounds production. *Escherichia coli* BL21 was used for the protein expression of ProG, ProE, ProI and ProL. While *E. coli* XL-1 was employed for cloning. All *E. coli* strains were cultured at 37 °C for cloning or 16 °C for protein expression.

1.3 Plasmids construction

The plasmids and primers utilized in this study are summarized in Tables S2–3, respectively. The amplified DNA fragments from gDNA of *F. proliferatum* by specific primers and polymerase were cloned into vectors utilizing the yeast homologous recombination in *S. cerevisiae* BJ5464-NpgA to generate heterologous expression and knockout plasmids. In order to express protein *in E. coli*, intron-free genes

were amplified from the cDNA of *F. proliferatum* and then digested with specific restriction enzymes. Subsequently, the digested fragments were ligated to the digested vector pQ8, which is a protein expression vector with an MBP tag on the *N*-terminus. The resulting plasmids were confirmed by DNA sequencing and then used for expression or knockout purposes. The schematic diagram of plasmids used in this study are shown in Figure S1.

To generate knockout strains of *proA–L, A, C, D, E, I, K* and *L* in *Fusarium proliferatum* CGMCC 3.4710, plasmids pIM 2501–2507 were gained through yeast homologous recombination in *S. cerevisiae* BJ5464-NpgA. These plasmids contain the *hygB* resistance gene, which was amplified from gDNA of *hygB* resistance knockout strain by PCR using primer pairs pYEU-*hyg*-F/pYEU-*hyg*-R. The upstream and downstream (~2,000 bp) of *proA* were obtained from the gDNA of *F. proliferatum* CGMCC 3.4710 by PCR using primer pairs pYEU-KO*proA*-UP-F/pYEU-KO*proA*-UP-R and pYEU-KO*proA*-DN-F/pYEU-KO*proA*-DN-R and cloned to *Spe* I-*Pml* I digested vector pYEU with *hygB* resistance gene to get plasmid pIM 2501. Plasmids pIM 2502–2507 were gained by the same method.

To construct the expression plasmids of *proF, G, E, I, J, K and L* genes for *A. nidulans*, each gene with its terminator was amplified from the gDNA of *F. proliferatum* CGMCC 3.4710 by PCR. The *glaA*, *gpdA*, *amyB* promoters were amplified from vectors pANU, pANR, pANP by using primer pairs *glaA*-F/*glaA*-R, *gpdA*-F/*gpdA*-R and *amyB*-F/*amyB*-R, respectively. Plasmid pANU was digested with *Not* I and plasmids pANR, pANP were digested with *BamH* I to use as vectors for insert genes. The specific construction methods of plasmids pIM 2508–2516 are as follows: The gene *proF* was amplified by PCR using primer pairs pANU-*proF*-F1/pANU-*proF*-R1, pANU-*proF*-F2/pANU-*proF*-R2, pANU-*proF*-F3/pANU-*proF*-R3, and was cloned into vector pANU with *glaA* promoter yielding plasmid pIM 2508. The genes *proF* and *proG* were amplified by PCR using primer pairs pANU-*proF*-F1/pANU-*proF*-R1, pANU-*proF*-F2/pANU-*proF*-R2, pANU-*proF*-F3/pANU-*proFG*-R3, pANU-*proG*-F/pANU-*proG*-R, and were cloned into vector pANU yielding plasmid pIM 2509, the promoters *glaA* and *gpdA* were used for *proF* and *proG* respectively. The genes *proG, proI* and *proJ* were amplified by PCR using primer pairs pANU-*proG*-F/pANR-*proG*-R, pANR-*proI*-F/ANR-*proIJ*-R, and were cloned into vector pANR yielding plasmid pIM 2510, the promoters *gpdA, glaA, and amyB* (amplified by PCR using primer pairs pANR-*amyB*-*proJ*-F/*amyB*-R) were used for *proG, proI* and *proJ* respectively. The genes *proE, proI* and *proJ* were amplified by PCR utilizing the following primer pairs: pANR-*proE*-F/pANR-*proE*-R, pANR-*proI*-F/pANR-*proIJ*-R and were cloned into vector pANR resulting in the construction of the plasmid pIM 2511, the promoters *gpdA, glaA, and amyB* (amplified by PCR using primer pairs pANR-*amyB*-*proJ*-F/*amyB*-R) were used for *proE, proI* and *proJ* respectively. The gene *proJ* was amplified by PCR using primer pairs pANR-*proJ*-F/pANR-*proJ*-R, and was cloned into vector pANR with *gpdA* promoter yielding plasmid pIM 2512.The gene *proK* was amplified by PCR using primer pairs pANP-*proK*-F/pANP-*proK*-R, and was cloned into vector pANP with *amyB* promoter resulting in the plasmid pIM 2513. The gene *proL* underwent a similar PCR amplification process using the primer pairs pANP-*proL*-F/pANP-*proL*-R, and was cloned into vector pANP with *amyB* promoter yielding plasmid pIM 2514. The gene variants *proL(D266A)* and *proL(E493A)* were amplified by

PCR using primer pairs pANP-*proL*-F/pANP-*proL*-D266A-R, pANP-*proL*-D266A-F/pANP-*proL*-R and pANP-*proL*-F/pANP-*proL*-E493A-R, pANP-*proL*-E493A-F/pANP-*proL*-R respectively, and were cloned into vector pANP with *amyB* promoter yielding plasmids pIM 2515 and pIM 2516.

To express *proG, E, I and L* in *E. coli* BL21, intron-free *proG* was amplified from the cDNA of *F. proliferatum* CGMCC 3.4710 using the primer pairs pGEX 4t-1-ProG-F/pGEX 4t-1-ProG-R. The fragments were digested with *EcoR* I/*Not* I and then was ligated into the correspondingly digested vector pGEX 4T-1 (protein expression vector with GST tag on *N*-terminus) yielding plasmid pIM 2517. Intron-free *proE* and *proI* were amplified from the cDNA of *F. proliferatum* CGMCC 3.4710 using the primer pairs pQ8-ProE-F/pQ8-ProE-R for ProE, pQ8-ProI-F/pQ8-ProI-R for ProI. The PCR products were subsequently subjected to enzymatic digestion with the restriction enzymes *EcoR* I/*Hind* III for ProE and *EcoR* I/*Not* I for ProI, and then were ligated into the correspondingly digested pQ8 (protein expression vector with MBP tag on *N*-terminus) yielding plasmids pIM 2518 and pIM 2519. Intron-free *proL* was amplified from the cDNA of *F. proliferatum* CGMCC 3.4710 using pColdI-ProL-F/pColdI-ProL-R and the PCR products were digested with *Nde* I/*Xho* I, and subsequently inserted into the correspondingly digested pColdI (protein expression vector with His tag on *N*-terminus) resulting in the plasmid pIM 2520.

1.4 Extraction of gDNA and the synthesis of cDNA

The strain *Fusarium proliferatum* CGMCC 3.4710 was cultured on solid glycerol medium (20 mL/L glycerol, 2 g/L tryptone, 2 g/L yeast extract, 20 g/L agar) at 25 °C for a duration of 4 days. Post incubation, the mycelium was harvested for the subsequent extraction of genomic DNA (gDNA) and the synthesis of complementary DNA (cDNA). CTAB (20 g/L CTAB, 81.8 g/L NaCl, 186.1 g/L $Na_2EDTA{\cdot}2H_2O$, 0.1 M Tris-HCl pH 8.0) method was used for the extraction of gDNA. RNA was prepared with TRLZOL® Reagent (Ambion), strictly adhering to the protocols provided by the manufacturer. To ensure the purity of the RNA and to eliminate any potential gDNA contamination, an RNAase-free DNase I treatment from New England Biolabs (NEB) was applied. Following this, the reverse transcription of RNA into cDNA was performed using the MonScript™ RTIII All-in-One Mix with dsDNase, facilitating a comprehensive removal of any residual gDNA and enabling the synthesis of high-quality cDNA.

1.5 Preparation and transformation of *F. proliferatum* or *A. nidulans* protoplasts

To obtain protoplasts, spores of *F. proliferatum* were germinated in 50 mL of PDB medium (26 g/L potato dextrose) at 25 °C, 220 rpm for 11 hours, while *A. nidulans* spores were inoculated in 40 mL of liquid CD medium and cultured at 37 °C, 220 rpm for 9 hours. Following spore germination, the precipitation was washed twice with 15 mL of Osmotic buffer (1.2 M $MgSO_4{\cdot}7H_2O$, 10 mM sodium phosphate, pH 5.8) and centrifuged at 4 °C, 3,750 rpm for 8 minutes to remove the supernatant. The precipitate was resuspended in 10 mL of Osmotic buffer containing 30 mg of *Lysing Enzymes* (Sigma) and 20 mg *Yatalase* (Takara), and then cultured at 28 °C, 80 rpm for 20 hours (*F. proliferatum*) or 12 hours (*A. nidulans*). The culture fluid was gently overlaid with 10 mL of Trapping buffer (0.6 M sorbitol, 0.1 M Tris-HCl, pH 7.0), then centrifuged at 4 °C, 4,000 rpm for 20 minutes. The protoplasm layer was transferred and fully dispersed into two volumes of STC buffer (1.2 M sorbitol, 10 mM $CaCl_2$, 10 mM

Tris-HCI, pH 7.5), and centrifuged at 4 °C, 3,750 rpm for 8 minutes. The supernatant was removed and STC buffer was added to resuspend the protoplasts for transformation.

1.6 Knockout of *pro* cluster in *F. proliferatum*

The knockout mutants were generated through double crossover homologous recombination, utilizing a hygromycin resistance split-marker for targeted gene knockout. The gene knockout cassettes containing the *hygB* resistance gene were amplified from the knockout plasmid, and then transformed into *F. proliferatum* protoplasts mediated via PEG solution (60% PEG4000, 50 mM $CaCl_2$, 10 mM Tris-HCl, pH 7.5). The transformed protoplasts were plated on solid SD-PDA medium (PDB medium with 1.2 mM sorbitol and 2g/L agar) containing 50 μg/mL hygromycin B at 25 °C for 3–4 days, and the resulting gDNA of transformants was verified by PCR. The knockout mutants were cultured on solid glycerol medium at 25 °C for 4 days. The fermentation mixture was extracted with a mixture of ethyl acetate and acetone (*v*:*v*=3:1) and dissolved in methanol for further LC–MS analysis.

1.7 Heterologous expression of *pro* cluster in *A. nidulans*

To gain stains of heterologous expression in *A. nidulans*, the expression plasmids were added to protoplasts of *A. nidulans* and the mixture was cultured on the regeneration dropout solid medium (CD medium with 1.2 mM sorbitol and appropriate supplements, CD-SD medium) at 37 °C. After 2–3 days, the transformants were respectively transferred to solid CD and cultivated at 37 °C for 3–4 days for sporulation. The spores were then inoculated on solid CD-ST medium and culture at 25 °C for 3 days, followed by analysis using LC–MS.

1.8 Chemical feeding of compounds 1 and 2 into *Fp-*Δ*proA-L* strain

The *Fp-*Δ*proA-L* strain was inoculated into 1 mL of glycerol medium at 25 °C for 4 days. Subsequently, the culture was supplemented with 100 μM of either compound **1** or compound **2**. Following a 12-hour incubation post-addition, the mycelia and medium were extracted with equal volume of ethyl acetate/acetone (*v*:*v*=3:1). The organic phase was then evaporated to dryness and was re-dissolved in methanol for further LC–MS analysis. As a control, the *Fp-wild type* and the *Fp-*Δ*proA-L* strain untreated with compounds **1** or **2** were processed under the same conditions.

1.9 Compound 1 in PBS buffer under different pH values

The compound **1** (100 μM) was introduced into 50 μL PBS buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4) under different pH values (3, 7, 8, 9, 10, 11, 12, 13). The reactions were conducted at a controlled temperature of 25 °C for a duration of 20 minutes. Then reaction mixtures were extracted with twice the volume of ethyl acetate, and then evaporated to dryness and re-dissolved in 50 μL of methanol for the further LC–MS analysis.

1.10 ECD calculations for compound 7

The absolute configuration of **7** was determined by quantum chemical calculations of their theoretical ECD spectra. Both *R*- and *S*- enantiomers were chosen for theoretical studies. Initially, conformational analyses were conducted employing the Molecular Merck force field (MMFF) calculations using

Spartan'14 program (Wavefunction Inc., Irvine, CA). The resulting conformers for **7** were then refined through geometry optimization and frequency calculations by Density functional theory (DFT) at the B3LYP/6-31G(d) level using the Gaussian16W C.01 program package (Gaussian Inc., Wallingford, CT). Subsequently, the optimized conformers were applied for theoretical ECD calculation. The TD-DFT calculations were performed at the M062X/TZVP level in solution (MeOH) state. The final ECD spectra were generated using the SpecDis 1.71 program (Berlin, Germany), where the simulated spectra were averaged according to the Boltzmann distribution theory and their respective Gibbs free energy (Table S15). A comparative analysis was conducted between the theoretical ECD curves of compound **7** and the experimental data.

1.11 Protein expression and purification in *E. coli*

The protein expression plasmids were individually transformed into *E. coli* BL21 strain by heat shock transformation. The stains were cultured in 1L of liquid LB medium (25 g/L LB broth) with 50 μg/mL kanamycin or 100 μg/mL ampicillin at 37 °C, 220 rpm until reaching an OD_{600} of 0.4–0.6. Subsequently, the cultures were induced with 0.2 mM isopropylthio-β-D-galactoside (IPTG) at 16 °C, 220 rpm for 20 hours. The cells were collected by centrifugation at 4 °C, 4,000 rpm for 15 minutes. To purify soluble ProG with *N*-GST-tag, the cells were resuspended in PBS buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3) and lysed by sonication on ice. Then, the supernatant was obtained by centrifugation at 4°C, 14,000 rpm for 40 min, and incubated with Glutathione Sepharose 4B resin (GE Healthcare). The resin was washed with PBS buffer three times to remove miscellaneous protein, the GST-tag protein was eluted by GST elution buffer (50 mM Tris-HCl, 10 mM reduced glutathione, pH 8.0). To purify soluble ProE or ProI with *N*-MBP-tag, the cells were resuspended in binding buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, pH 7.4). Subsequently, the cells were lysed through sonication on ice and centrifugated at 4 °C, 14,000 rpm for 40 minutes to obtain soluble fraction. The protein was purified by Dextrin Sepharose resin, and the protein-bound resin was washed with MBP elution buffer (10 mM maltose in binding buffer). The pooled fraction was concentrated and exchanged into buffer C (50 mM Tris-HCl, 50 mM NaCl, 5% glycerol, pH 7.5). The purified enzyme was analyzed by SDS-PAGE and the concentration was measured with BCA protein quantification kit (Beijing Dingguo Changsheng Biotechnology Co., Ltd).

1.12 Preparation and characterization of *A. nidulans* microsome and crude enzymes

The *A. nidulans* mutants were inoculated in 100 mL CD-ST medium and cultured at 25 °C, 220 rpm for 2.5 days. The culture broth was filtered out and the cells were collected. The cells were lysed by grinding, and cellular debris was resuspended in buffer C for the preparation of crude enzymes. The microsome fractions were then harvested by centrifugation at 4 °C, 14,000 rpm for 30 minutes and resuspended in buffer C. Substrates and cofactors were added to the microsome or crude enzymes to accomplish reaction at 25 °C. The reaction mixtures were extracted with twice the volume of ethyl acetate and the extracted ethyl acetate layer was evaporated to dryness The residues were re-dissolved in methanol for the subsequent LC–MS analysis.

1.13 *In vitro* characterization of ProG towards compound 1

For α, β-hydrolase ProG *in vitro* assay, the reaction was carried out in 50 μL PBS (pH7.3), containing 5 μM ProG (with GST tag), and 100 μM of the substrate (compound **1**). The *in vitro* assays were incubation at 25 °C for 5 hours, and the reactions were stopped by the addition of twice the volume of

ethyl acetate. The extracted ethyl acetate layer was then evaporated to dryness and re-dissolved in 50 μL of methanol. Subsequently, the products were analyzed by LC–MS. The control assays were performed without ProG.

1.14 *In vitro* characterization of ProE towards compound 8

For the *in vitro* assay of B-V monooxygenase ProE, the reaction was conducted in 50 μL buffer C (pH7.5), containing 5 μM of ProE (with MBP tag), 200 μM FAD, 2 mM NADPH and 100 μM of the compound **8**. The *in vitro* assays were incubation at 25 °C for 5 hours then stopped by the addition of twice the volume of ethyl acetate. The extracted ethyl acetate layer was then evaporated to dryness and re-dissolved in 50 μL of methanol. Subsequently, the products were analyzed by LC–MS. The control assays were performed without ProE.

1.15 Time-course assays for ProI towards compound 7

The *in vitro* assays of FMO ProI were conducted in 50 μL buffer C (pH7.5), containing 5 μM of ProI, 200 μM FAD, 1 mM NADPH and 100 μM of the compound **7** at 25 °C. After 1 minute, 20 minutes, 40 minutes, 80 minutes and 5 hours, the reactions were stopped by the addition of twice the volume of ethyl acetate. The extracted ethyl acetate layer was then evaporated to dryness and re-dissolved in 50 μL of methanol for the further LC–MS analysis. The control assays were performed without ProI.

1.16 *In vitro* characterization of microsome fractions of ProJ and ProI

For mGT ProJ and FMO ProI *in vitro* assay: microsome fractions of ProJ, 10 μM ProI, 100 μM compound **7**, 1 mM uridine diphosphate glucose (UDPG), 200 μM FAD and 2 mM NADPH were conducted in 100 μL buffer C at 25 °C for 5 hours. After the incubation, twice the volume of ethyl acetate was added, the ethyl acetate layer was evaporated to dryness and redissolved in 50 μL of methanol for further LC–MS analysis. The control assays were performed without ProJ and ProI.

1.17 Time-course assays for crude enzymes of *AN-proL* towards compound 1

The *in vitro* assays of the *AN-proL* crude enzymes were conducted in 50 μL buffer C (pH7.5), containing crude enzymes of ProL and 100 μM of the compound **1** at 25 °C. After 30 s, 2 minutes, 5 minutes and 10 minutes, the enzymatic reactions were quenched by the addition of equal volume of methanol. The supernatant was obtained by centrifugation at 13,300 rpm for 5 min for the further LC–MS analysis. As a control, the *AN-wild type* crude enzymes treated with compound **1** was processed under the same conditions.

1.18 *In vitro* characterization of crude enzymes of *AN-proK*, *AN-proL* and its mutants

The *in vitro* assays of *AN-proK, AN-proL*, *AN-proL*-D266A or *AN-proL*-E493A crude enzymes were systematically executed in 50 μL buffer C (pH7.5) at 25 °C. Each reaction mixture was composed of the respective crude enzymes and 100 μM of the compound **1**. After a duration of 10 minutes, the enzymatic reactions were halted by introducing equal volume of methanol. The supernatant was obtained by centrifugation at 13,300 rpm for 5 min for the further LC–MS analysis. As a control, the *AN-wild type* or *AN-proL* crude enzymes treated with compound **1** was processed under the same conditions.

1.19 *In vivo* $H_2{}^{18}O$ labeling experiments and LC–MS detection

For isotope incorporation experiments with $H_2{}^{18}O$, the *F. proliferatum* was cultivated on solid glycerol medium containing 90% $H_2{}^{18}O$ at 25 °C for 4 days. The fermentation mixture was then extracted with a mixture of ethyl acetate and acetone (*v*:*v*=3:1) and reconstituted in methanol for further LC–MS analysis. The LC–MS analysis were performed on Agilent Technologies 1260 Infinity II–6125 system in conjunction with a C18 column (Poroshell 120 EC-C18, 1.9 μm, 2.1 mm × 50 mm, Agilent). LC–MS metabolite profiles were generated utilizing the following method: Chromatographic separation was achieved with a linear gradient of 10–100% MeCN-H_2O (the water components of the mobile phase contained 0.1% *v/v* formic acid) over 3 minutes, followed by 100% MeCN for 30 s, then returning to 10% MeCN-H_2O for 30 s, all at a flow rate of 0.3 mL/min. The MS data were collected in the *m/z* range 100–1000 in positive ion modes simultaneously.

1.20 Acid hydrolysis of compounds 1 and 9 and HPLC detection

To analyze the saccharide polyol configuration of compounds **1** and **9**, 5 mg of each compound was added to a solution of 4 M trifluoroacetic acid (TFA) in MeOH. The mixture for compound **1** was agitated at a temperature of 110 °C for a duration of 1 hour, whereas the mixture for compound **9** was subjected to the same conditions for a shorter period of 20 minutes to control the further hydrolysis of the disaccharide units. The resulting mixture was then dried by vacuum to obtain the crude hydrolysate of compounds **1** and **9**. This hydrolysate was dissolved in 20 mL of distilled water and extracted three times with an equivalent amount of ethyl acetate. The aqueous fraction was freeze-dried and dissolved in 200 μL of distilled water to yield saccharide polyol sample. The saccharide polyol sample of compounds **1** and **9**, along with standards of D-glucose, D-galactose, D-mannose, gentiobiose, cellobiose, and laminaribiose were analyzed by HPLC equipped with an amino column (Asahipak NH2P-50 4E, 4.6 mm I.D. x 250 mm L, Shodex) and an evaporative light scattering detector for disaccharides or a refractive index detector for monosaccharides. The analysis was performed at a column temperature of 40 °C and an isocratic program of 75% MeCN/25% H_2O with a flow rate of 0.5 mL/min.

1.21 The procedure for 3 and 5 configuration determination

To determine the absolute configuration of compounds **3–6**, **3** and **5** were used as the examples. A sufficient amount of **3** and **5** were isolated through multiple large-scale fermentations of *F. proliferatum* wild type. 5 mg compound **3** or **5** was derived with 5 equimolar *R*-MTPA derivatization reagent and 1 equimolar 4-dimethylaminopyridine (DMAP) in anhydrous dichloromethane at 25 °C for 2 hours, which yield the reaction mixture by vacuum to dryness. The mixtures were then purified by semi-preparative HPLC using isocratic program of 66%/34%-MeCN/H_2O to gain *S*-MTPA ester **3*S*** (1.5 mg, t_R=13 min, *R*-MTPA derivative) and **5*S*** (1.8 mg, t_R=18.5 min, *R*-MTPA derivative). Similarly, 5 mg **3** or **5** was derived with *S*-MTPA derivatization reagent to form *R*-MTPA ester **3*R*** (1.2 mg, t_R=13 min, *S*-MTPA derivative) and **5*R*** (1.6 mg, t_R=18.5 min, *S*-MTPA derivative). The absolute configuration of C11 of **3** and C10 of **5** were confirmed as *R* according to Mosher rule.[1]

1.22 Purification and structural characterization of compounds

The compounds were separated by medium pressure liquid chromatography (MPLC) Reveleris® X2 (BUCHI, Inc) with a gradient of MeOH and H_2O at a flow rate of 25 mL/min on 80 g reversed phase silica gel (C18) following. The semi-preparation high performance liquid chromatograph (HPLC) used a Shimadzu LC-20AR Prominence UFLC system and a YMC ODS-A 5 μm 120A (10×250 mm) column was used at a flow rate of 2.5 mL/min, and 0.02% formic acid was added to the aqueous phase. MCI Column chromatography was performed on MCI gel CHP 20P/P120 (37-75 μm, Mitsubishi Chemical Corporation, Japan).

Isolation of compounds **1**–**6**

For isolation of compounds **1**–**6***, F. proliferatum* CGMCC 3.4710 was cultured on 14 L solid glycerol medium at 25 °C for 5 days, the cultures were then extracted with a mixture of ethyl acetate and acetone (*v:v=3:1*) three times. The resulting organic solvent was evaporated to dryness under vacuum to obtain the residues, which were subjected to MCI gel column chromatography using various MeOH-H_2O eluents (*v:v=4:6*, *5:5*, *6:4*, *7:3*, *8:2*, *9:1, 10:0*) to yield seven fractions. Fraction 6 (MeOH: H_2O, 9:1) was purified by semi-preparative HPLC eluded with 65%/35%-MeCN/H_2O (0.02% formic acid) to gain compound **1** (11.0 mg, t_R=14.5 min). Fraction 7 (MeOH: H_2O, 10:0) was purified by semi-preparative HPLC eluded with 85%/15%-MeOH/H_2O (0.02% formic acid) to gain compound **2** (7.2 mg, t_R=19.5 min). Fraction 5 (MeOH: H_2O, 8:2) was purified by semi-preparative HPLC eluded with 40%/60%- MeCN /H_2O (0.02% formic acid) to gain compounds **3**–**6** (8.1 mg, 7.0 mg, 7.3 mg, 5.6mg, t_R=18.0 min t_R=22.0 min, t_R=21.0min, t_R=25.0min).

Isolation of compound **7**

Transform the pIM2509 into *A. nidulans* protoplasts to construct *AN-proFG* strain for the isolation of compound **7**. The *AN-proFG* strain was cultivated in 5 L CD-ST agar medium at 25 °C for 3.5 days. Utilizing the above established method, the crude residues were obtained. The residues then underwent a purification regimen via MPLC with a linear gradient of 60% to 100% MeOH in H_2O over a period of 60 minutes at a flow rate of 25 mL per minute. The fraction containing compound **7** was purified by semi-preparative HPLC using isocratic program of 80% MeOH-20% H_2O (0.02% formic acid) to afford compound **7** (7.0 mg, t_R=24.0 min).

Isolation of compound **8**

Transform the pIM2508 and pIM2510 into *A. Nidulans* protoplasts to construct *AN-proFGIJ* strain for the isolation of compound **8**. The *AN-proFGIJ* strain was cultivated in 8 L CD-ST agar medium at 25 °C for 3.5 days. The residues were purified by MPLC with a linear gradient of 40% to 100% MeOH in H_2O over a period of 60 minutes at a flow rate of 25 mL per minute. The fraction containing compound **8** was purified by semi-preparative HPLC using isocratic program of 60% MeCN-40% H_2O (0.02% formic acid) to afford compound **8** (14.0 mg, t_R=13.0 min).

Isolation of compound **9**

For isolation of **9**, *Fp*-Δ*proK* strain was cultivated in 14 L solid glycerol medium at 25 °C for 5 days, The residues were subjected to MCI gel column chromatography using various MeOH-H_2O eluents (*v/v=4:6*, *5:5*, *6:4*, *7:3*, *8:2*, *9:1, 10:0*) to yield seven fractions. The fraction 6 (MeOH:H_2O, 9:1) containing compound **9** was purified by semi-preparative HPLC using isocratic program of 62% MeCN-38% H_2O (0.02% formic acid) to afford compound **9** (7.6 mg, t_R =14.0 min).

1.23 Root growth inhibition assay of *Arabidopsis thaliana*

For the growth inhibition test of compounds **1**–**9** on *Arabidopsis thaliana*, MS (2.2 g/L Murashige and Skoog basal medium, 10 g/L sucrose, 8 g/L agar) medium was utilized. Compounds **1**–**9** are dissolved in dimethyl sulfoxide and added to the medium at a final concentration of 50 μg/mL prior to inoculation of growing plants, the control treatment medium contained the same amount of dimethyl sulfoxide. Before planting, *A. thaliana* seeds were surface sterilized with 75% ethanol for 2 minutes, followed by 5% sodium hypochlorite for 5 minutes and subsequently rinsed five times with distilled water. The seeds were then kept in the dark at 4 °C for 2 days for stratification. Then plants were grown under long day conditions (16/8 h light/dark) at 25 °C with cool-white fluorescence bulbs serving as the light resource (15,000 lx). The seven-day-old *A. thaliana* plants were used to for experimental analysis. The total root length of *A. thaliana* was measured by ImageJ software, and the statistical data are plotted with GraphPad Prism 8.0.

1.24 Bioinformatics analysis

To elucidate the *pro* gene cluster from *F. proliferatum* CGMCC 3.4710, we used protein sequence hrPKS (FFUJ_11199) from *F. fujikuroi* IMI 58289 to retrieve FFUJ_11199 orthologues by using local-blast in public fungal genome database from NCBI and the private database of our lab. The biosynthetic gene cluster annotation was conducted using the 2ndFind program to predict open reading frames and introns. The gene function was assigned based on a BlastP search. The domains of ProF were analyzed by *interpro* website. TMHMM-2.0 transmembrane predictor was used to predict the protein transmembrane helices in ProJ. Active sites analysis of the ProL was performed using DNAMAN8.0 software. The enzyme function initiative-enzyme similarity tool (EFI-EST) was used to generate protein sequence similarity networks (SSNs) of ProK, the SSN was generated using the UniRef90 database with alignment score of 16. Cytoscape software platform for visualizing complex networks of ProK.

2nd Find: http://biosyn.nih.go.jp/2ndFind/

NCBI BLAST: https://blast.ncbi.nlm.nih.gov/Blast.cgi

interpro website: http://www.ebi.ac.uk/interpro/search/sequence/

EFI-ENZYME SIMILARITY TOOL: https://efi.igb.illinois.edu/efi-est/

TMHMM-2.0: https://services.healthtech.dtu.dk/service.php?TMHMM-2.0

2. Supplementary Tables

Table S1. Stains and plasmids used in this study.

Stains or plasmids	Characteristics	References or sources
Fusarium proliferatum CGMCC 3.4710	Host for knockout of the *pro* gene cluster and the isolation of compound **1**–**6**	Obtained from the China General Microbiological Culture Collection Centre (CGMCC)
Aspergillus nidulans LO8030	Host for heterologous expression	[2]
Saccharomyces cerevisiae BJ5464-NpgA	Host for heterologous recombination to construct the *A. nidulans* overexpression plasmids	[3]
Escherichia coli BL21	Host for protein expression	Novagen
Escherichia coli XL-1	General cloning host	Stratagene
pColdI	Protein expression vector used in *E. coli*, encoding *N*-terminal 6×His-tag, ampicillin resistance	Takara
pQ8	Protein expression vector used in *E. coli*, encoding *N*-terminal His×6-MBP-His×6-tag, kanamycin resistance	[4]
pGEX-4T-1	Protein expression vector used in *E. coli*, encoding *N*-terminal GST-tag, ampicillin resistance	GE Healthcare
pYEU	*E. coli-Saccharomyces* shuttle vector, ampicillin resistance, *Ura3*	[5]
pANU	*E. coli-Saccharomyces-A. nidulans* shuttle vector for heterologous expression, ampicillin resistance, *Ura3* and *pyrG89*	[5]
pANR	*E. coli-Saccharomyces-A. nidulans* shuttle vector for heterologous expression, ampicillin resistance, *Ura3* and *riboB2*	[5]
pANP	*E. coli-Saccharomyces-A. nidulans* shuttle vector for heterologous expression, ampicillin resistance, *Ura3* and *pyroA4*	[5]

Table S2. Primers used in this study.

Primer name	Primer sequence (5'→3')
pYEU-*hyg*-F	TACAACGACCATCAAAGTCGTATAG
pYEU-*hyg*-R	AGGATTACCTCTAAACAAGTGTACC
hyg-F	CCATCCAAGAACCTTTATTTCC
hyg-R	GTTTCCACTATCGGCGAGTAC
KO-left-YZ-R	TTTAGGCTCAAGTCATGACC
KO-right-YZ-F	TGGGTTCGCAAAGATAATTGC
pYEU-KO*proA*-UP-F	ATTATAAGGATGATGATGATAAGACTAGTATTAATGTTAATAA GTAGTACTATTGTCAC
pYEU-KO*proA*-UP-R	CACTGGTAGCTATACGACTTTGATGGTCGTTGTACCTCTTTTGT AGAGACTCACAG
pYEU-KO*proA*-DN-F	GAATGCACAGGTACACTTGTTTAGAGGTAATCCTTGAGGTGAT AAATGAGGTGATTTG
pYEU-KO*proA*-DN-R	CATTTAAATTAGTGATGGTGATGGTGATGCACGTGAGAAGACG AAGATAAAGGCATTG
KO*proA*-left-YZ-F	GAGTGAATGCAGGTGATGGAAG
KO*proA*-right-YZ-R	CCGATAGCCCAGAACATTTGC
pYEU-KO*proL*-UP-F	GGCTAGCGATTATAAGGATGATGATGATAAGACTAGATGTCTG TGGCCAGAACGAG
pYEU-KO*proL*-UP-R	CTGGTAGCTATACGACTTTGATGGTCGTTGTACTTAGGCACGG AGTATTCAGTCTAAG
pYEU-KO*proL*-DN-F	CCCAGAATGCACAGGTACACTTGTTTAGAGGTAATCCTCCAAG ACACCGGTGTCC
pYEU-KO*proL*-DN-R	ATTTAAATTAGTGATGGTGATGGTGATGCACGTGCTCCTCTTTT ATAATACAAAAGGGC
KO*proL*-left-YZ-F	GGTTGGCCAGGAATGTACG
KO*proL*-right-YZ-R	GGCCTACTTTTAGCTATTTAATT
pYEU-KO*proC*-UP-F	CTAGCGATTATAAGGATGATGATGATAAGACTAGCGTTTGAAC TCACAAATGAGGATTC
pYEU-KO*proC*-UP-R	CACTGGTAGCTATACGACTTTGATGGTCGTTGTAACGTAATATC GAGAGATCTGCAAG
pYEU-KO*proC*-DN-F	GCACAGGTACACTTGTTTAGAGGTAATCCTGTTGACCAGTGAT GTATGTCTATCC
pYEU-KO*proC*-DN-R	GTCATTTAAATTAGTGATGGTGATGGTGATGCACGTGGCAGCC TCTGGGTTATGAGTG
KO*proC*-left-YZ-F	GTTATGGTAGCAGCCCAGGTC
KO*proC*-right-YZ-R	GCGAGAAAGGTATGGTTCTAGG
pYEU- KO*proD*-UP-F	GGCTAGCGATTATAAGGATGATGATGATAAGACTAGTGTCTTG TTTGTCTGCCCGAC
pYEU-KO*proD*-UP-R	CTGGTAGCTATACGACTTTGATGGTCGTTGTAGTTAAAAAAGG TAACGAGCGATGAAG
pYEU-KO*proD*-DN-F	CCAGAATGCACAGGTACACTTGTTTAGAGGTAATCCTTTGACG GATTCGCGCATG
pYEU-KO*proD*-DN-R	GTCATTTAAATTAGTGATGGTGATGGTGATGCACGTGAACTCG CCGAGAAGGCGG
KO*proD*-left-YZ-F	CAGTTCTCCTTGGTTCTTAATGTC
KO*proD*-right-YZ-R	GCTACTACGACTATCACAACGG
pYEU- KO*proE*-UP-F	CATATGGCTAGCGATTATAAGGATGATGATGATAAGACTAGCG CGTCATCAAGGCTTTC
pYEU-KO*proE*-UP-R	CACTGGTAGCTATACGACTTTGATGGTCGTTGTAAATATCGAT AGGGGCATCGTTG
pYEU-KO*proE*-DN-F	GCACAGGTACACTTGTTTAGAGGTAATCCTTTTGGTTCTATTTC TGGAGGTCTATC
pYEU-KO*proE*-DN-R	GTCATTTAAATTAGTGATGGTGATGGTGATGCACGTGTGCGCG AGCTCATTGATG
KO*proE*-left-YZ-F	CCTTGATGGAACACCTCG
KO*proE*-right-YZ-R	CATCATCGGGTCCCTTCAC

Primer name	Primer sequence (5'→3')
pYEU- KO*proI*-UP-F	CTAGCGATTATAAGGATGATGATGATAAGACTAGTCAGTTGAA CAGACAGCTAAGGC
pYEU-KO*proI*-UP-R	CTCACTGGTAGCTATACGACTTTGATGGTCGTTGTAGATGAGG TTTGTGCTGGTGAAC
pYEU-KO*proI*-DN-F	CCAGAATGCACAGGTACACTTGTTTAGAGGTAATCCTCTTCGA GCCCAAGCTACTC
pYEU-KO*proI*-DN-R	CATTTAAATTAGTGATGGTGATGGTGATGCACGTGAACCTTGA CTTATTTCACTCAGA
KO*proI*-left-YZ-F	GCTCGGCATTGCGTATCATAAC
KO*proI*-right-YZ-R	CGAATTGCTGGCGCAGAG
pYEU- KO*proK*-UP-F	GGCTAGCGATTATAAGGATGATGATGATAAGACTAGTGAAGA CACCCATGGACAGC
pYEU-KO*proK*-UP-R	CCTCACTGGTAGCTATACGACTTTGATGGTCGTTGTAATTTCTA GCCTGGCGTGC
pYEU-KO*proK*-DN-F	CAGAATGCACAGGTACACTTGTTTAGAGGTAATCCTAGAAGCG GTTTTCTAGGCAC
pYEU-KO*proK*-DN-R	CATTTAAATTAGTGATGGTGATGGTGATGCACGTGTTGAACTG AGTTGTCATCTGCC
KO*proK*-left-YZ-F	GCGCTTATAGCTTCCATCTCC
KO*proK*-right-YZ-R	CTTTGCCAATTCGAAGAGTCTTC
amyB-F	GATTAAAGGTGCCGAACGAGC
amyB-R	AAATGCCTTCTGTGGGGTTTATTG
glaA-F	CCTGATCTTCCGAACTGGTCG
glaA-R	TGCTGAGGTGTAATGATGCTG
gpdA-F	ACTCCGGTGAATTGATTTGGG
gpdA-R	TGTTTAGATGTGTCTATGTGGC
pANU-*proF*-F1	CTGAGCTTCATCCCCAGCATCATTACACCTCAGCAATGTCGGC ATACACCAAAGATATC
pANU-*proF*-F2	GGTCTGCAGACTCTCTGTGCTCCCCGTGCCACTGCC
pANU-*proF*-F3	GGCGGACTTTGCGTCCATCAATCGTAAGAACCCTTC
pANU-*proF*-R1	GGCAGTGGCACGGGGAGCACAGAGAGTCTGCAGACC
pANU-*proF*-R2	GAAGGGTTCTTACGATTGATGGACGCAAAGTCCGCC
pANU-*proF*-R3	GGAGGACATACCCGTAATTTTCTGGGCATTTAAATTTTCCAGTT GGCCGGCCGTAGTCA
pANU-*proFG*-R3	TCACCCAAATCAATTCACCGGAGTTTTCCAGTTGGCCGGCCGT AGTCAAACCACCACAG
pANU-*proG*-F	CTAACCATTACCCCGCCACATAGACACATCTAAACAATGAGGT TTCTTTGCTTGCATG
pANU-*proG*-R	GGAGGACATACCCGTAATTTTCTGGGCATTTAAATTCGAACTT CACTACTCCGTTTCAA
pANR-*proG*-R	TAGGTCGCCAGGTACGACCAGTTCGGAAGATCAGGTCGAACTT CACTACTCCGTTTCAA
pANR-*proI*-F	CTGAGCTTCATCCCCAGCATCATTACACCTCAGCAATGACAGA CAAGAAACCCATTCGT
pANR-*proIJ*-R	CCTTCTCTGAACAATAAACCCCACAGAAGGCATTTATGCCTAG TTCCGAGTTGGCGCAC
pANR-*amyB-proJ*-F	AGGGTATCATCGAAAGGGAGTCATCCAATTTAAATGATTAAAG GTGCCGAACGAGCTAT
pANR-*proE*-F	CCATTACCCCGCCACATAGACACATCTAAACAATGGACACATT CGACGTCATAATTG
pANR-*proE*-R	CATAGGTCGCCAGGTACGACCAGTTCGGAAGATCAGGAGGAT AGGGAAGGGATTGAG
pANR-*proJ*-F	CTAACCATTACCCCGCCACATAGACACATCTAAACAATGCCTA GTTCCGAGTTGG
pANR-*proJ*-R	CTAAAGGGTATCATCGAAAGGGAGTCATCCAATTTAAATCTTC GAGCCCAAGCTACTC
pANP-*proK*-F	CCTTCTCTGAACAATAAACCCCACAGAAGGCATTTATGGCTAG GGAGTATTCCGG
pANP-*proK*-R	GACCCAACAACCATGATACCAGGGGATTTAAATCTGTACCACG TCAGTAACCCTC

Primer name	Primer sequence (5'→3')
pANP-*proL*-F	CTTCTCTGAACAATAAACCCCACAGAAGGCATTTATGTATCGACTCAGTTTGGCATC
pANP-*proL*-R	GACCCAACAACCATGATACCAGGGGATTTAAATGTGTATGGGCAACAATATCCCCAAC
pANP-*proL*-D266A-F	GCTTGGTTTTGAGGGCTATGTCATGTCGGCTTTTCTGGCTACGCCGCCTGGACTTGG
pANP-*proL*-D266A-R	GCCAGAAAAGCCGACATGAC
pANP-*proL*-E493A-F	GCCACCGCTGGTTTTGATC
pANP-*proL*-E493A-R	GTCTTGCGATCAAAACCAGCGGTGGCGAAAGATTTGAGGAAGACTATACAGACCG
proA-YZ-F	CTGTCAAAATCAGGTGCCCAAC
proA-YZ-R	GCAGTTCTCTATGTCGATCAGAG
proC-YZ-F	ATGGCAGTTACTTCAATTTCAC
proC-YZ-R	TCAGGCCTCAATATCACCAATTTTG
proD-YZ-F	ATGTTTACCGCTTTGTTTCTTTCTG
proD-YZ-R	TCAGACATGGCGGTGG
proF-YZ-F	TGCCCTAGCCGCTCTGGCAGAAGCTCTC
proF-YZ-R	CAGATCAACGAGCACACGCGGCTGTGCTG
proJ-YZ-F	ATGCCTAGTTCCGAGTTGGCGC
proJ-YZ-R	CTACTTGATCAGGTTCCTCAACC
proK-YZ-F	GGCTAGGGAGTATTCCG
proK-YZ-R	CTACCAAAAGTACTCGTCAAATCCCTC
pGEX 4t-1-ProG-F / *proG*-YZ-F	CCGGAATTCATGAGGTTTCTTTGCTTGCATGG
pGEX 4t-1-ProG-R / *proG*-YZ-R	AAGGAAAAAAAGCGGCCGCCTATACAGCGGTCATGGCCTTATC
pQ8-ProE-F / *proE*-YZ-F	CCGGAATTCATGGACACATTCGACGTCAT
pQ8-ProE-R / *proE*-YZ-R	CCCAAGCTTTCACAGCTGCTTTCCCTC
pQ8-ProI-F / *proI*-YZ-F	CCGGAATTCATGACAGACAAGAAACCCATTC
pQ8-ProI-R / *proI*-YZ-R	AAGGAAAAAAGCGGCCGCTCATACGTCCTTCTTCATCTCAC
pColdI- ProL-F / *proL*-YZ-F	GGGAATTCCATATGATGTATCGACTCAGTTTGGCATC
pColdI- ProL-R / *proL*-YZ-R	CCGCTCGCGTTATTTTACCAGTTGAACGGTTGTC

Table S3. Plasmids used in this study.

Name	Description	Enzyme site	Aim
pIM2501	*proA* upstream 2,000 bp, *proA* downstream 2,000 bp, and *hygB* resistance gene in pYEU	*Spe* I-*Pml* I	*F. proliferatum* knockout
pIM2502	*proL* upstream 2,000 bp, *proL* downstream 2,000 bp, and *hygB* resistance gene in pYEU	*Spe* I-*Pml* I	*F. proliferatum* knockout
pIM2503	*proC* upstream 2,000 bp, *proC* downstream 2,000 bp, and *hygB* resistance gene in pYEU	*Spe* I-*Pml* I	*F. proliferatum* knockout
pIM2504	*proD* upstream 1,998 bp, *proD* downstream 2,000 bp, and *hygB* resistance gene in pYEU	*Spe* I-*Pml* I	*F. proliferatum* knockout
pIM2505	*proE* upstream 2,000 bp, *proE* downstream 2,000 bp, and *hygB* resistance gene in pYEU	*Spe* I-*Pml* I	*F. proliferatum* knockout
pIM2506	*proI* upstream 2,000 bp, *proI* downstream 2,000 bp, and *hygB* resistance gene in pYEU	*Spe* I-*Pml* I	*F. proliferatum* knockout
pIM2507	*proK* upstream 2,000 bp, *proK* downstream 2,000 bp, and *hygB* resistance gene in pYEU	*Spe* I-*Pml* I	*F. proliferatum* knockout
pIM 2508	*proF* gDNA with downstream 500 bp, in pANU	*Not* I	*A. nidulans* overexpression
pIM 2509	*proF* gDNA with downstream 500 bp, *proG* gDNA with downstream 144 bp, in pANU	*Not* I	*A. nidulans* overexpression
pIM 2510	*proG* gDNA with downstream 144 bp, *proI* gDNA with downstream 193 bp, *proJ* gDNA in pANR	*BamH* I	*A. nidulans* overexpression
pIM 2511	*proE* gDNA with downstream 92 bp, *proI* gDNA with downstream 193 bp, *proJ* gDNA in pANR	*BamH* I	*A. nidulans* overexpression
pIM 2512	*proJ* gDNA with downstream 193 bp, in pANR	*BamH* I	*A. nidulans* overexpression
pIM 2513	*proK* gDNA with downstream 500 bp, in pANP	*BamH* I	*A. nidulans* overexpression
pIM 2514	*proL* gDNA with downstream 488 bp, in pANP	*BamH* I	*A. nidulans* overexpression
pIM 2515	*proL-mutation(D266A)* gDNA with downstream 488 bp, in pANP	*BamH* I	*A. nidulans* overexpression
pIM 2516	*proL-mutation(E493A)* gDNA with downstream 488 bp, in pANP	*BamH* I	*A. nidulans* overexpression
pIM 2517	*proG* cDNA in pGEX 4t-1 with *N*-GST tag	*EcoR* I-*Not* I	*E. coli* overexpression
pIM 2518	*proE* cDNA in pQ8 with *N*-MBP tag	*EcoR* I-*Hind* III	*E. coli* overexpression
pIM 2519	*proI* cDNA in pQ8 with *N*-MBP tag	*EcoR* I-*Not* I	*E. coli* overexpression
pIM 2520	*proL* cDNA in pColdI with *N*-His tag	*Nde* I-*Xho* I	*E. coli* overexpression

Table S4. The protein sequence of *pro* gene cluster.

Protein name	Protein sequence
ProA (551 aa)	MKSIPSILVAAAIPSLASGQSFQSTPIMGWNSYNQVSCSPTNKTITTAIEALSSR GFVNAGYKFFQIDCGWASRDAQRDPTTGALKIDSDAFPDGLKPLSDLARSK GMKWTMYSDAGVRMCDPQYPSPVLGSLGHEAVDAAFFKSLNTEYLKYDN CYADSASNNAPKDPRTDFVTRFGTMWSELQKVGIPGMLICQWGVPYSSSSG LEGPAEWTQNVSTSFRLSDDIAEGWGNVYRISNQAIHIAHRGLSGPGHIADA DLLEVGNSGMTFDEQATHFALWAMLKSALMISTDITALSDQTVAVLQNKDL ISINQDAAVKPVSLVQRWTSDRDLWAGPLANGDVAVLYVDQSNSARTLSLQ LSNLGYQSADIKDLWTGKTTTGASSFSKQVNGHGSVALRLSNIKKLSSTANY KYTSVSTGSLSSGAKTASCSGCTSSTKVGNVGGSANGQVVLSNISTSKATQN VLFDYINGDVGFSFTGSTNDRLASIKVNGGTAQIVSFPLTGYNWDKDVYKG YAVELSGFNTSGPNTITISGVNSGWAPDFDRLAVVA
ProC (539 aa)	MAVTSISLLVLGIVAGATIFYLSSPTRKDRKRPLKLPIIGDIHNSPIEKPLLRWD AWVKENGAIATSKLFGIMPVVVINTAEAATELLGKRGAWYSNRPRSVGME MITGAGPGQSRFTLMHDMDAHLKLHHRILSPSLGGVAAPRYQPVMELEAKQ LVKDLVELSEHKSVVVTSDDVFPFLERAQASIILALHYSVRVPTLDYPLYQR VRETQAKVTSYASRPGLPDIFPFLAKLPAAISPWRRAADKLFNEQKDLNLHL LSLGDDSPGWNATKQARSLAAKYAKEPIPDIDLAFTLATSVQGGIETSTRTIL WLFIAAMTANRSFMKRAHDVLDAVVGRDRLPCFADRSSLCYIDAIVSELLR WRPISPGGVPRRADKQDSFKGINIVKNAMVLTNAWSIGRDEAVFDQSLGDL DQFIPERWLDGESIGVDIKNQGELRTSLPLPVFGHGRRSCLGKRVAVDGTFA QVATMIWAFDFEPTQDVDEMEMEVVWFMTEPKQFKFKLKPRGSWVSKVIE EEWRTADKDLGKIMGKIGDIEA
ProD (312 aa)	MFTALFLSALAWSQMANAHGTITRVIGANGVVMPGLTILDGTPRSSTSAASG AQVDTSVIRDPELGTSKASALGRTSKGPVDGARVIKAFMHGLKGRSLADTIL GGGEEATREAVSFVTGNAGAVVNGVQDGIESSPVGGLALGAEHGVNGLLD DFFQTAKGVPSPRGYIEDGVQNSTGVGAKSGLPTTASDGTLKLIYHQVNEDG AGPLLVDIDFTSGGTDPKAFKSAEVVQNIIGVLGFSTVSSTDFPVVVKVPTGQ VCTGKVAGVSGICIARVRNSATAGPFGGAAAFTHNPEAAKGKTSSAKFRHR HV
ProE (489 aa)	MDTFDVIIVGAGLAGINAAYRLQIALPNLSYAILEARDDIGGTWDLFRYPGIR SDSDLYTFGFAWQPWDQGTALGEGGAILNYMKKCAQTYGIDRHVHCQRRL KHASWSTPQQKWTLEVEDIGAQKELQYHARFVMLCTGYYDYHNGRDTAIP GLENFKGQVVHPQFWPEDLDYTNKQVAVIGSGATAITLIPKLAEKAARTTM VQRSPSYILSIPNGGKKAPWLVRLFPARWAHAYTRLSFLIWSRLIWLFCQTFP DRARQRLRNGVEKELPHHLPYDPHFSPRYNPWDQRVCLTPNADFFKCLHTG KADVKTGNIKEVVADGIVLENAPQDRIPAAIIITATGLKLQLAGGATIDVDGV PIKPSEQYFWNGTMLQDVPNLSLVIGYTTISWTLGVDTAAMLVCRLLQKMQ KAKLSSATPRAEQGLALTPRRLFSLSSTYVTTAESELPRAAAQAPWQPRTNY LSDYWFVKLGRLDRGLQFVREGKQL

Protein name	Protein sequence
ProF (2649 aa)	MSAYTKDIMNNGDAIPEPLVQAVLQQDKSMPIAIVGMACKLPGEATNPDKL WELCAERKAAWSEVPKERMNIDSFWHPDGERGGNTNAKGGHFLKQDVAA FDASFFSIPPTEAKSIDPQQRILMETVYEAMESGGMSMEDMAGSGTSCYVGS FSRDYYEIIDRDPETAPLYSVTGNGAAILSNRISYFYDLKGPSLTLDTACSSSL VALHLACQSLRSGESHCSIVGATNLILNPDIITAMSNMHFFSPDSRCYSFDAR ANGYSRGEGIACLILKPLADAIRDNDTIRAVIRGTACNQDGKTSGIMLPSREA QEELIKTAYRDSGCDFATTAYFEAHGTGTQAGDPLESGAIGSTFGPHRPVDE NGCKLPLYIGSVKTNVGHTEGTSGLAGVIKAVLSLERGAIAPNVYFETPNPQI DLEGWNLRVPTKVIAWPTKGQRRISVNSFGFGGTNGHVILDDAFHYLNGRG LKGAHRTAPRPLLDLQTGIAMSNAIAPVLSNGNGVNGHTNGHTNGPTNGYT NGHSNGHSLSSAKRHRVFIWSTHEENIAGKSAAVYAEHLAHRKEADEEAFL DNLSYTLCSRRSMLPWKSFLVAQSLPDLAEKVAATRQKPVRASISPLKIGFVF TGQGAQWFAMGRELIQTYPLFQQRLEYSDRFVKSLGAKWSLIDELNKDEET SRINESLVSQTACTALQLALVDLLASWGIRPQKVIGHSSGEIAAAYASGILPTE SALKAAYFRGLHSSRVQEKLPSADGAMMAVGLAEEDANNYISALAPELGK ASVACINSPTNVTVSGDRPALAALAEALQKDSVFARLLKVSTAYHSHHMDA VASDYEHSLAGMEVNTPSETIEMVSAVTGELLKSSDILGPKYWGENMVSPV RFNTGLQTLCAPRATAKRTRRLAANQASVDVLVEVGPHAALAGPIKQILGV PVLEKSGIVYKSVLSRGQDACATALEAAAFLFAKGCTTIDLYKVNSPAGSHT AQPRVLVDLPPYFWNHTKSHWMESRLSIEHRFRRHARTDFLGYPVNDWNP MDPRWRNLIRLREQPWLKGHIVQGSYVYPGTGYLCMALEAMYHLRQMPEY TAPQGDFVGYRLKDVKLIRALMVPASEEGVETLFSLRHYKESTASYSNNWY EFRVFSWSAADGWAEHAHGMITAVYDAPNTAHLNRIPFNPTMDIAVDLENF AGTRTAENLYDIVDAVGLVYEEPFRNLTGDLVSNEGTAKGVVTVPDTKSLM PFEFEYPYLVHPATMDGFVQMVFPALLHAQTSIASPYLPVFFADTFVRGDIA QAAGHRFDCIATAGYTGFREVTTNVLVRDEGTGDAVVSFKDVKCSGVDSGE NANDGLSGREAIKKLCFHSTWHPDPELLSQDKGDELMRSFIPVPEDPQRVAN LESIAYYYYYRVLQTVTEDQVPSMKPHHQKFFRFMQYQSDLVLAHKSPHQT PEWEQLDDPVISQKMENLAVRLEPSGTDAQLICRVGRKLDKILTGVVDPLAL MLEDELLYKYYESVMGYDIIYHYAEILANKNPNMQVLEIGGGTGGATAPILS AFGGNNGKYPKFTSYTFTDISSGFFEEAETKFKDWEGLVEYRRLNIEEDPVD QGFEAGKYDLVVAASVLHATANITRTLENTRKLLKPGGRLILVEISNILNQAF LLFGCLPGWWMSEESYRPWGPTMDQDMWAEHLKRVGFSDLTLAAPDSEDP KDELGRVFSCVAVEQPAPPSVDPSISSVVLIVDDGAPSDLEQCIDEKFSKLGV PVRRTRLAEAADTPLENTCCVSLAEVNRSVIKDMTLAEFEGIKTIFKSSRALL WVTEGAFNNASRPESALFHGLARSLRAENETFPLTTADFASINRKNPSETAQ QLLELLKQVLRNPGTQEPEYHYEDGSWKVCRCLESVDASAQIHGCLHHESA VSDKTELQPFYQPGRPLKLSIKTPGLLDTLRFEDDPLPAEPLGTDEVEVEVKA SGVNFRDIMICTGQMSDPSLGLECAGIVHRVGDNVSNVKVGQRVVAWTRY NYSNYARTPACVVQPIPDDMSYATAASLPIVYNTVIYGLQHMARLRKGESIL IHAAAGGVGQAAITLAQRLGADIFVTVGTQDKRDLMKSEFGIPDERIFSSRGI SFVQDIKKATNGRGVDVVLNSLAGEMLSATWDCIATFGRFIEIGKKDMIDNR RLDMAPFLRNVMFASIDLITVYEQNISLAGELLHETMDLIRTKAIKPIPRIKIFS FSQFEEAFRFMQQGKHVGKIVMVPEENDMVPAIPASSGPFTFQEDASYLITGF GGLGRSMARWMASRGAKNLIFASRSGDSRPEVRELIDELAEVGTRVKALPV DITNGPALDAAIRGIADSFPPLRGVVQAAMVLDDAIFDNMSLKSFNNGIRPK VQGTWNLHQSTLNQPLDFFVIMASSVGVLGNSGQANYSAGNAYEDALAQY RRSQGLPAISVDLGMILSVGAVAQDSTGMIRNNLESKGFVGIEEDEFLAILET SIRESSSSSAASQMITGIQTQSTVPGDDGVPVDEPFWKSSPVFSHLPKLGARSS GQSNNAGEQSIQSLLKGALLLTDAVTVIIDAMTIKLSRSLMMDVAELDPTRP TSAFGIDSLVAVELRNWFQKHMKADIAVFEILQTNSLQTLAFRVAEKSTLVE GSLADGQ

Protein name	Protein sequence
ProG (253 aa)	MRFLCLHGAGTNSDVLDIQTGPIRDSLDSNATFKFYDGFWDVEPVEEIKNIFA GPFYTWYSPGLGGRTLTEAKAELLDLIATEGPFDACMGFSQGAALLAAVIID HQVQNPFGPNLFKFAAFICGGSPLLVTKALEQDHLDYQPTVDRMAPLTEPW LGPYVPGHEPHPDEQWNMLVAHRVREAGLTIRIPTAHIYGAKDHTVKESLN LRDMCDPRRRVEFDHGGRHEVPRATRVVQQMAMTLRRGIDKAMTAV
ProI (439 aa)	MTDKKPIRVAIIGAGPGGLVLAQTLRQDPRFSVTVYERGVRDGSGVSSLVGF RILLPPSILDNLRSQLPASVATLIDDAIGVPQAQGNRVAFMDEQCGIICRLDVQ QSRDMCSVSRWKLREALLHDAEEIVQFGKQFSSYEQLGGESGDVKVRFADG DEIECDVLVGADGAGSKVRKLLLPNSQRSASGLTVVYFKAPFTPETEAMIPW KSGCVAITPRRSMVVAYYKDRRRPYGPYDLEKIDPADSFLMFGLGCYTNEF VNQSKHPDEMTPEELKDECLARAKDWSPLLRALIALSVPSSVFVSHVKTQDP IKPWESGRVTLLGDAAHSMTPYLGKGASSAMIDAMSLAKALKSEPKQGQG DFLKAQLSIYEEAMLKHGFEAARQSMTAQKFTFNAGDTPWKCWWRNLALK AWDWWMSHPPAMEENFPVSYSEMKKDV
ProJ (549 aa)	MPSSELAHPSRKILLVVTTGGFTHASPVFEIGKILAVRGHTIEFATLEGQEAW TKGYEFITKVHTLGPGPTHDQMNAHYLRMRTWDISKGISGTMPSKYLWDSF WPQTYRGLKAIMDDPKTRPSMMIADFFVDAVKDIHVEYNLPITQVWPQMPF LMMPCSYIPGQPGFQLEGTLTSEHASLWHRIKNELVIFFDLPVIVKWMKWTK KMRLENGVKYPPHKIQKPDYLIFVNSFLGLEIPRDLPPTCAPVGPLISDTYPPL NEECKQFLTKHTKVIYIALGTHIILTNADAAKIINGLLLLLEGSLLDGVIWSIP KSGRQDLDVNTTYKTGNKTLRLGDILDGKNPDWLCSTFVPQRAILDHPSTKL YYTHGGGSSANEGLFHGVPMLSMGVFMDQISNTARLVDGGVAEPLNKFRFT SQEIYIKAKKILLDGDGSYKRNSLRLMRIAHVAARRKKHAADLVEELIYDTE LRFKDGKELRPMHLQTADMRMPVWKAKNWDIWAVSLLGIGAVFGGLGIG GRMLWLHRVWLTGSVKGFVGSQEWLRNLIK
ProK (142 aa)	MAREYSGTCITAVDTPYLQASKSLSEVAFAGDEWIQSIQAINGLMDLAPAFD TDTLYVDDWTKVSMEEMNFGQEQHMFCQPLEVEALPVRNWCMIYSAKAA NAGDANSLGYQVQVVVLKGRAAELINGIVTDLQEGFDEYFW
ProL (775 aa)	MYRLSLASLLSLALPVVAGSQLSSAIDPSQWLNARGKADSLTSKMTLEEKSS MVTGTFDGTCIEHIAPLKRLGFGGLCIQDGPIGLRLGDLVSVFPSGVTTAATW DRQLMALRGEAMAEEFKAKGAHVILGPVAGALGRSPYGGRNWEGFSPDPY LTGIAMGETIRAIQDTGVQATAKHLVGNEQETQRKPTLINGKMVDAVSSNID DRTMHELYMWPFADAVHAGVSSVMCGYNRVNGTYSCENKHLINNLLKKE LGFEGYVMSDFLATPPGLGPVKAGLDMNQPGPVNPLSPVDTYWGDNLVEC VKNKTLSESDLNGMVRRILTPYFYLGQDKDYPSKDPSSQPLVYNGFGYPYPG PSPVGRDVRGNHSALIREIAAAGTVLLKNQGSILPLNKSLTNIGLFGNDAADP SVGTLFSDHDGIDIGTLISGGGSGSGRPSYVISPLDTFKSYAKANGKRLQYVT NNTAILSIMPGLYPWPAVCIVFLKSFATEGFDRKTLVADDNSVQVVNSIASRC PRRTVVVTHSGGPDVMPWATNPNVSAIVAAHYPGQESGNSILDVLIGKVNPS GKLPYTIAKKEEDYNGKITNITGSAAEDSSNWQSEFSEGLFIDYRHFENKGLE PLYEFGYGLSYTTFKLSSTLVVSSANKISARASPSNATLTLGGNPHLWETVIK CHVEVSNTGRVAGATVIQLYASLPKNNIPANSPVRMLRGFEKVYLDPGEAK RVSFALRRRDLSYWDVTIQDWMVPKGGIKLSVGFSSKDLRSSTTVQLVK

Table S5. HR-MS data of compounds in this study.

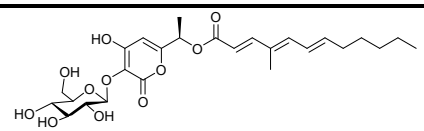
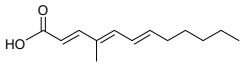
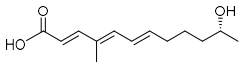
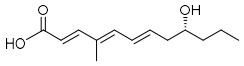
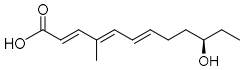
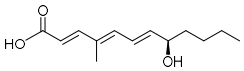
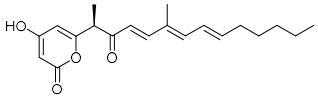
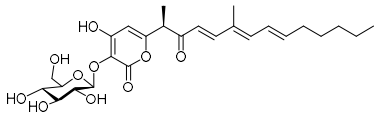
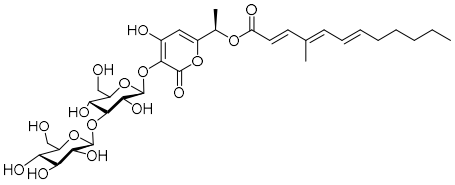
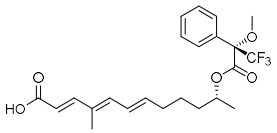
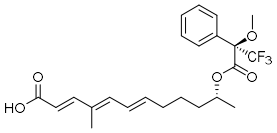
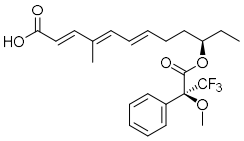
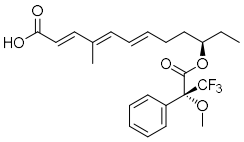
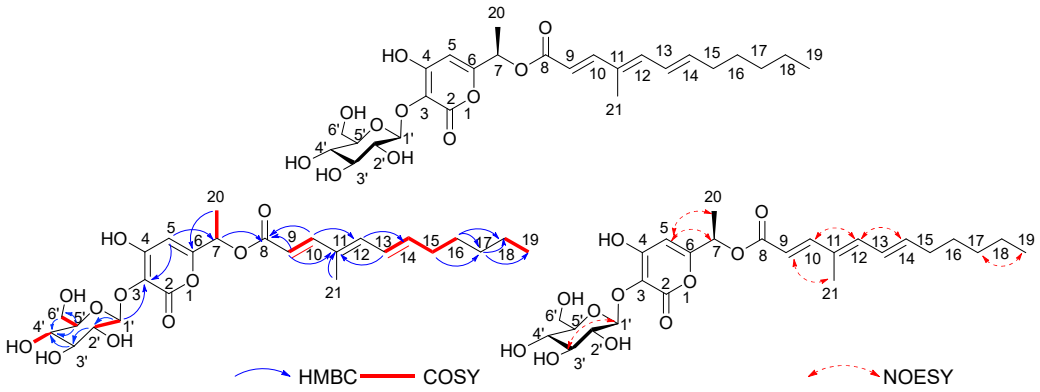
Com.	Structure	Ion Formula	Meas. *m/z*	Calc. *m/z*	Err (ppm)
1		$C_{26}H_{36}NaO_{11}$	$[M+Na]^+$ 547.2158	$[M+Na]^+$ 547.2150	-1.4
2		$C_{13}H_{20}NaO_2$	$[M+Na]^+$ 231.1357	$[M+Na]^+$ 231.1356	-0.8
3		$C_{13}H_{21}O_3$	$[M+H]^+$ 225.1487	$[M+H]^+$ 225.1485	-0.9
4		$C_{13}H_{21}O_3$	$[M+H]^+$ 225.1486	$[M+H]^+$ 225.1485	-0.2
5		$C_{13}H_{21}O_3$	$[M+H]^+$ 225.1487	$[M+H]^+$ 225.1485	-0.6
6		$C_{13}H_{19}O_2$	$[M+H-H_2O]^+$ 207.1380	$[M+H-H_2O]^+$ 207.1380	-0.3
7		$C_{20}H_{25}O_4$	$[M-H]^-$ 329.1773	$[M-H]^-$ 329.1758	-4.4
8		$C_{26}H_{37}O_{10}$	$[M+H]^{+\,-}$ 509.2383	$[M+H]^{+\,-}$ 509.2381	-0.4
9		$C_{32}H_{45}O_{16}$	$[M-H]^-$ 685.2685	$[M-H]^-$ 685.2713	4.0
3*S*		$C_{23}H_{26}F_3O_5$	$[M-H]^-$ 439.1744	$[M-H]^-$ 439.1738	-1.4
3*R*		$C_{23}H_{26}F_3O_5$	$[M-H]^-$ 439.1724	$[M-H]^-$ 439.1738	3.2
5*S*		$C_{23}H_{26}F_3O_5$	$[M-H]^-$ 439.1739	$[M-H]^-$ 439.1738	-0.2
5*R*		$C_{23}H_{26}F_3O_5$	$[M-H]^-$ 439.1726	$[M-H]^-$ 439.1738	2.7

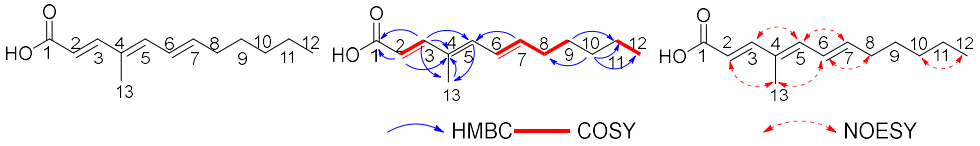
Table S6. NMR data of compound **1** in DMSO-d_6.



(400 MHz for ^{1}H NMR, 100 MHz for ^{13}C NMR)

No.	δ_H, (mult, *J* in Hz)	δ_C, type	HMBC	COSY
2		157.9, C		
3		122.8, C		
4		160.5, C		
5	6.23, s	101.2, CH	C3, C7	
6		159.4, C		
7	5.58, q (6.62)	67.9, CH	C5, C8, C20	H20
8		165.8, C		
9	5.93, d, (15.5)	115.6, CH	C8, C10, C11	H10
10	7.34, d, (15.5)	150.6, CH	C8, C9, C11, C12, C21	H9
11		131.5, C		
12	6.56, d, (11.3)	140.4, CH	C10, C14, C21	
13	6.48, m	127.1, CH	C11, C15	H14
14	6.03, m	142.0, CH	C12, C15, C16	H13, H15
15	2.17, q (7.2)	33.1, CH_2	C13, C14, C16, C17	H14, H16
16	1.39, m	28.7, CH_2	C14, C17, C18	H15
17	1.26, m	31.3, CH_2	C15, C18, C17	
18	1.27, m	22.4, CH_2	C17, C19	H19
19	0.86, t, (6.8)	14.3, CH_3	C17, C18	H18
20	1.45, m	18.4, CH_3	C6, C7	H7
21	1.85, s	12.6, CH_3	C10, C11, C12	
1'	4.81, d (7.3)	103.3, CH	C3, C2'	H2'
2'	3.24, m	74.0, CH	C1', C3'	H1'
3'	3.20, m	76.8, CH	C2', C4'	
4'	3.15, m	70.0, CH		OH4'
5'	3.12, m	77.7, CH	C4', C6'	H6'
6'	3.44, 3.61 m	61.2, CH_2	C4', C5'	H5'
4'-OH	4.94, s			H4'

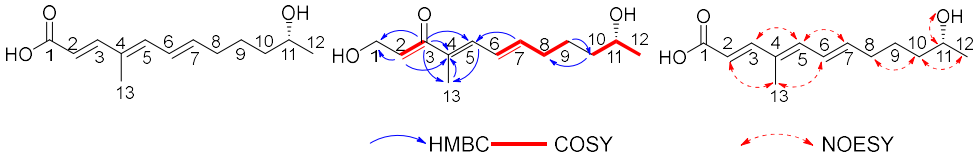
Table S7. NMR data of compound **2** in $CDCl_3$.



(400 MHz for ^{1}H NMR, 100 MHz for ^{13}C NMR)

No.	δ_H, mult (J in Hz)	δ_C, type	HMBC	COSY
1		171.3, C		
2	5.84, d (15.5)	114.9, CH	C1, C4	H3
3	7.42, d (15.5)	151.8, CH	C1, C2, C4, C5, C13	H2
4		131.3, C		
5	6.39, m	140.3, CH	C3, C4, C13	
6	6.40, m	126.6, CH	C7	H7
7	5.99, m	141.9, CH	C5, C8	H6, H8
8	2.18, m	33.5, CH_2	C6, C7, C9, C10	H7, H9
9	1.43, m	28.9, CH_2	C7, C8, C11	H8
10	1.29, m	31.6, CH_2	C11, C12	
11	1.31, m	22.7, CH_2	C10	H12
12	0.89, t (6.9)	14.2, CH_3	C10, C11	H11
13	1.89, s	12.5, CH_3	C3, C4, C5	

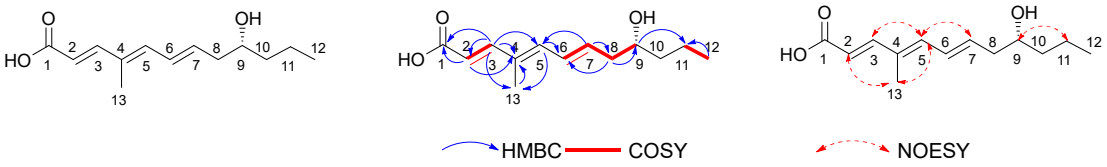
Table S8. NMR data of compound **3** in DMSO-d_6.



(400 MHz for ^{1}H NMR, 100 MHz for ^{13}C NMR)

No.	δ_H, mult (J in Hz)	δ_C, type	HMBC	COSY
1		167.8, C		
2	5.79, d (15.5)	117.1, CH	C1, C4	H3
3	7.22, d (15.5)	148.4, CH	C1, C2, C4, C5, C13	H2
4		131.0, C		
5	6.46, m	138.4, CH	C3, C13	
6	6.44, m	126.6, CH	C7	H7
7	5.99, m	140.5, CH	C5, C8, C9	H6, H8
8	2.16, m	32.7, CH_2	C6, C7, C9	H7, H9
9	1.46, m	24.9, CH_2	C7, C10	H8
10	1.31, m	38.9, CH_2	C8, C9, C11	H11
11	3.58, m	65.6, CH		H10, H12
12	1,03, d (6.2)	23.6, CH_3	C10, C11	H11
13	1.83, s	12.2, CH_3	C3, C4, C5	

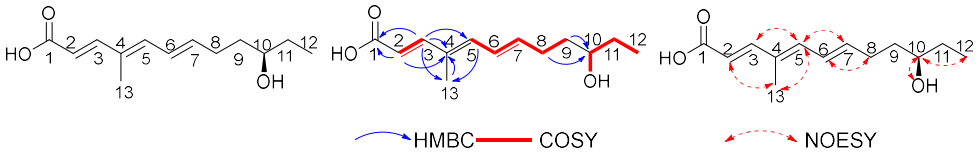
Table S9. NMR data of compound **4** in DMSO-d_6.



(400 MHz for ^{1}H NMR, 100 MHz for ^{13}C NMR)

No.	δ_H, mult (J in Hz)	δ_C, type	HMBC	COSY
1		167.9, C		
2	5.79, d (15.5)	117.3, CH	C1, C4	H3
3	7.22, d (15.5)	148.3, CH	C1, C2, C4, C5, C13	H2
4		131.1, C		
5	6.48, m	138.4, CH	C3, C6, C7, C13	
6	6.46, m	127.9, CH	C5, C8	H7
7	6.02, m	137.7, CH	C5, C8	H6, H8
8	2.23, m	41.2, CH_2	C6, C7, C9	H7, H9
9	3.52, m	69.3, CH	C11	H8
10	3.17, m	48.6, CH_2	C8, C9, C11	
11	1.39, m	18.4, CH_2	C9	H12
12	0.86, m	14.0, CH_3	C11	H11
13	1.83, s	12.2, CH_3	C3, C4, C5	

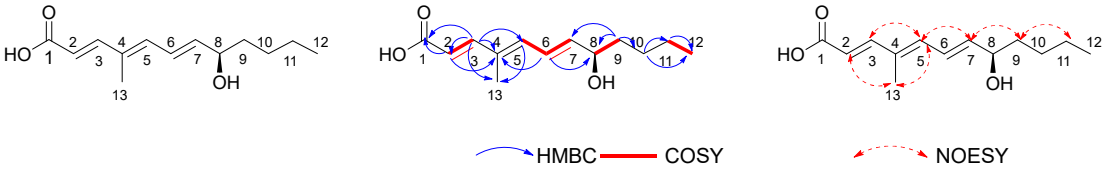
Table S10. NMR data of compound **5** in DMSO-d_6.



(400 MHz for ^{1}H NMR, 100 MHz for ^{13}C NMR)

No.	δ_H, mult (J in Hz)	δ_C, type	HMBC	COSY
1		167.8, C		
2	5.78, d (15.5)	117.0, CH	C1, C4	H3
3	7.21, d (15.5)	148.5, CH	C1, C2, C4, C5, C13	H2
4		131.0, C		
5	6.48, m	138.6, CH	C3, C4, C13	H6
6	6.45, m	126.5, CH	C7	H5, H7
7	6.00, m	140.8, CH	C5, C8, C9	H6, H8
8	1.45, m	36.0, CH_2	C7, C9, C10	H7, H9
9	2.19, m	29.1, CH_2	C7, C8, C10	H8
10	3.35, s	70.4, CH		H11, OH10
11	1.32, m	29.8, CH_2	C10, C12	H10, H12
12	0.85, t (7.4)	10.0, CH_3	C10, C11	H11
13	1.83, s	12.2, CH_3	C3, C4, C5	
10-OH	4.37, s			H10

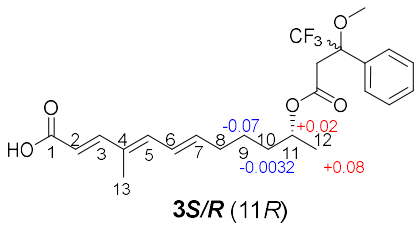
Table S11. NMR data of compound **6** in DMSO-d_6.



(400 MHz for ^{1}H NMR, 100 MHz for ^{13}C NMR)

No.	δ_H, mult (J in Hz)	δ_C, type	HMBC	COSY
1		168.4, C		
2	5.82, d (15.5)	118.3, CH	C1, C4	H3
3	7.22, d (15.5)	148.5, CH	C1, C2, C4, C5, C13	H2
4		132.7, C		
5	6.49, m	138.3, CH	C3, C7, C13	H6
6	6.57, m	124.9, CH	C4	H5, H7
7	5.97, dd (5.9, 14.4)	144.0, CH	C5, C8, C9	H6
8	4.07, q (6.1)	70.8, CH	C7, C9, C10	H9
9	1.42, m	37.3, CH_2	C7, C8, C10	H8
10	1.36, m	27.7, CH_2	C8, C9, C11, C12	
11	1.28, m	22.6, CH_2	C10, C12	H12
12	0.86, t (6.5)	14.4, CH_3	C10, C11	H11
13	1.85, m	12.7, CH_3	C3, C4, C5	

Table S12. The ^{1}H NMR spectroscopic data of compounds **3*S*** (*R*-MTPA derivative) and **3*R*** (*S*-MTPA derivative) in DMSO-d_6.

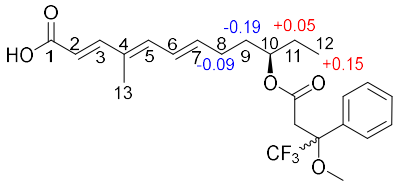


3*S*/*R* (11*R*)

(400 MHz for ^{1}H NMR)

No.	**3*S***	**3*R***	$\Delta\delta^{S-R}$
1			
2	5.7976	5.7951	0.0025
3	7.2209	7.2153	0.0056
4			
5	6.4411	6.4713	-0.0302
6	6.4062	6.4567	-0.0505
7	5.8741	5.9638	-0.0897
8	2.1769	2.1803	-0.0034
9	1.5419	1.6107	-0.0688
10	1.4475	1.4507	-0.0032
11	3.4939	3.4715	0.0224
12	1.2998	1.2155	0.0843
13	1.8274	1.8266	0.0008

Table S13. The ^{1}H NMR spectroscopic data of compounds **5*S*** (*R*-MTPA derivative) and **5*R*** (*S*-MTPA derivative) in DMSO-d_6.

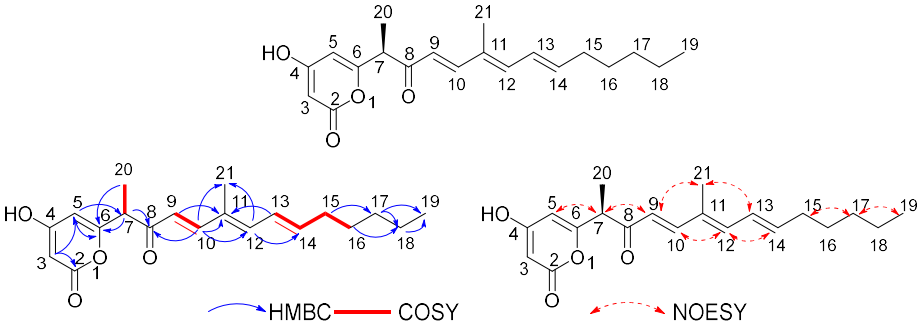


5*S*/*R* (10*R*)

(400 MHz for ^{1}H NMR)

No.	5*S*	5*R*	$\Delta\delta^{S-R}$
1			
2	5.8038	5.8062	-0.0024
3	7.2048	7.2194	-0.0146
4			
5	6.4380	6.4833	-0.0453
6	6.4098	6.4585	-0.0487
7	5.8887	5.9800	-0.0913
8	1.6806	1.7658	-0.0852
9	2.0133	2.2034	-0.1901
10	3.5038	3.4978	0.0060
11	1.6277	1.5760	0.0517
12	0.8866	0.7380	0.1486
13	1.8237	1.8321	-0.0084

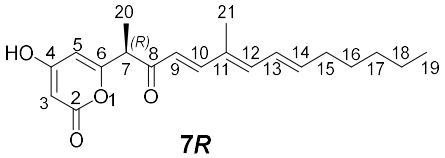
Table S14. NMR data of compound **7** in DMSO-d_6.



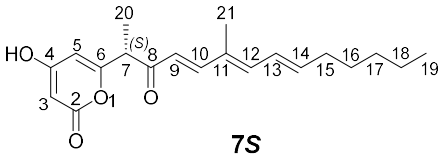
(400 MHz for ^{1}H NMR, 100 MHz for ^{13}C NMR)

No.	δ_H, (mult, *J* in Hz)	δ_C, type	HMBC	COSY
2		163.4, C		
3	5.25, d (2.1)	89.0, CH	C2, C4, C5	
4		170.2, C		
5	6.11, d (2.1)	100.9, CH	C3, C4, C6, C7	
6		164.5, C		
7	4.12, q, (7.0)	48.0, CH	C5, C6, C8, C20	H20
8		195.9, C		
9	6.31, d, (15.6)	122.4, CH	C8, C11	H10
10	7.35, d, (15.6)	148.3, CH	C8, C9, C11, C12, C21	H9
11		131.5, C		
12	6.61, d, (11.3)	141.1, CH	C10, C14, C21	
13	6.50, m	126.8, CH	C11, C15	H14
14	6.05, m	141.8, CH	C12, C15, C16	H13, H15
15	2.17, q, (7.2)	32.6, CH_2	C13, C14, C16, C17	H14, H16
16	1.40, m	28.2, CH_2	C14, C15, C18	H15
17	1.26, m	30.9, CH_2	C18, C19	
18	1.28, m	21.9, CH_2	C16, C17	
19	0.86, t, (6.8)	13.9, CH_3	C17, C18	
20	1.29, m	13.9, CH_3	C6, C7, C8	H7
21	1.84, s	12.1, CH_3	C10, C11, C12	

Table S15. Energy (298.15 K) analysis for **7*R*** and **7*S***.



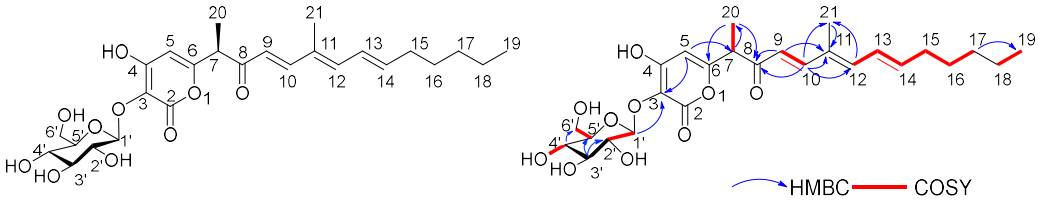
7*R*



7*S*

Conformers	G (Hartree)	Boltzmann Distribution (%)
7*R*1	-1078.5439941	24.00
7*R*2	-1078.5440875	26.49
7*R*3	-1078.5409769	0.98
7*R*4	-1078.5416929	2.09
7*R*5	-1078.5431765	10.09
7*R*6	-1078.543357	12.21
7*R*7	-1078.5431769	10.09
7*R*8	-1078.543208	10.43
7*R*9	-1078.541643	1.98
7*R*10	-1078.5414591	1.63
7*S*1	-1078.543994	24.38
7*S*2	-1078.5440875	26.92
7*S*3	-1078.5410144	1.04
7*S*4	-1078.541693	2.13
7*S*5	-1078.5431765	10.25
7*S*6	-1078.543357	12.41
7*S*7	-1078.5431769	10.25
7*S*8	-1078.543208	10.60
7*S*9	-1078.541643	2.02

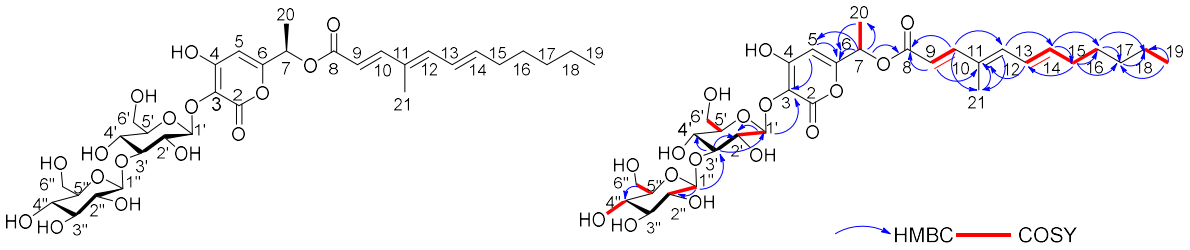
Table S16. NMR data of compound **8** in DMSO-d_6.



(400 MHz for ^{1}H NMR, 100 MHz for ^{13}C NMR)

No.	δ_H, (mult, *J* in Hz)	δ_C, type	HMBC	COSY
2		159.3, C		
3		122.0, C		
4		160.7, C		
5	6.15, m	102.0, CH	C3, C6, C7	
6		159.8, C		
7	4.08, m	47.7, CH	C5, C6, C8, C20	H20
8		196.2, C		
9	6.30, d, (15.5)	122.5, CH	C8, C11	H10
10	7.34, d, (15.5)	148.5, CH	C8, C11, C12, C21	H9
11		131.6, C		
12	6.60, d, (11.3)	141.2, CH	C10, C14, C21	H13
13	6.48, m	126.9, CH	C15	H12, H14
14	6.05, m	142.0, CH	C12, C15, C16	H13, H15
15	2.17, m	32.7, CH_2	C13, C14, C16, C17	H14, H16
16	1.38, m	28.2, CH_2	C14, C17, C18	H15, H17
17	1.26, m	30.9, CH_2	C15, C16, C18, C19	H16
18	1.27, m	22.0, CH_2	C19	H19
19	0.85, t, (6.7)	14.1, CH_3	C17, C18	H18
20	1.28, m	14.0, CH_3	C6, C7, C8	H7
21	1.83, s	12.2, CH_3	C10, C11, C12	
1'	4.74, m	103.1, CH	C3	H2'
2'	3.22, m	73.7, CH	C1', C3'	H1'
3'	3.20, m	76.3, CH	C4'	
4'	3.18, m	69.6, CH	C3'	OH4'
5'	3.11, m	77.2, CH	C4'	H6'
6'	3.44, 3.60, m	60.7, CH_2	C4', C5'	H5'
4'-OH	4.96, s			H4'

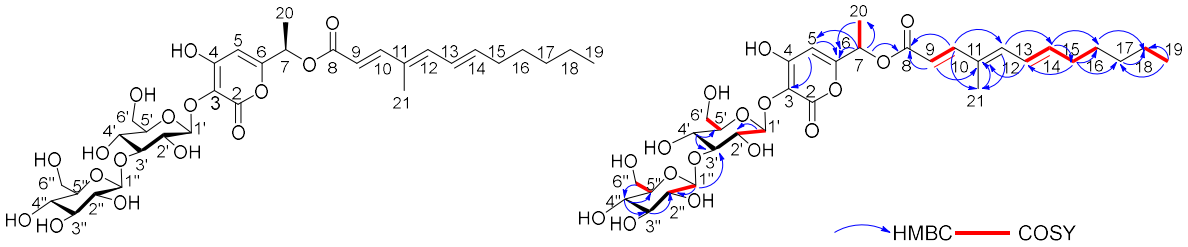
Table S17. NMR data of compound **9** in DMSO-d_6.



(400 MHz for ^{1}H NMR, 100 MHz for ^{13}C NMR)

No.	δ_H, (mult, *J* in Hz)	δ_C, type	HMBC	COSY
2		156.4, C		
3		123.8, C		
4		162.2, C		
5	5.69, s	106.9, CH	C3, C6, C7	
6		160.8, C		
7	5.42, q (6.7)	68.1, CH	C5, C6, C8, C20	H20
8		165.5, C		
9	5.92, d, (15.5)	115.5, CH	C8, C11	H10
10	7.32, d, (15.5)	149.8, CH	C8, C9, C11, C12, C21	H9
11		131.1, C		
12	6.56, m	139.7, CH	C10, C14, C21	
13	6.47, m	126.7, CH	C15	H14
14	6.03, m	141.4, CH	C12, C15, C16	H13, H15
15	2.17, q (7.2)	32.6, CH_2	C13, C14, C16, C17	H14, H16
16	1.39, m	28.2, CH_2	C14, C15, C17, C18	H15
17	1.26, m	30.9, CH_2	C15, C16, C18, C19	
18	1.27, m	21.9, CH_2	C16, C17, C19	H19
19	0.86, t, (6.8)	13.9, CH_3	C17, C18	H18
20	1.41, m	18.2, CH_3	C6, C7	H7
21	1.85, s	12.1, CH_3	C10, C11, C12	
1'	4.22, d (7.7)	107.6, CH	C3, C2'	H2'
2'	3.26, m	73.1, CH	C1', C3'	H1'
3'	3.39, m	86.7, CH	C1', C2', C4'	
4'	3.23, m	68.0, CH	C3'	
5'	3.17, m	76.4, CH		H6'
6'	3.48, 3.67 m	60.7, CH_2		H5'
1"	4.40, d (7.7)	104.2, CH	C3', C2"	H2"
2"	3.03, m	74.4, CH	C1"	H1"
3"	3.17, m	76.8, CH		
4"	3.04, m	70.2, CH	C6"	OH4"
5"	3.17, m	76.9, CH		H6"
6"	3.48, 3.67 m	61.2, CH_2	C4"	H5"
4"-OH	4.95, s			H4"

Table S18. NMR data of compound **9** in pyridine-d_5.



(400 MHz for ^{1}H NMR, 100 MHz for ^{13}C NMR)

No.	δ_H, (mult, *J* in Hz)	δ_C, type	HMBC	COSY
2		157.7, C		
3		123.4, C		
4		162.8, C		
5	6.54, m	104.6, CH	C3, C6, C7	
6		159.0, C		
7	5.90, m	68.8, CH	C5, C6, C8, C20	H20
8		166.7, C		
9	6.06, d, (15.6)	116.5, CH	C8, C11	H10
10	7.64, d, (15.6)	151.0, CH	C8, C12, C21	H9
11		132.0, C		
12	6.46, m	140.7, CH	C10, C14, C21	
13	6.51, m	127.5, CH	C11, C15	H14
14	6.00, m	142.2, CH	C12, C15, C16	H13, H15
15	2.12, q (7.3)	33.9, CH_2	C13, C14, C16	H14, H16
16	1.35, m	29.5, CH_2	C14, C15, C18	H15
17	1.22, m	32.1, CH_2	C18, C19	
18	1.24, m	23.2, CH_2	C16, C17, C19	H19
19	0.85, t, (6.6)	14.6, CH_3	C17, C18	H18
20	1.52, d (6.7)	18.9, CH_3	C6, C7	H7
21	1.86, s	12.8, CH_3	C10, C11, C12	
1'	5.60, d (7.3)	103.8, CH	C2'	H2'
2'	3.94, m	74.3, CH	C1'	H1'
3'	4.19, m	88.5, CH	C1"	
4'	4.02, m	69.9, CH	C3', C5'	
5'	3.89, m	78.7, CH	C4'	H6'
6'	4.37, 4,21 m	62.2, CH_2		H5'
1"	5.27, d (7.8)	106.3, CH	C3', C2"	H2"
2"	4.05, m	76.2, CH	C1", C3"	H1"
3"	4.24, m	78.7, CH	C2", C4"	
4"	4.18, m	72.1, CH	C3", C5", C6"	
5"	4.00, m	79.1, CH	C4"	H6"
6"	4.55, 4.28, m	63.0, CH_2	C4"	H5"

3. Supplementary Figures

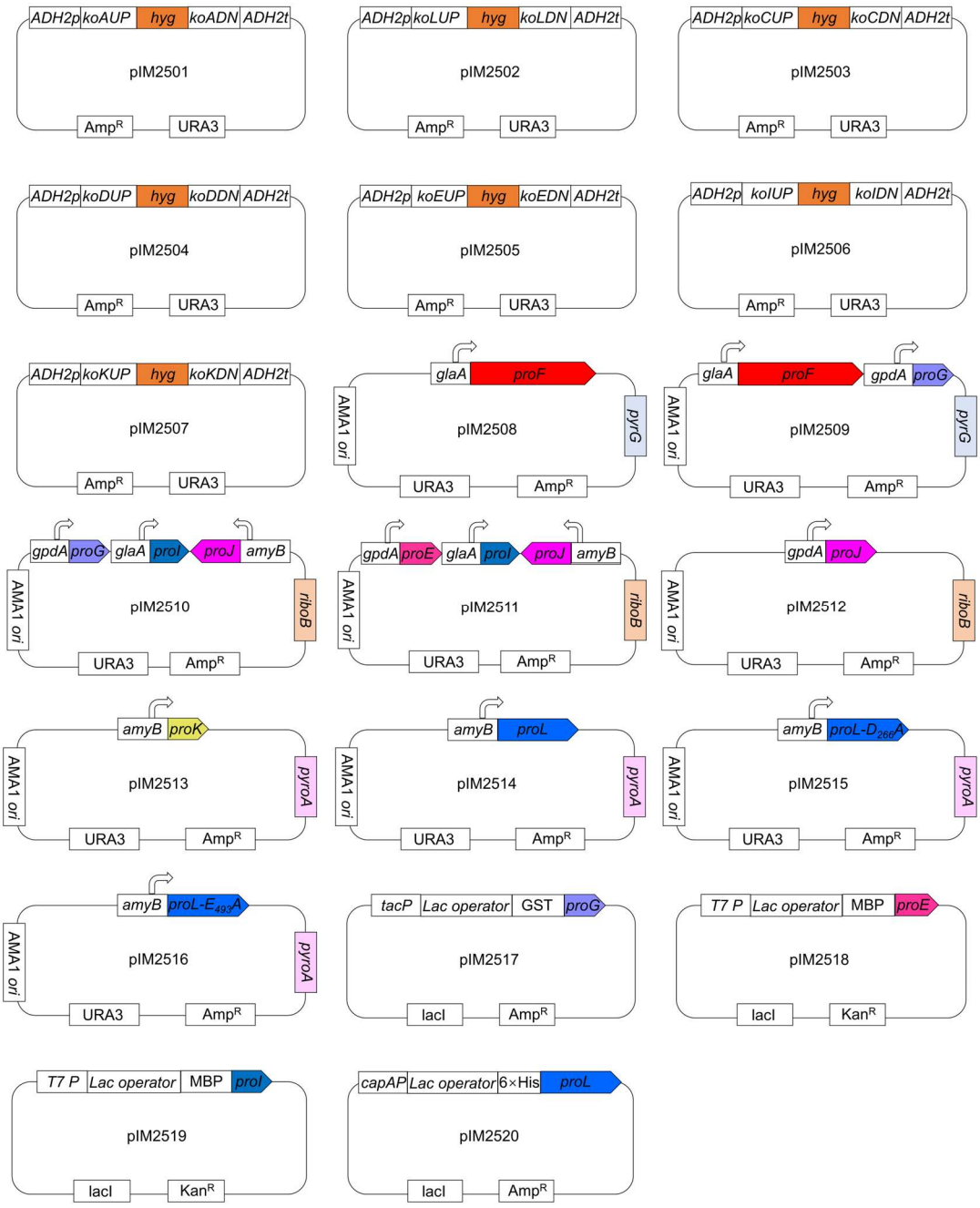


Figure S1. Schematic diagram of plasmids used in this study.

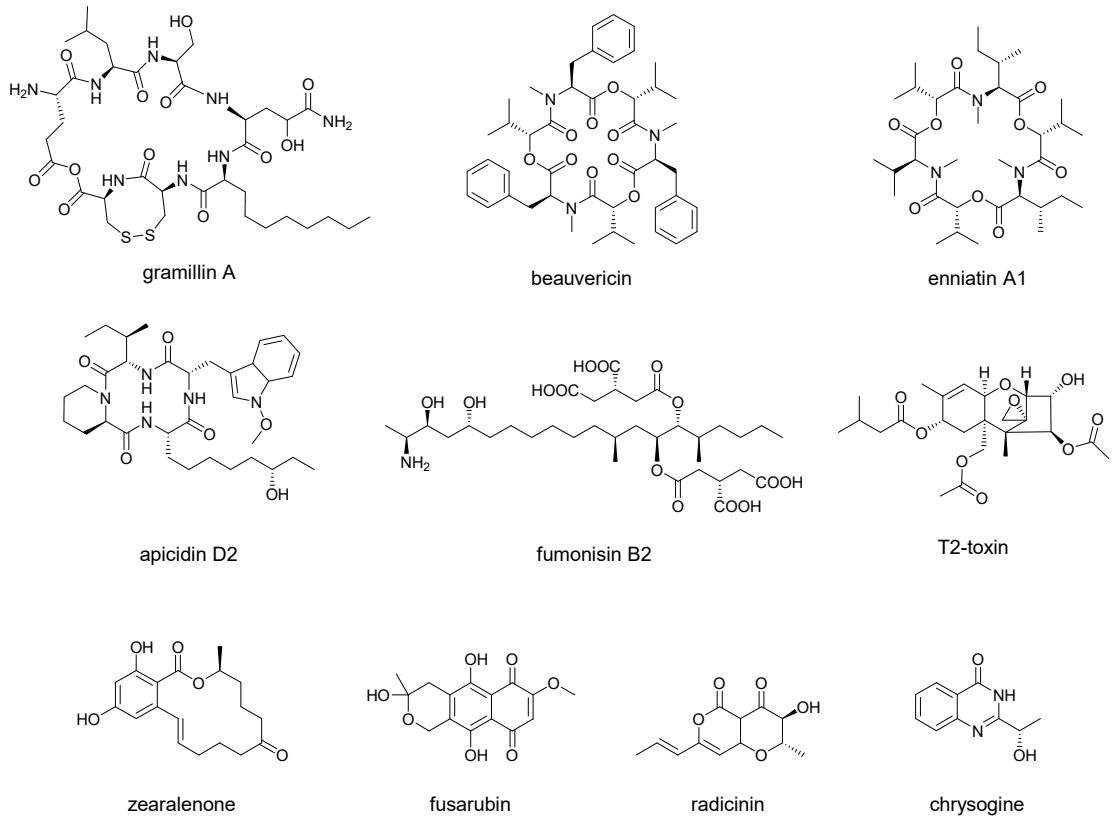


Figure S2. The well-known fungal phytotoxins produced by *Fusarium* species.

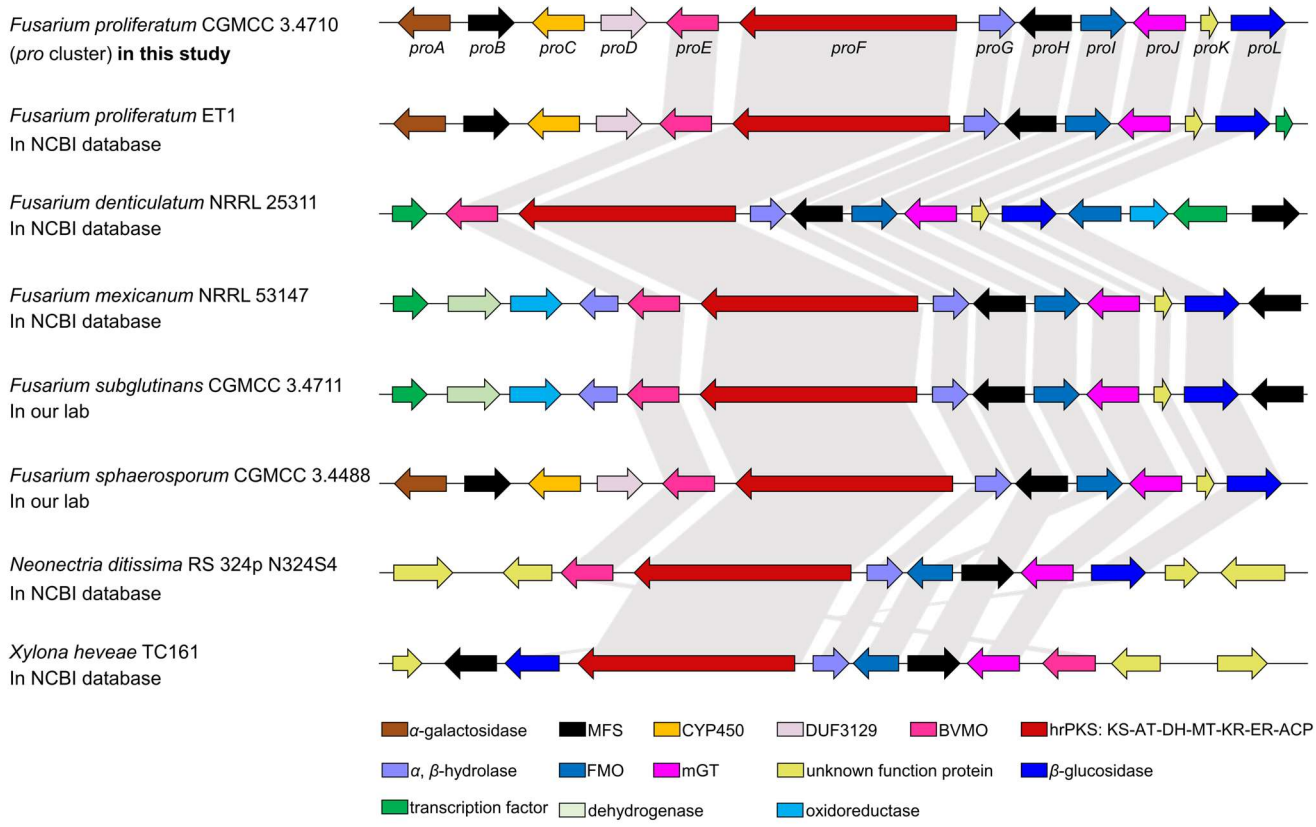


Figure S3. The genome-mined homologue gene clusters of PKS16 in public and our lab database.

```
# ProJ Length: 549
# ProJ Number of predicted TMHs:  1
# ProJ Exp number of AAs in TMHs: 23.78525
# ProJ Exp number, first 60 AAs:  4.16467
# ProJ Total prob of N-in:        0.19967
ProJ    TMHMM1.0        outside     1   500
ProJ    TMHMM1.0        TMhelix     501 519
ProJ    TMHMM1.0        inside      520 549
```

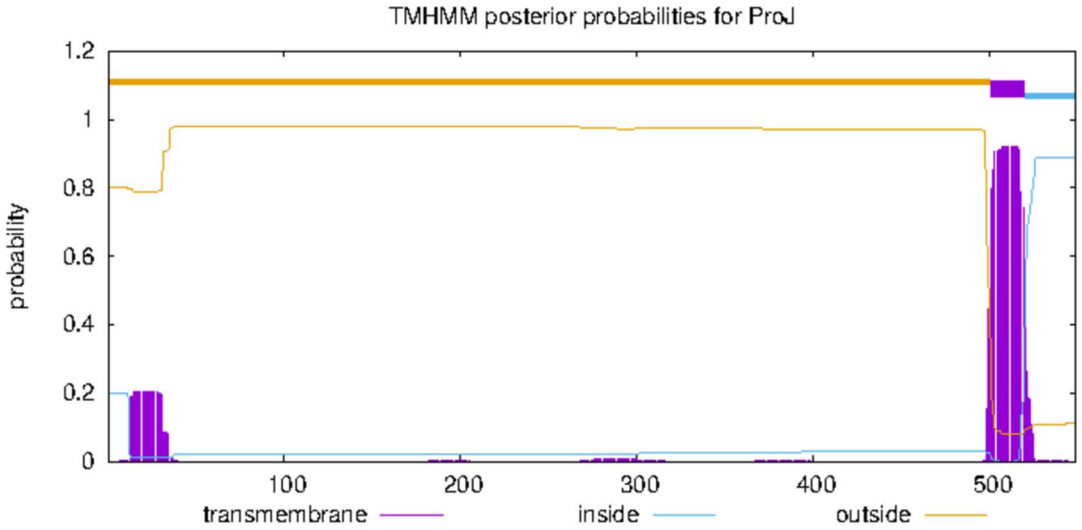


Figure S4. The prediction of transmembrane helices of mGT ProJ. ProJ is a membrane-bound protein, a transmembrane sequence was predicted at the *C*-terminal.

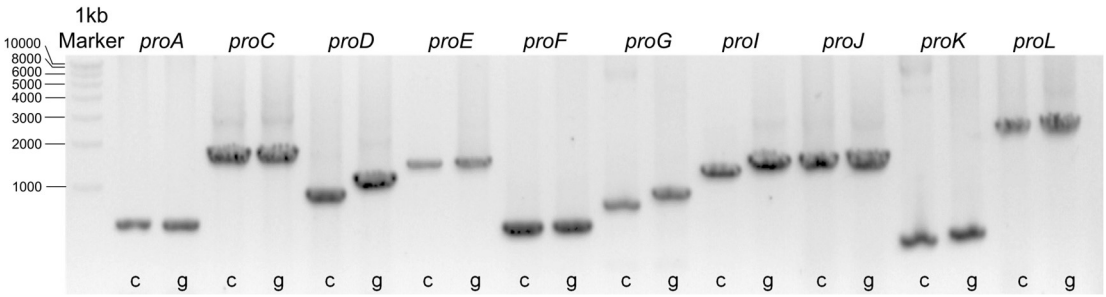


Figure S5. The *pro* genes are all under transcription when *F. proliferatum* is incubated on glycerol medium. c, complementary DNA; g, genomic DNA. PCR was conducted utilizing a set of validation primers that were strategically designed to span the regions containing introns.

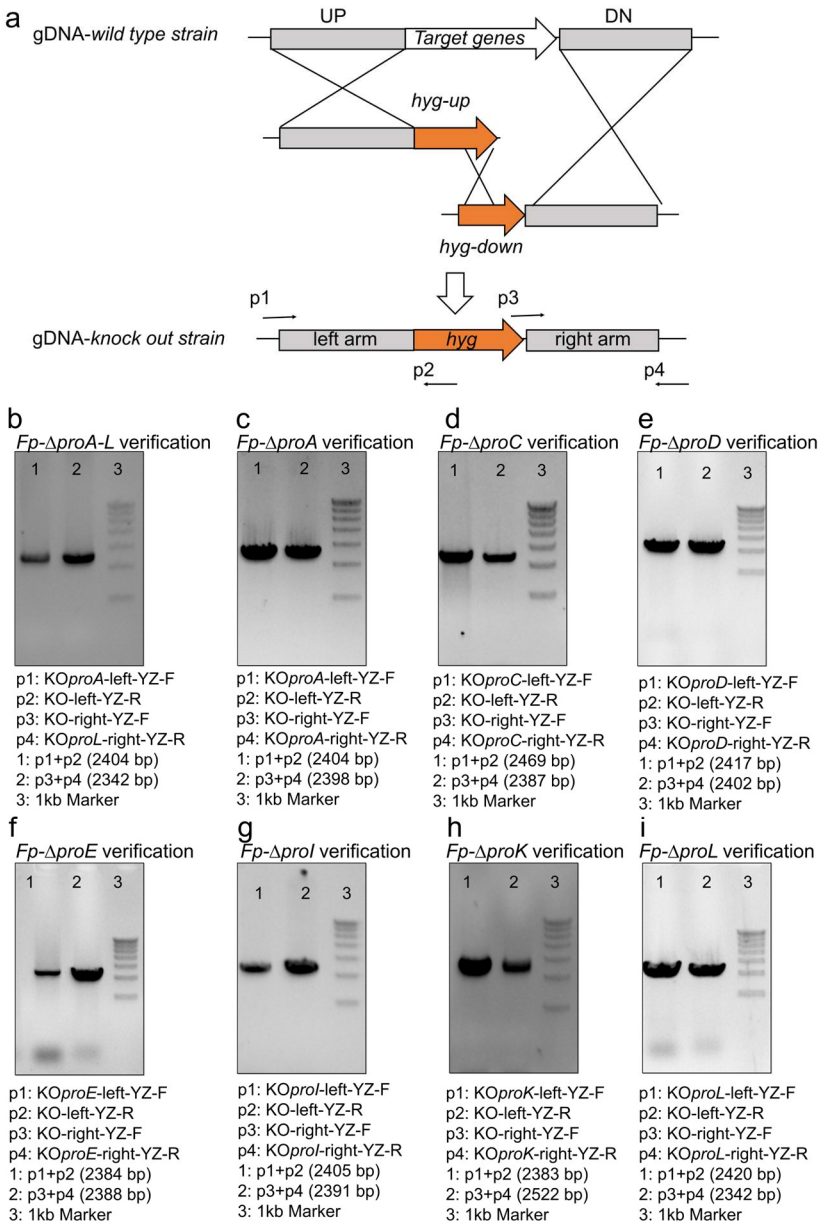


Figure S6. PCR verification of knockout mutants of *F. proliferatum* in this study. a Schematic diagram of the constructed knockout mutants. b-i PCR verification of *Fp*-Δ*proA-L* (b), *Fp*-Δ*proA* (c), *Fp*-Δ*proC* (d), *Fp*-Δ*proD* (e), *Fp*-Δ*proE* (f), *Fp*-Δ*proI* (g), *Fp*-Δ*proK* (h) *and Fp*-Δ*proL* (i) mutants.

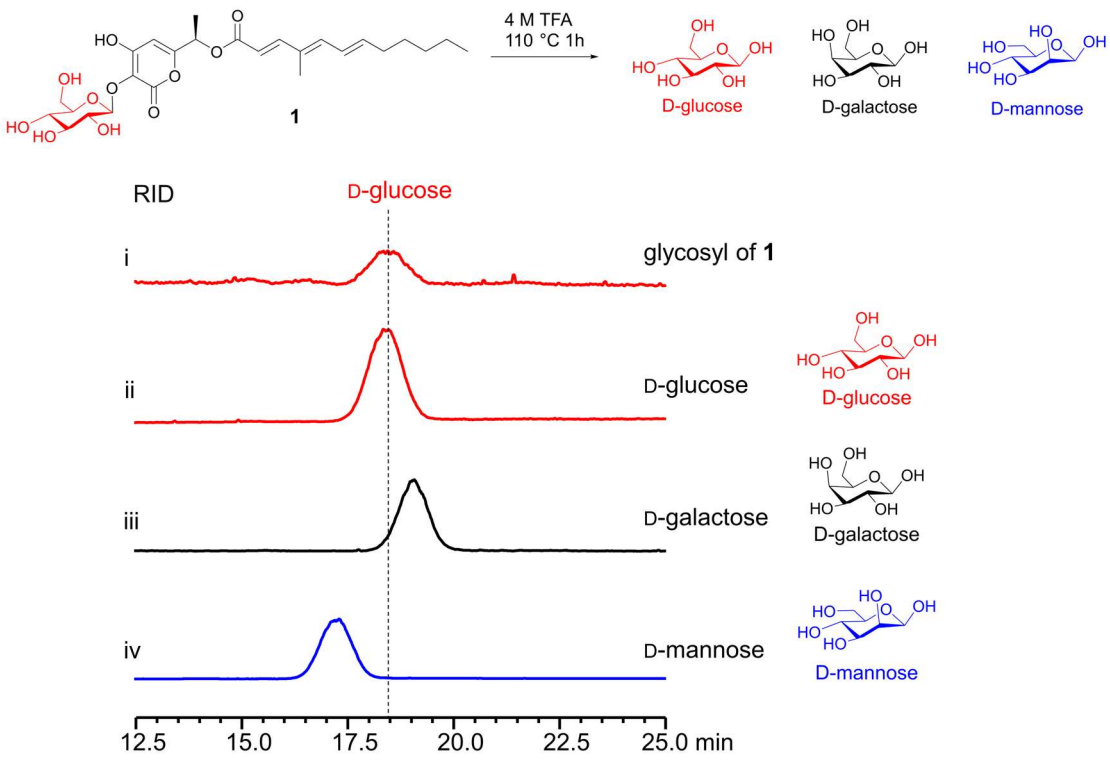


Figure S7. HPLC (refractive index detector) analysis of the glycosyl of compound **1**. By comparison with the standards of monosaccharides as well as NMR analysis (Table S6 and Figures S36–42), the glycan moiety of **1** is confirmed to be D-glucose.

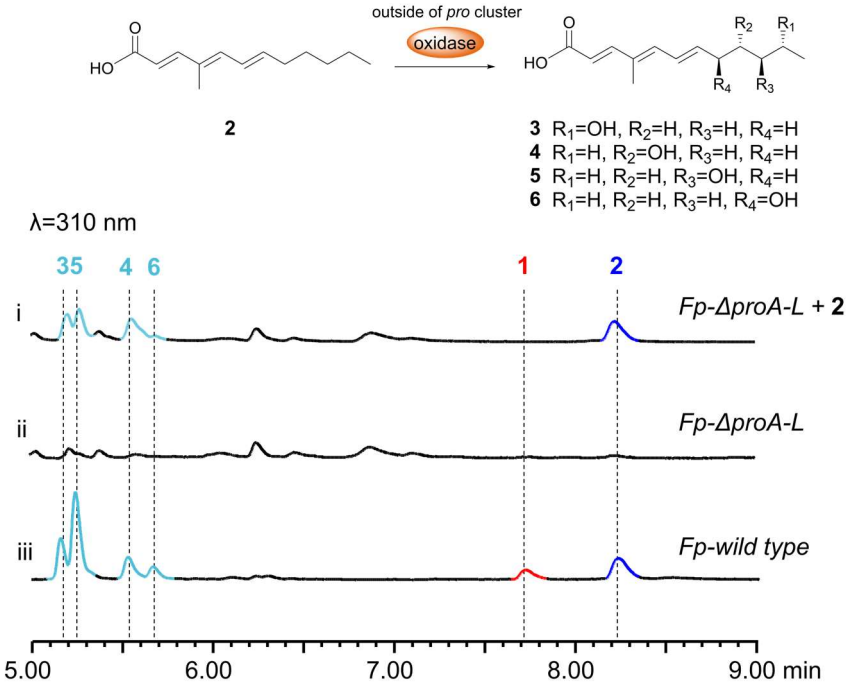


Figure S8. LC–MS analysis of the feeding assay of *Fp*-Δ*proA-L* towards compound **2**. Compounds **3**–**6** are over-oxidized off-pathway derivates of **2** by *F. proliferatum*.

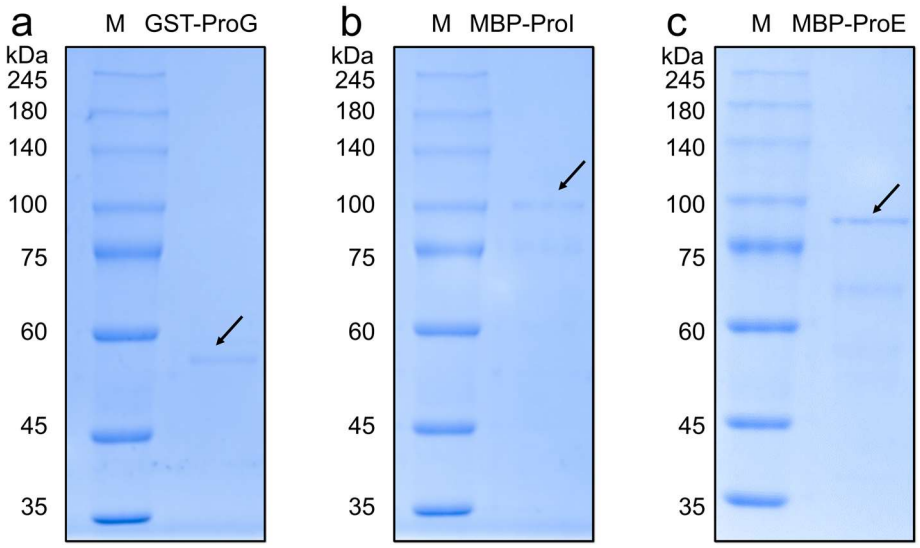


Figure S9. SDS-PAGE of proteins purified from *E. coli* BL21 in this study. a GST-ProG (54.9 kDa). b MBP-ProI (96.1 kDa). c MBP-ProE (98.8 kDa).

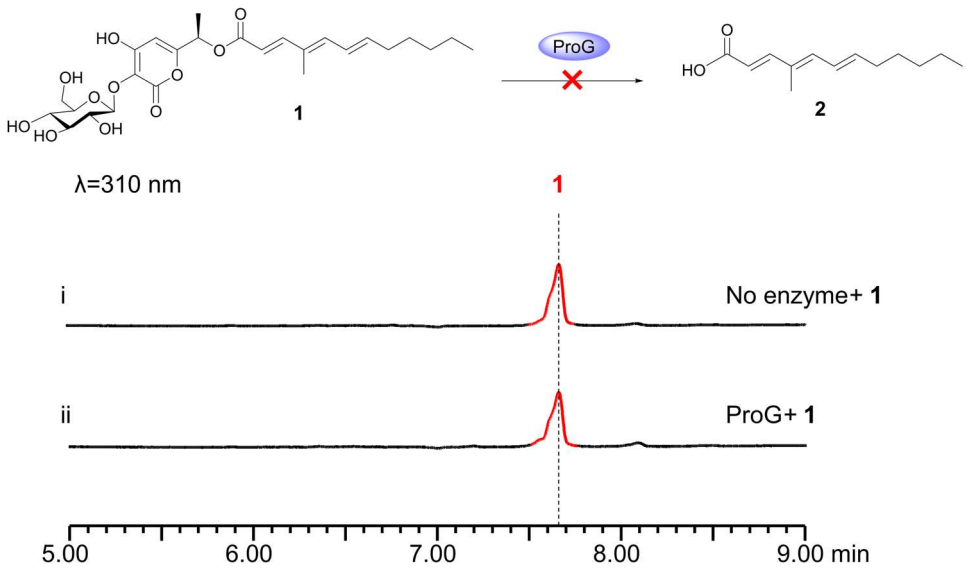


Figure S10. LC–MS analysis of the *in vitro* biochemical assays of α, β-hydrolase ProG towards compound **1**. The conversion of **1** to **2** is not catalyzed by α, β-hydrolase ProG.

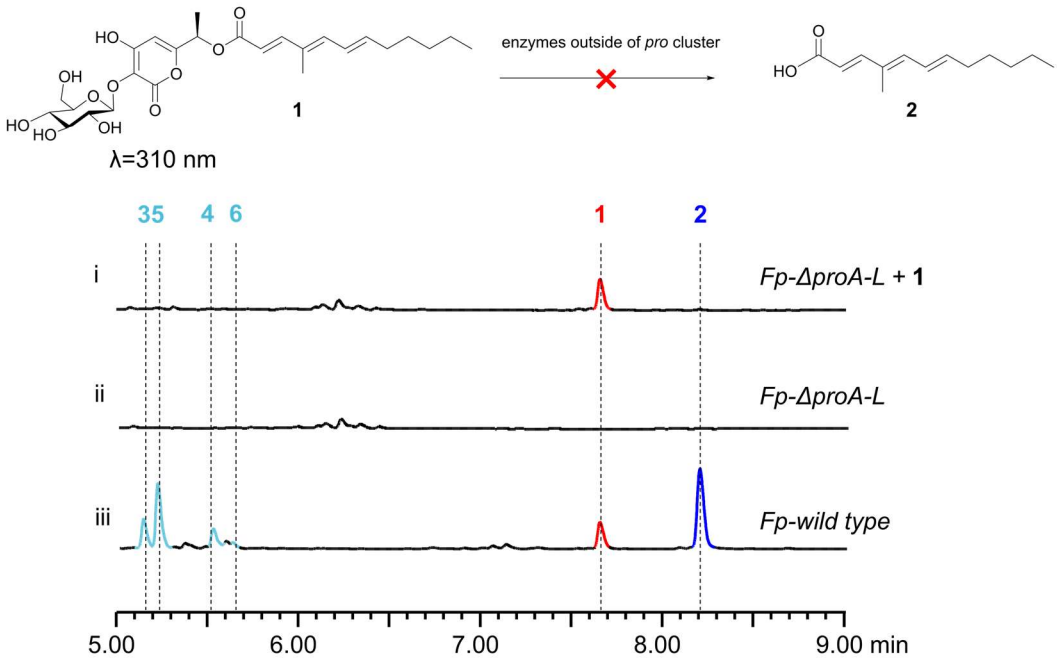


Figure S11. LC–MS analysis of the feeding assay of *Fp*-Δ*proA-L* towards compound **1**. Enzymes outside of *pro* cluster do not have the ability to catalyze the hydrolysis of **1** to **2**, indicating that it is an on-pathway conversion in *F. proliferatum.*

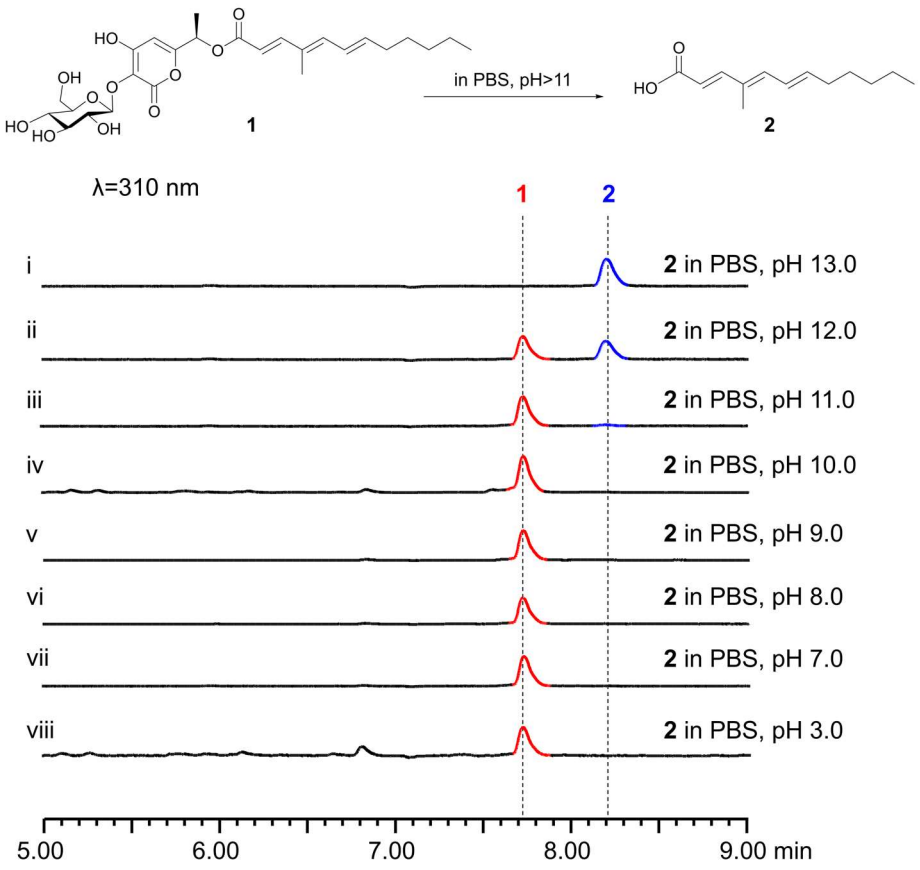


Figure S12. LC–MS analysis of compound **1** in PBS buffer under different pH values. The formation of **2** is only detected under strong base conditions (pH>11).

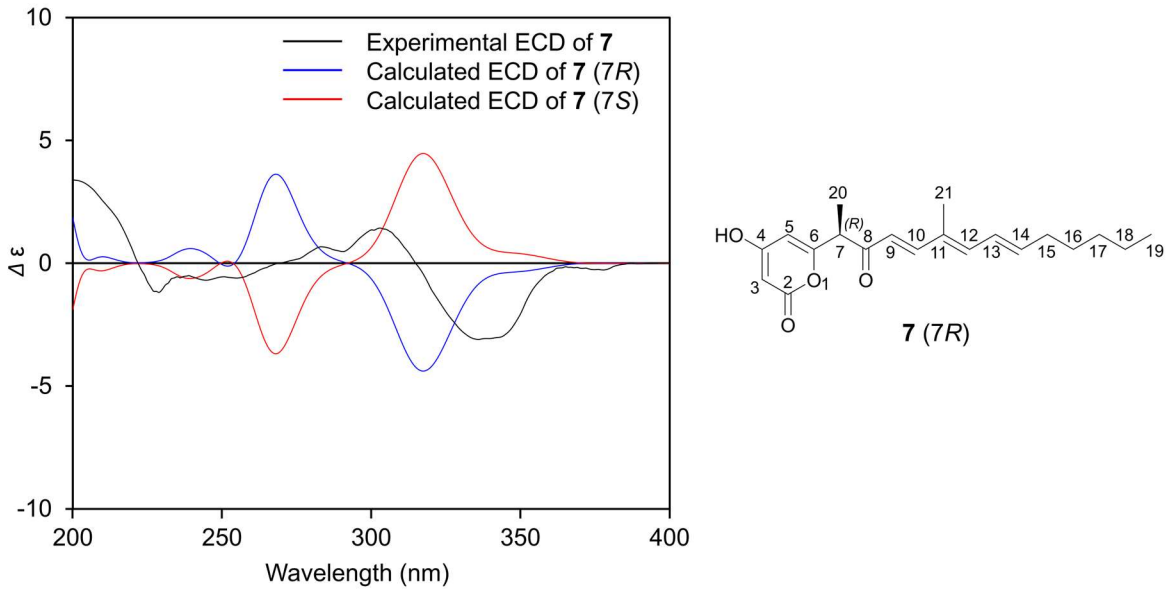


Figure S13. Comparison of calculated ECD spectra of (7*S*)- and (7*R*)-isomers with the experimental spectra of **7** (*c* 0.1 mg/mL, MeOH). The C7-methyl stereochemistry of **7** is confirmed to be *R* configuration.

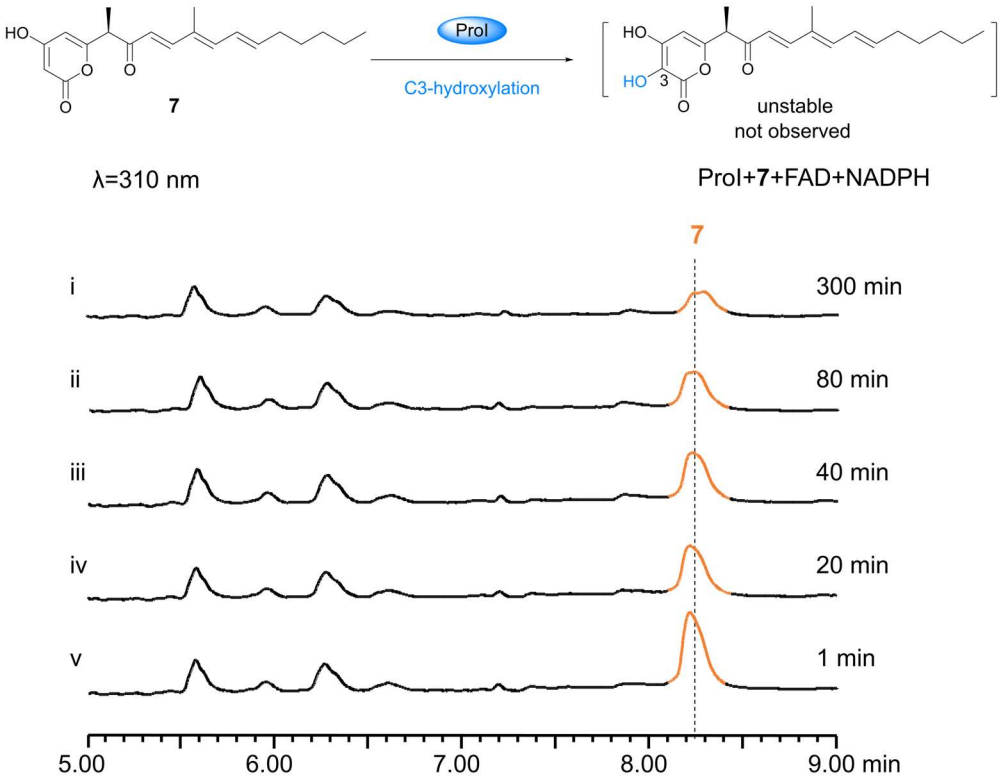


Figure S14. Time-course assays for FMO ProI towards compound **7**. Time-dependent decrease of **7** can be observed when it is incubated with ProI, the generation of the proposed product C3-hydroxyl **7** is not observed, implying hydroxylation occurring at the C3 of the pyrone ring of **7** possibly makes it unstable.

Figure S15. Bioinformatic analysis of β-glucosidase ProL. a Conserved domain analysis of ProL indicates that it belongs to the glycoside hydrolase family 3 (GH3) of β-glucosidase proteins. b Sequence alignments of ProL with homologous β-glucosidase proteins showed that it harbors the classical catalytic amino acid residues, the red box shows the conserved catalytic residues aspartate and glutamate in β-glucosidase that involved in the hydrolysis of glycosidic bonds. The sequences of β-glucosidase are downloaded from NCBI database.

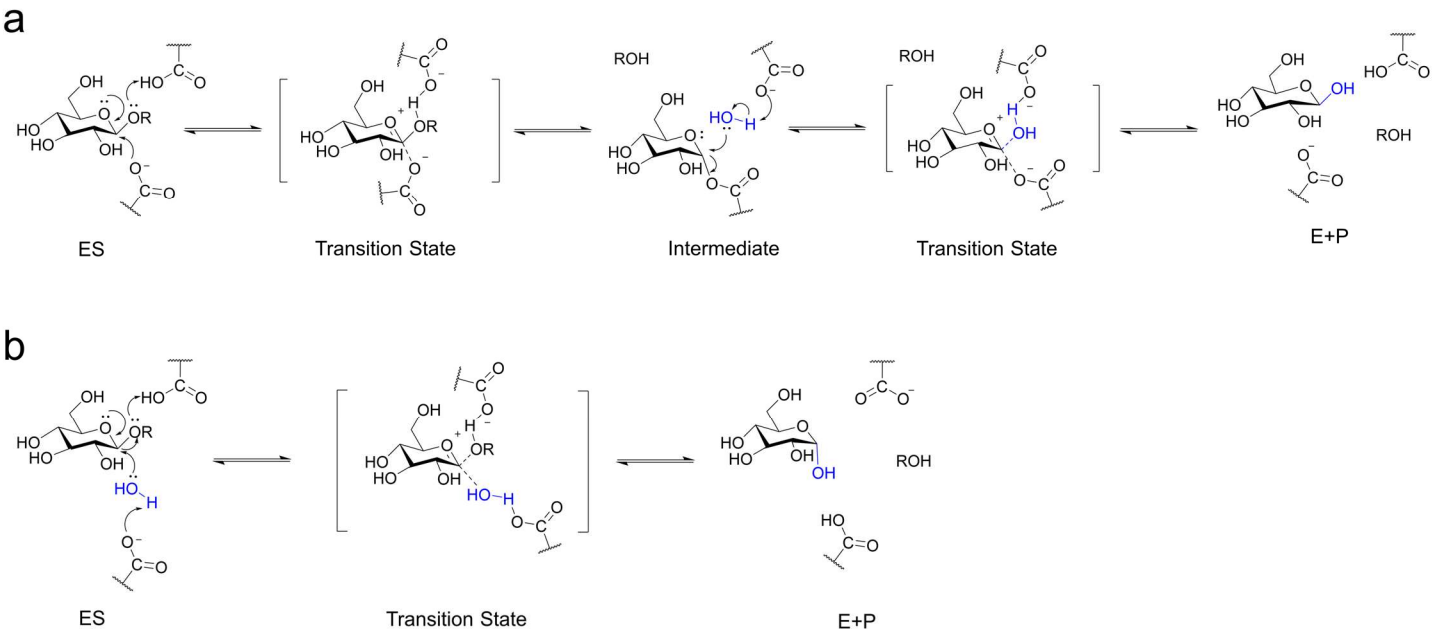


Figure S16. Traditional catalytic mechanism of β-glucosidase that hydrolyze glycosidic bonds. a Mechanism of action of retaining β-glucosidase. b Mechanism of action of inverting β-glucosidase.

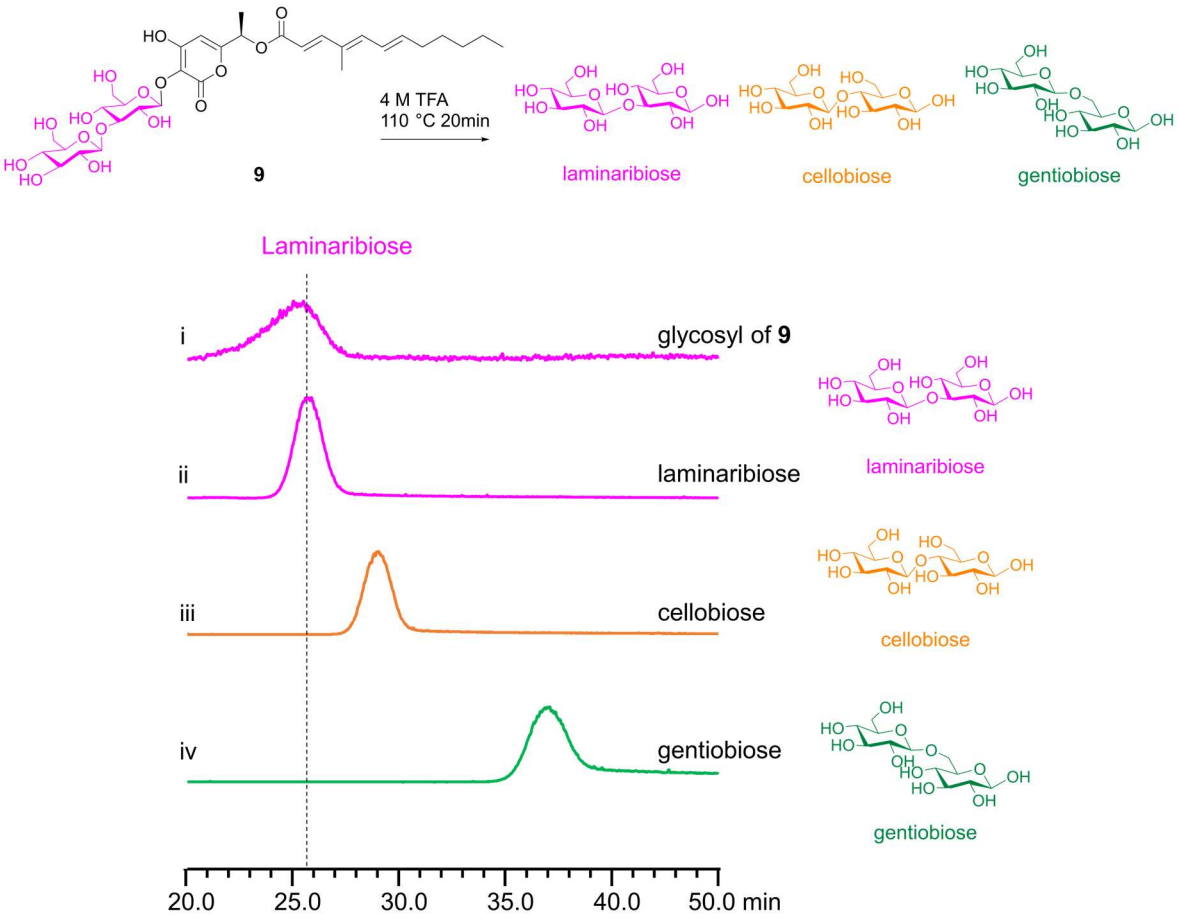


Figure S17. HPLC (evaporative light scattering detector) analysis of the glycosyl of compound **9**. By comparison with the standards of disaccharides as well as NMR analysis (Tables S17–18 and Figures S96–109), the di-glucose moiety of **9** is confirmed to be laminaribiose.

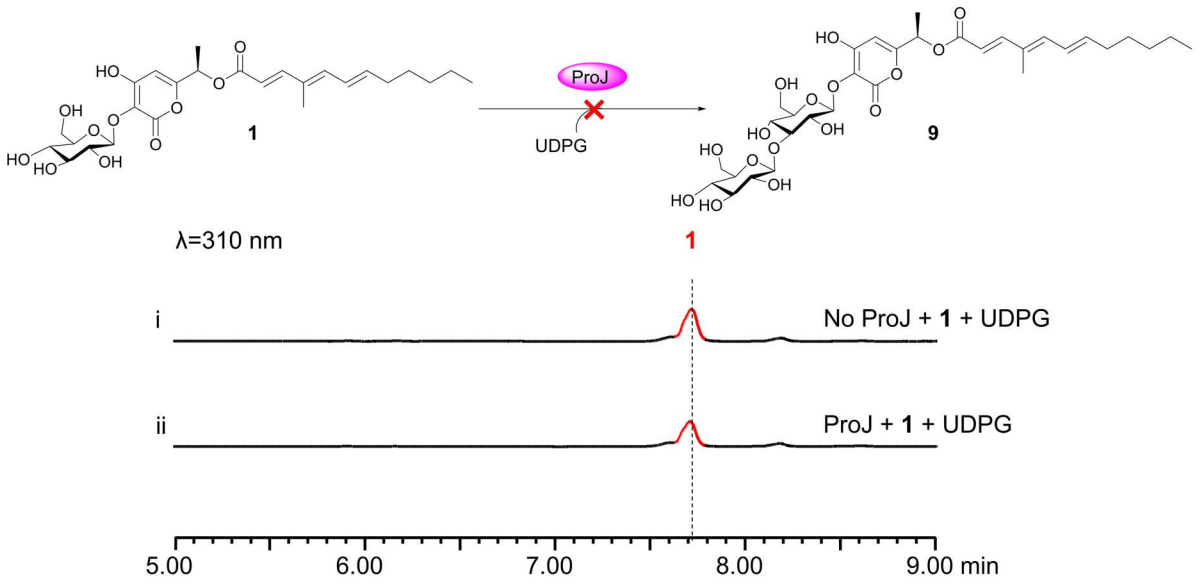


Figure S18. LC–MS analysis of the *in vitro* biochemical assays of mGT ProJ towards compound **1**. mGT ProJ is not responsible for catalyzing the glycosylation of **1** to produce **9**.

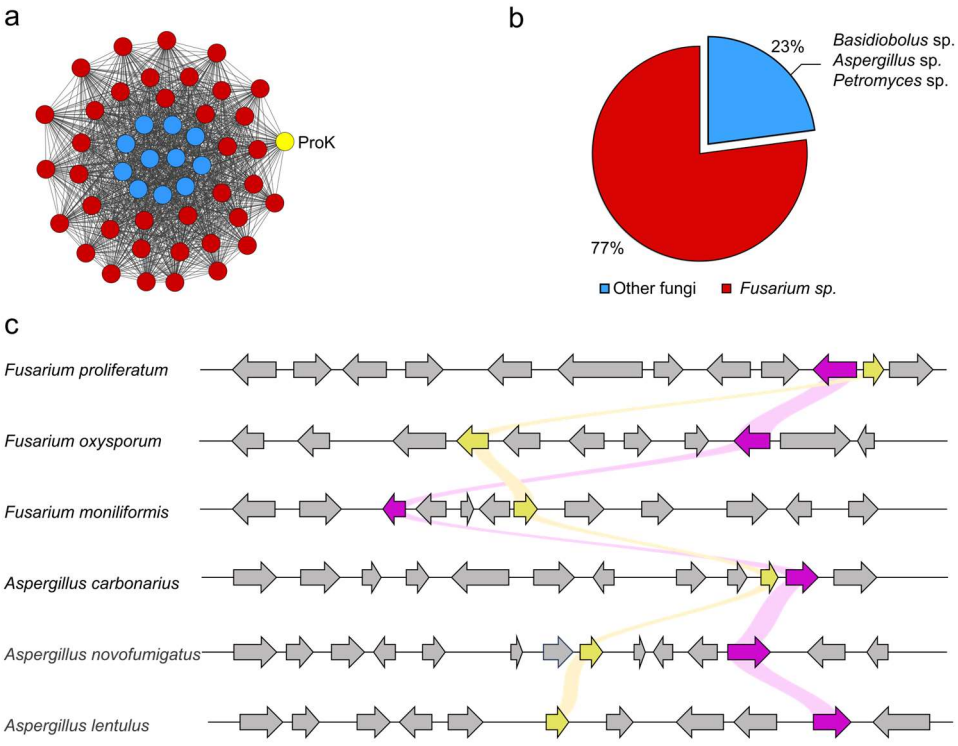


Figure S19. Sequence similarity network (SSN) analysis of unknown function protein ProK. a The homologue proteins of unknown function protein ProK are a very small family clade in fungi (n = 47). ProK is shown in yellow, proteins from *Fusarium* sp. and other fungi are shown in red and blue, respectively. b Distribution of homologue proteins of ProK in fungi, the homologous genes of ProK are mainly distributed in the *Fusarium* sp. (77%). c Clusters containing ProK and its homologues protein from diverse fungi analyzed by EFI-GNT (EFI Genome Neighborhood Tool). Genes encoding ProK homologues are shown in yellow, and genes encoding glycosyltransferase are shown in purple.

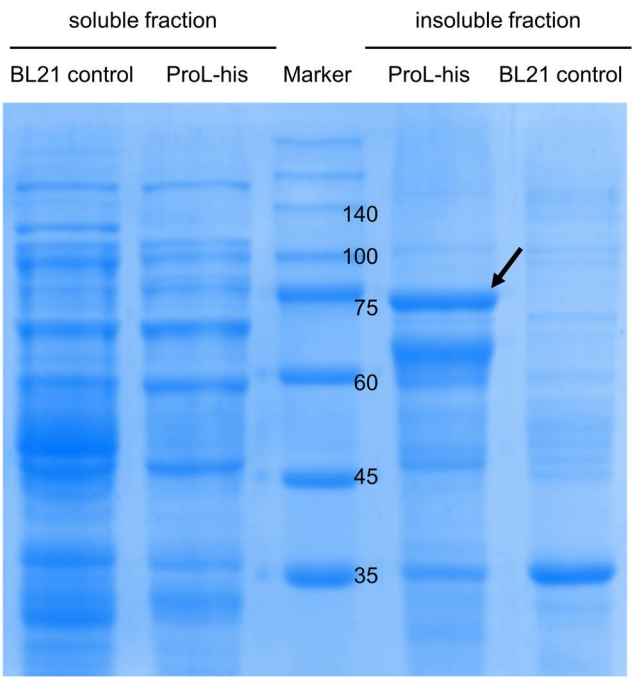


Figure S20. The β-glucosidase ProL with His-tag in *E. coli* BL21 is insoluble.

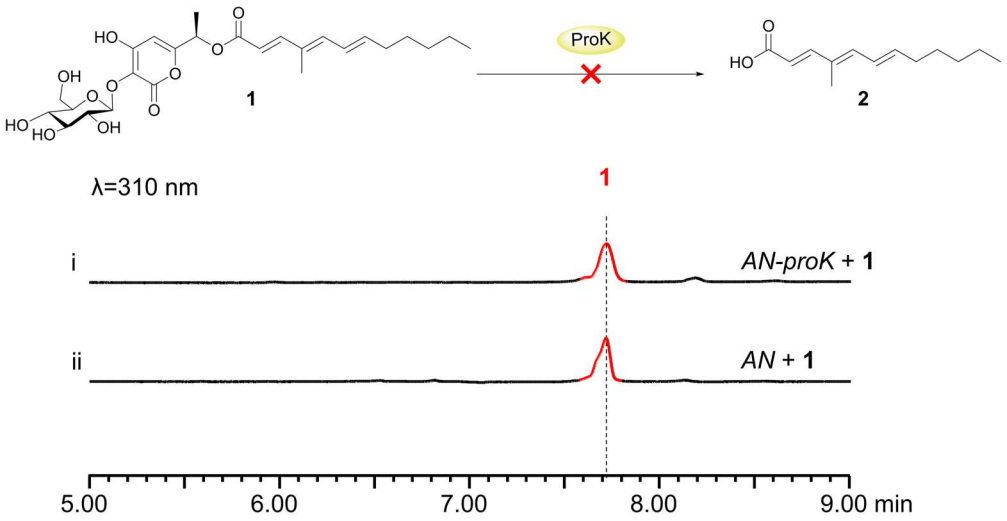


Figure S21. LC–MS analysis of *in vitro* biochemical assays of crude enzymes of *AN-proK* towards compound **1**. Unknown function protein ProK is not responsible for catalyzing the hydrolysis of **1** to **2**.

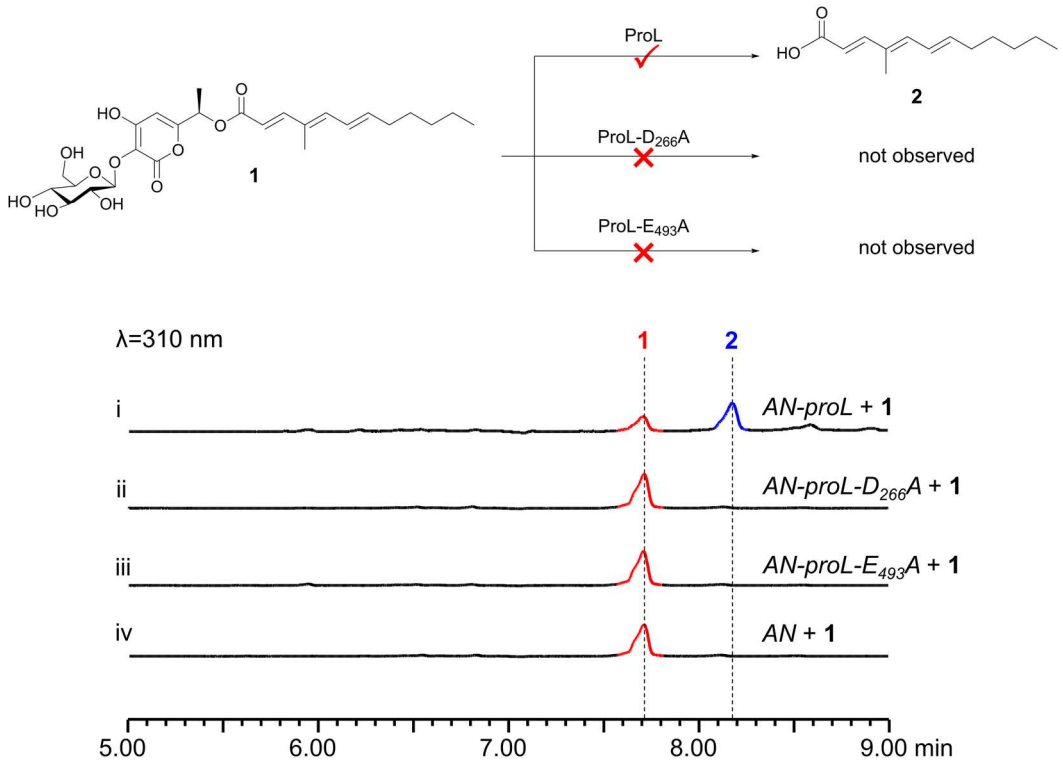


Figure S22. LC–MS analysis of mutagenesis experiments on the conventional active sites D266 and E493 of β-glucosidases in ProL towards compound **1**. ProL harbors the classical catalytic amino acid residues (D266 and E493) of β-glucosidases to mediate the hydrolysis of the ester bond of **1** to form **2**.

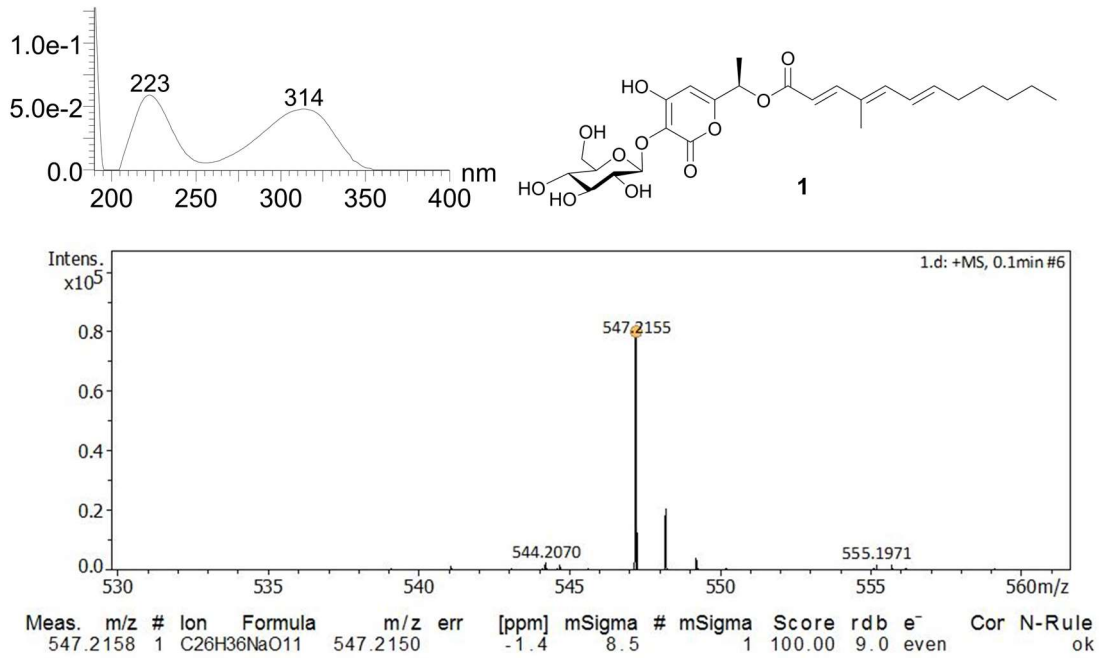


Figure S23. UV absorption and HRMS spectrum (positive ionization) analysis of compound **1**.

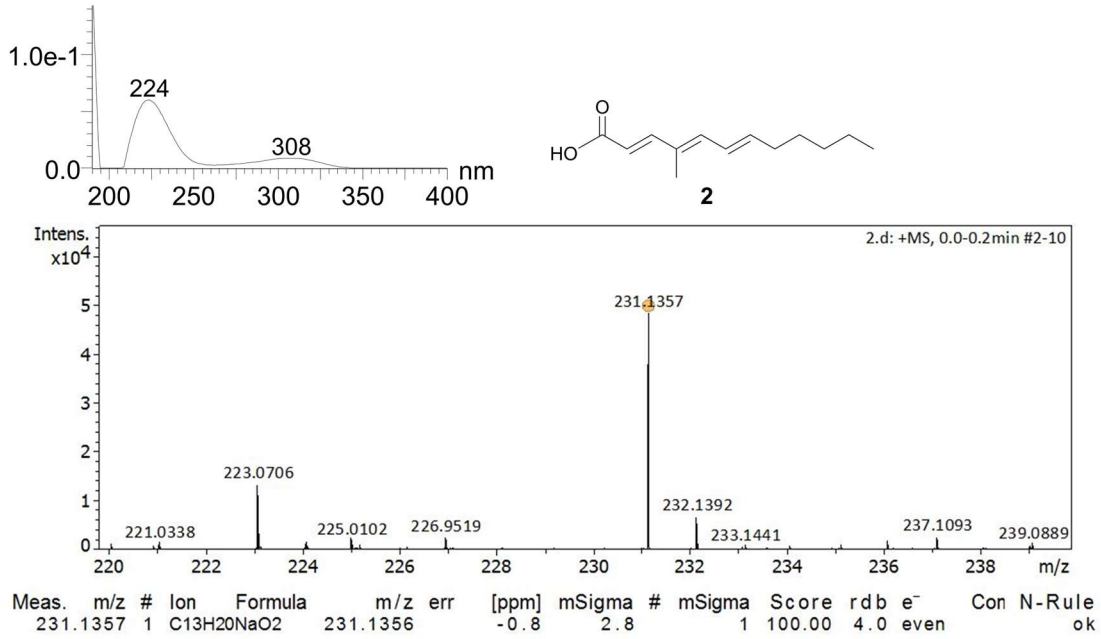


Figure S24. UV absorption and HRMS spectrum (positive ionization) analysis of compound **2**.

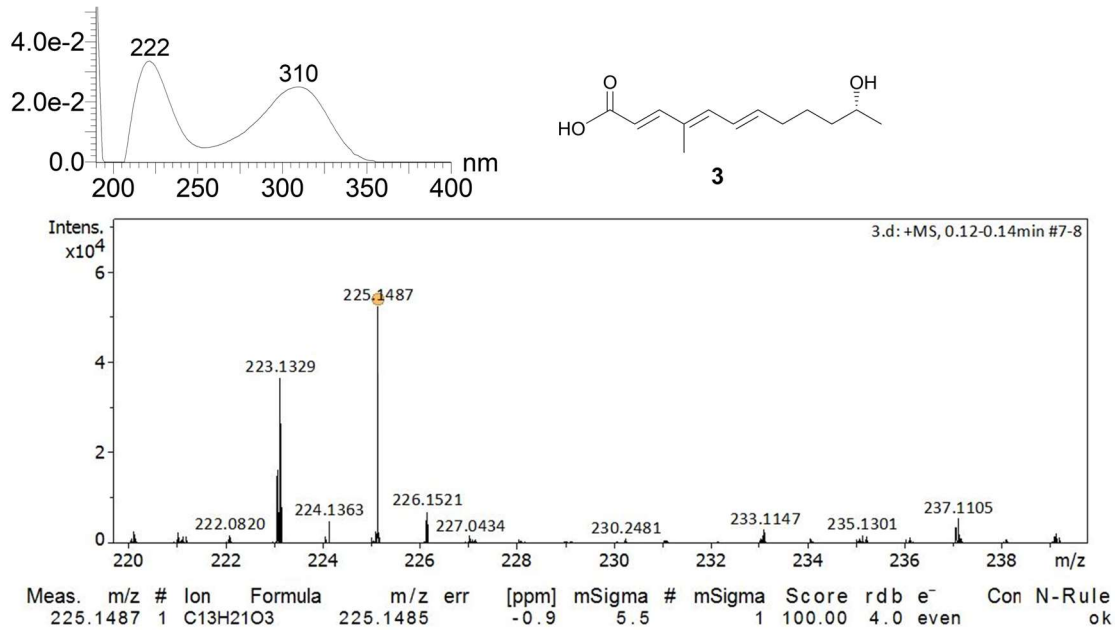


Figure S25. UV absorption and HRMS spectrum (positive ionization) analysis of compound **3**.

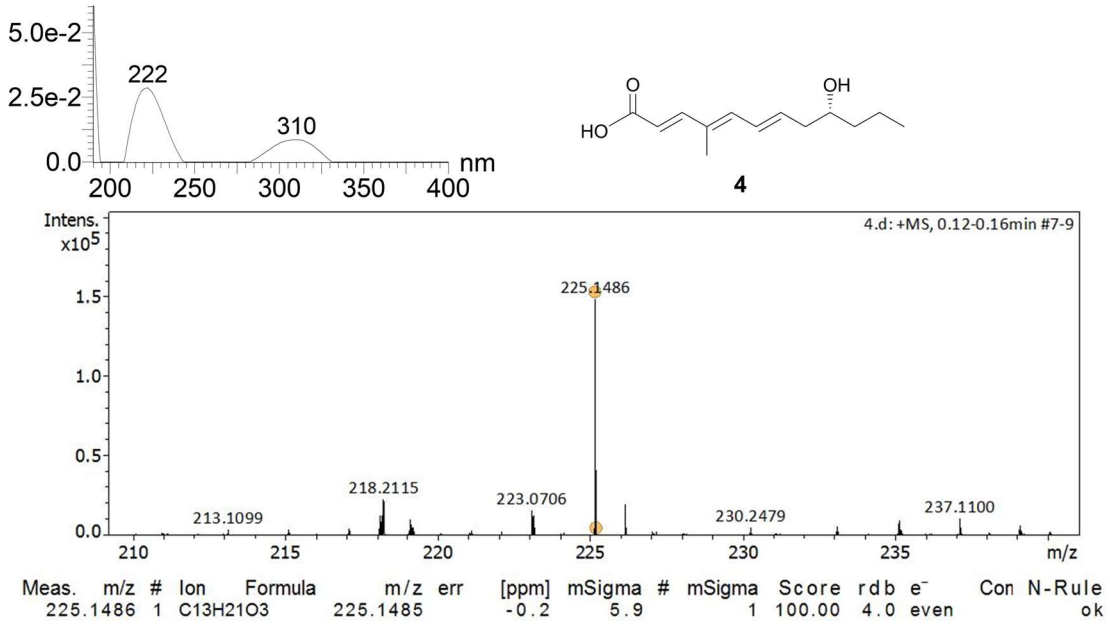


Figure S26. UV absorption and HRMS spectrum (positive ionization) analysis of compound **4**.

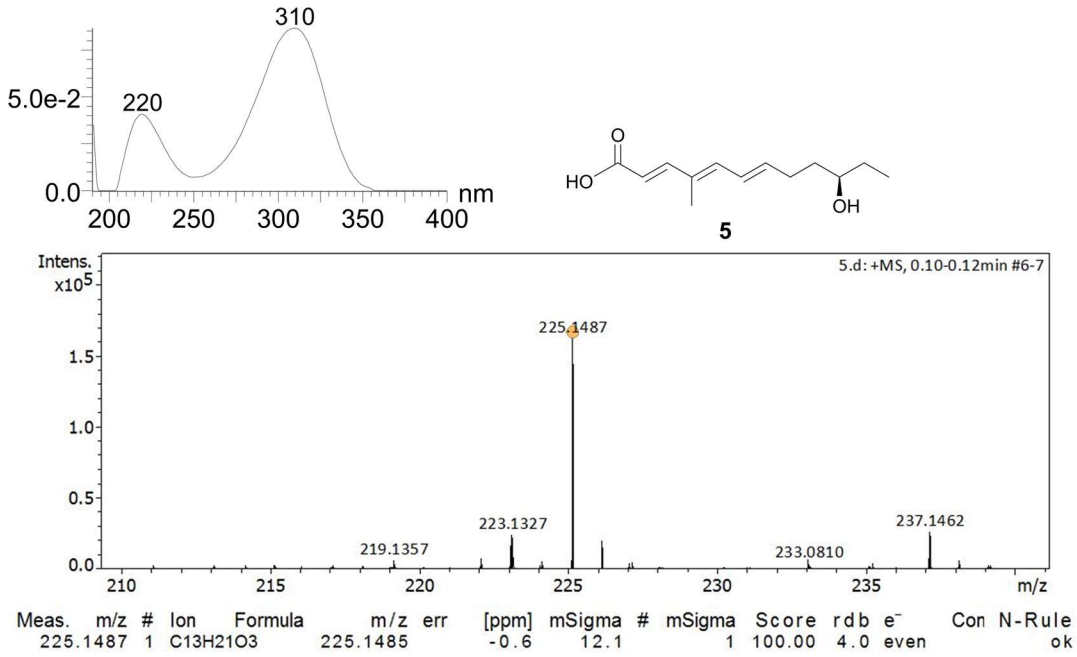


Figure S27. UV absorption and HRMS spectrum (positive ionization) analysis of compound **5**.

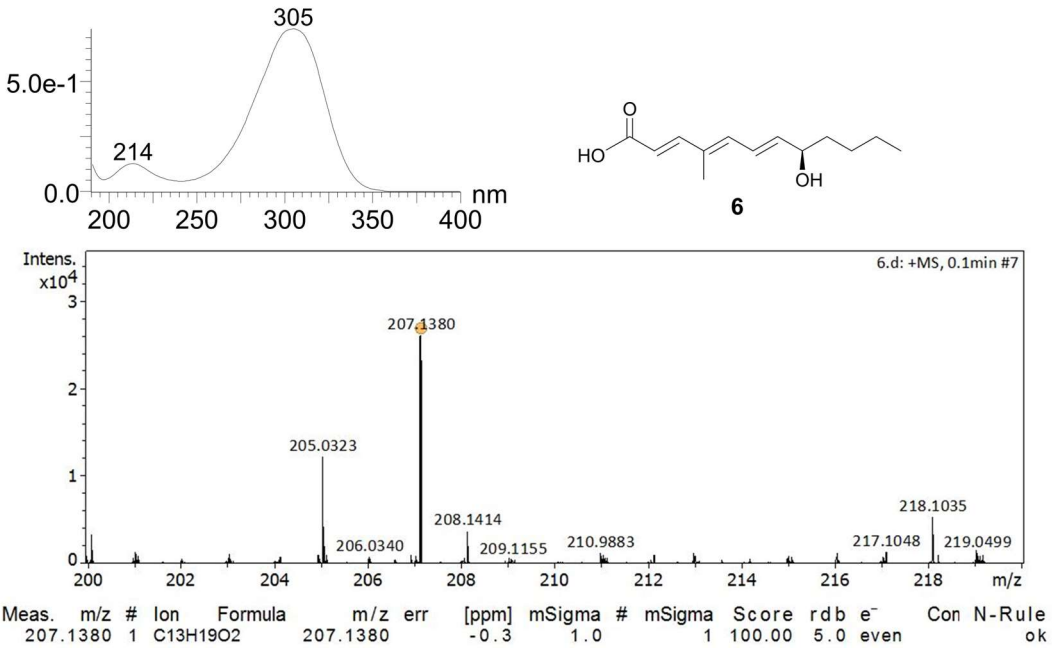


Figure S28. UV absorption and HRMS spectrum (positive ionization) analysis of compound **6**.

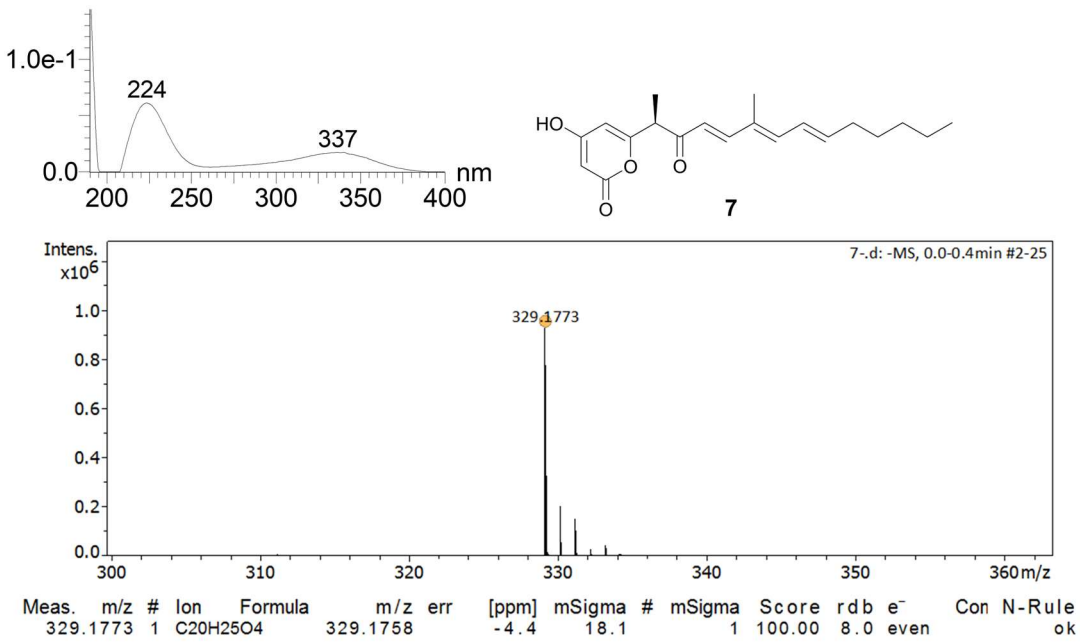


Figure S29. UV absorption and HRMS spectrum (negative ionization) analysis of compound **7**.

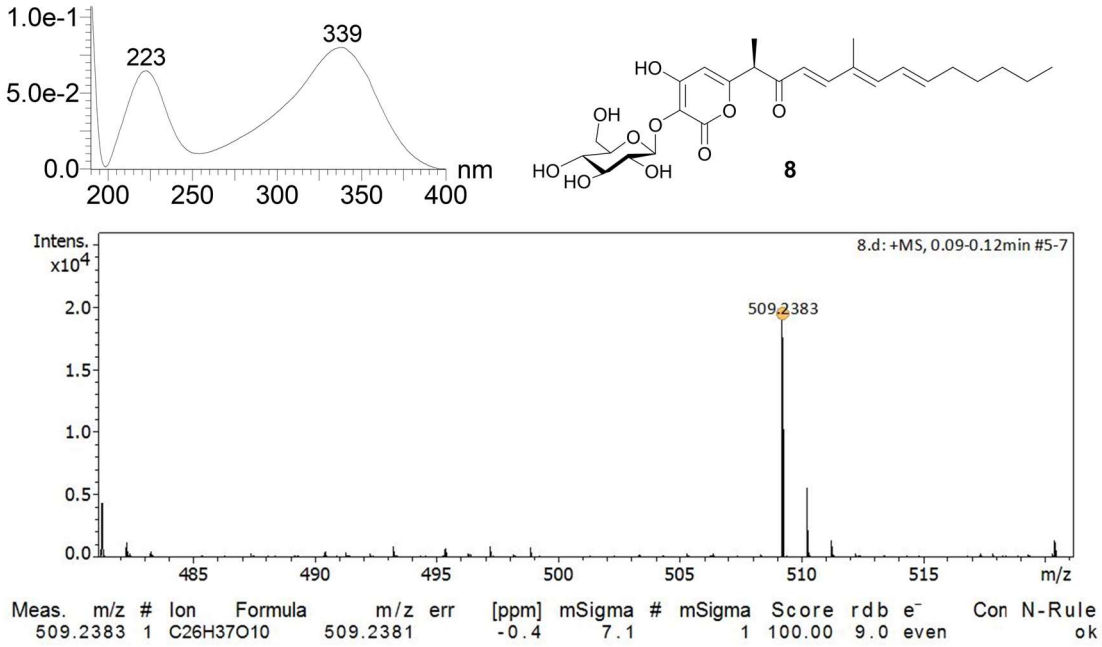


Figure S30. UV absorption and HRMS spectrum (positive ionization) analysis of compound **8**.

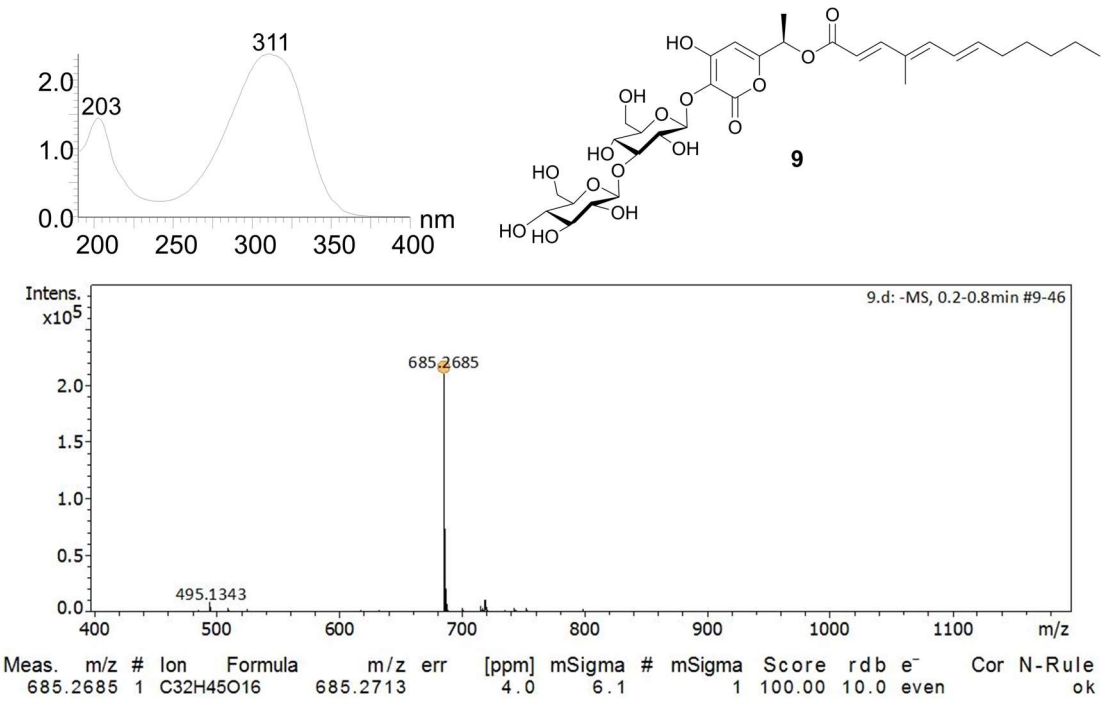


Figure S31. UV absorption and HRMS spectrum (negative ionization) analysis of compound **9**.

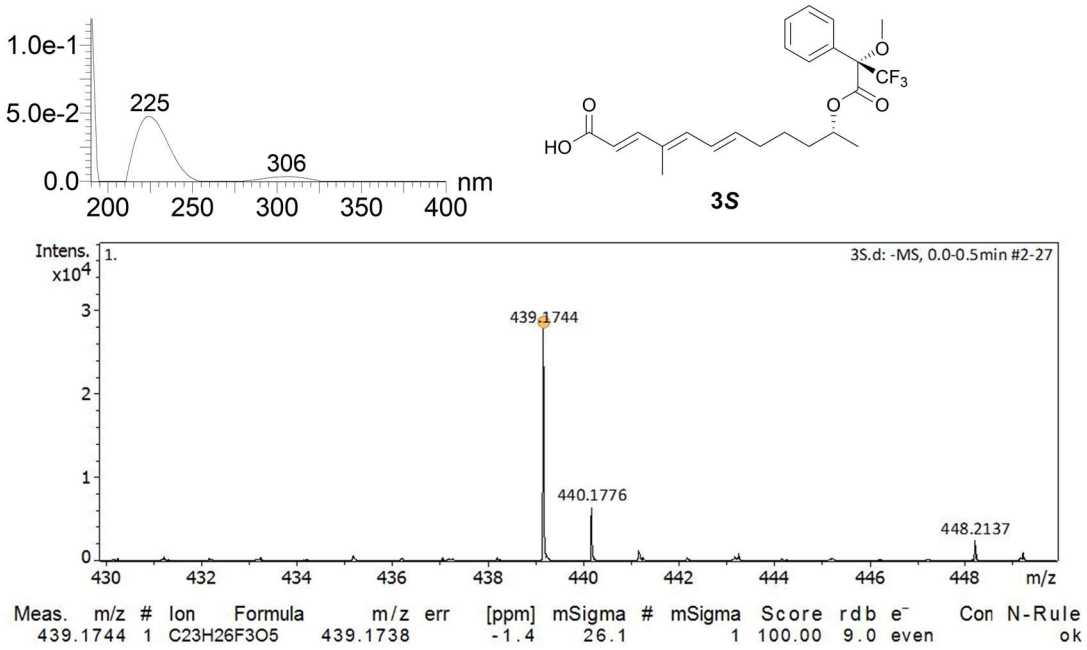


Figure S32. UV absorption and HRMS spectrum (negative ionization) analysis of compound **3*S***.

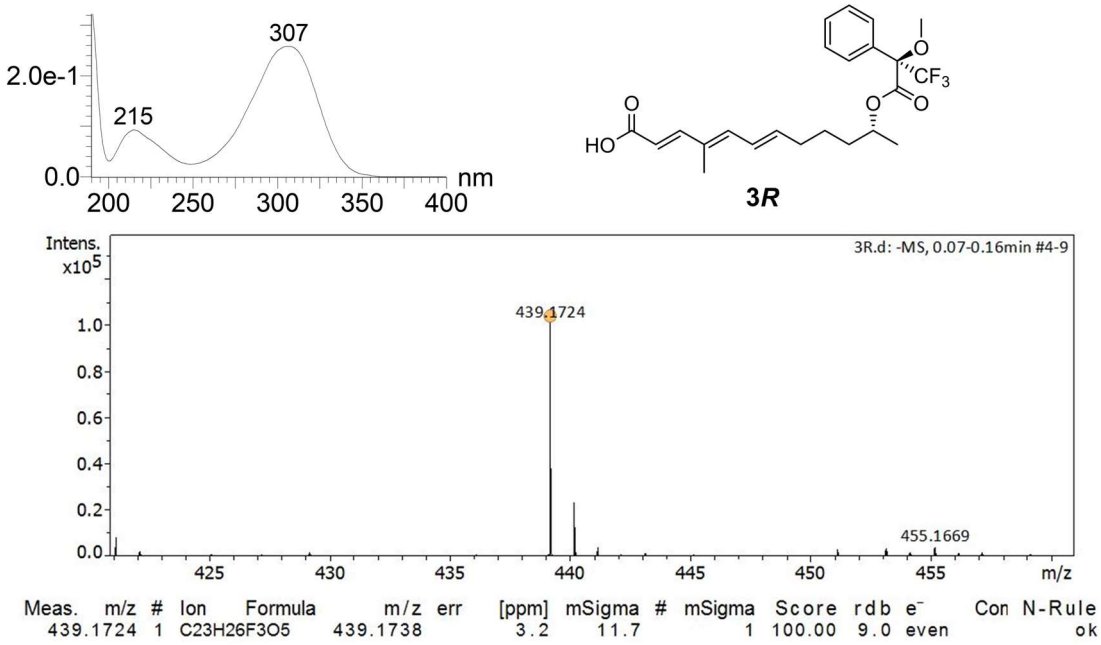


Figure S33. UV absorption and HRMS spectrum (negative ionization) analysis of compound **3*R***.

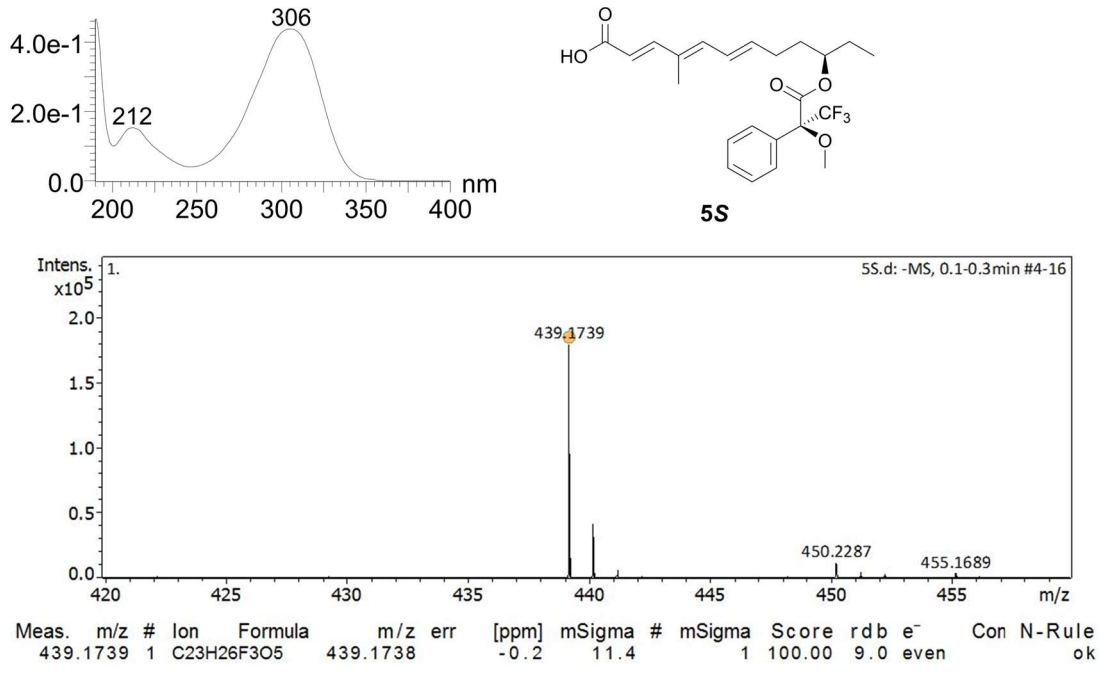


Figure S34. UV absorption and HRMS spectrum (negative ionization) analysis of compound **5*S***.

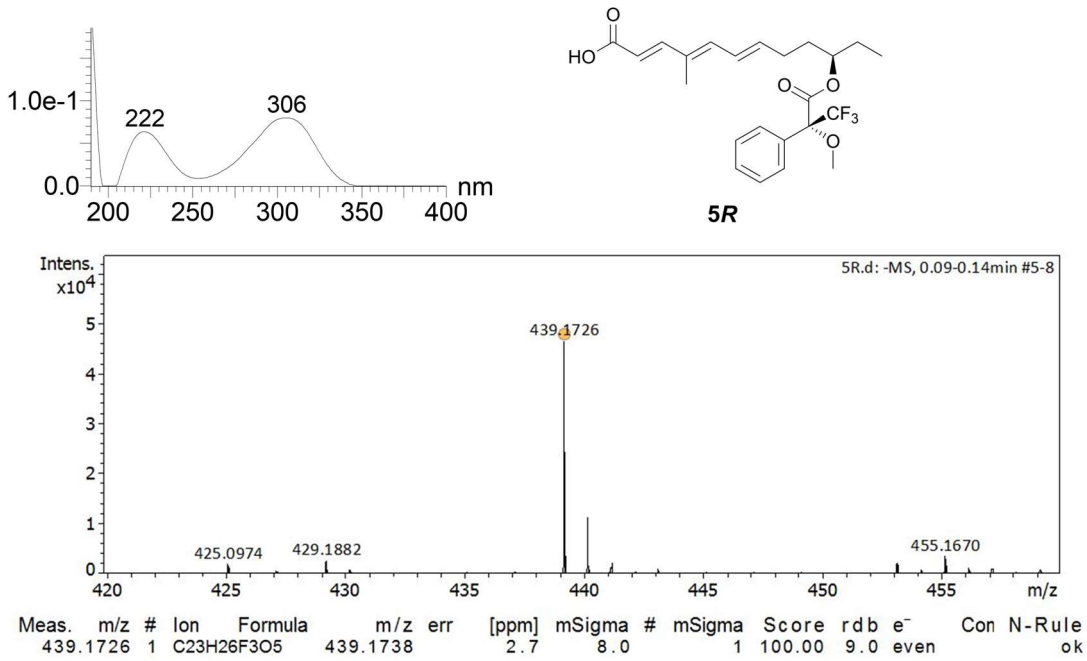


Figure S35. UV absorption and HRMS spectrum (negative ionization) analysis of compound **5*R***.

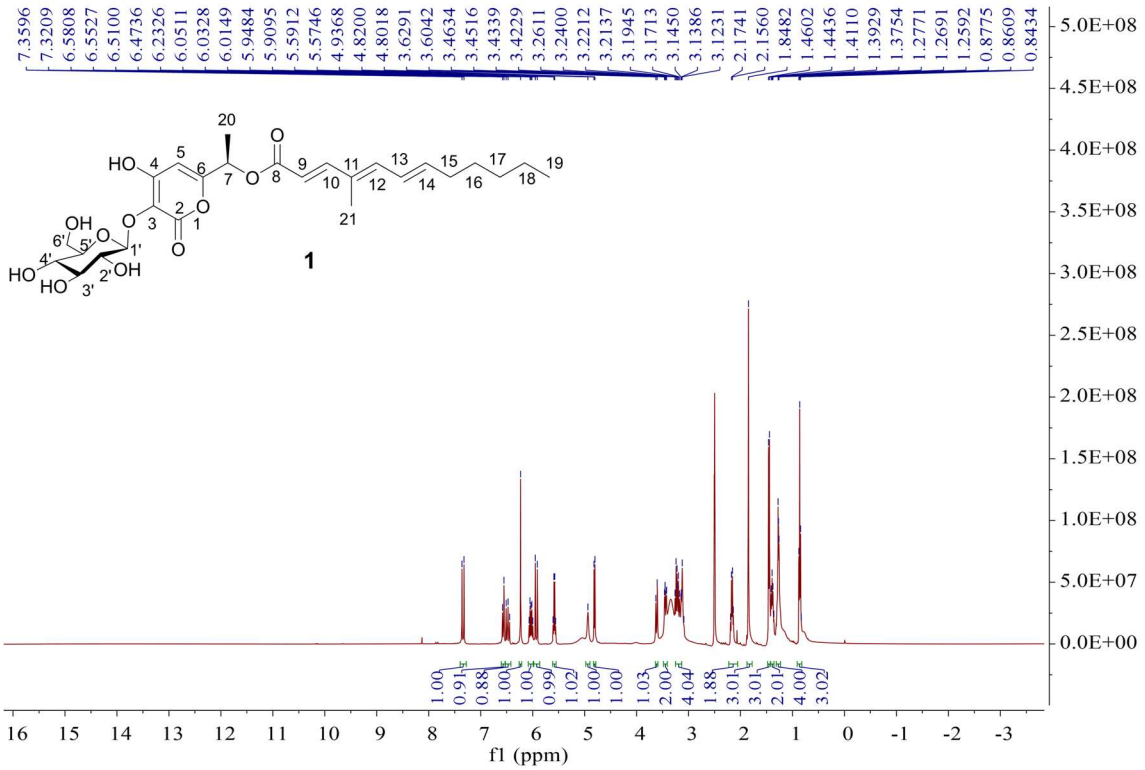


Figure S36. ^{1}H NMR spectrum of compound **1** in DMSO-d_6 (400 MHz).

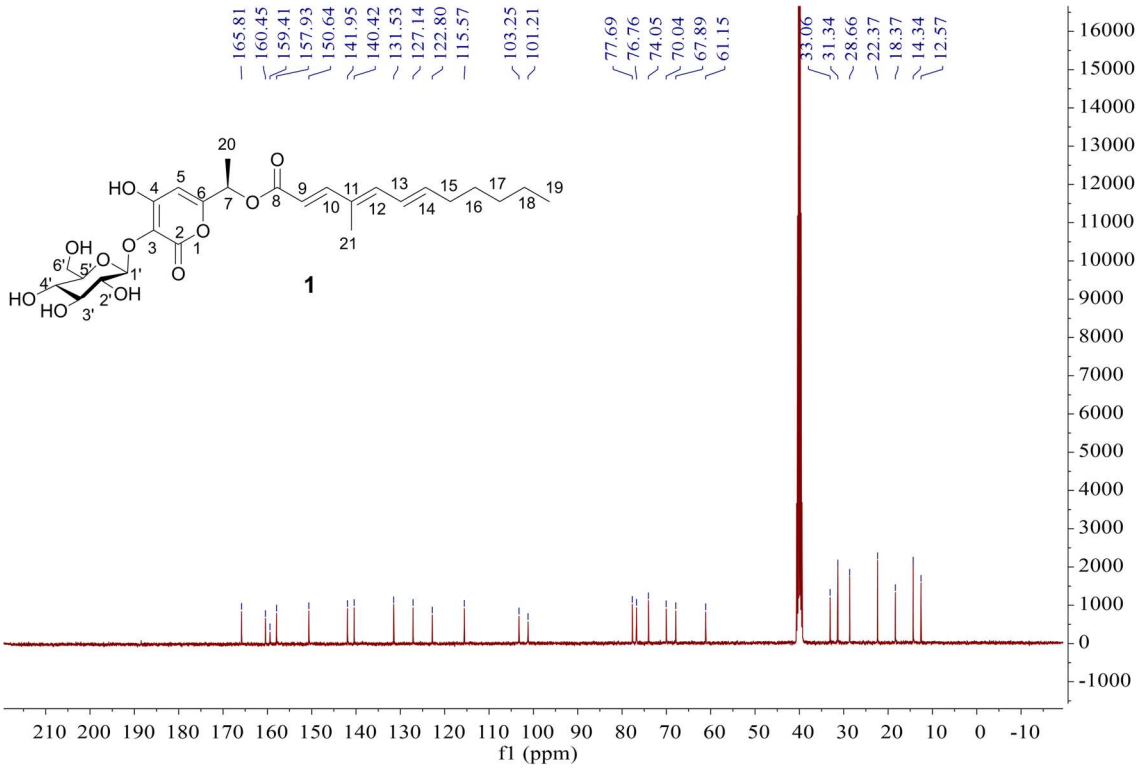


Figure S37. ^{13}C NMR spectrum of compound **1** in DMSO-d_6 (100 MHz).

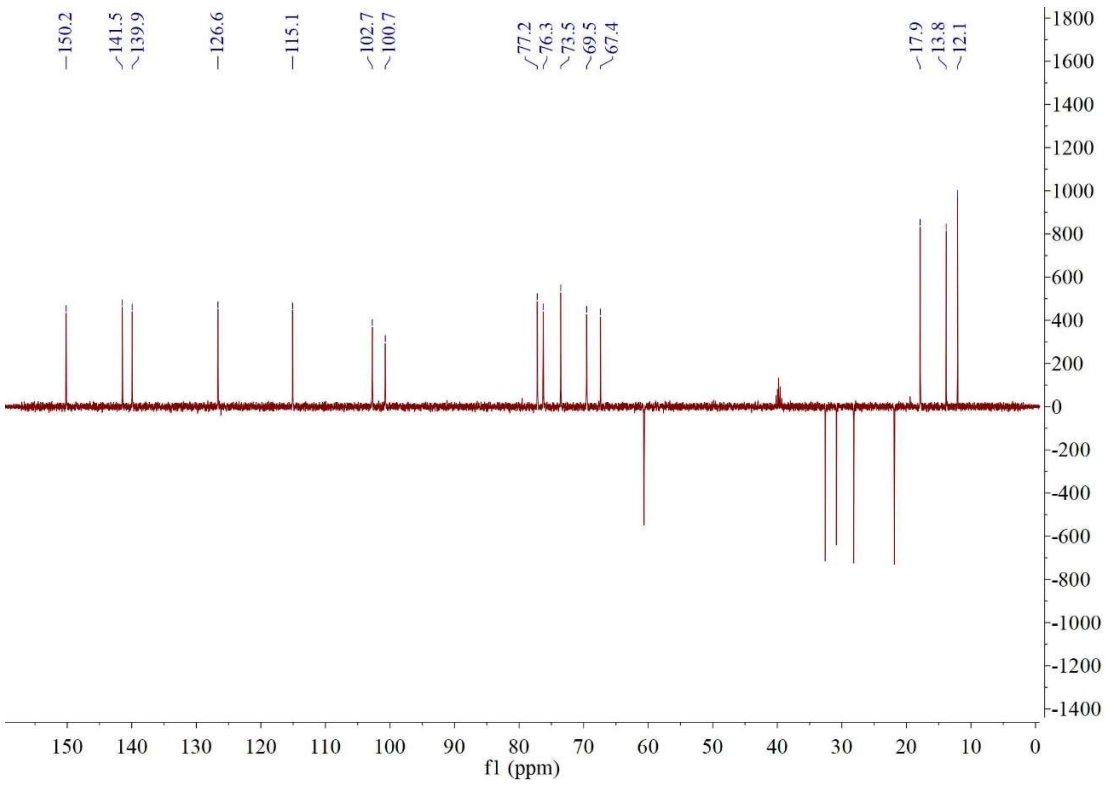


Figure S38. DEPT-135° spectrum of compound **1** in DMSO-d_6.

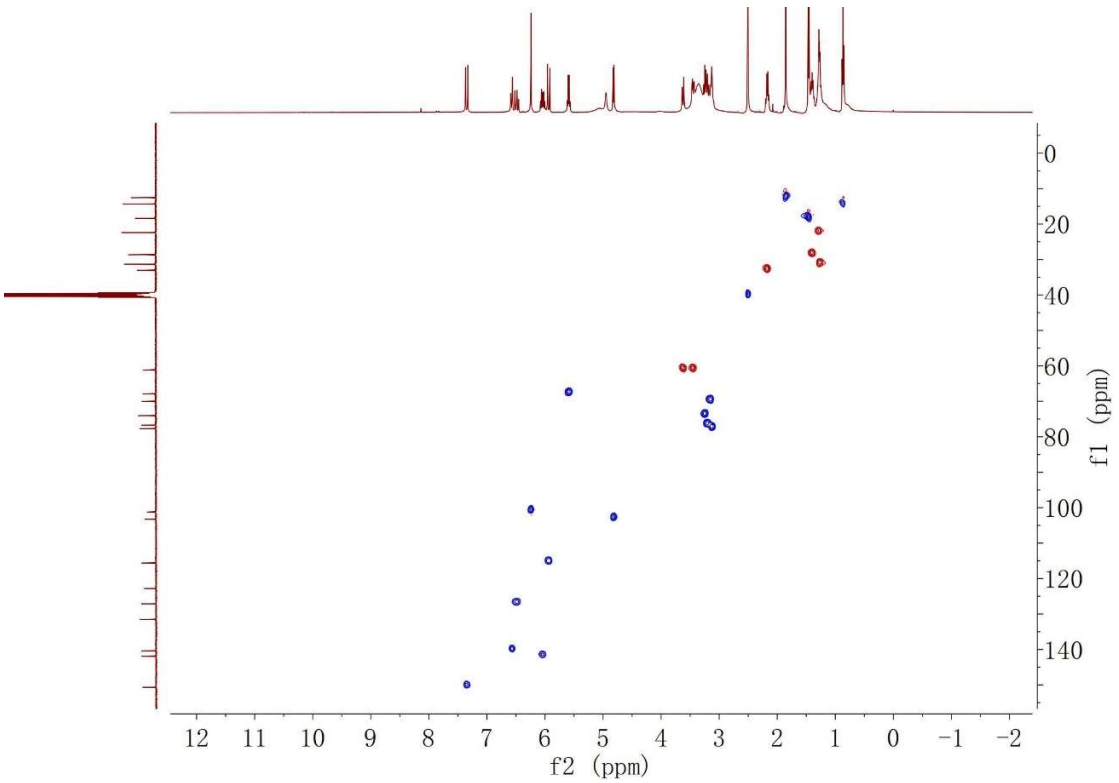


Figure S39. HSQC spectrum of compound **1** in DMSO-d_6.

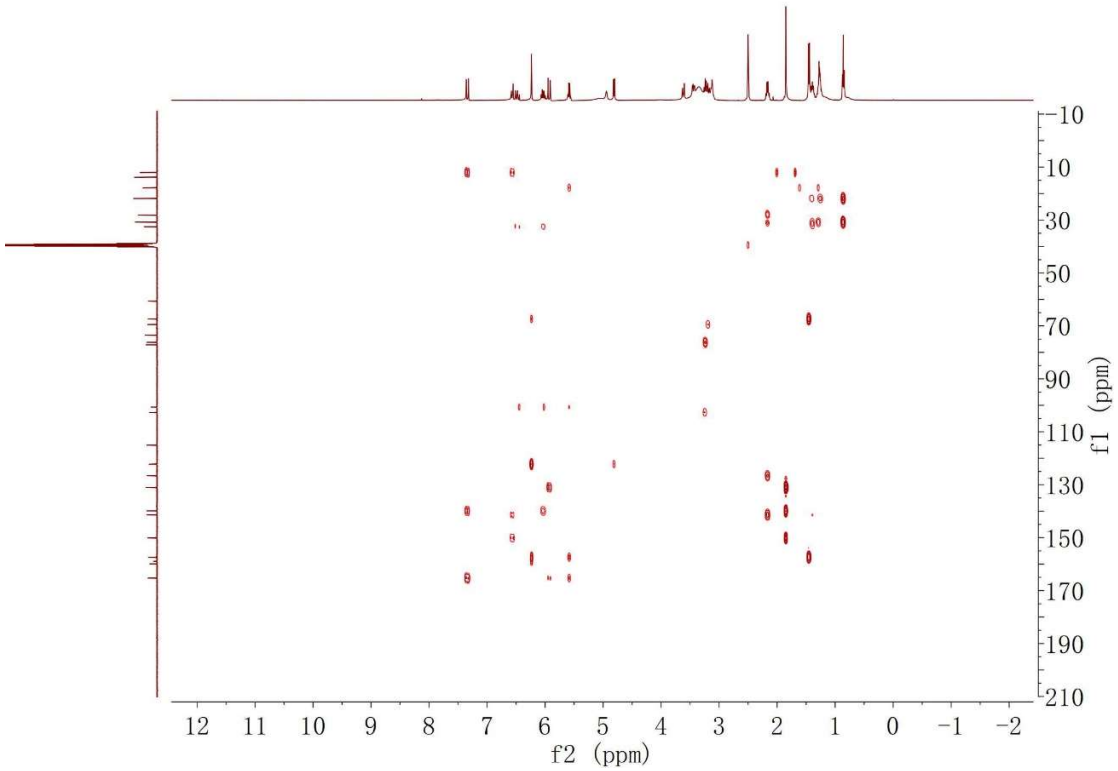


Figure S40. HMBC spectrum of compound **1** in DMSO-d_6**.**

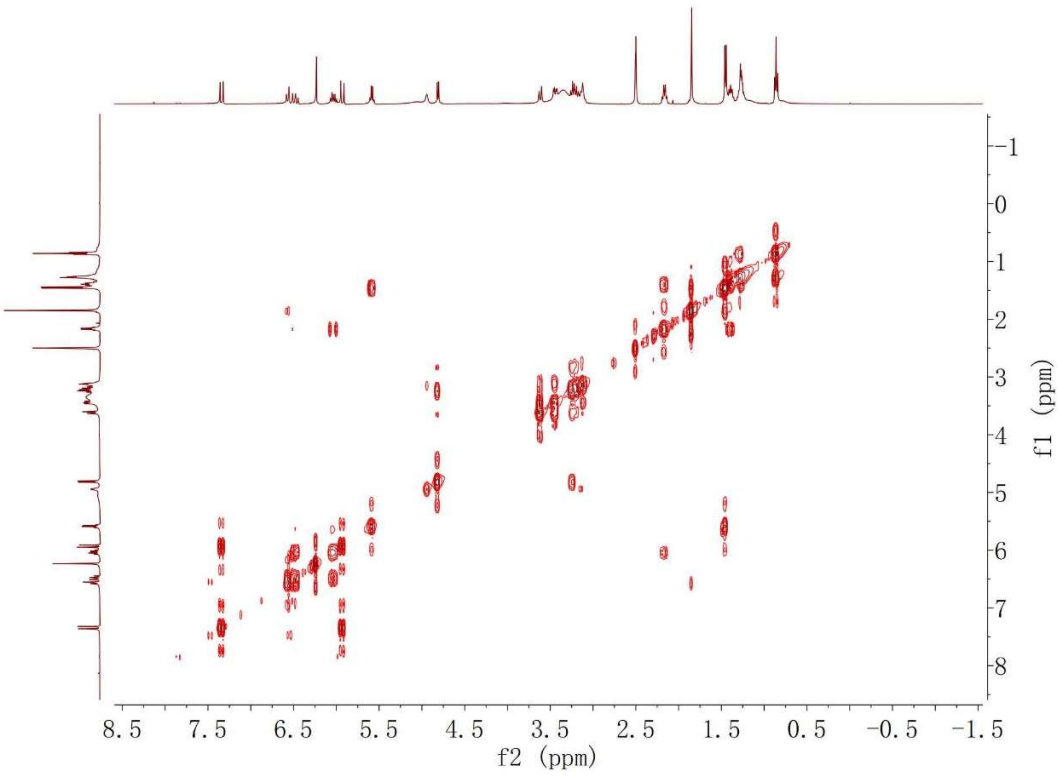


Figure S41. ^{1}H-^{1}H COSY spectrum of compound **1** in DMSO-d_6.

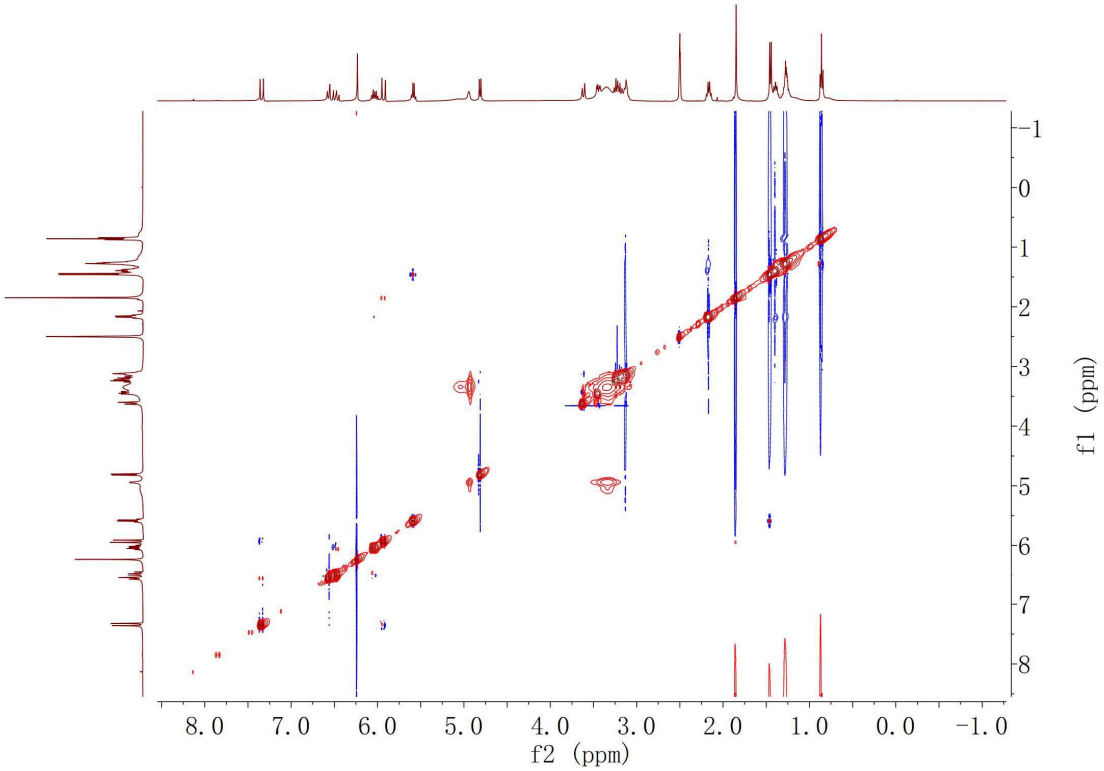


Figure S42. ^{1}H-^{1}H NOESY spectrum of compound **1** in DMSO-d_6.

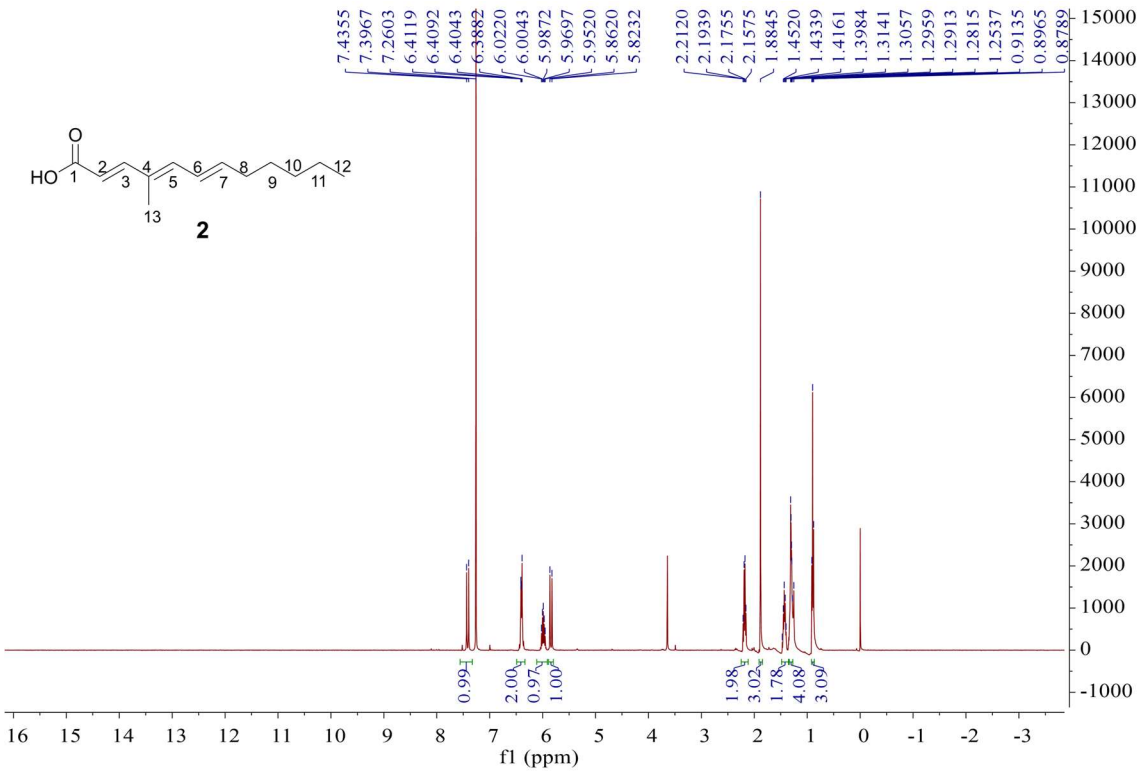


Figure S43. ^{1}H NMR spectrum of compound **2** in $CDCl_3$ (400 MHz).

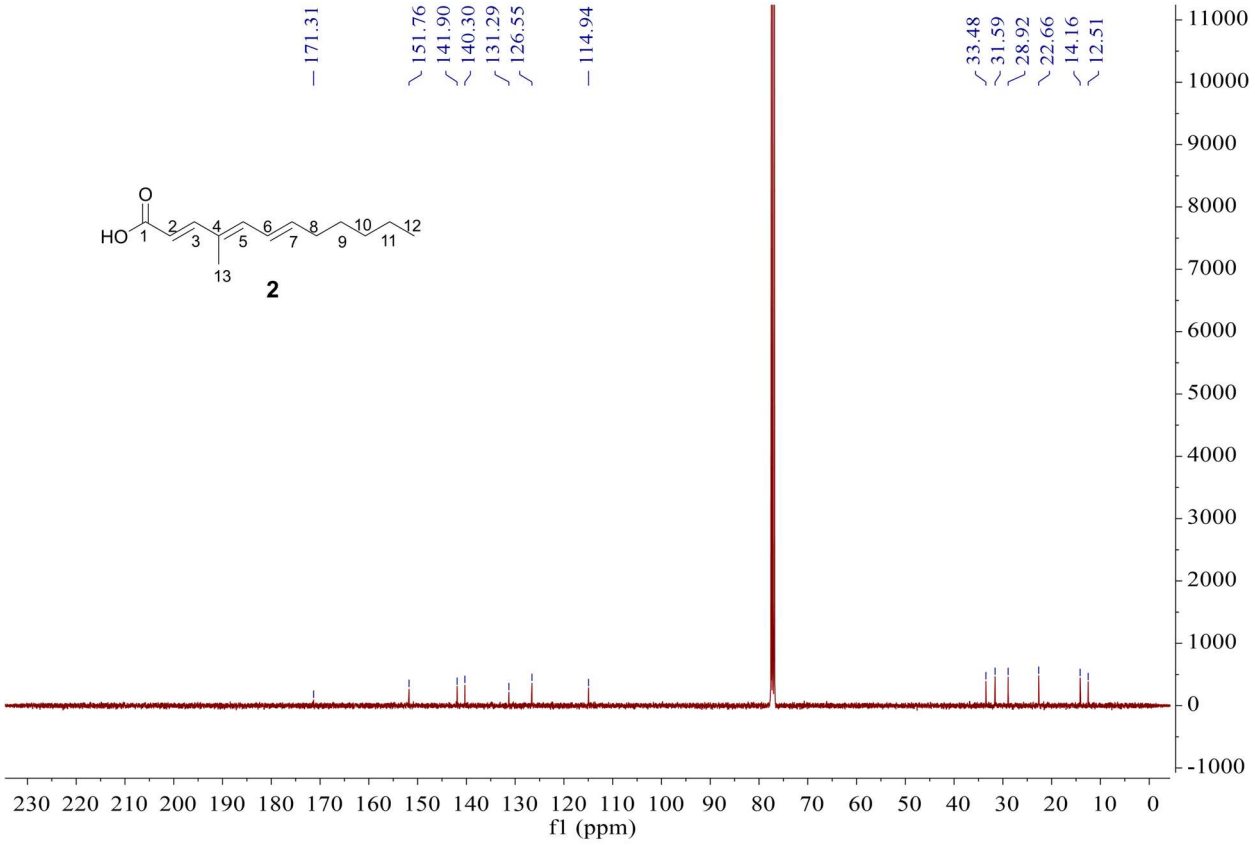


Figure S44. ^{13}C NMR spectrum of compound **2** in $CDCl_3$ (100 MHz).

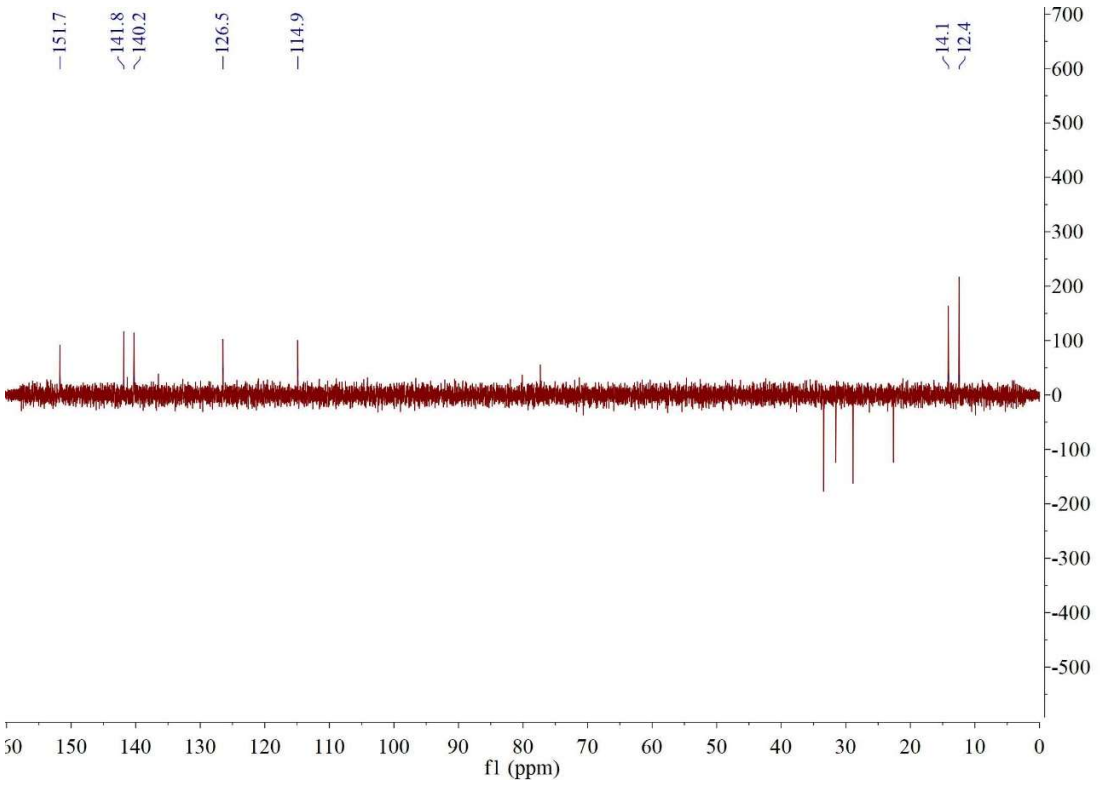


Figure S45. DEPT-135° spectrum of compound **2** in $CDCl_3$.

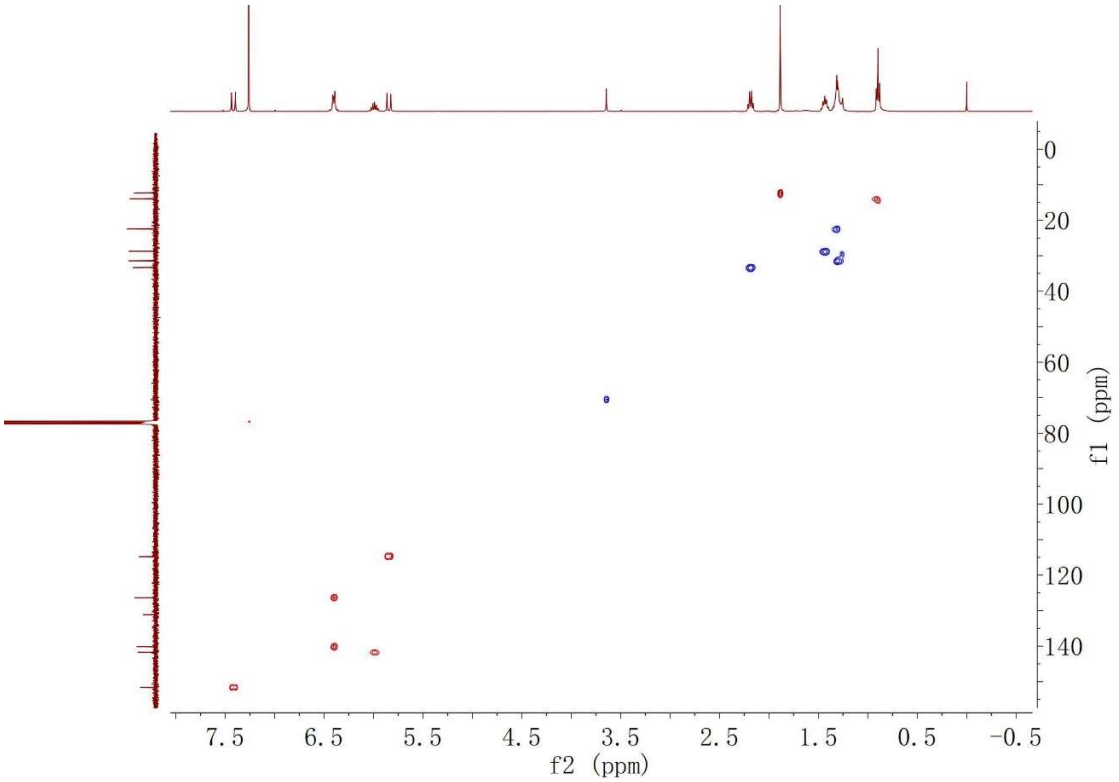


Figure S46. HSQC spectrum of compound **2** in $CDCl_3$.

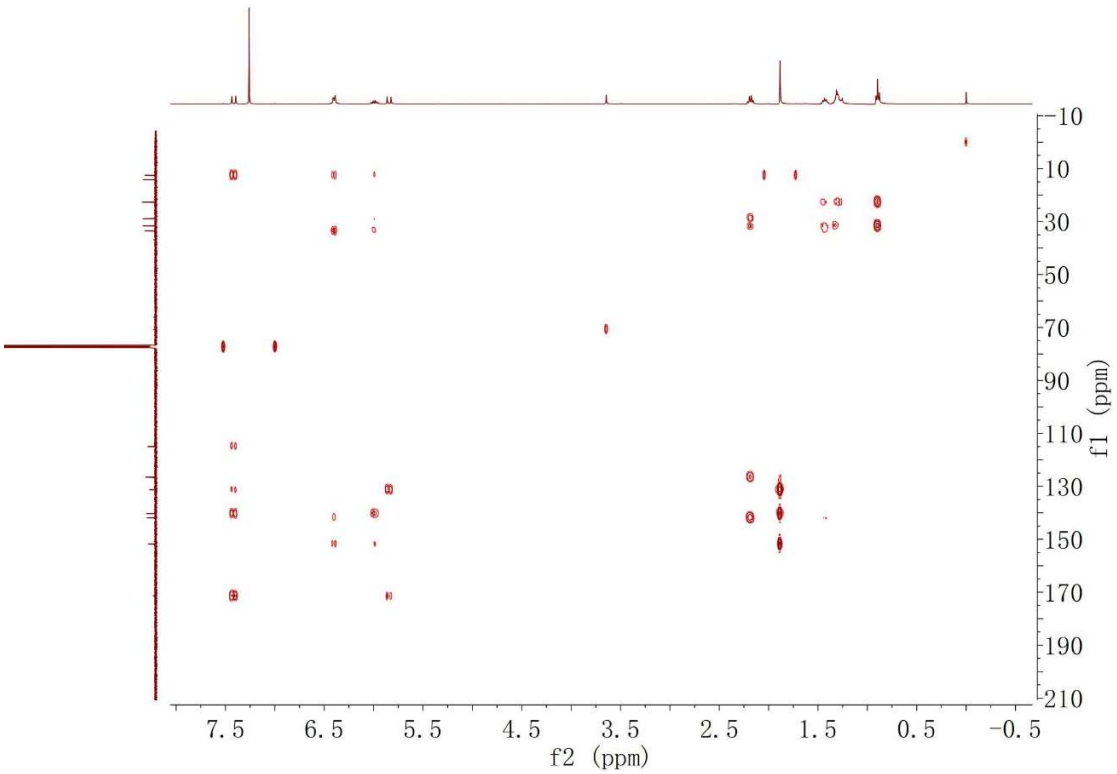


Figure S47. HMBC spectrum of compound **2** in $CDCl_3$.

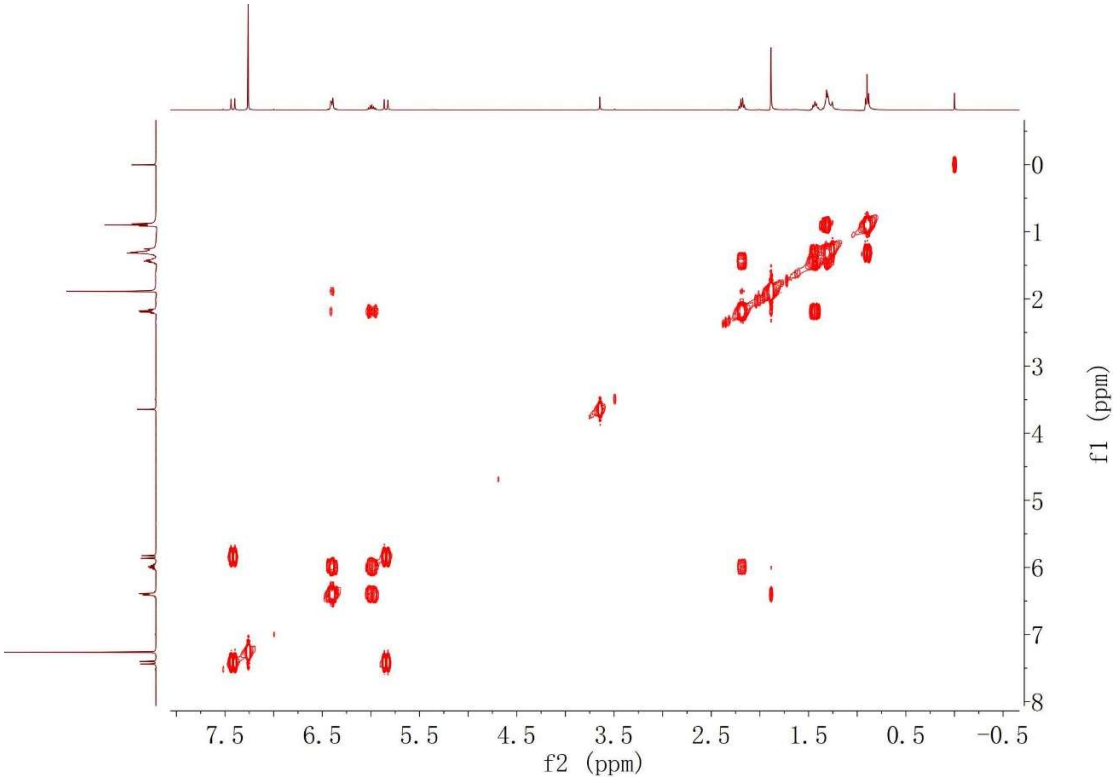


Figure S48. ^{1}H-^{1}H COSY spectrum of compound **2** in $CDCl_3$.

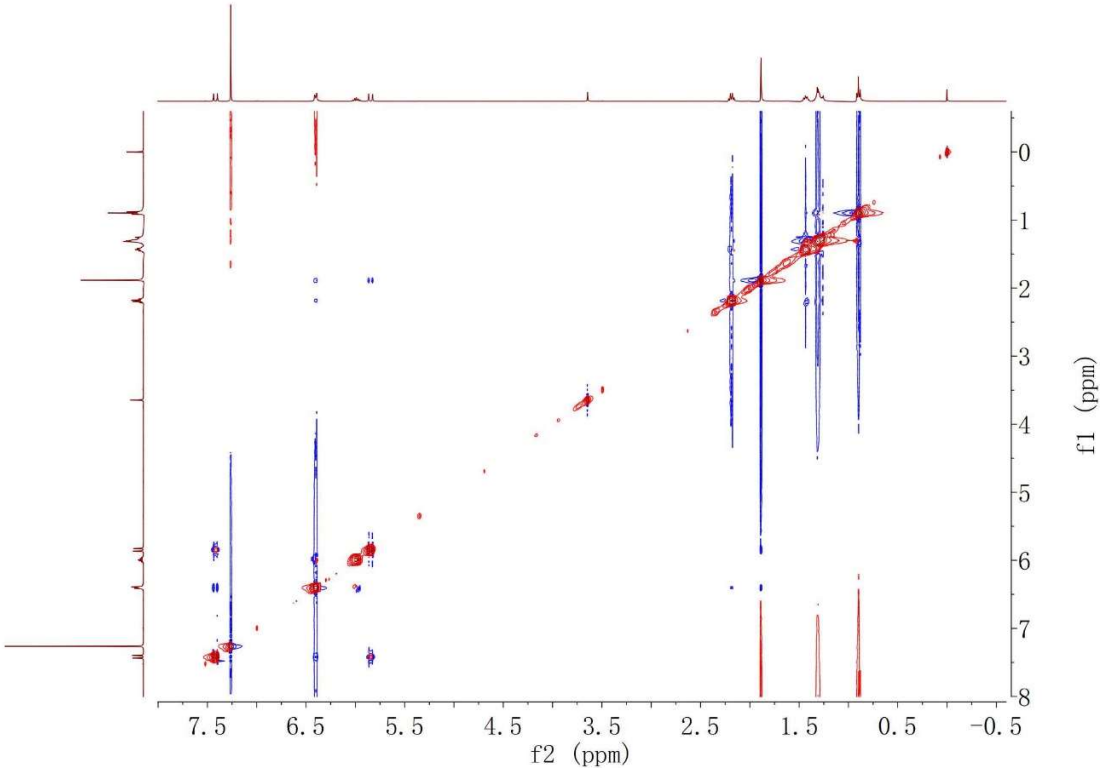


Figure S49. ^{1}H-^{1}H NOESY spectrum of compound **2** in $CDCl_3$.

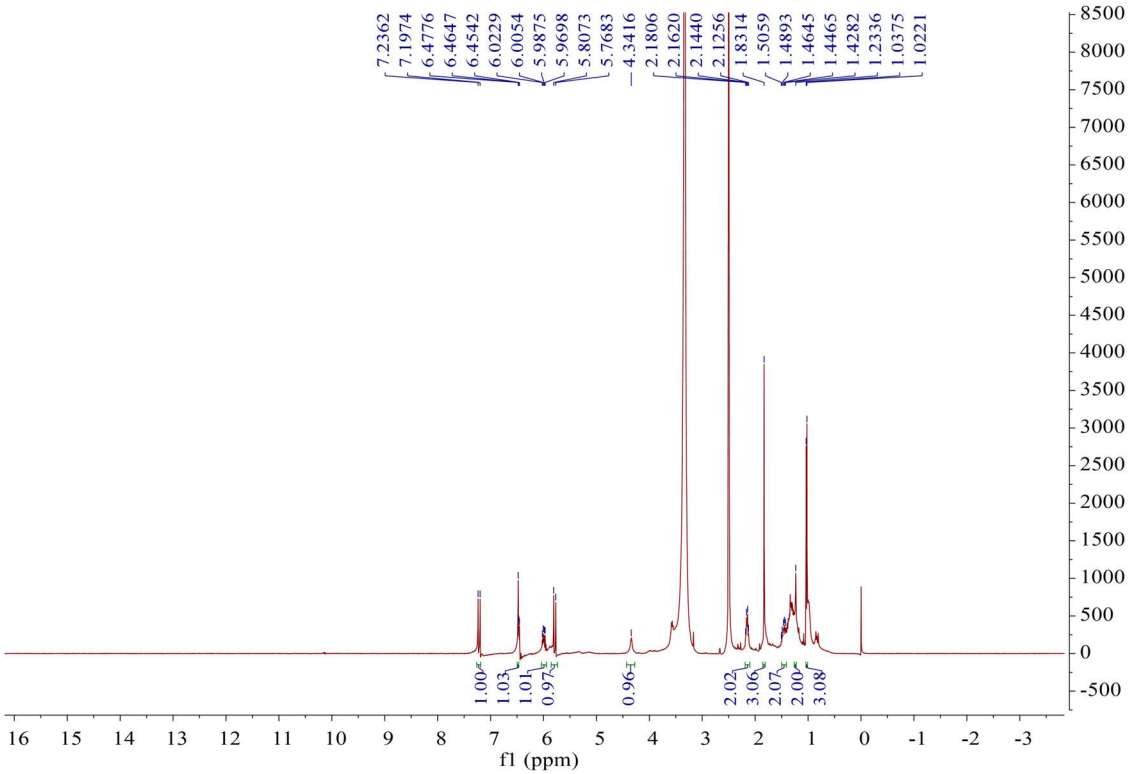


Figure S50. ^{1}H NMR spectrum of compound **3** in DMSO-d_6 (400 MHz).

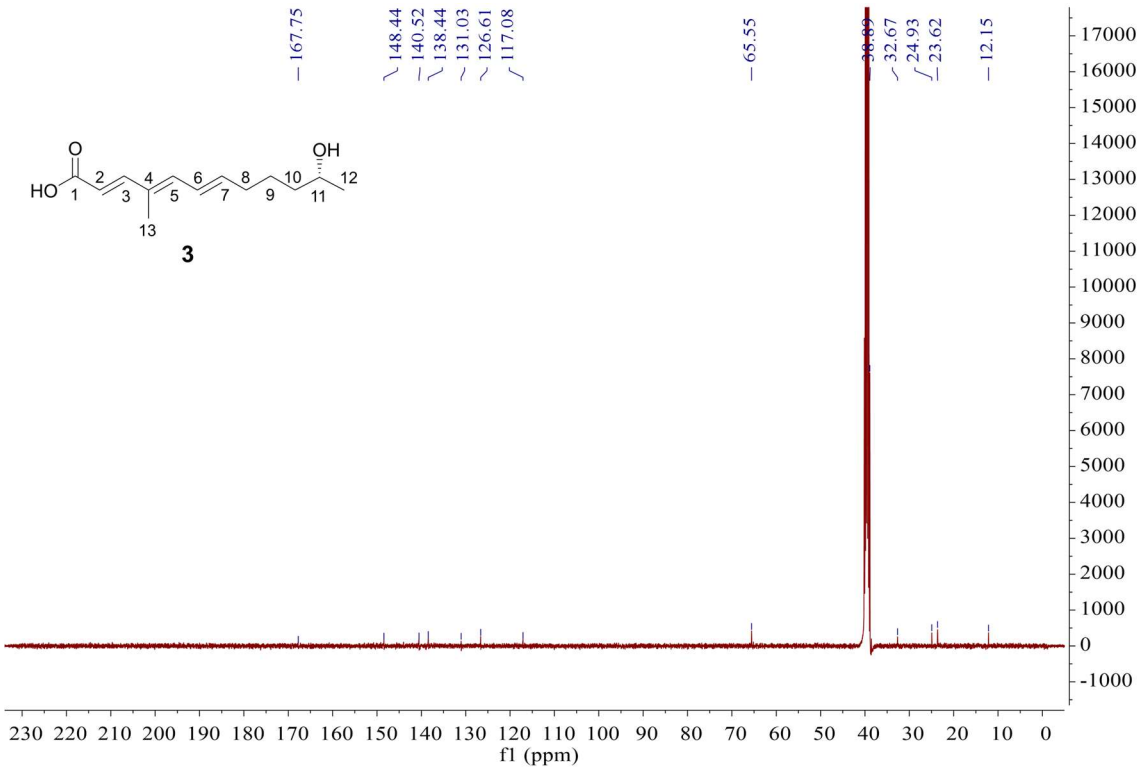


Figure S51. ^{13}C NMR spectrum of compound **3** in DMSO-d_6 (100 MHz).

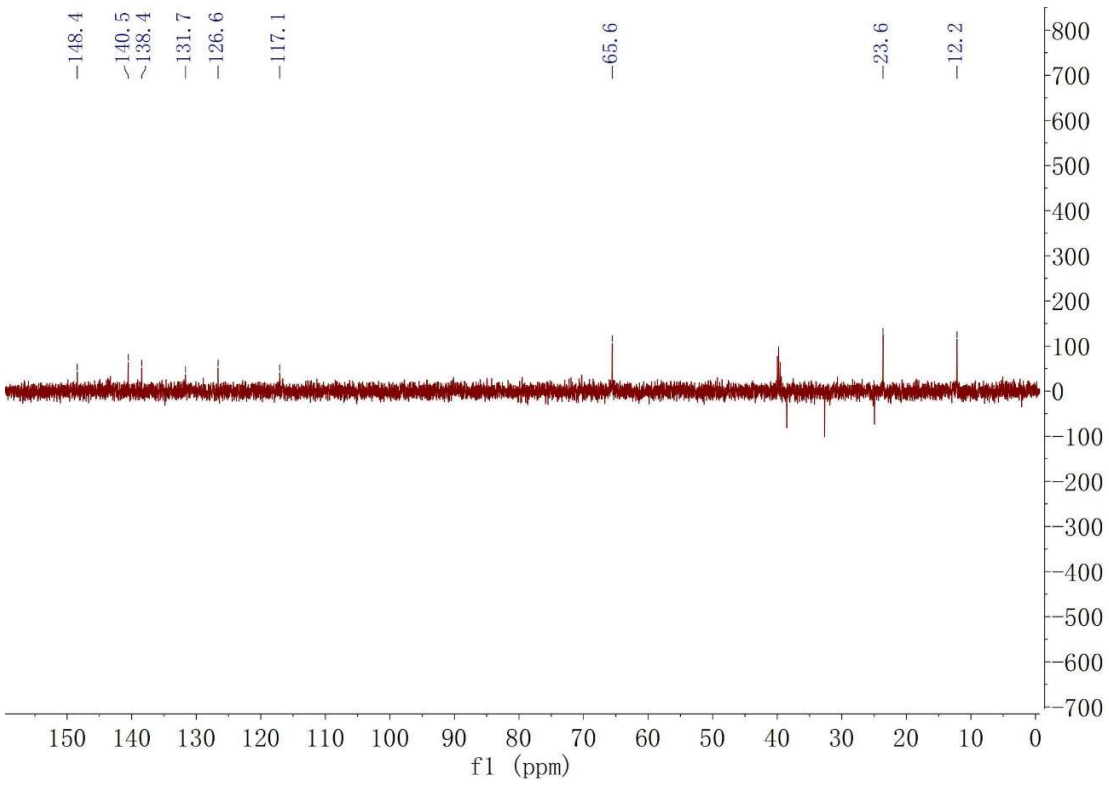


Figure S52. DEPT-135° spectrum of compound **3** in DMSO-d_6.

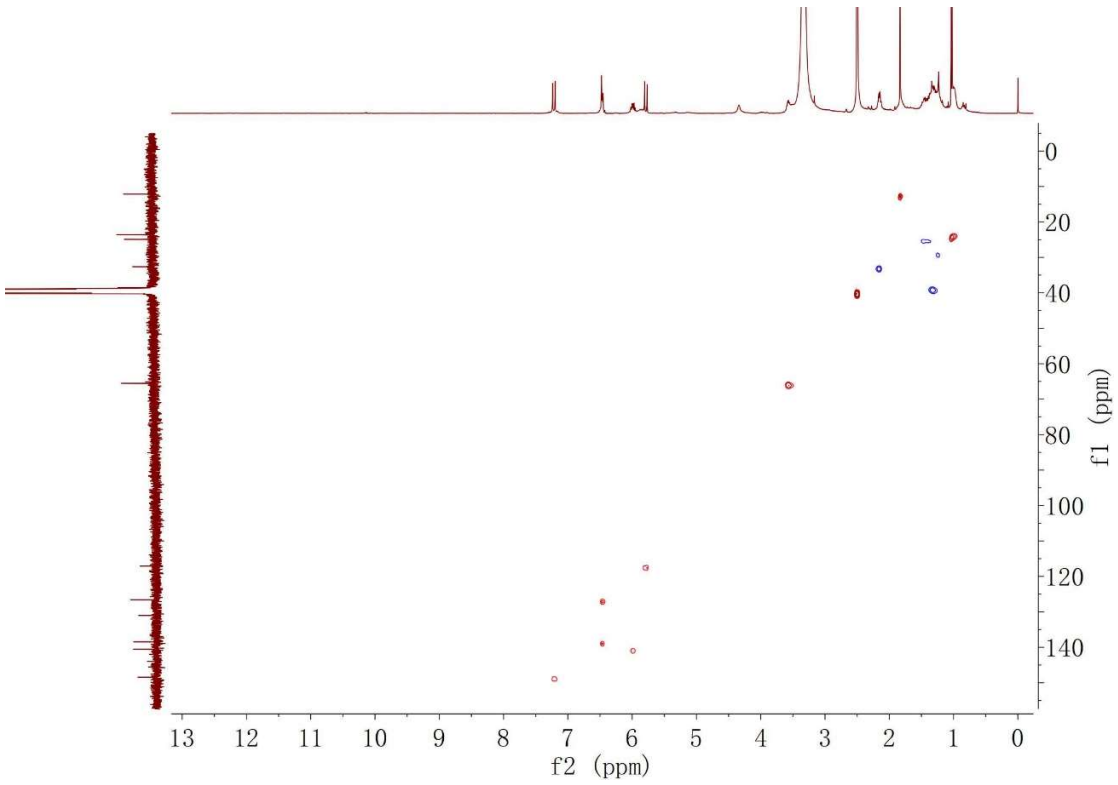


Figure S53. HSQC spectrum of compound **3** in DMSO-d_6.

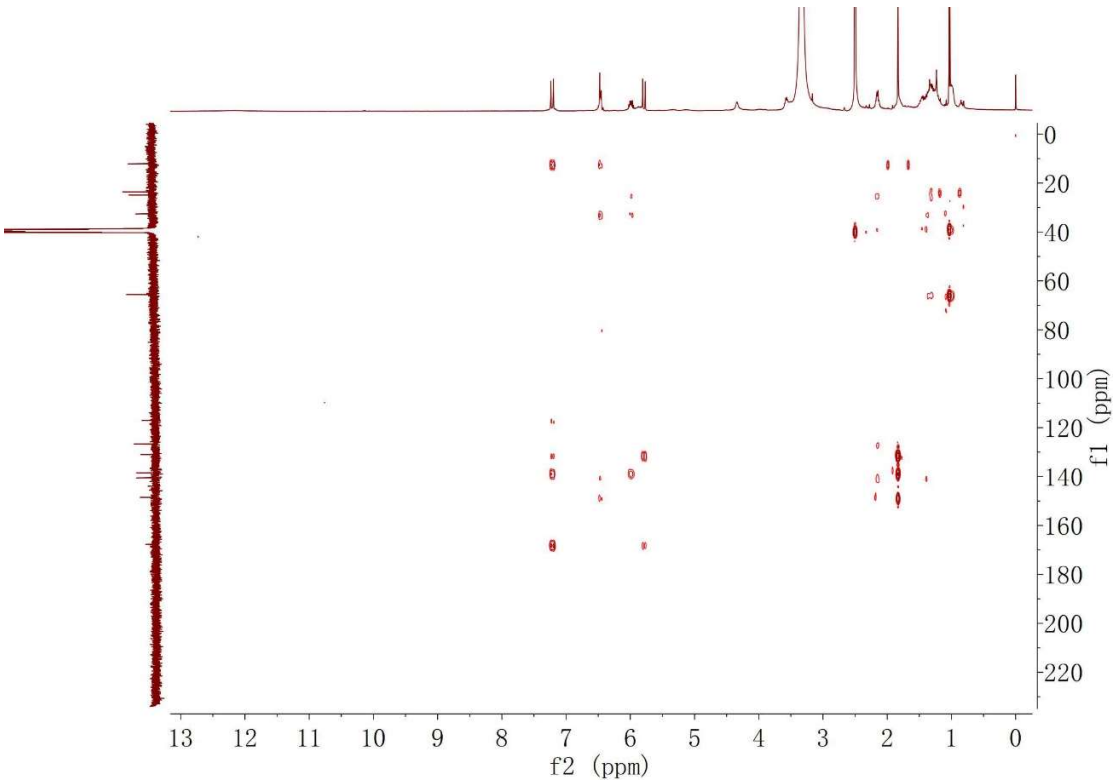


Figure S54. HMBC spectrum of compound **3** in DMSO-d_6.

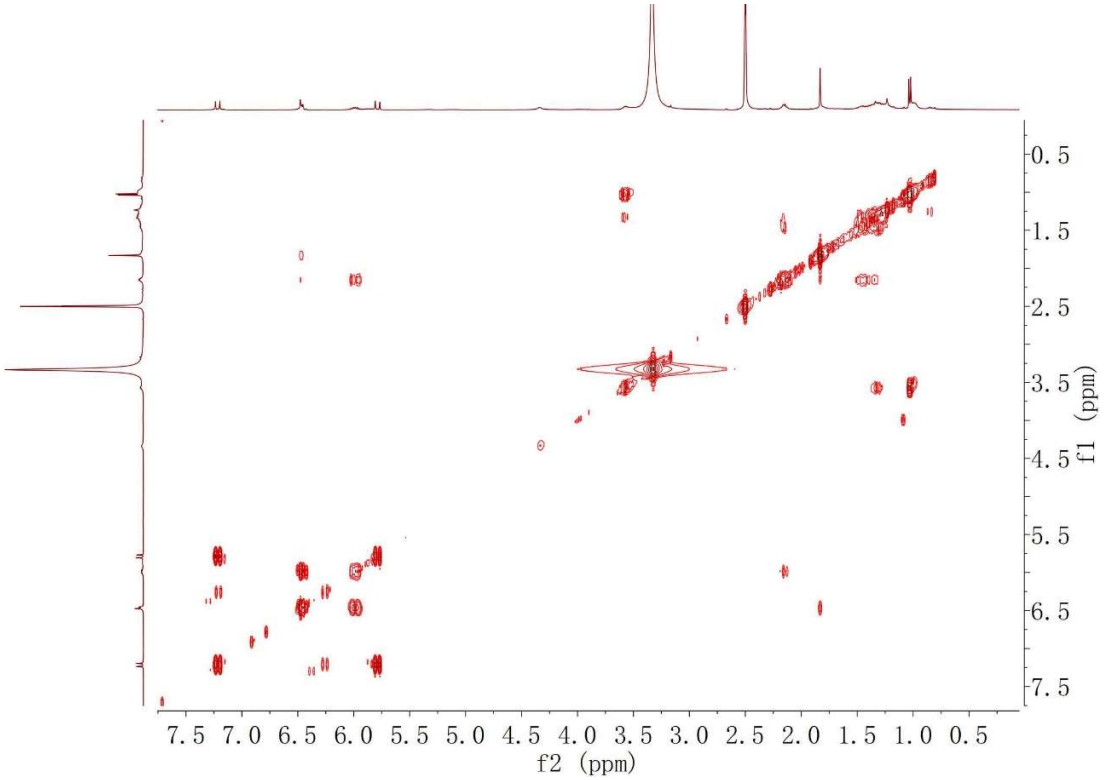


Figure S55. ^{1}H-^{1}H COSY spectrum of compound **3** in DMSO-d_6.

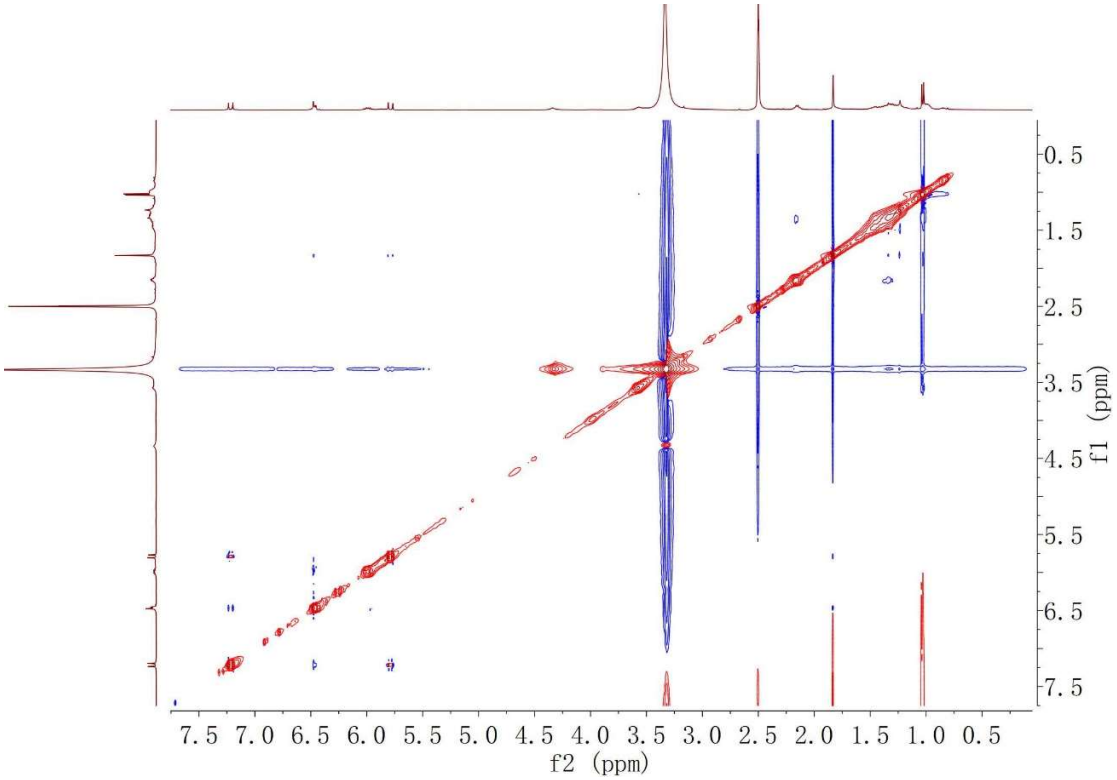


Figure S56. ^{1}H-^{1}H NOESY spectrum of compound **3** in DMSO-d_6.

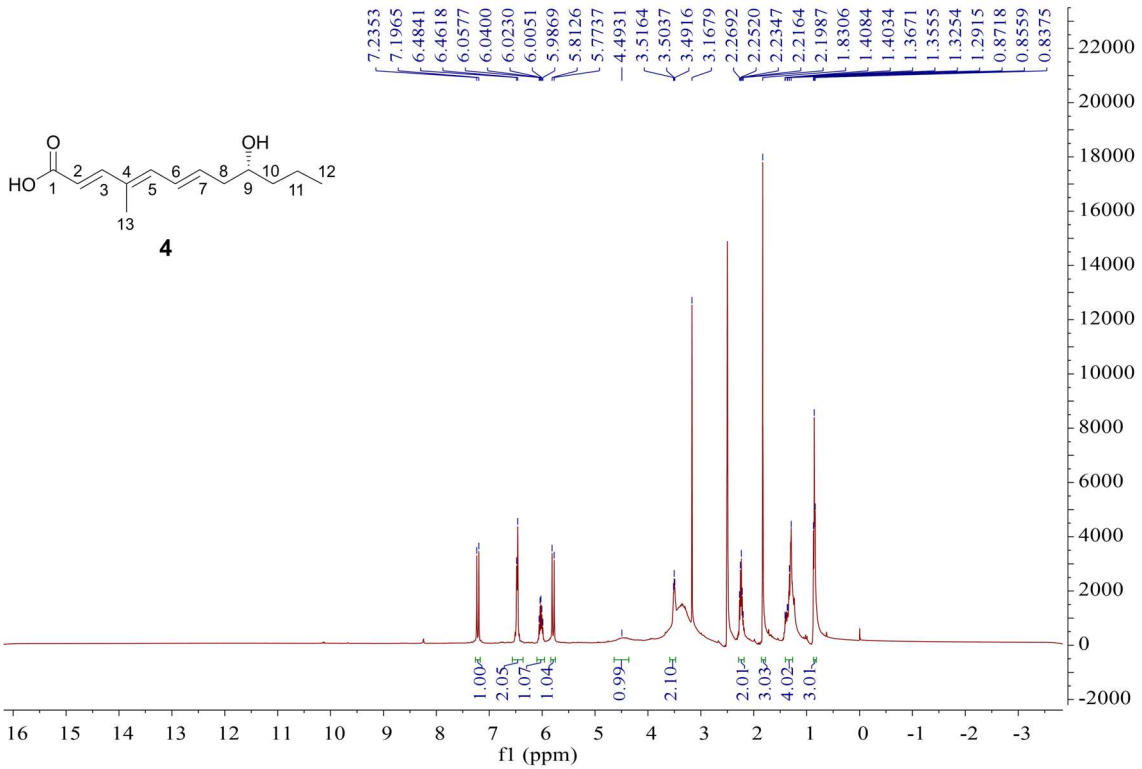


Figure S57. [1]H NMR spectrum of compound **4** in DMSO-d_6 (400 MHz).

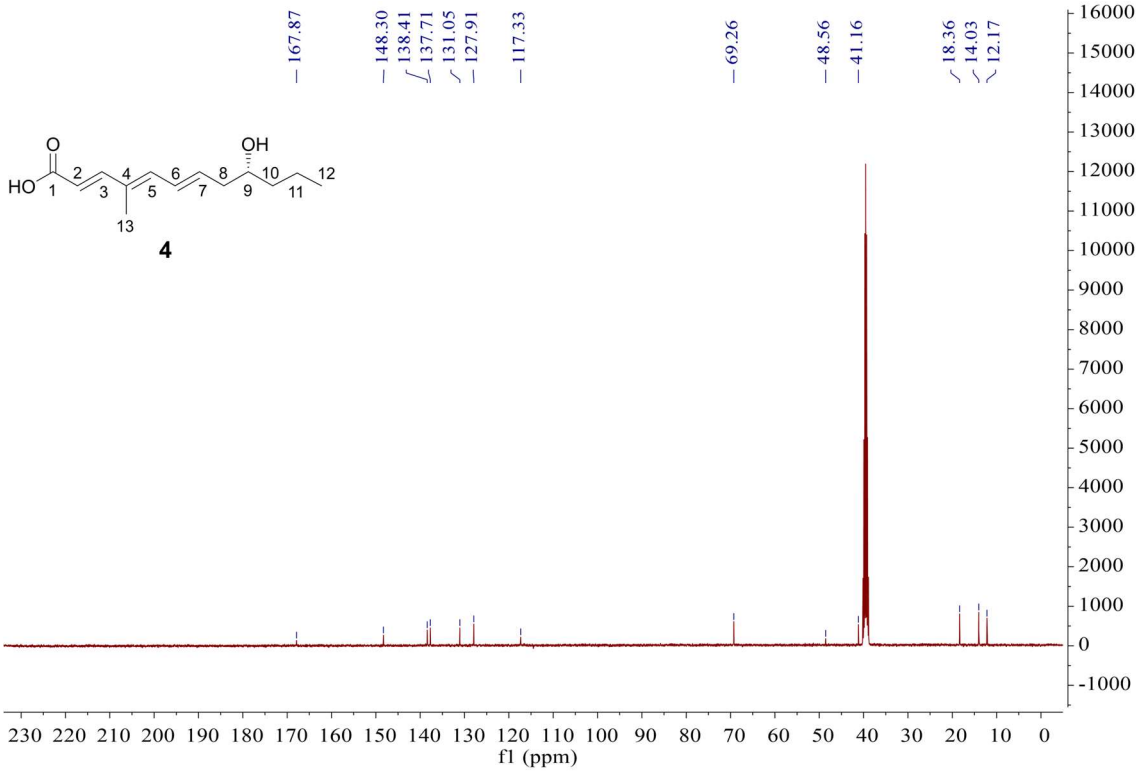


Figure S58. ^{13}C NMR spectrum of compound **4** in DMSO-d_6 (100 MHz).

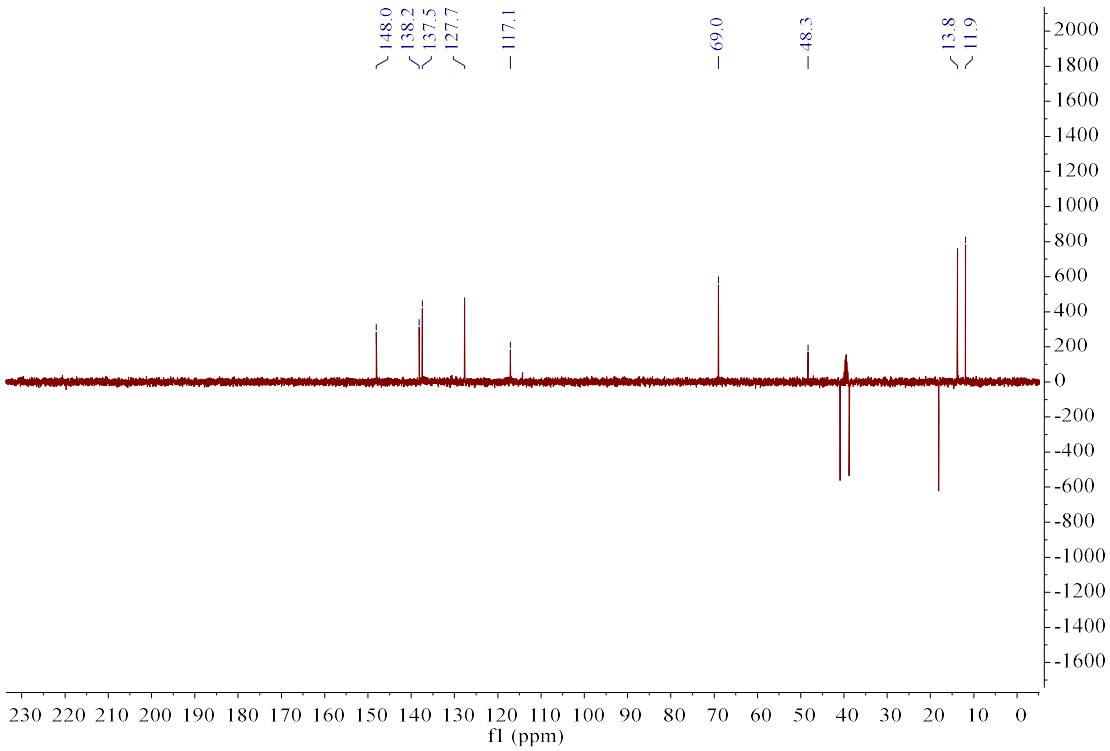


Figure S59. DEPT-135° spectrum of compound **4** in DMSO-d_6.

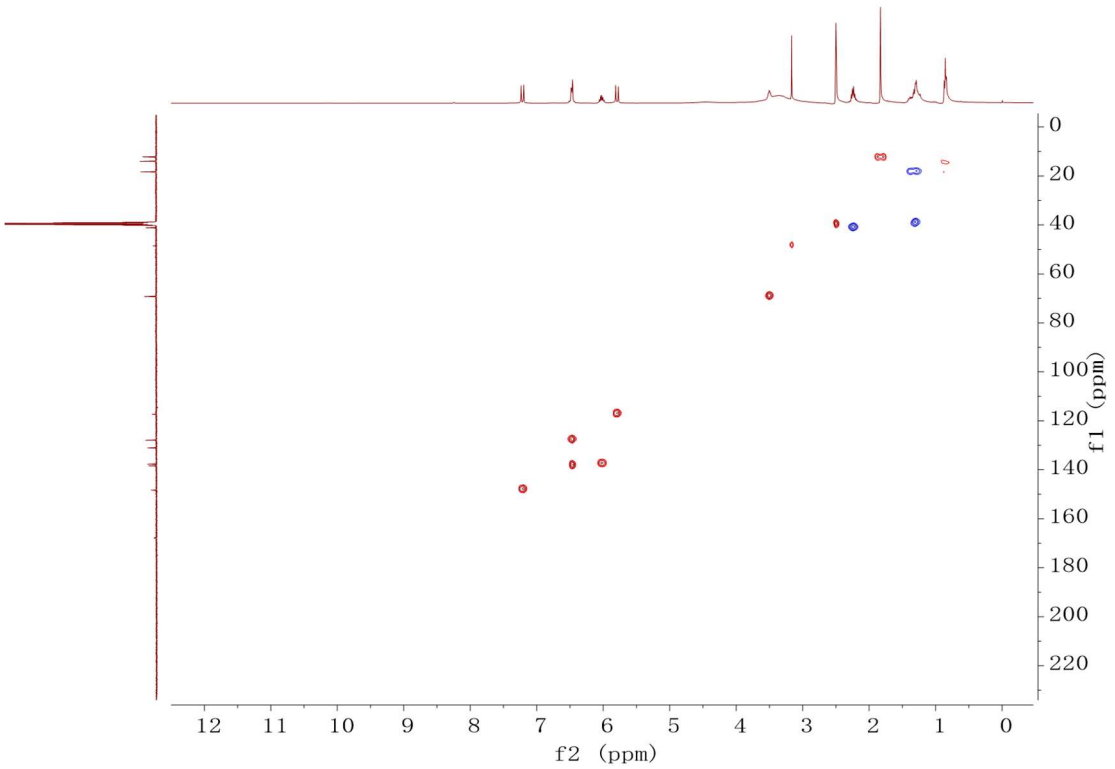


Figure S60. HSQC spectrum of compound **4** in DMSO-d_6.

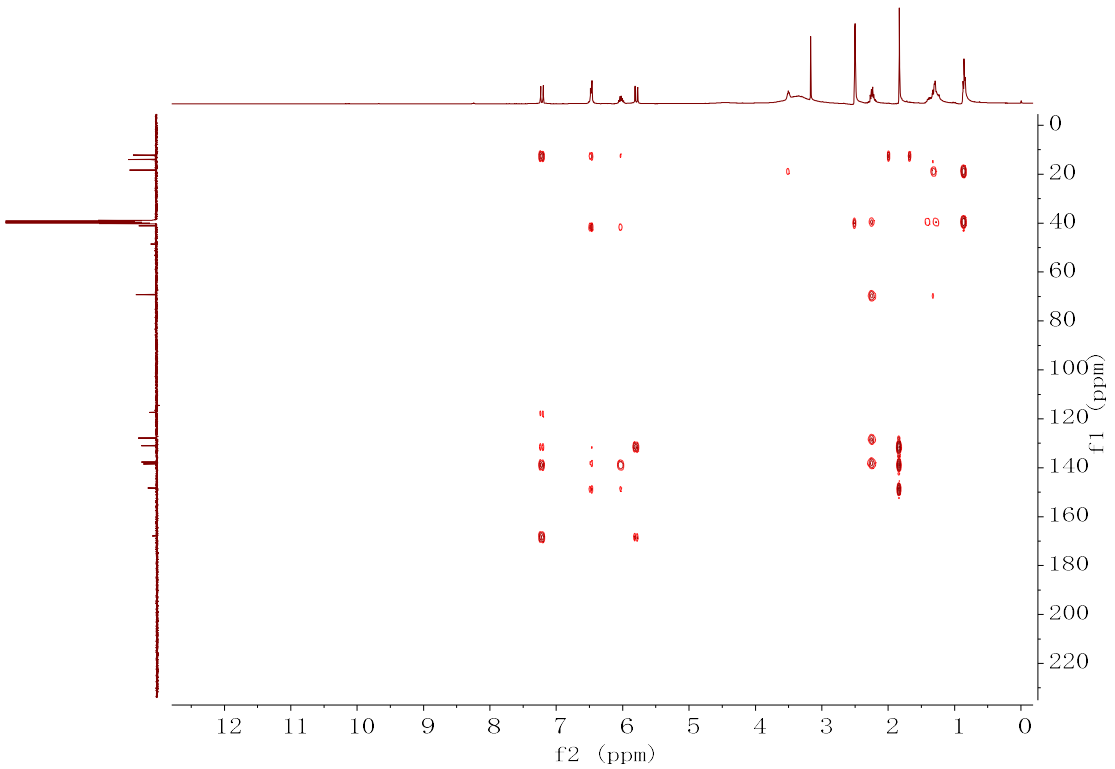


Figure S61. HMBC spectrum of compound **4** in DMSO-d_6.

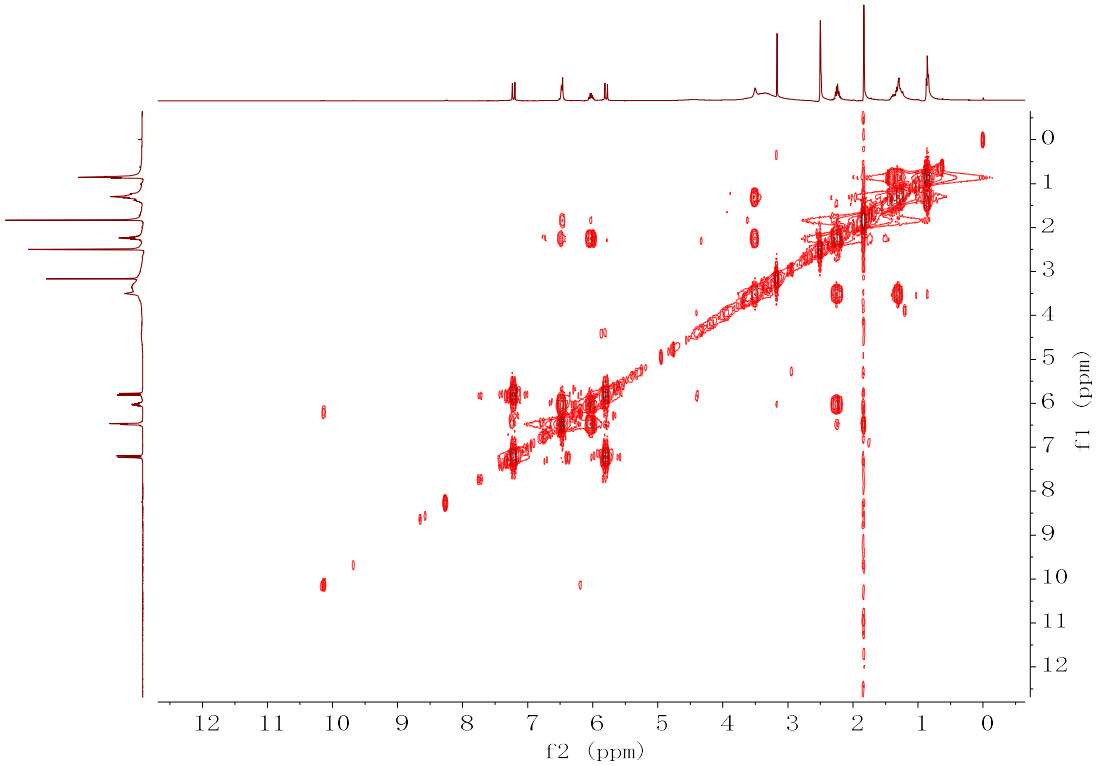


Figure S62. ^{1}H-^{1}H COSY spectrum of compound **4** in DMSO-d_6.

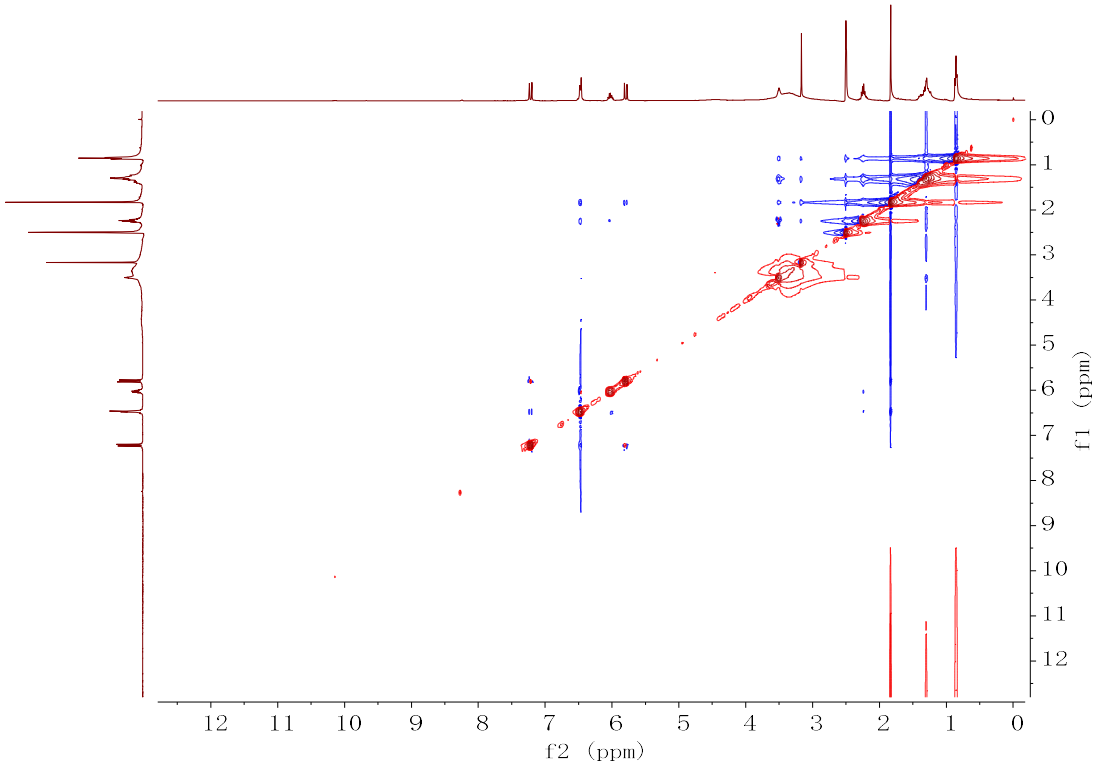


Figure S63. ^{1}H-^{1}H NOESY spectrum of compound **4** in DMSO-d_6.

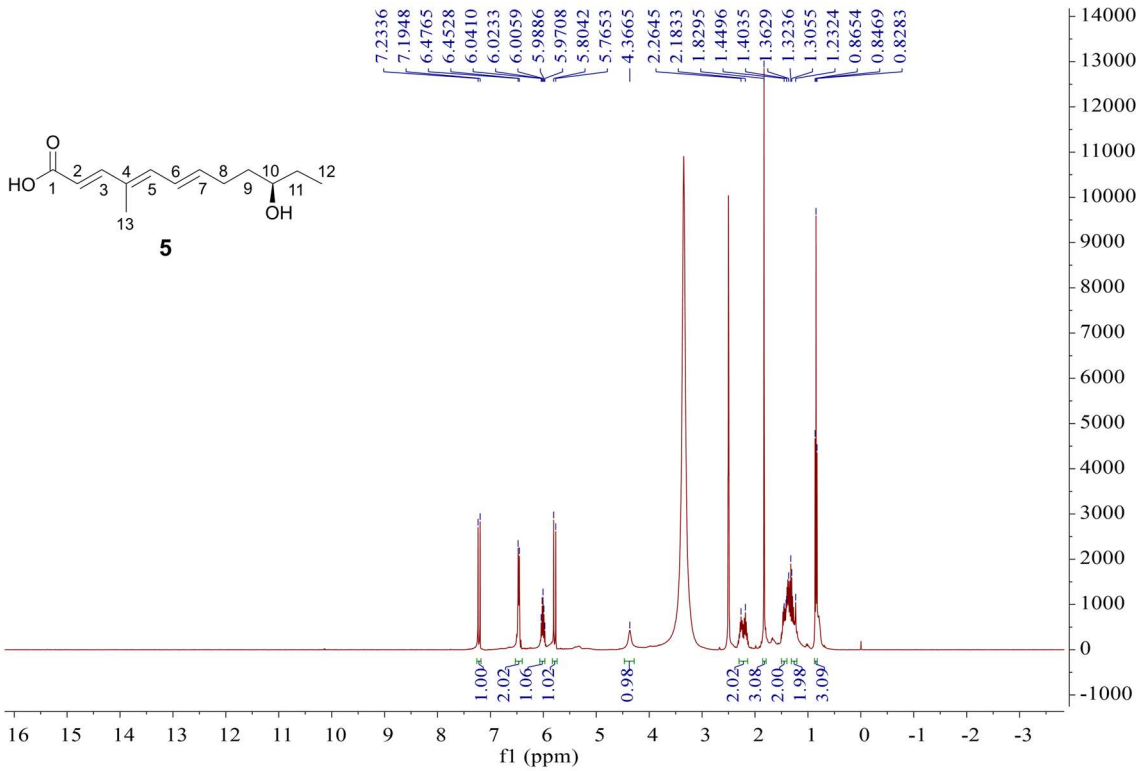


Figure S64. ^{1}H NMR spectrum of compound **5** in DMSO-d_6 (400 MHz).

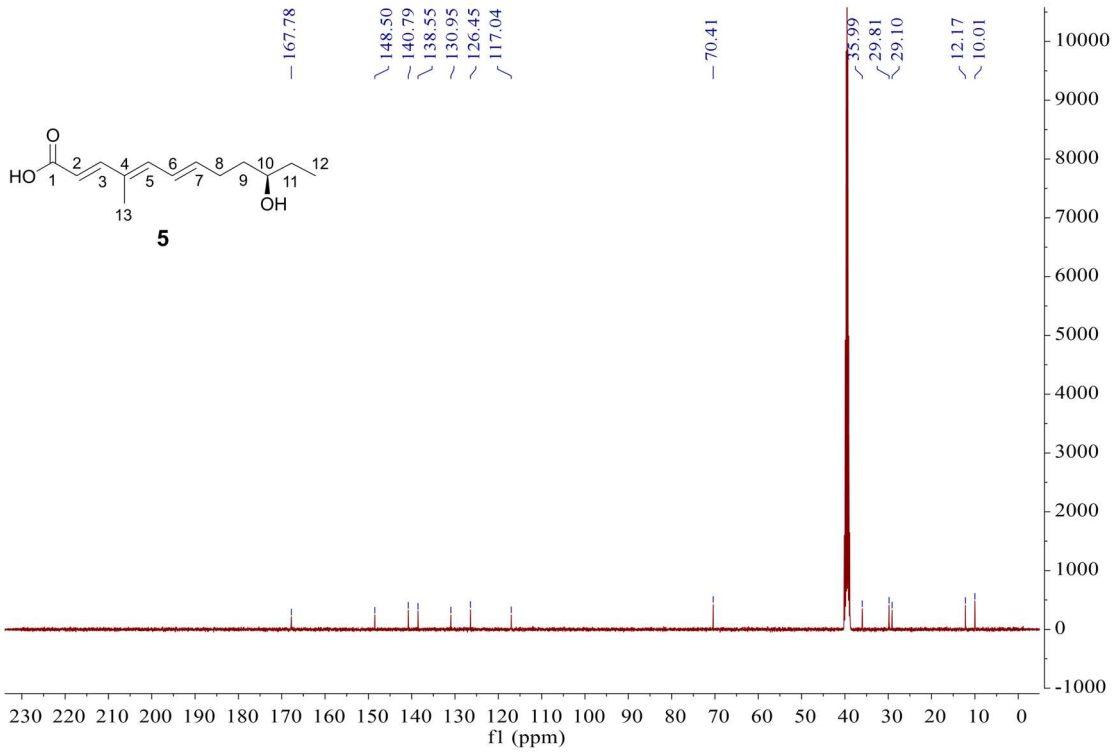


Figure S65. ^{13}C NMR spectrum of compound **5** in DMSO-d_6 (100 MHz).

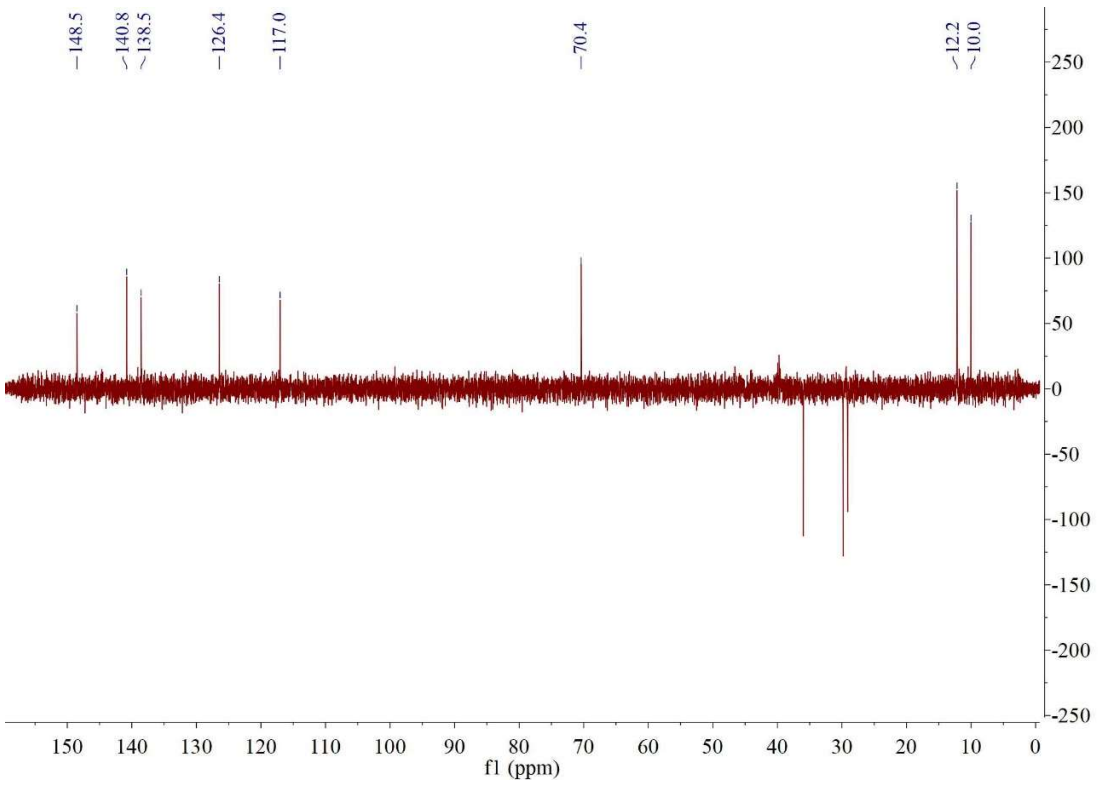


Figure S66. DEPT-135° spectrum of compound **5** in DMSO-d_6.

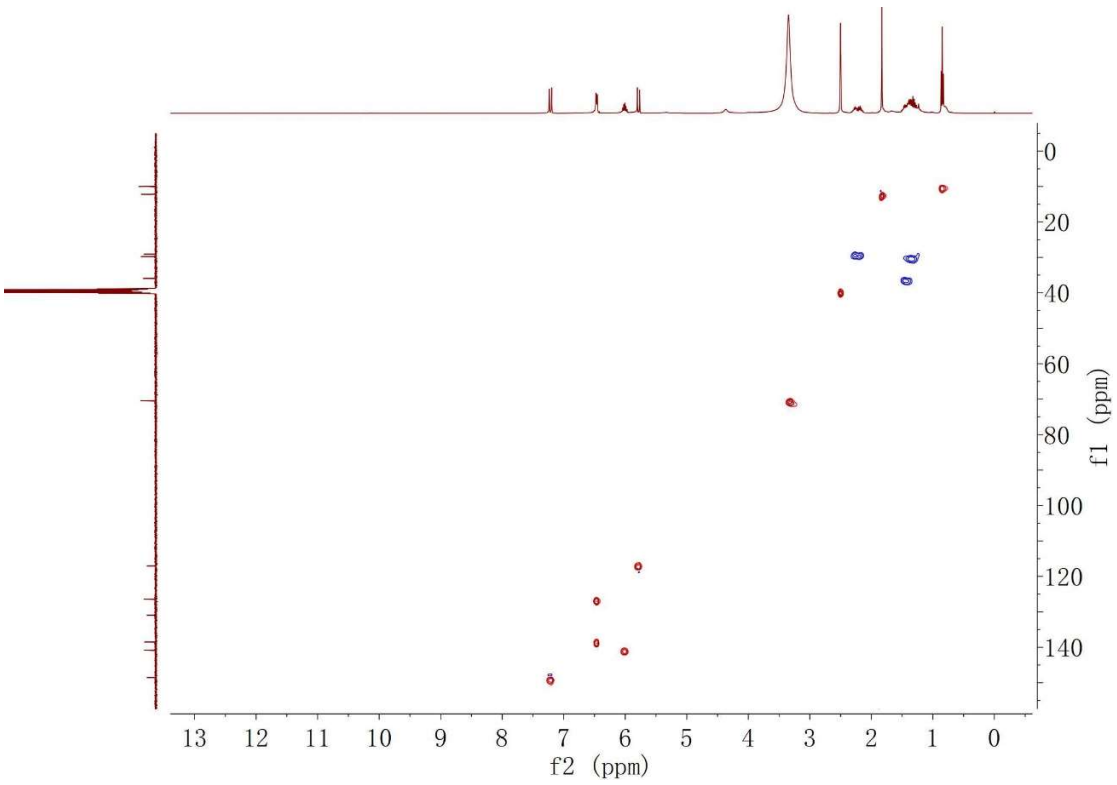


Figure S67. HSQC spectrum of compound **5** in DMSO-d_6.

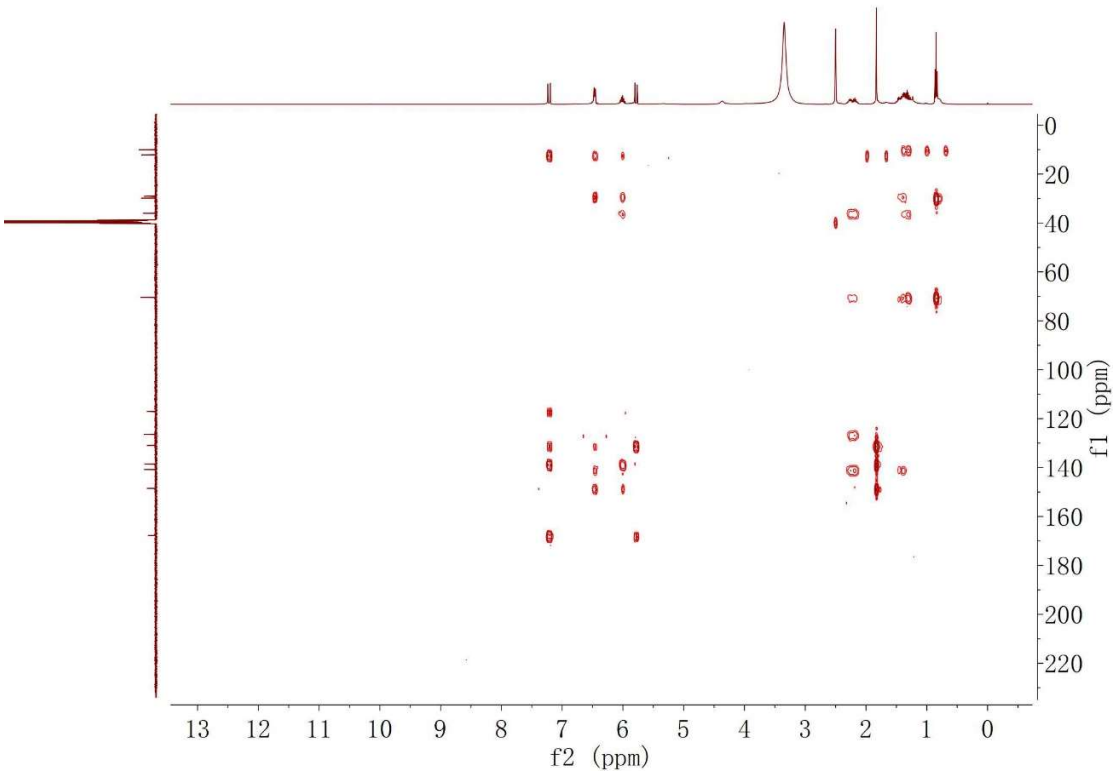


Figure S68. HMBC spectrum of compound **5** in DMSO-d_6.

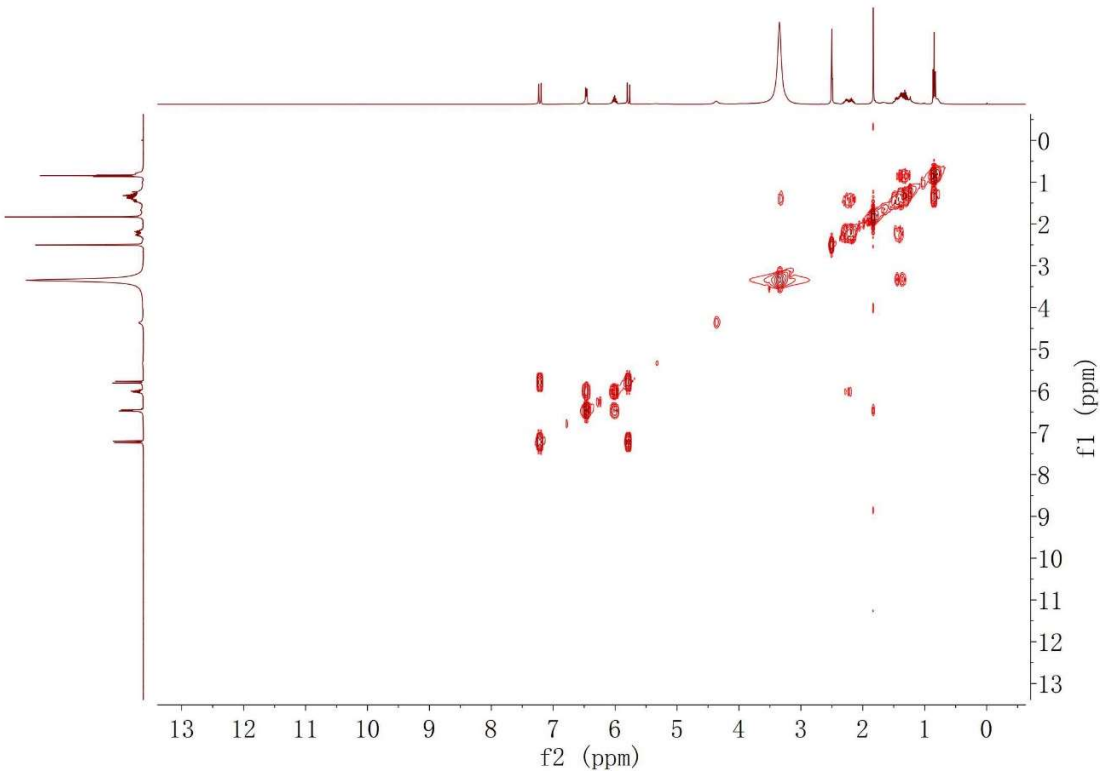


Figure S69. ^{1}H-^{1}H COSY spectrum of compound **5** in DMSO-d_6.

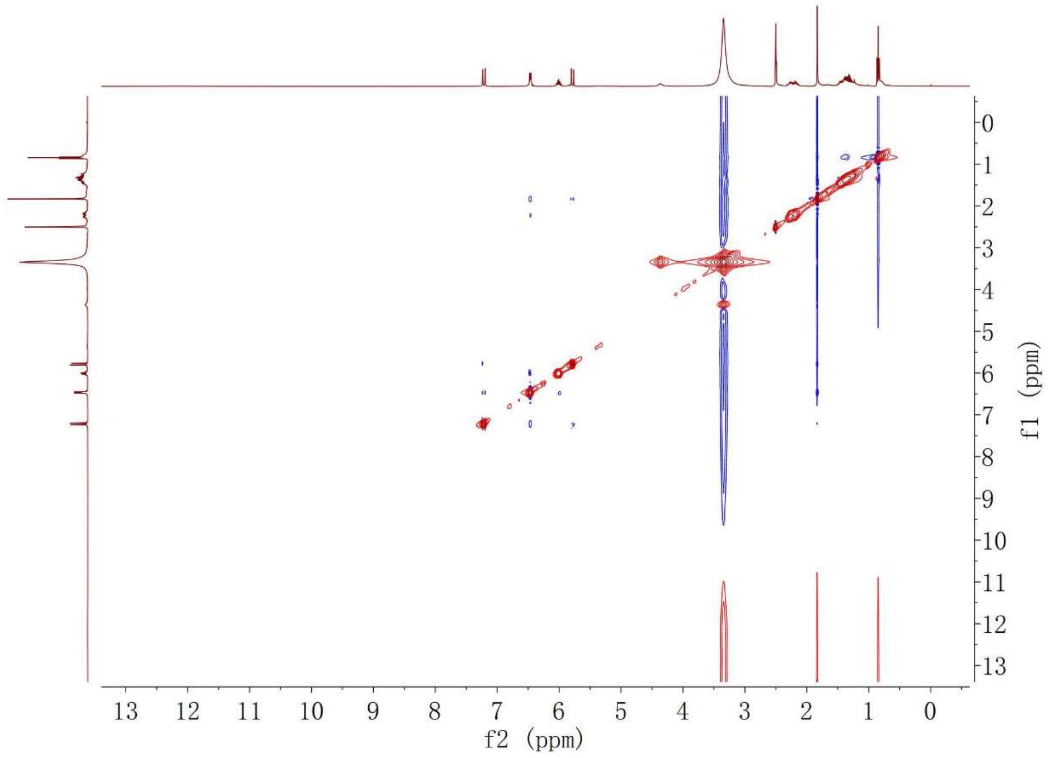


Figure S70. ^{1}H-^{1}H NOESY spectrum of compound **5** in DMSO-d_6.

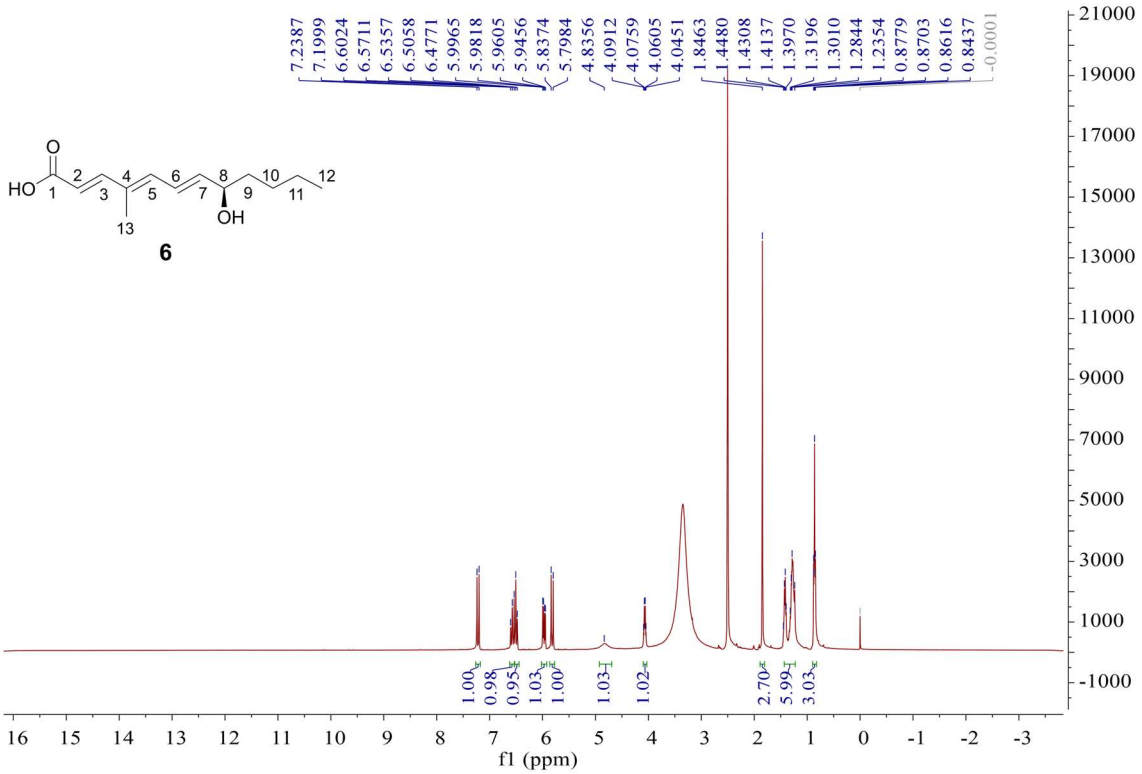


Figure S71. [1]H NMR spectrum of compound **6** in DMSO-d_6 (400 MHz).

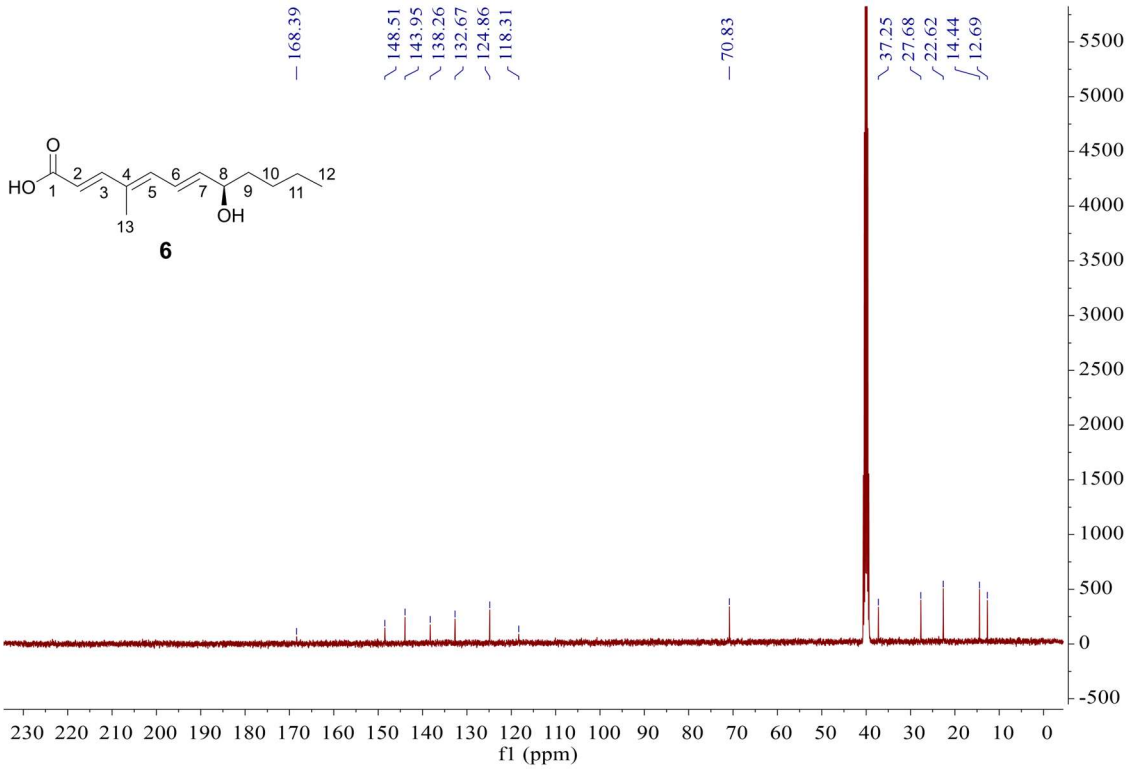


Figure S72. ^{13}C NMR spectrum of compound **6** in DMSO-d_6 (100 MHz).

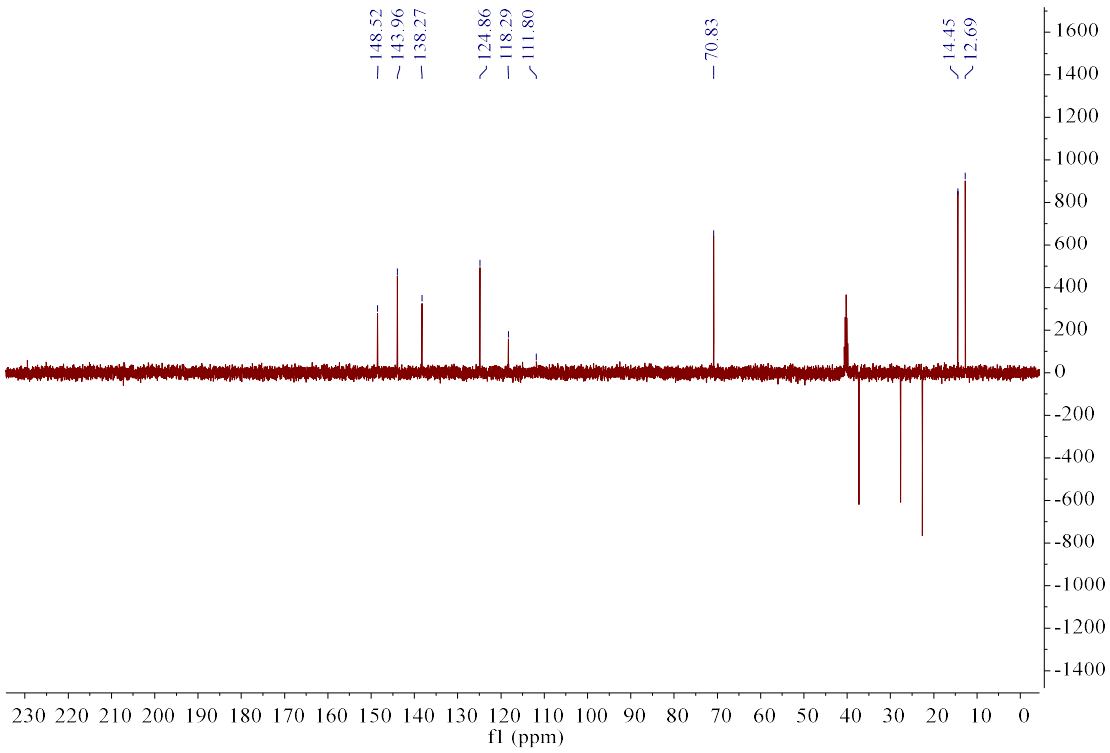


Figure S73. DEPT-135° spectrum of compound **6** in DMSO-d_6.

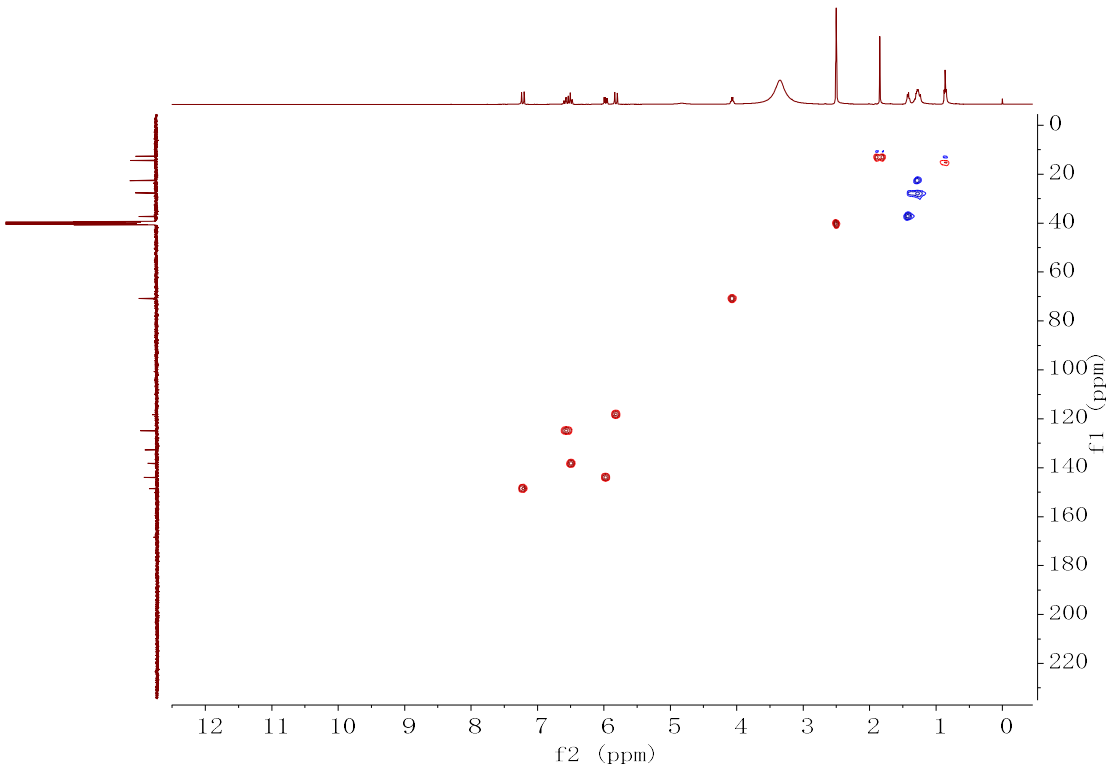


Figure S74. HSQC spectrum of compound **6** in DMSO-d_6.

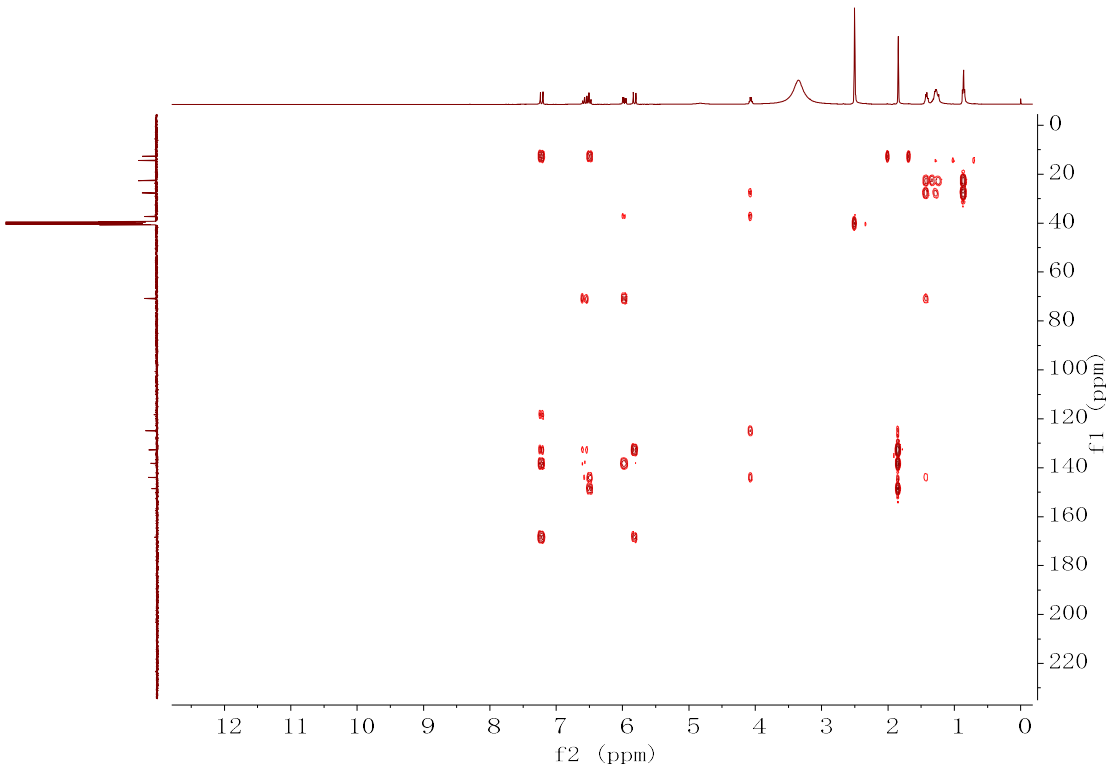


Figure S75. HMBC spectrum of compound **6** in DMSO-d_6.

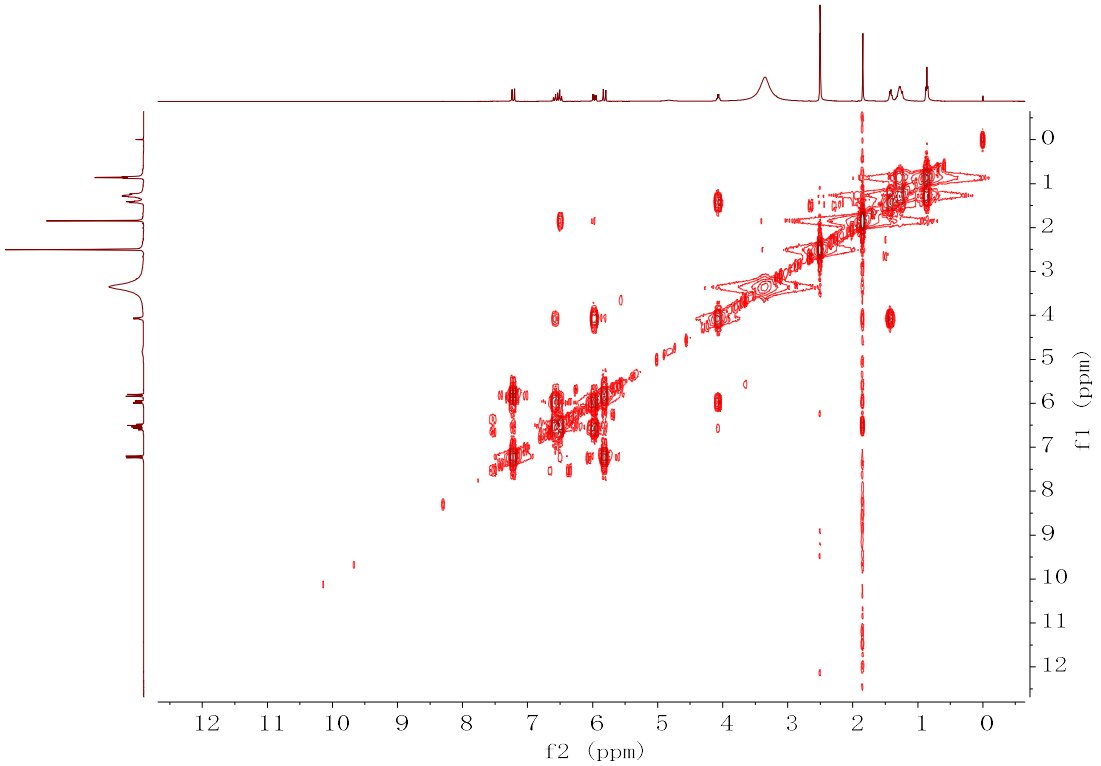


Figure S76. ^{1}H-^{1}H COSY spectrum of compound **6** in DMSO-d_6.

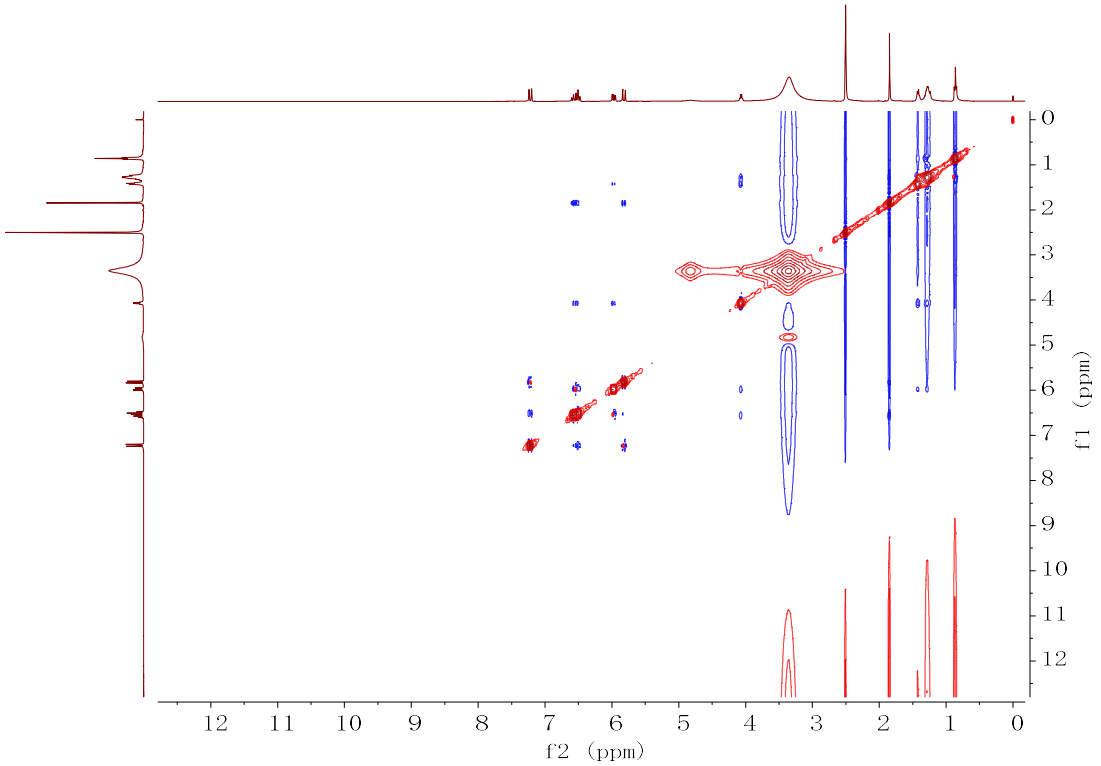


Figure S77. ^{1}H-^{1}H NOESY spectrum of compound **6** in DMSO-d_6.

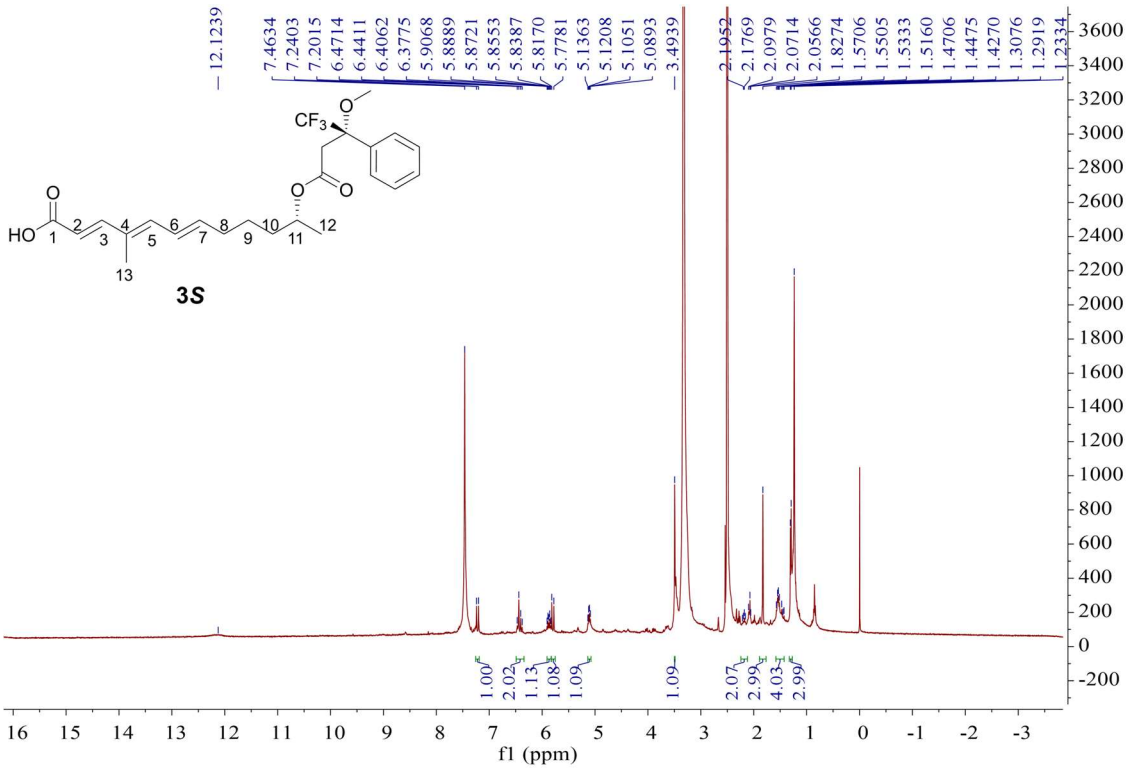


Figure S78. ^{1}H NMR spectrum of compound **3*S*** in DMSO-d_6 (400 MHz).

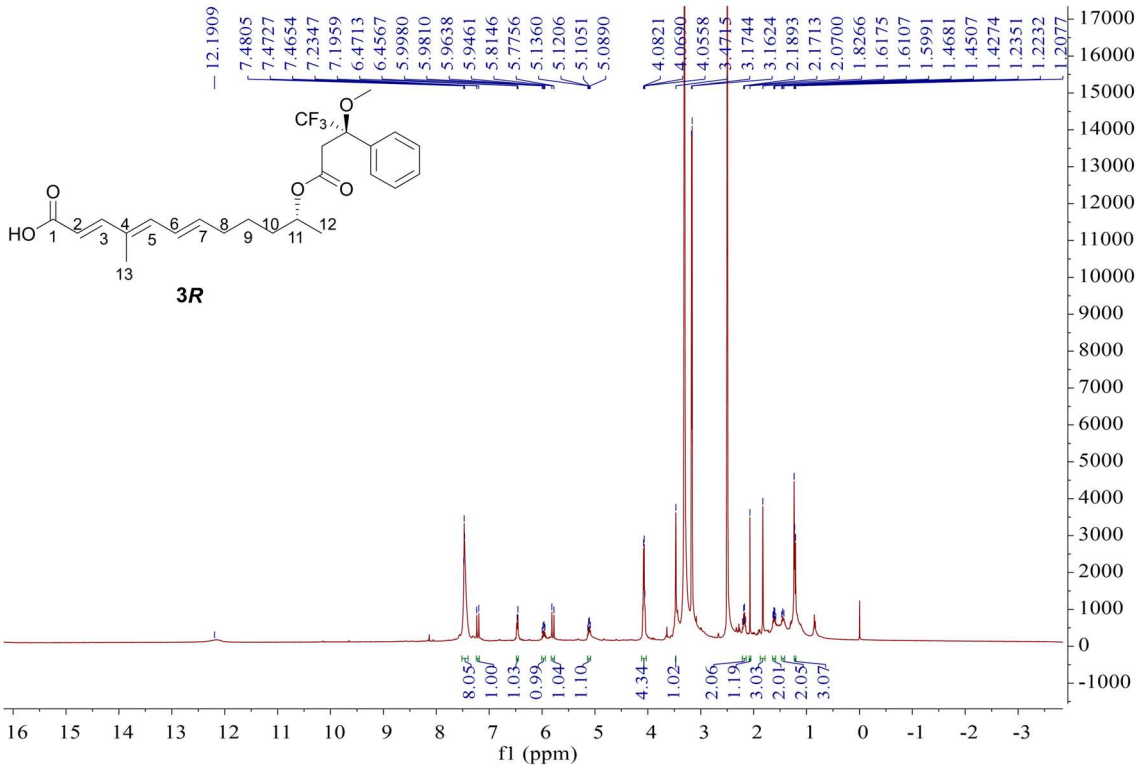


Figure S79. ^{1}H NMR spectrum of compound ***3R*** in DMSO-d_6 (400 MHz).

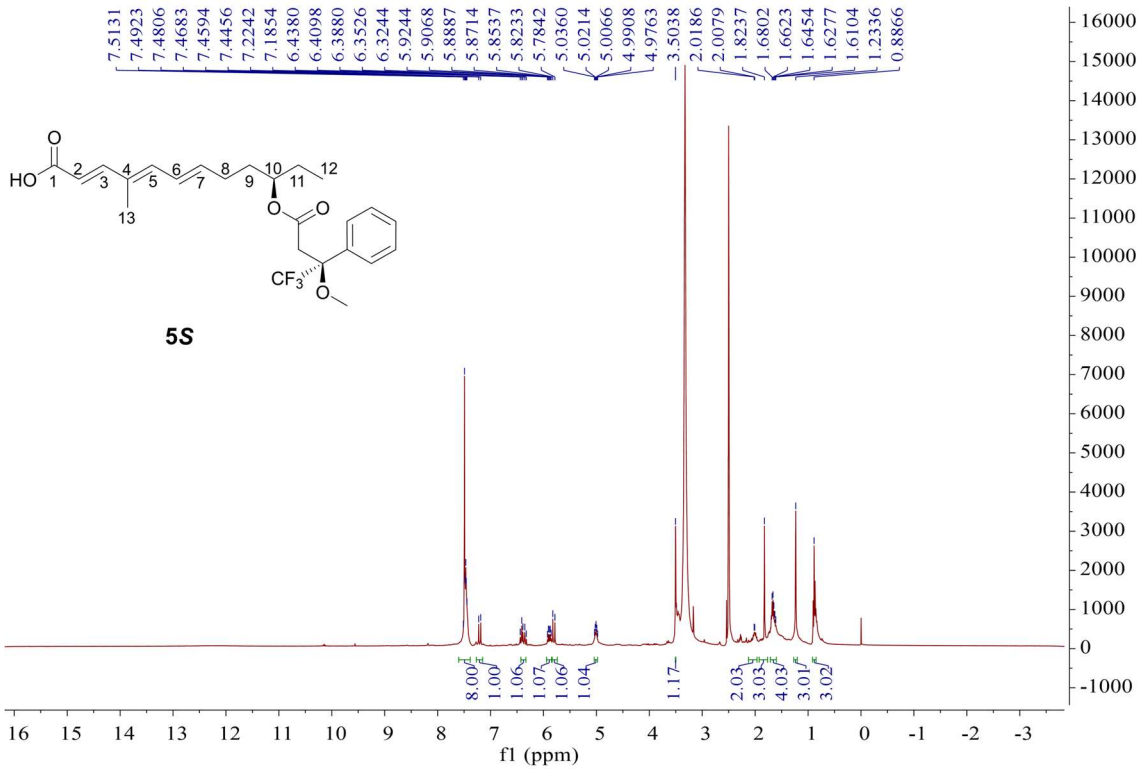


Figure S80. ^{1}H NMR spectrum of compound **5*S*** in DMSO-d_6 (400 MHz).

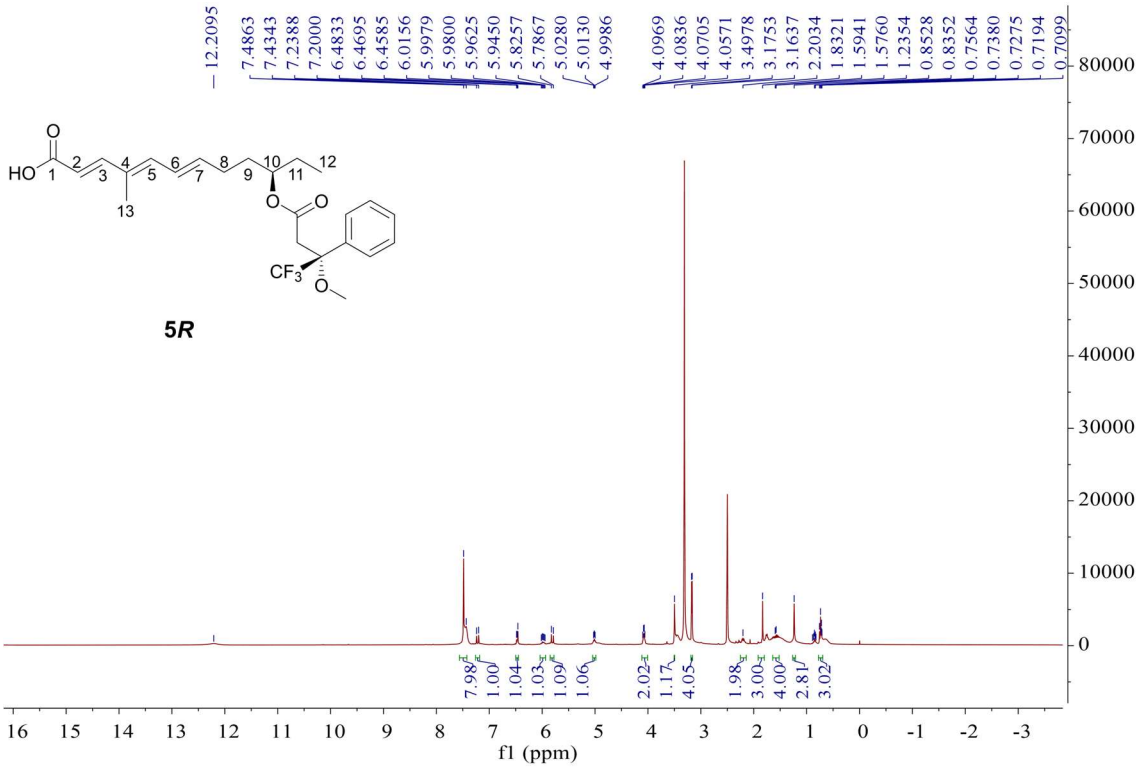


Figure S81. [1]H NMR spectrum of compound **5*R*** in DMSO-d_6 (400 MHz).

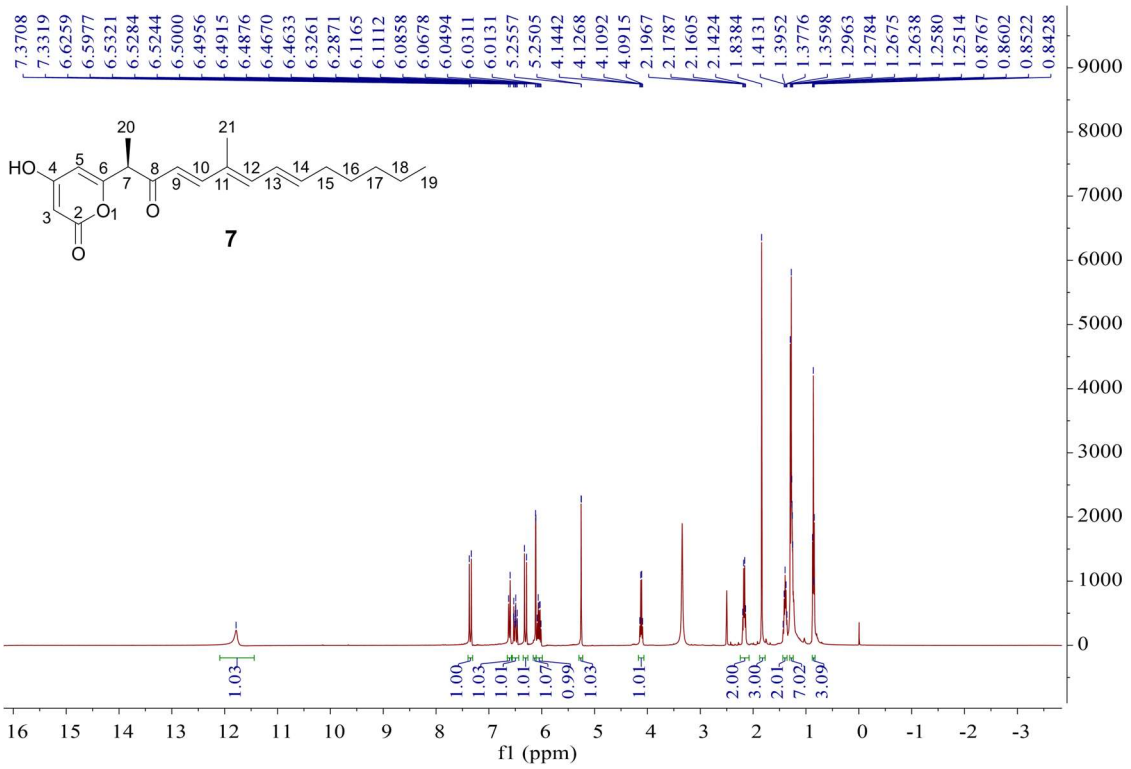


Figure S82. [1]H NMR spectrum of compound **7** in DMSO-d_6 (400 MHz).

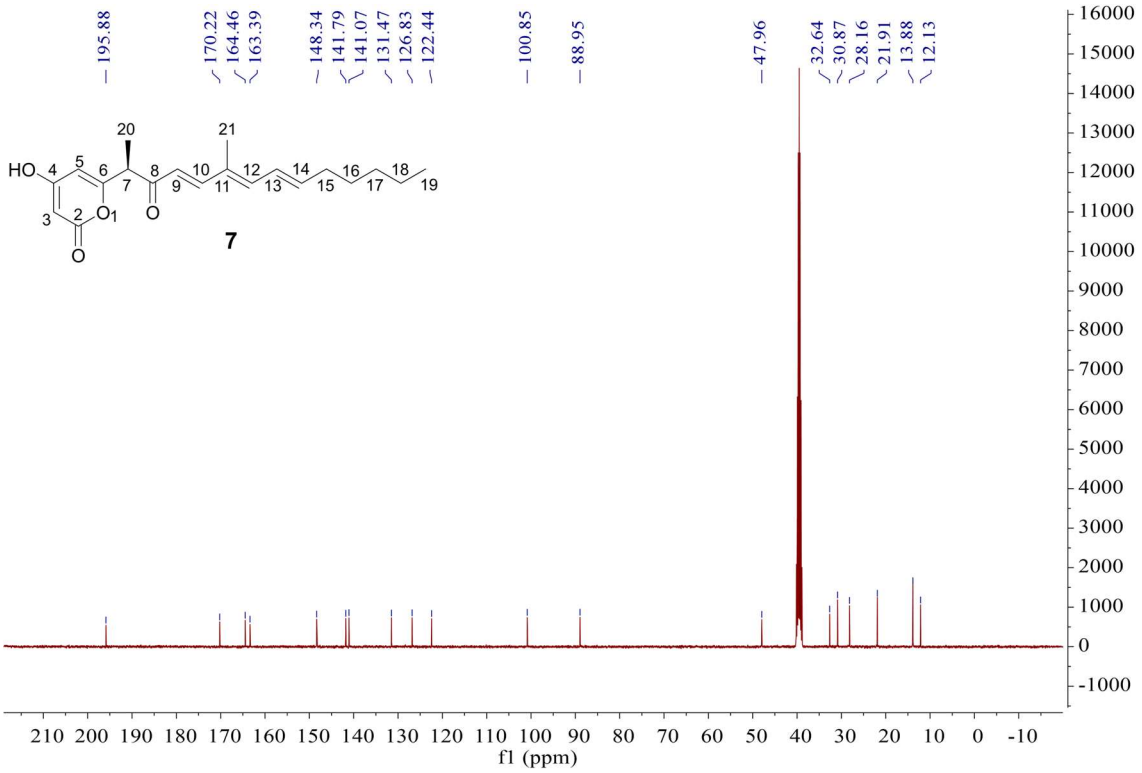


Figure S83. ^{13}C NMR spectrum of compound **7** in DMSO-d_6 (100 MHz).

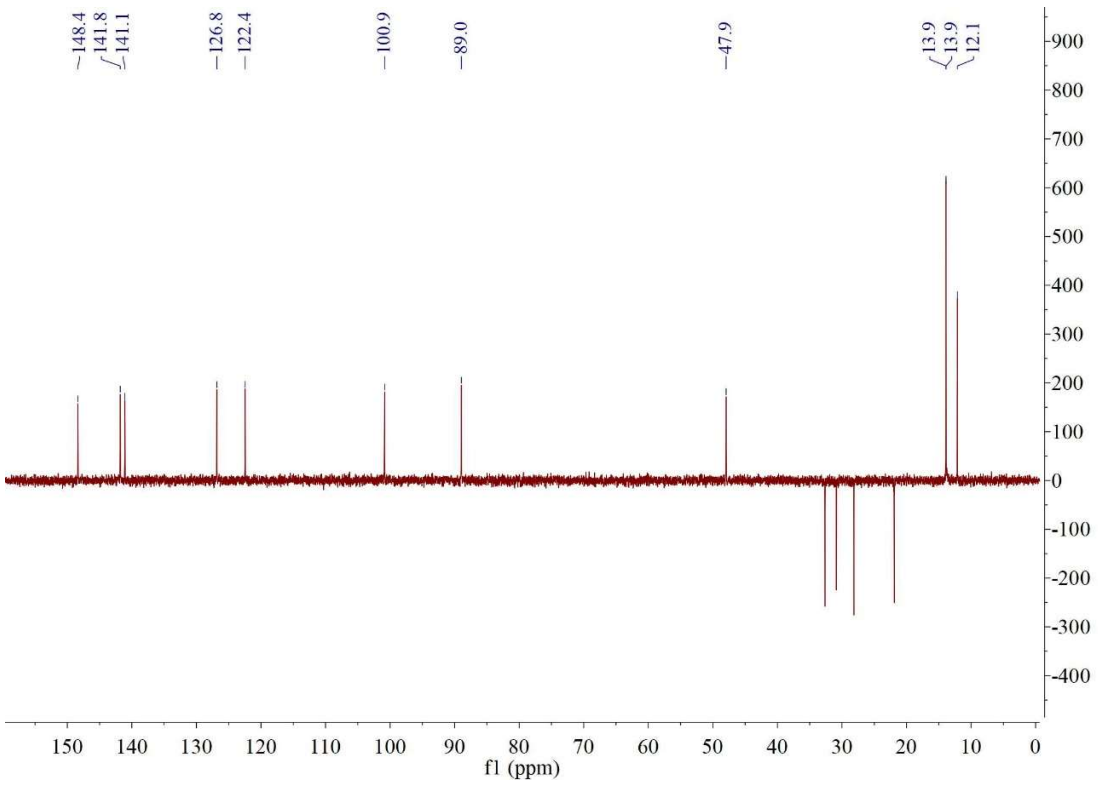


Figure S84. DEPT-135° spectrum of compound **7** in DMSO-d_6.

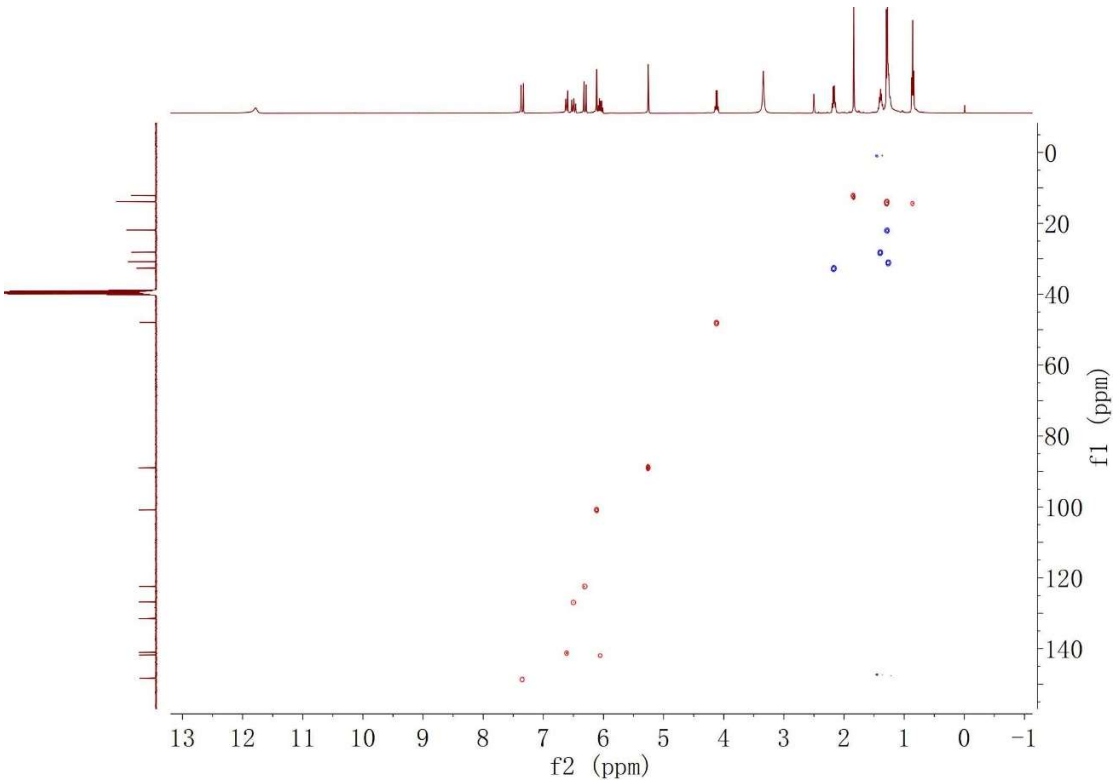


Figure S85. HSQC spectrum of compound **7** in DMSO-d_6.

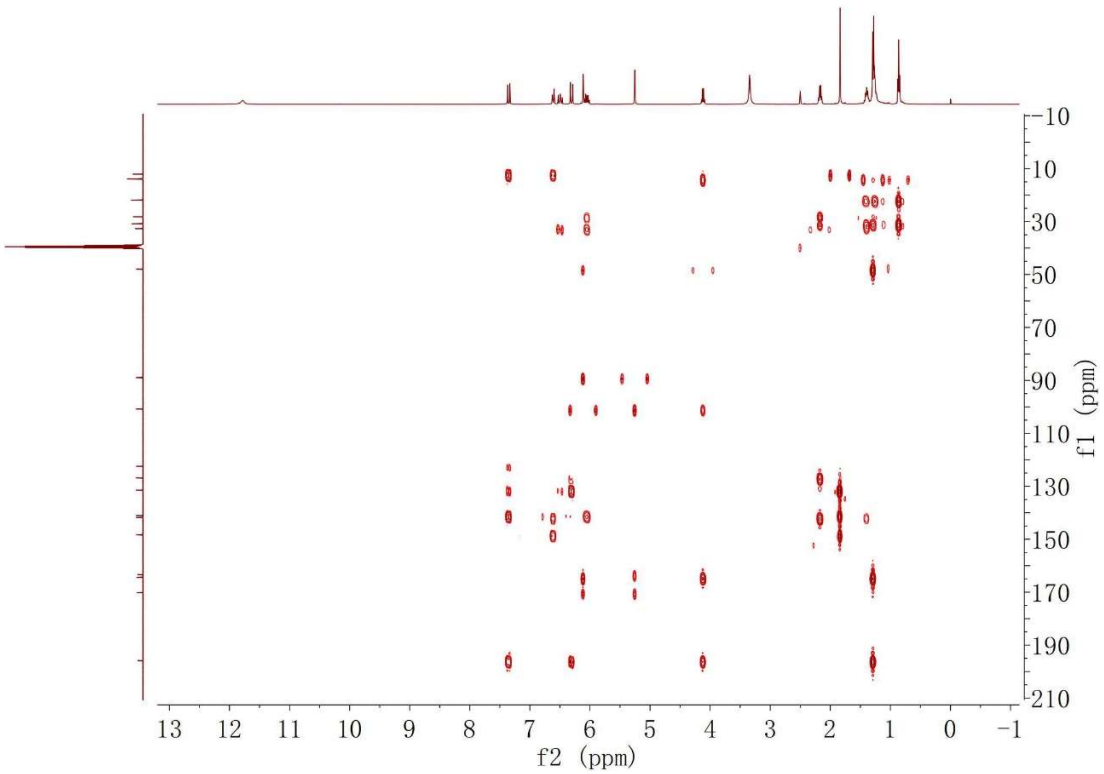


Figure S86. HMBC spectrum of compound **7** in DMSO-d_6.

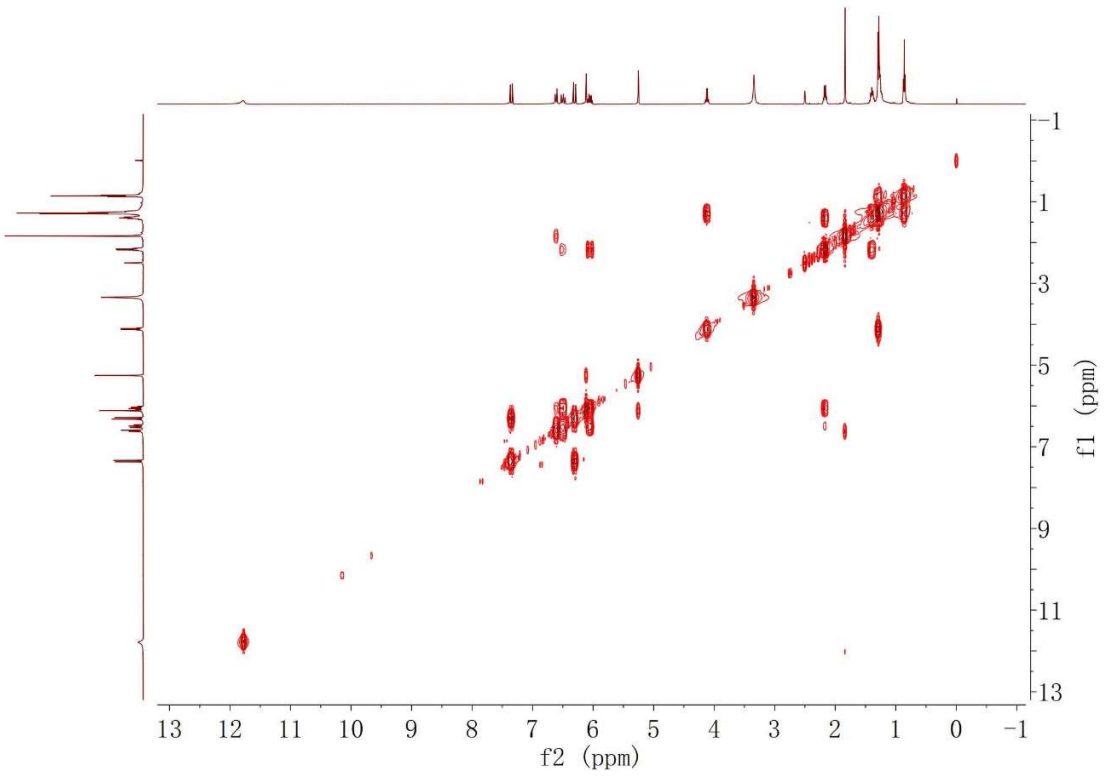


Figure S87. ^{1}H-^{1}H COSY spectrum of compound **7** in DMSO-d_6.

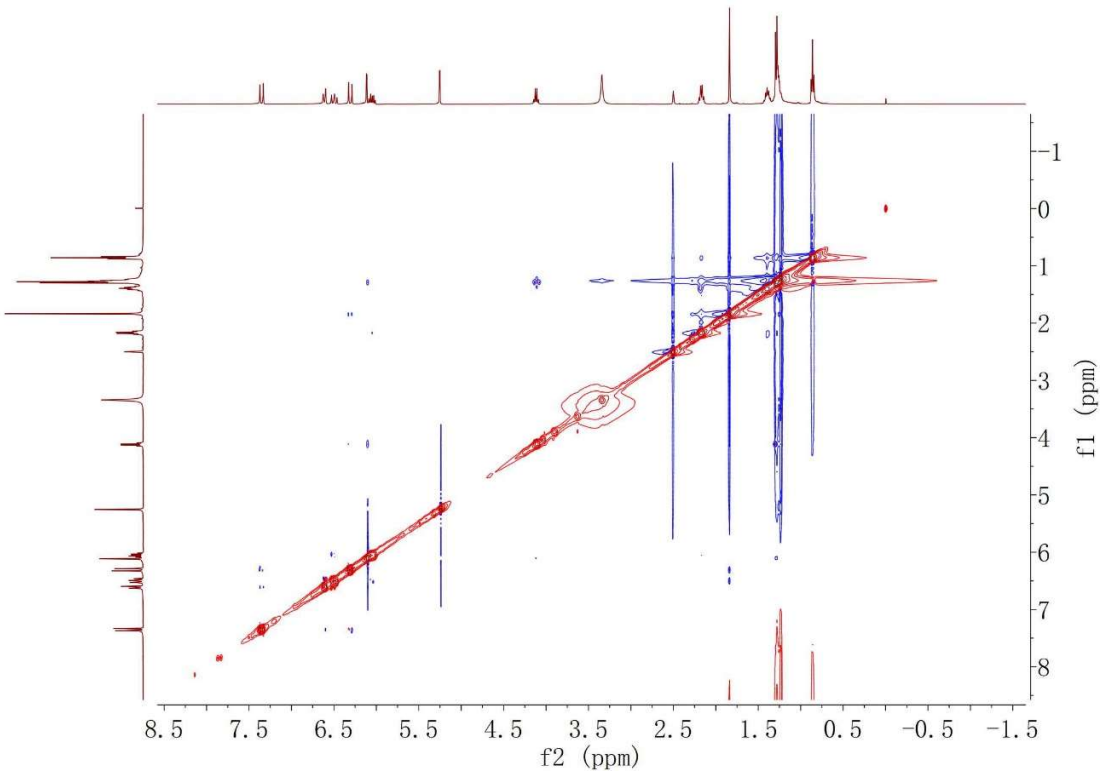


Figure S88. ^{1}H-^{1}H NOESY spectrum of compound **7** in DMSO-d_6.

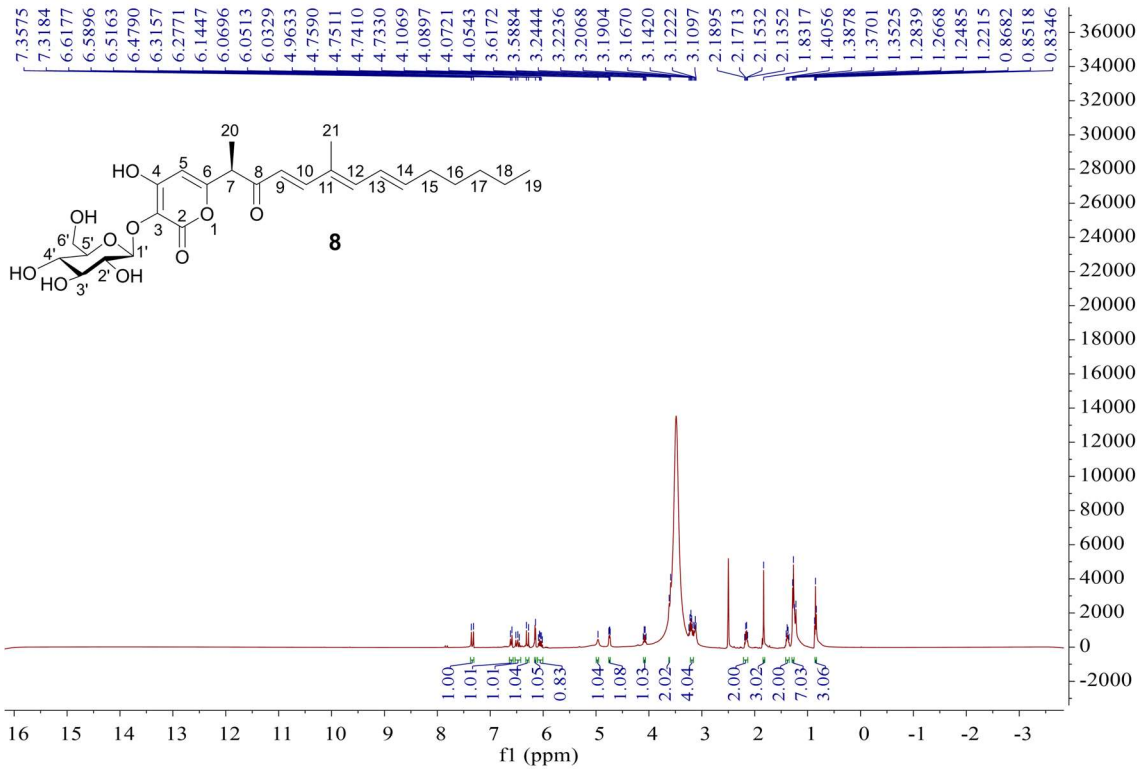


Figure S89. [1]H NMR spectrum of compound **8** in DMSO-d_6 (400 MHz).

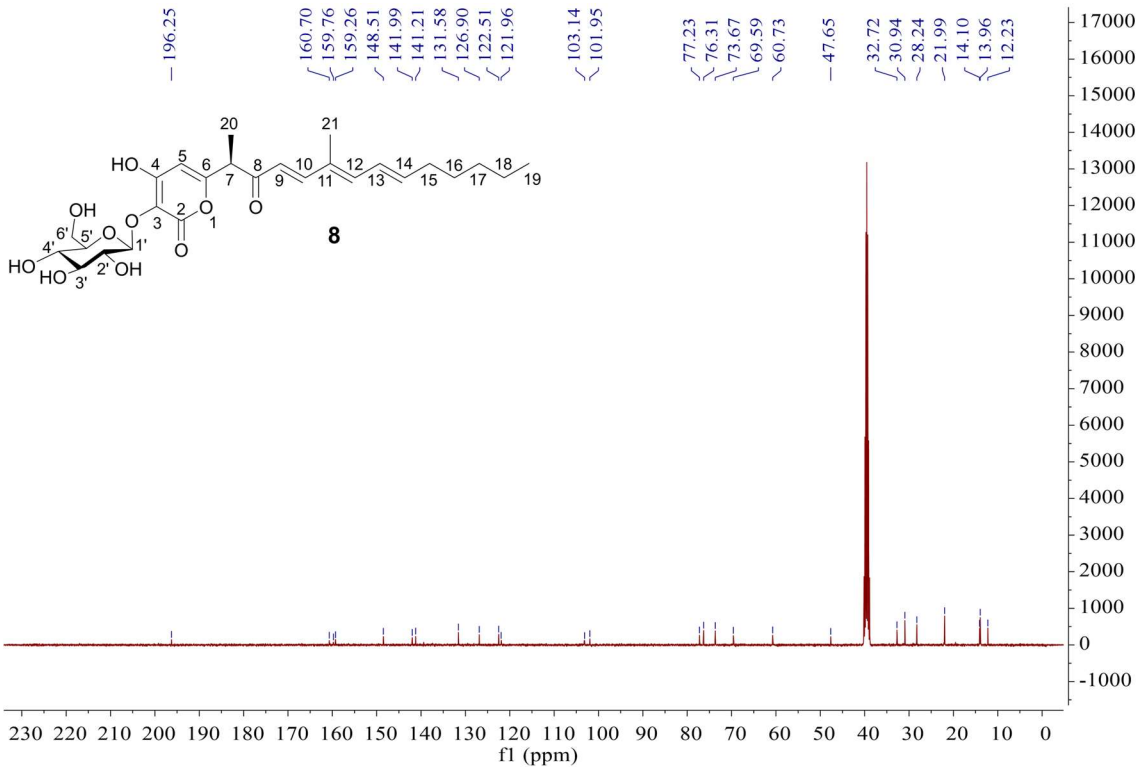


Figure S90. ^{13}C NMR spectrum of compound **8** in DMSO-d_6 (100 MHz).

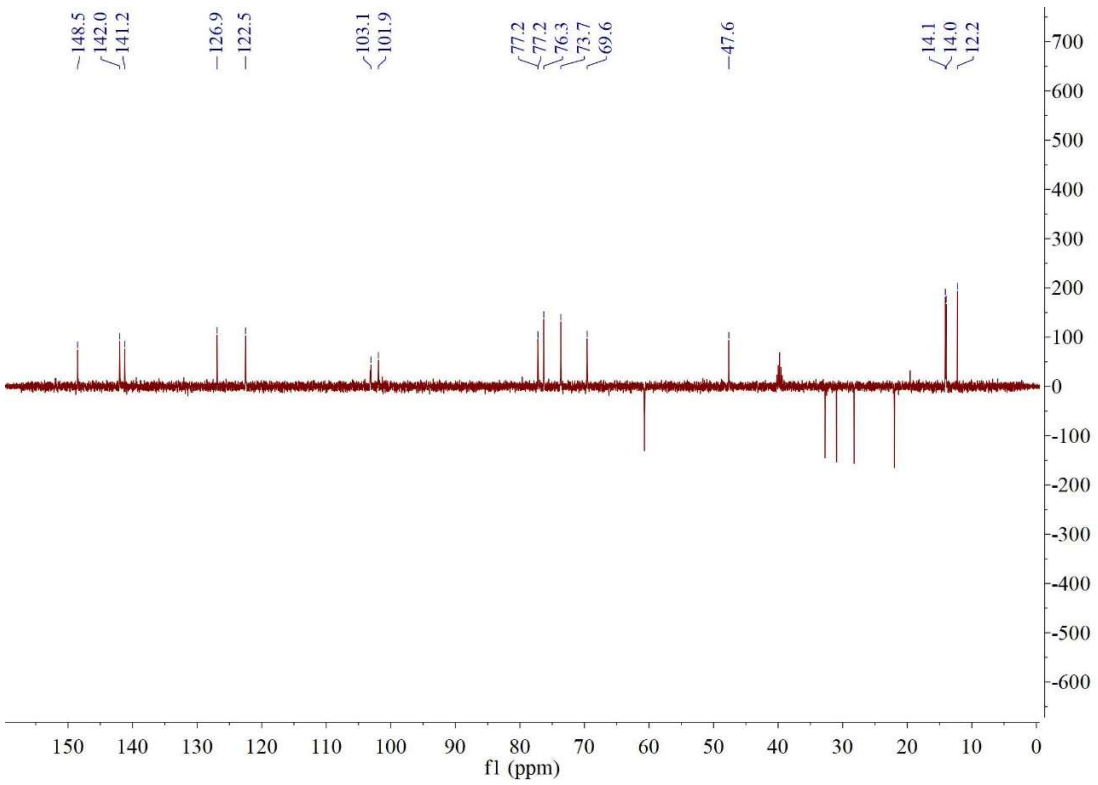


Figure S91. DEPT-135° spectrum of compound **8** in DMSO-d_6.

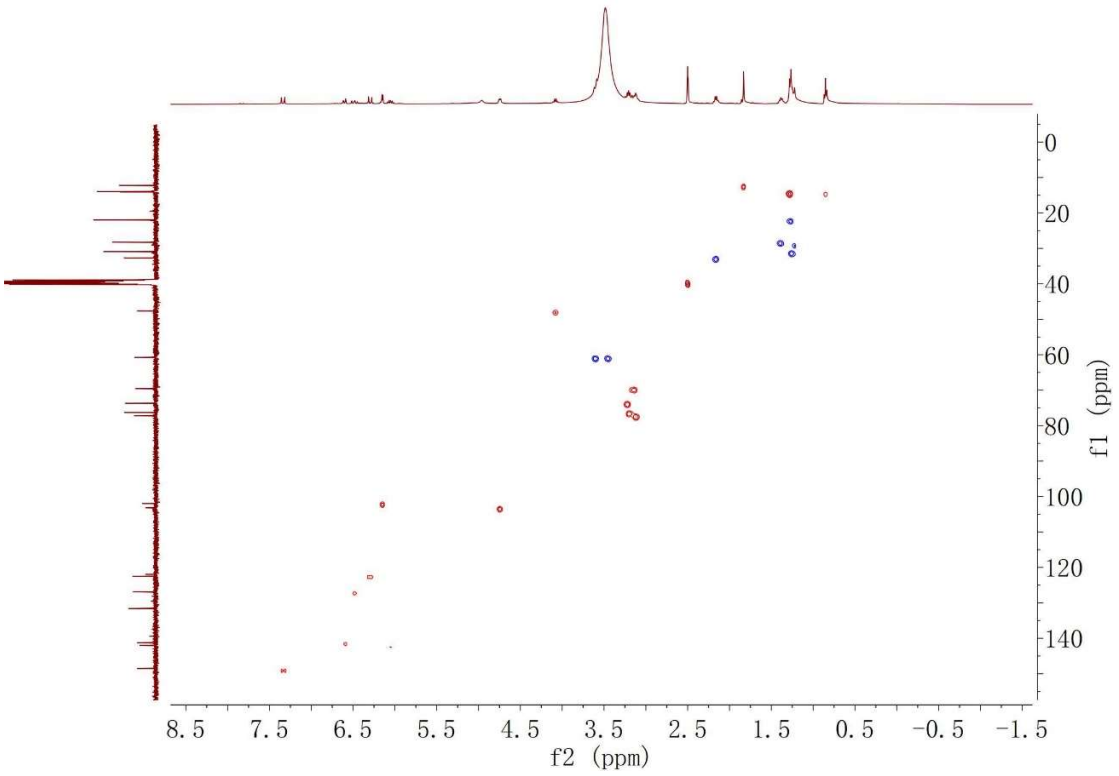


Figure S92. HSQC spectrum of compound **8** in DMSO-d_6.

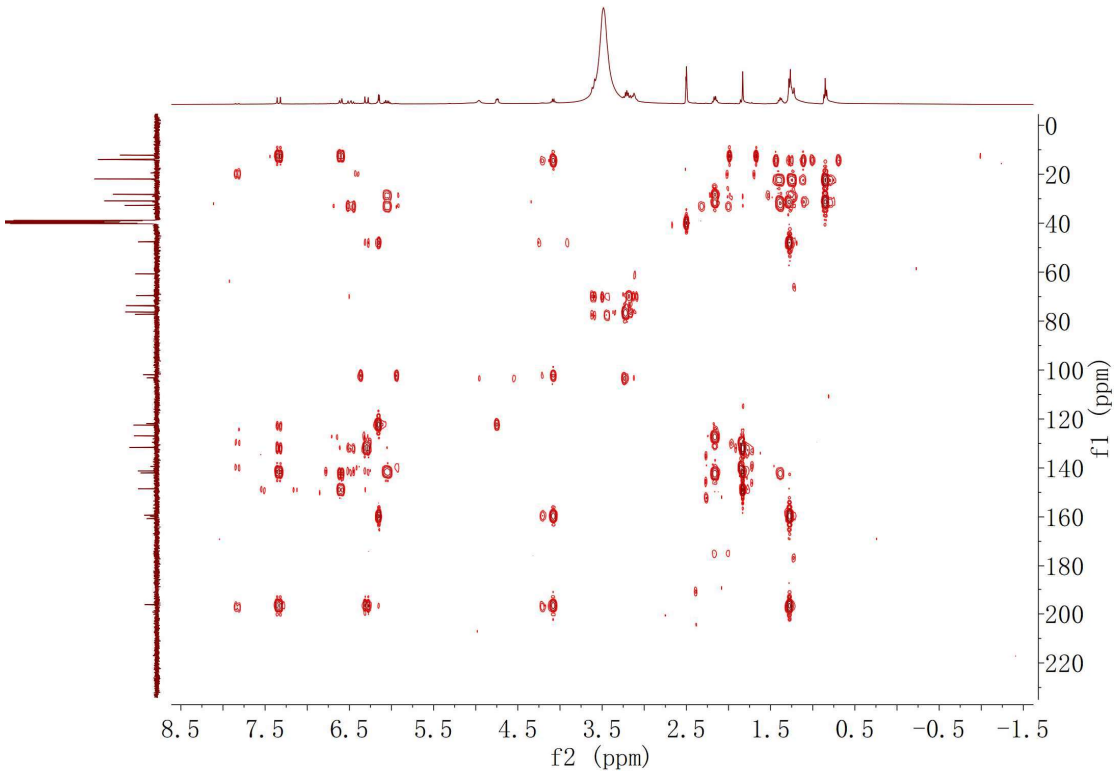


Figure S93. HMBC spectrum of compound **8** in DMSO-d_6.

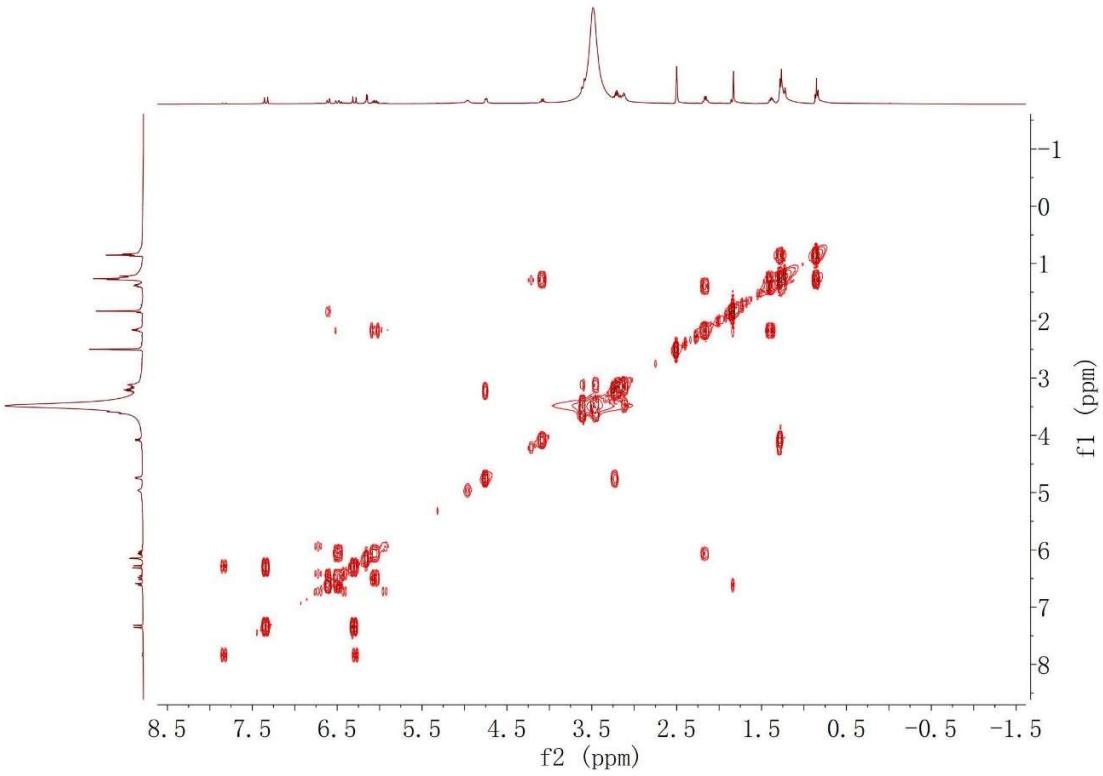


Figure S94. ^{1}H-^{1}H COSY spectrum of compound **8** in DMSO-d_6.

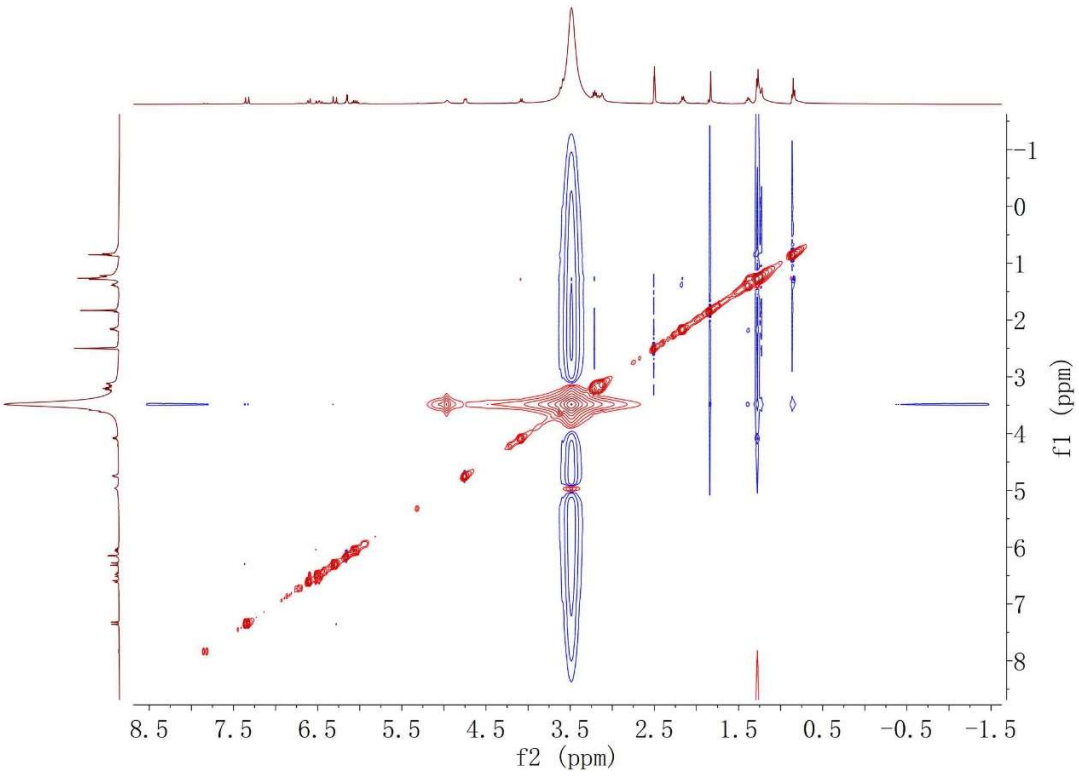


Figure S95. ^{1}H-^{1}H NOESY spectrum of compound **8** in DMSO-d_6.

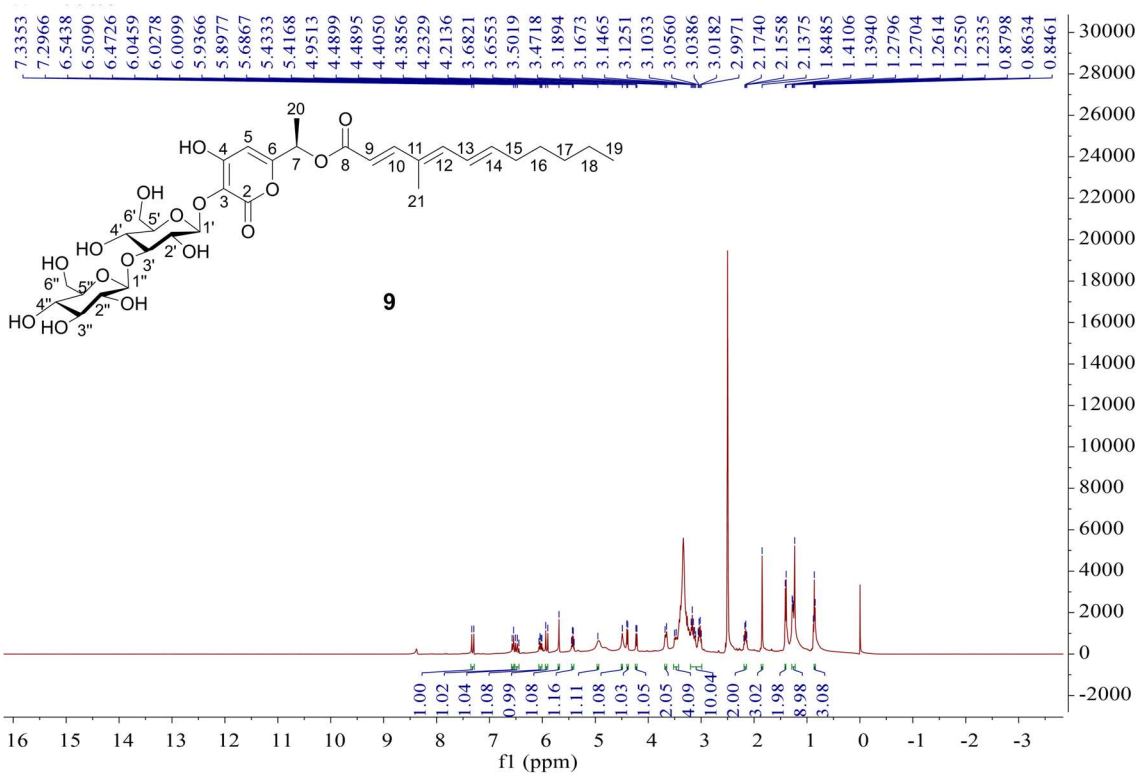


Figure S96. [1]H NMR spectrum of compound **9** in DMSO-d_6 (400 MHz).

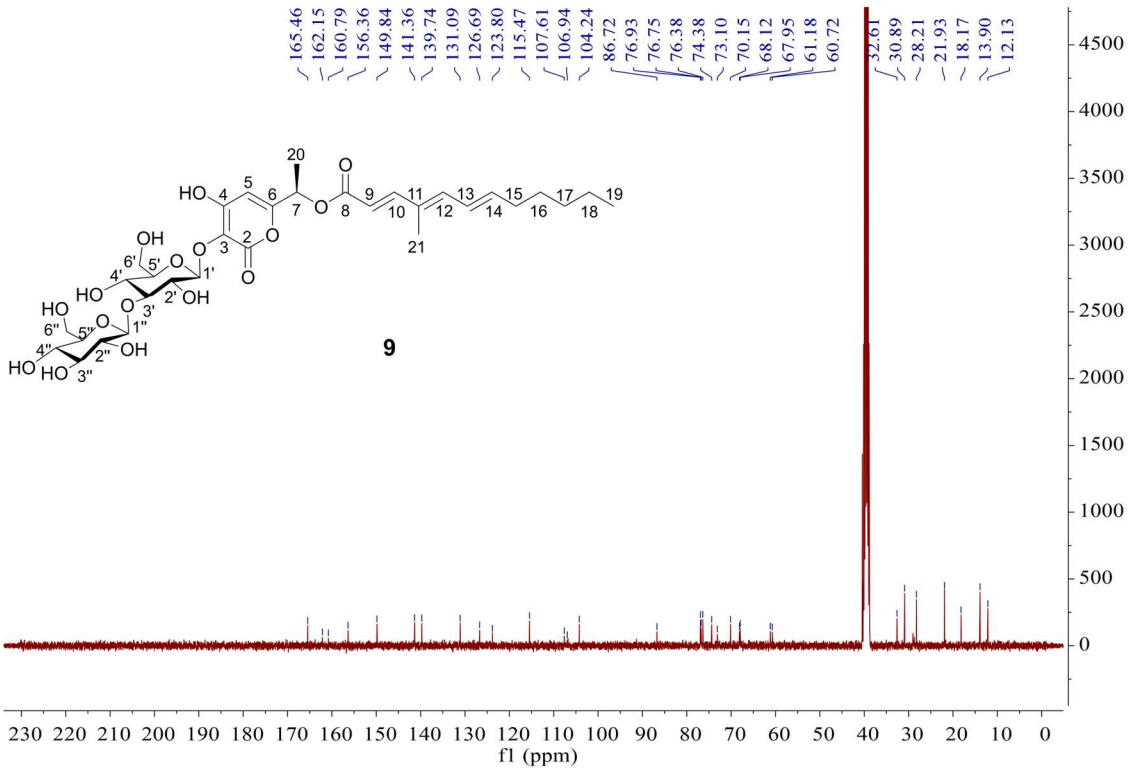


Figure S97. ^{13}C NMR spectrum of compound **9** in DMSO-d_6 (100 MHz).

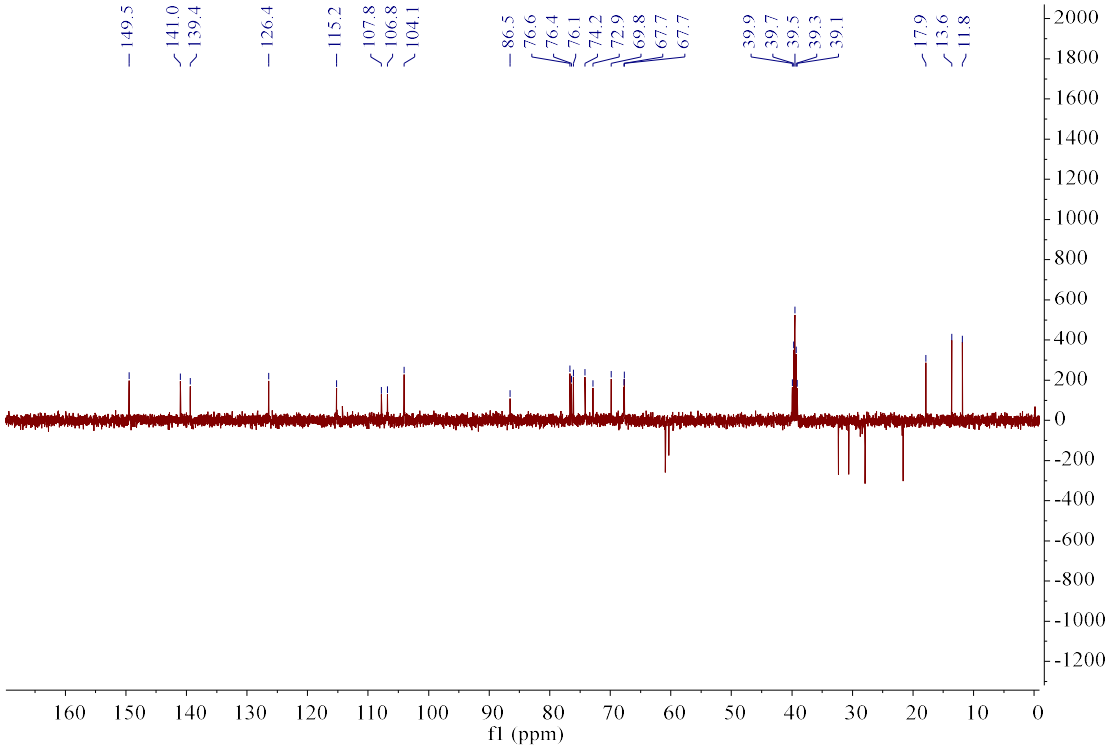


Figure S98. DEPT-135° spectrum of compound **9** in DMSO-d_6.

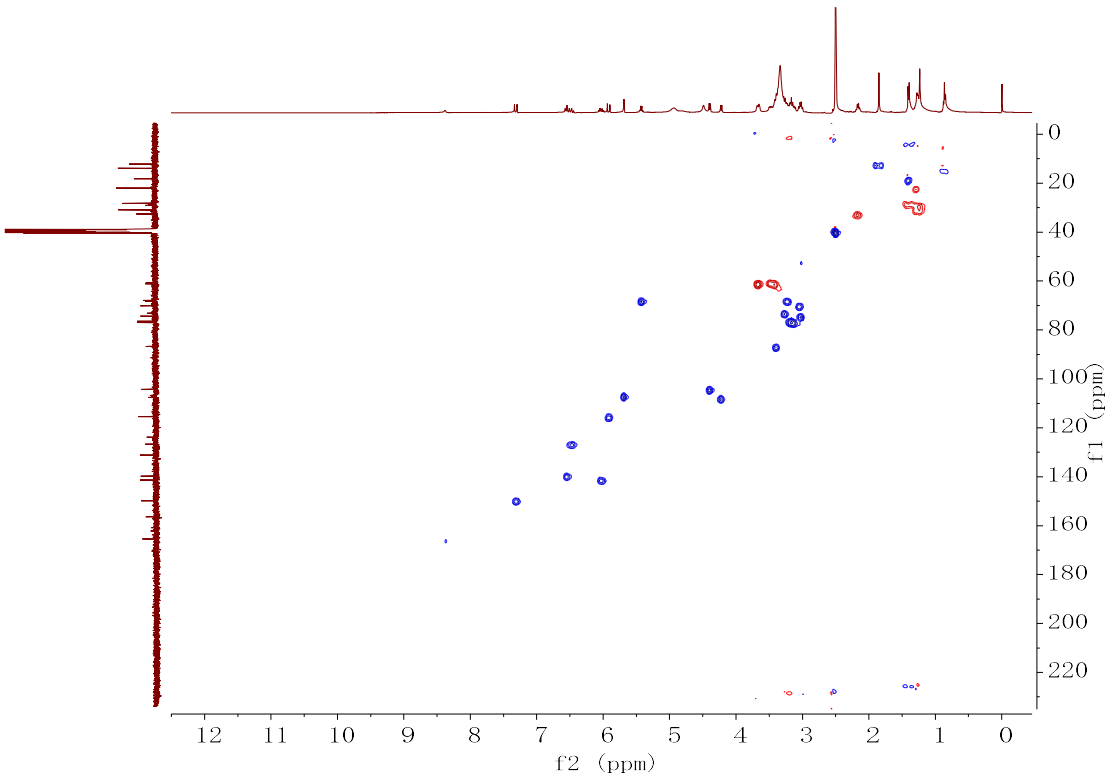


Figure S99. HSQC spectrum of compound **9** in DMSO-d_6.

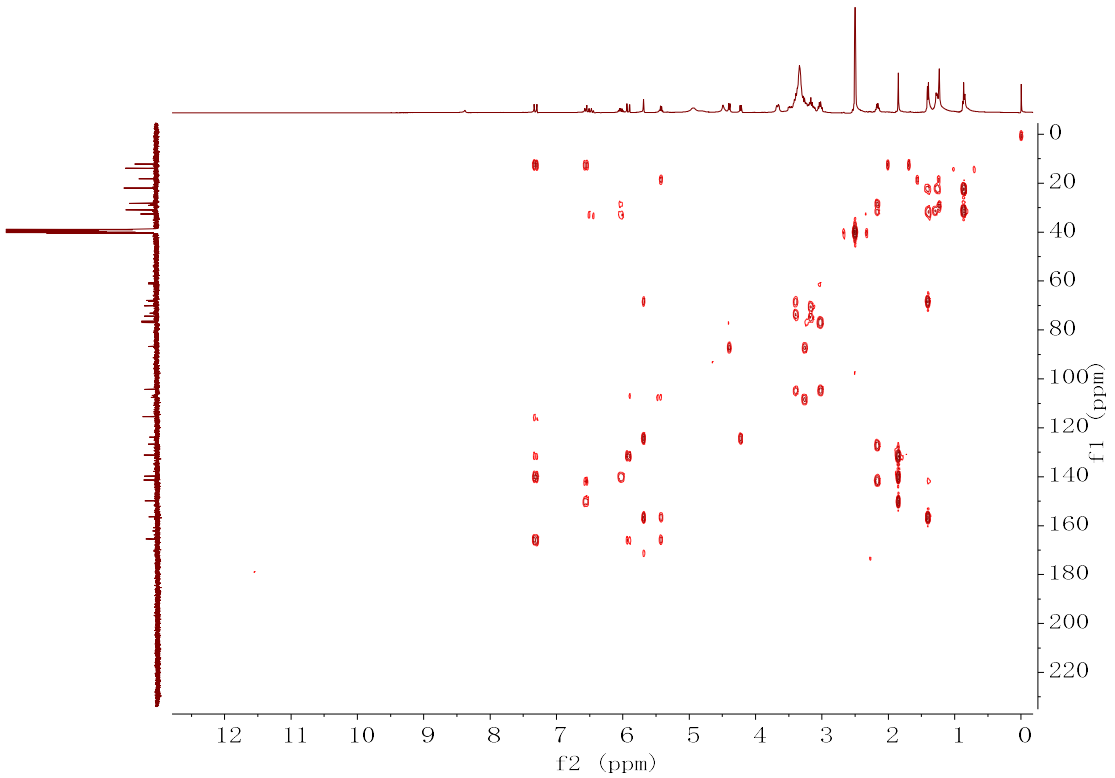


Figure S100. HMBC spectrum of compound **9** in DMSO-d_6.

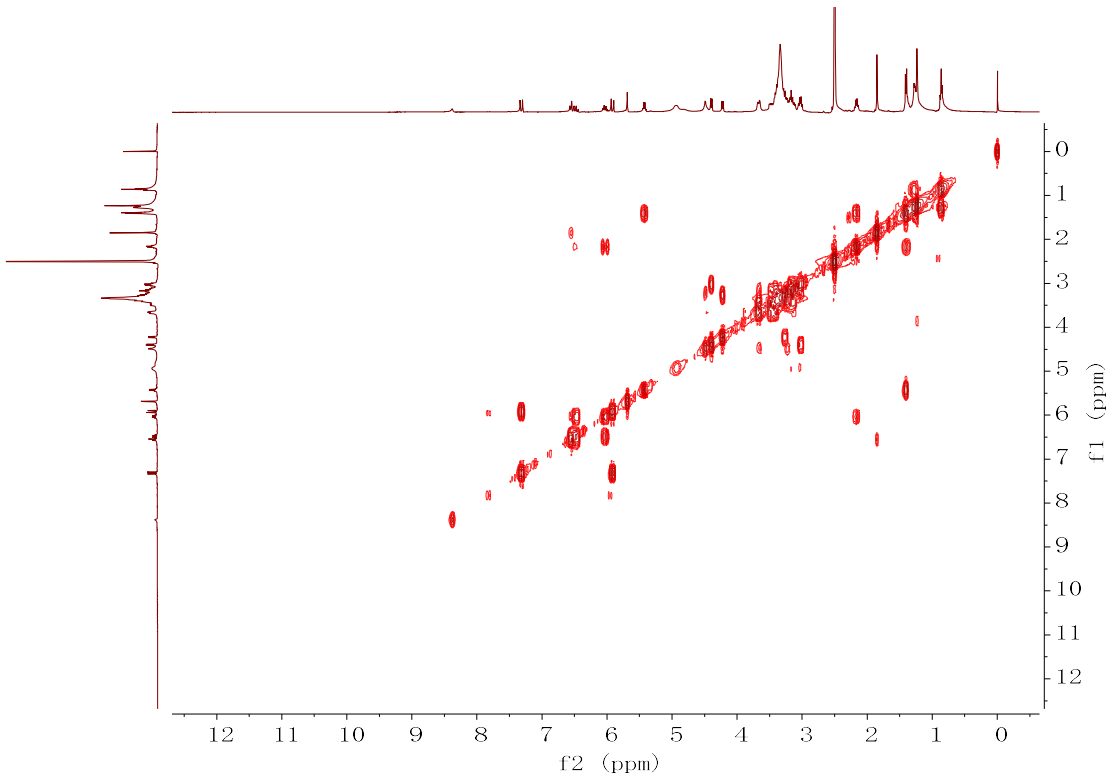


Figure S101. ^{1}H-^{1}H COSY spectrum of compound **9** in DMSO-d_6.

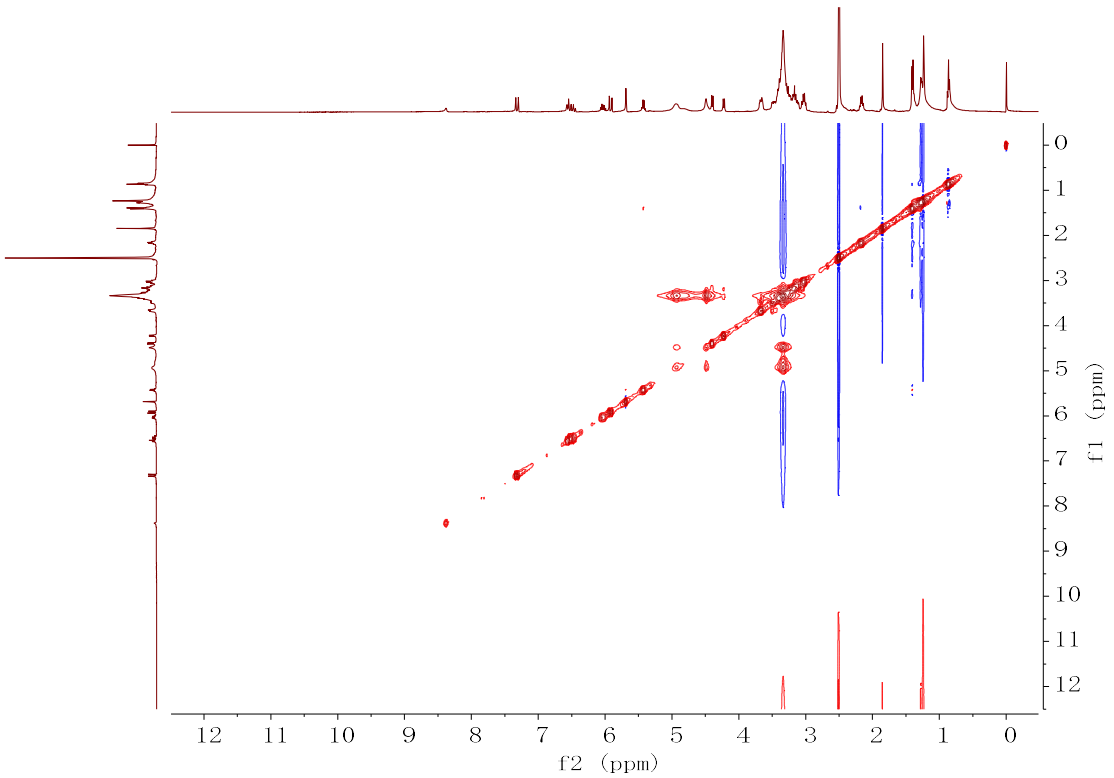


Figure S102. ^{1}H-^{1}H NOESY spectrum of compound **9** in DMSO-d_6.

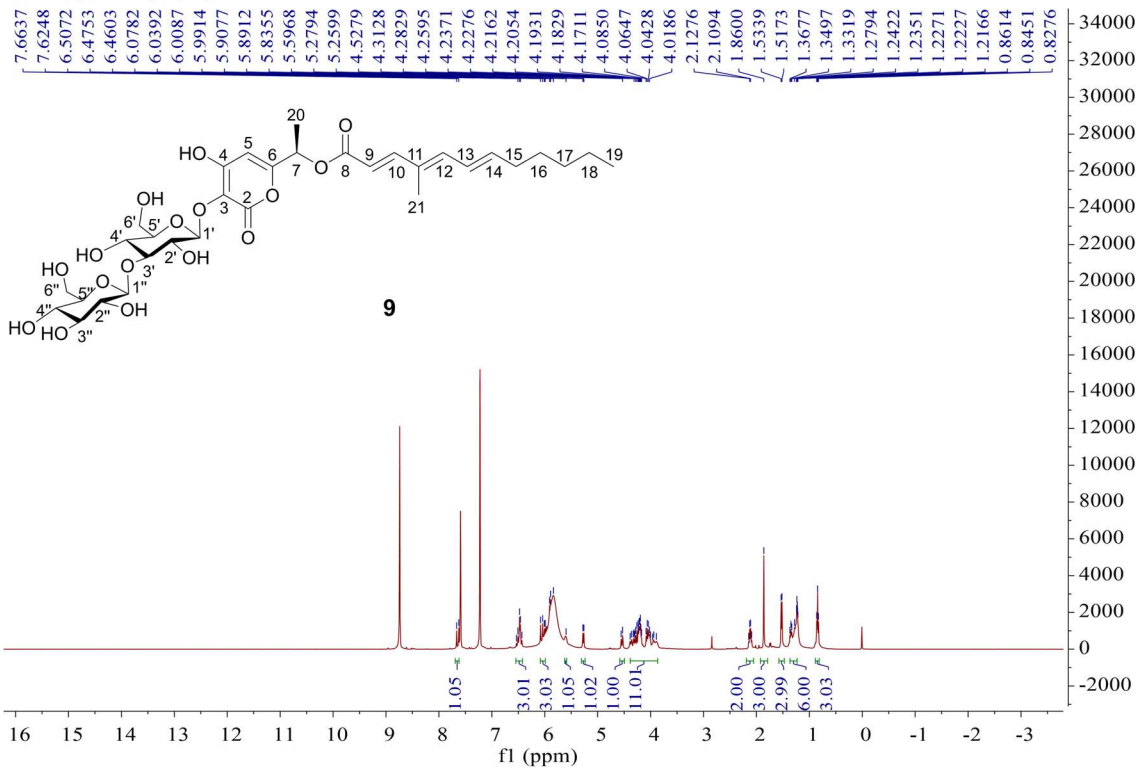


Figure S103. ^{1}H NMR spectrum of compound **9** in pyridine-d_5 (400 MHz).

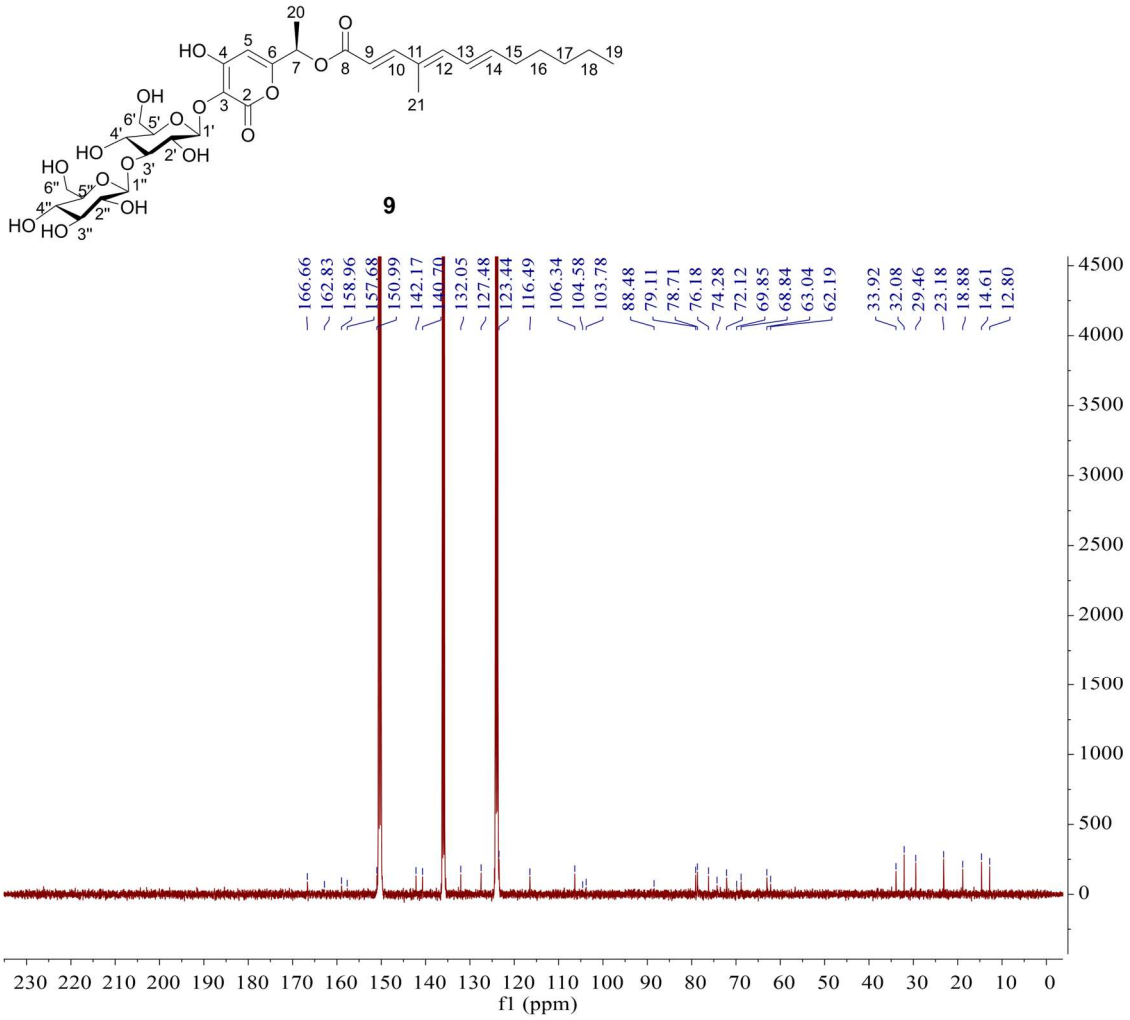


Figure S104. ^{13}C NMR spectrum of compound **9** in pyridine-d_5 (100 MHz).

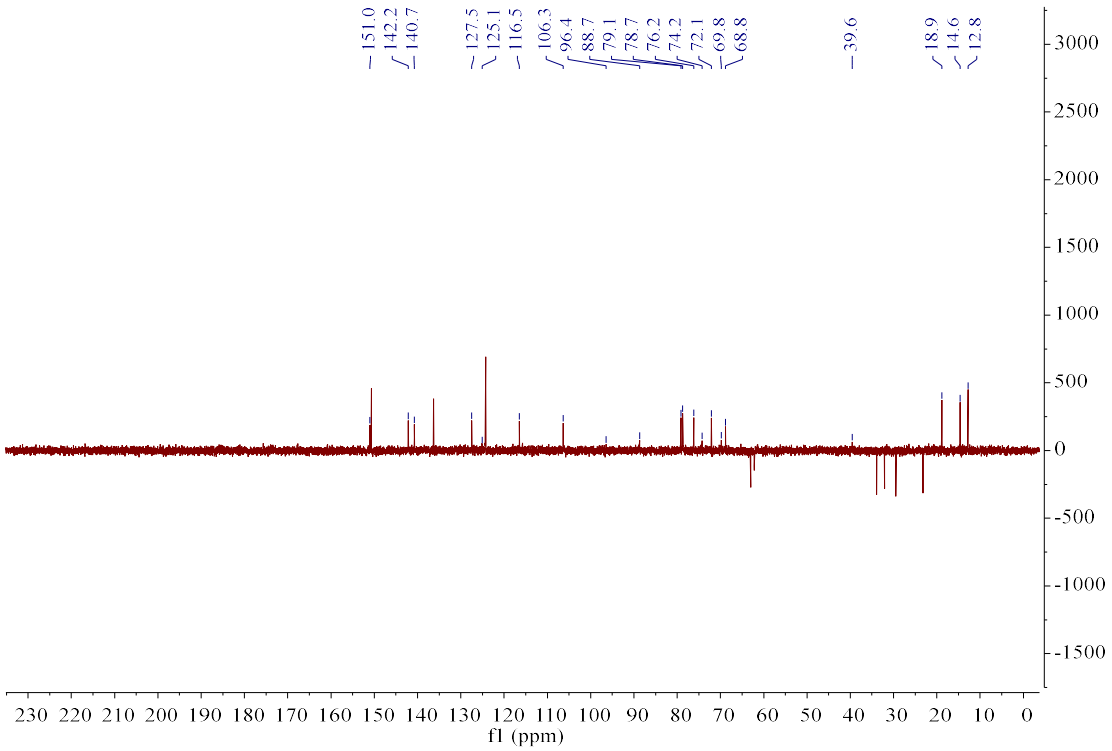


Figure S105. DEPT-135° spectrum of compound **9** in pyridine-d_5.

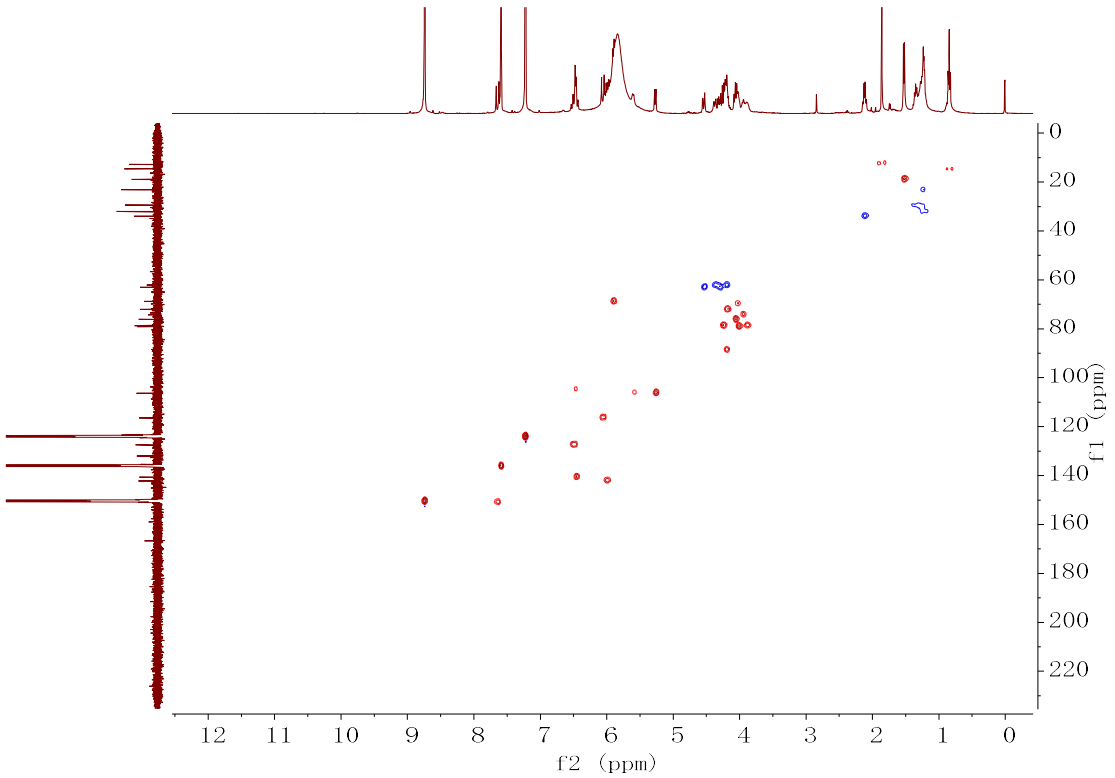


Figure S106. HSQC spectrum of compound **9** in pyridine-d_5.

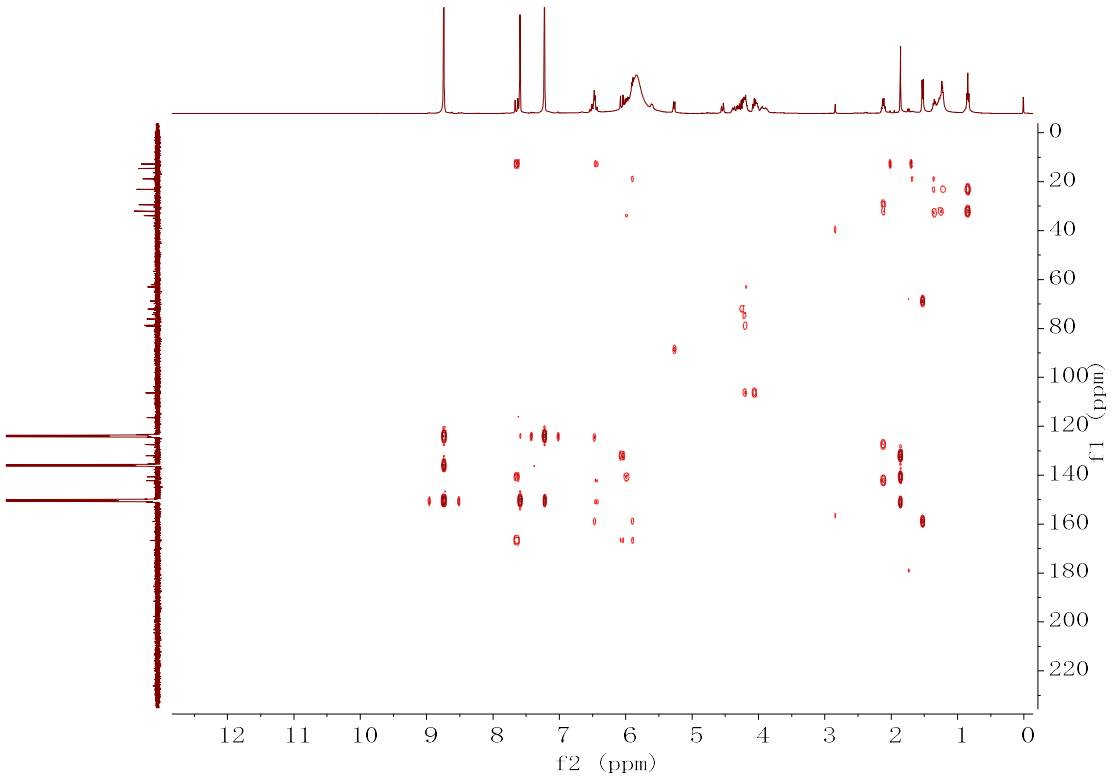


Figure S107. HMBC spectrum of compound **9** in pyridine-d_5.

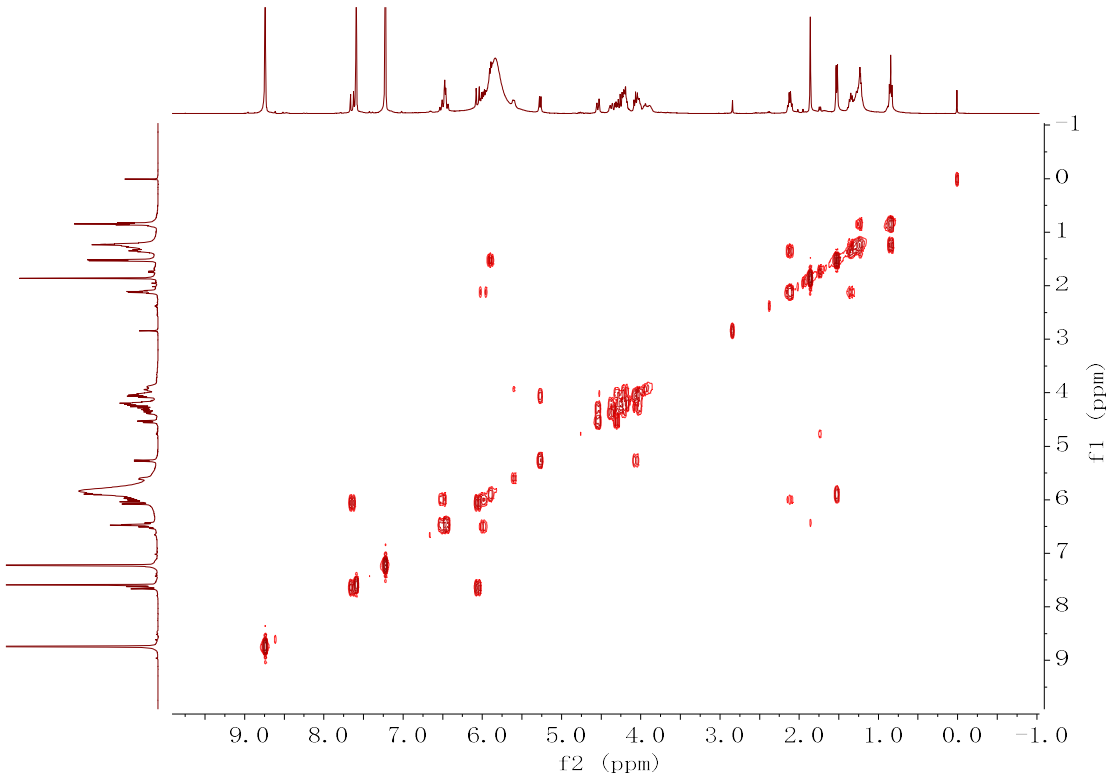


Figure S108. ^{1}H-^{1}H COSY spectrum of compound **9** in pyridine-d_5.

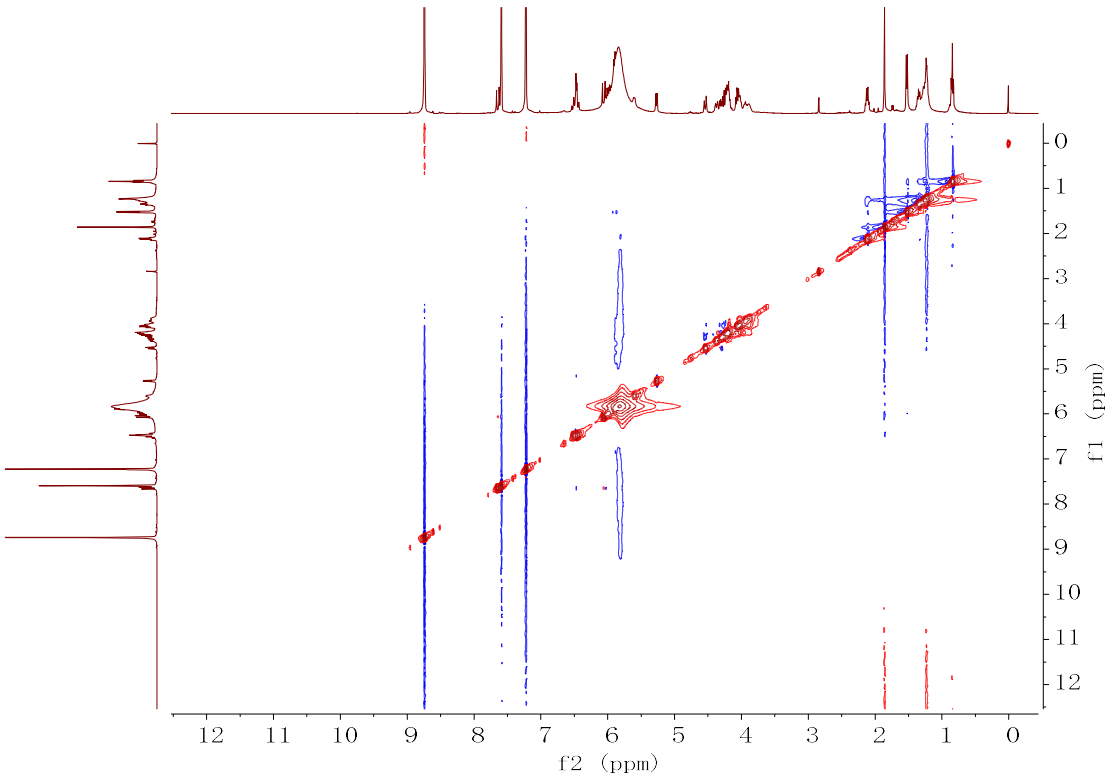


Figure S109. ^{1}H-^{1}H NOESY spectrum of compound **9** in pyridine-d_5.

4. Sequence information of *pro* gene cluster

ProA gDNA nucleotide sequence

ATGAAGTCCATCCCTTCAATCCTGGTGGCTGCCGCCATACCTTCCCTGGCCTCCGGTCAATCTT
TCCAGTCTACCCCGATTATGGGCTGGAACTCATATAATCAGGTTTCGTGCAGCCCCACCAACAA
AACCATCACCACAGCTATTGAGGCTCTCTCCAGCCGTGGTTTCGTCAATGCTGGCTACAAATTC
TTCCAGATCGATTGTGGCTGGGCCTCGCGAGATGCGCAGCGTGACCCTACCACGGGTGCGTTG
AAGATTGATTCCGATGCCTTCCCGGATGGACTTAAGCCTTTGAGCGACTTGGCCAGGTCCAAG
GGAATGAAGTGGACTATGTATTCAGACGCCGGTGTGCGTATGTGTGATCCCCAGTACCCATCAC
CTGTGCTTGGATCATTGGGTCACGAAGCCGTCGATGCCGCATTTTTCAAGTCGCTGAACACAG
AATATTTGAAATGTAGGTATTTCACATCCGAAACAAATAGCATCCTACTAACCAGGGCATGATTT
AGATGACAACTGCTACGCTGATAGCGCCAGCAACAATGCCCCCAAGGATCCACGAACTGACTT
TGTCACTCGTTTCGGCACAATGTGGAGCGAGTTGCAGAAGGTTGGGATTCCTGGAATGCTCAT
TTGCCAGTGGGGCGTTCCATATTCCTCATCCAGTGGCCTCGAAGGTCCCGCGGAATGGACTCA
GAACGTATCCACCTCTTTCCGCCTCTCAGACGATATAGCCGAAGGCTGGGGCAATGTGTATCGC
ATCTCCAACCAAGCCATCCACATCGCACATAGGGGCCTCTCAGGTCCTGGCCACATCGCCGAT
GCTGATCTTCTCGAGGTTGGTAACTCTGGCATGACATTTGACGAACAGGCAACCCATTTCGCTC
TATGGGCCATGCTCAAATCTGCTCTTATGATTTCCACGGATATCACGGCGCTGTCGGACCAGAC
CGTGGCTGTTCTTCAGAACAAGGATTTGATCTCAATCAACCAGGACGCAGCCGTTAAGCCCGT
CAGCTTGGTGCAGCGATGGACCAGTGACCGAGACTTGTGGGCGGGACCGCTTGCCAACGGGG
ATGTTGCAGTTCTCTATGTCGATCAGAGCAATAGTGCTCGAACCCTCAGCCTCCAGCTCTCTAA
CCTCGGCTACCAGTCAGCCGACATCAAGGATCTGTGGACGGGCAAAACTACCACTGGAGCTT
CTTCCTTTTCCAAGCAAGTCAACGGGCATGGCTCGGTCGCCTTACGTCTCTCAAACATCAAGA
AACTCAGTAGCACCGCCAATTACAAGTACACTTCTGTATCTACTGGGTCATTGTCTTCAGGGGC
TAAGACTGCCTCTTGTTCTGGTTGCACCTCCTCAACCAAGGTCGGAAATGTTGGTGGCAGCGC
CAACGGTCAGGTTGTCCTCTCGAACATCAGCACGTCAAAGGCCACTCAAAACGTTCTATTTGA
TTACATCAATGGCGATGTAGGTTTCAGCTTCACCGGTAGTACCAACGACCGCTTGGCTTCTATC
AAGGTCAATGGCGGCACTGCGCAGATTGTGAGCTTTCCACTGACAGGTTACAACTGGGACAA
GGATGTTTACAAGGGCTACGCCGTGGAGCTTTCTGGCTTCAATACGAGTGGACCGAATACCATT
ACCATTTCGGGCGTCAATAGTGGTTGGGCACCTGATTTTGACAGACTGGCCGTCGTTGCCTAA

ProC gDNA nucleotide sequence

ATGGCAGTTACTTCAATTTCACTTTTGGTGCTCGGTATTGTTGCTGGTGCAACAATTTTTTACTT
GTCGTCGCCAACTAGGAAAGACAGGAAGAGGCCCCTCAAACTACCCATCATCGGCGACATCC
ACAACTCCCCAATCGAGAAGCCACTGCTGCGATGGGATGCTTGGGTGAAAGAAAATGGCGCT
ATCGCAACGTCGAAACTGTTCGGGATAATGCCTGTTGTTGTTATCAACACAGCCGAGGCAGCG
ACAGAGCTGCTGGGCAAGAGGGGCGCTTGGTACAGCAATCGTCCAAGGTCCGTGGGCATGGA
GATGATCACAGGCGCTGGGCCAGGACAGTCACGATTCACGTTGATGCATGACATGGATGCGCA
TTTAAAGCTCCATCATCGCATCTTGTCTCCGTCGCTCGGTGGTGTCGCTGCTCCGCGCTATCAG
CCCGTCATGGAACTCGAAGCCAAGCAGCTCGTCAAGGACCTCGTCGAACTGAGCGAGCACAA
GTCGGTAGTCGTTACCAGCGACGATGTGTTCCCTTTCCTCGAACGCGCACAAGCGAGTATTATT
CTCGCATTGCACTATAGCGTTCGGGTTCCCACCCTTGATTATCCACTTTATCAACGGGTCCGTGA

GACGCAGGCTAAGGTCACCTCGTATGCGTCGAGGCCTGGCCTGCCCGATATCTTCCCATTCCTT
GCCAAGCTCCCAGCGGCAATTTCACCGTGGCGTAGAGCTGCAGATAAATTATTTAACGAGCAA
AAAGACCTCAACCTGCACTTGTTAAGCCTCGGTGATGACAGTCCCGGATGGAATGCAACGAA
ACAAGCACGCTCCCTTGCAGCGAAATACGCCAAGGAGCCGATTCCCGACATTGACCTCGCTTT
CACACTTGCGACGAGTGTCCAGGGTGGTATAGAGACTAGCACTCGTACAATACTCTGGTTGTT
CATCGCTGCGATGACAGCCAACAGGAGCTTTATGAAACGAGCGCATGACGTACTGGATGCTGT
TGTCGGGCGGGATCGACTCCCTTGCTTCGCGGACAGATCCTCCCTCTGTTACATTGACGCTATT
GTATCTGAGCTTCTGCGTTGGCGGCCTATTTCCCCGGGAGGTGTACCTCGTCGGGCAGACAAA
CAAGACTCTTTTAAGGGTATCAACATCGTCAAGAATGCGATGGTCTTGACAAATGCGTGGTCTA
TCGGCCGTGATGAAGCTGTGTTTGATCAGTCGCTGGGCGACCTCGACCAATTCATTCCCGAAA
GATGGCTAGATGGTGAAAGCATAGGCGTAGACATTAAGAACCAAGGAGAACTGAGGACTTCC
CTCCCGTTGCCAGTATTTGGTCACGGTCGGCGAAGCTGTCTTGGAAAACGTGTTGCTGTAGAT
GGCACCTTTGCCCAGGTCGCAACGATGATCTGGGCATTTGACTTTGAGCCTACCCAAGACGTA
GATGAAATGGAGATGGAGGTTGTCTGGTTTATGACGGAGCCCAAACAGTTCAAATTCAAGCTA
AAGCCGAGGGGTTCATGGGTATCCAAGGTCATCGAGGAGGAGTGGAGGACTGCTGATAAAGA
CCTTGGTAAGATCATGGGCAAAATTGGTGATATTGAGGCCTGA

ProD gDNA nucleotide sequence

ATGTTTACCGCTTTGTTTCTTTCTGCCCTGGCCTGGAGCCAAATGGCAAATGCCCATGGAACTA
TTACCAGAGTAATTGGTGCAAATGGTGTTGTCATGCCCGGACTTACAAGTAGGAACATCTCTG
GAAACTTGAGCCTGGTTGCTAACAGGTGGCAGTCCTTGATGGAACACCTCGATCTTCTACCTC
TGCTGCTTCCGGAGCACAGGTTGATACTAGCGTGATCCGTGACCCCGAACTTGGTACTAGTAA
AGCCTCCGCACTTGGTCGCACTAGCAAAGGCCCAGTTGACGGCGCTCGCGTCATCAAGGCTTT
CATGCATGGTCTCAAGGGCCGGTCCCTTGCTGACACTATCCTCGGGGGAGGCGAAGAAGGTAT
AGTGAAATGGCAAGCAACAATAGCTCTTGACAGCTAACACTCTATAGCTACGAGAGAAGCCGT
GTCTTTTGTAACTGGAAATGCTGGTGCAGTTGTAAACGGTAGGCTTTTGTTAATATACTTTAACT
CGAGTTACTGACCTATTCTAGGCGTGCAAGATGGAATTGAAAGCTCTCCAGTTGGCGGTTTGG
CCCTAGGTATGTCGACAGAAGCAGTTCAAAACTGGCATTATATGCTGATCGATAGTAGGGGCTG
AGCATGGAGTTAATGGTCTTCTGGATGACTTTTTCCAAACGGCAAAGGGTGTCCCATCTCCTCG
TGGATACATCGAAGACGGTGTCCAGAATTCCACTGGAGTTGGTGCCAAGTCTGGGTTGCCAAC
CACCGCATCCGATGGAACTCTCAAGTTGATCTACCACCAGGTTCGTGACCATGATTCTCATTTG
TGATCCATATTGACTTTTTGCCATTCAGGTCAATGAGGACGGTGCTGGCCCCTTGCTAGTAGAC
ATTGACTTCACTTCAGGAGGGACTGATCCGAAAGCTTTTAAATCGGCTGAAGTGGTTCAGAAC
ATCATTGGAGTATTGGGCTTTTCGACCGTGAGCTCGACTGACTTTCCCGTTGTTGTCAAAGTGC
CTACTGGTCAAGTTTGCACTGGAAAGGTAGCTGGCGTTTCTGGTATCTGTATTGCAAGGGTGC
GAAACAGCGCTACTGCTGGCCCCTTTGGAGGCGCTGCTGCATTCACTCATAACCCAGAGGCTG
CAAAAGGGAAGACGTCCAGCGCAAAATTTCGCCACCGCCATGTCTGA

ProE gDNA nucleotide sequence

ATGGACACATTCGACGTCATAATTGTCGGTGCTGGCTTGGCCGGTATTAACGCGGCGTACCGCC
TCCAAATTGCCTTGCCGAACCTCAGCTACGCCATACTCGAGGCCCGTGATGATATCGGCGGCAC
ATGGGACCTTTTTCGCTATCCAGGCATCCGCTCGGACTCTGATCTATACACCTTTGGCTTCGCAT

GGCAGCCGTGGGATCAGGGCACCGCGCTGGGTGAGGGCGGTGCTATCCTTAACTACATGAAG
AAATGTGCGCAGACTTACGGAATCGACAGGCATGTTCACTGCCAACGGCGACTAAAACATGCC
TCATGGTCTACACCCCAGCAGAAATGGACCCTTGAGGTAGAAGACATTGGCGCACAAAAGGA
GCTTCAGTATCATGCCCGGTTCGTCATGCTCTGCACTGGCTACTACGACTATCACAACGGCCGT
GACACGGCCATCCCCGGCCTAGAAAACTTCAAGGGCCAGGTAGTGCACCCCCAGTTTTGGCC
CGAGGATCTGGACTACACCAACAAGCAGGTAGCGGTGATCGGCAGCGGCGCCACGGCCATCA
CGCTGATCCCCAAACTCGCCGAGAAGGCGGCTCGAACTACCATGGTGCAGCGGAGCCCGTCG
TATATCTTGTCCATCCCCAACGGCGGCAAAAAGGCGCCATGGCTTGTACGGCTGTTCCCTGCGC
GGTGGGCTCATGCTTACACCCGCCTGTCGTTCCTGATATGGTCCCGCCTGATCTGGCTCTTTTG
TCAAACGTTTCCGGATCGGGCCCGTCAGCGGCTCCGGAACGGCGTCGAGAAGGAGTTGCCGC
ACCATCTGCCCTACGATCCACACTTTAGCCCACGCTACAACCCTTGGGATCAGCGTGTTTGCCT
GACCCCCAATGCCGACTTCTTCAAGTGTCTGCACACGGGCAAAGCTGACGTCAAGACGGGCA
ACATCAAAGAGGTAGTTGCCGACGGCATCGTGCTTGAGAATGCGCCCCAAGACAGGATCCCC
GCTGCTATAATCATCACGGCGACTGGCCTCAAGTTGCAACTCGCAGGTGGTGCCACCATTGAC
GTCGACGGCGTCCCTATCAAACCCTCTGAGCAGTACTTTTGGAACGGAACCATGCTGCAGGAT
GTGCCAAATCTGTCCCTTGTCATCGGCTACACCACTATCTCATGGACTCTTGGTGTTGATACCG
CCGCCATGCTCGTGTGCCGTCTACTGCAGAAGATGCAAAAGGCCAAGCTGTCGAGCGCCACG
CCTCGAGCTGAACAGGGCCTCGCCCTGACGCCAAGGCGTCTGTTTAGCCTCTCGTCGACGTAC
GTCACCACGGCAGAGAGCGAGCTCCCAAGAGCCGCAGCCCAAGCACCCTGGCAGCCTCGAA
CCAACTACCTCTCAGATTACTGGTTCGTCAAGCTGGGCAGATTAGACAGGGGTCTTCAGTTTG
TGCGGGAGGGAAAGCAGCTGTGA

ProF gDNA nucleotide sequence

ATGTCGGCATACACCAAAGATATCATGAACAATGGCGACGCAATTCCAGAGCCTCTCGTCCAG
GCTGTCTTGCAGCAGGACAAGTCCATGCCTATCGCCATTGTGGGCATGGCTTGTAAACTACCCG
GAGAGGCCACCAACCCAGACAAGCTATGGGAACTATGTGCCGAGAGGAAAGCCGCCTGGAG
CGAAGTTCCCAAGGAGCGTATGAACATTGATTCTTTCTGGCATCCTGATGGCGAACGAGGAGG
AAACGTATGATCTTTTGTCCTTGCCACATTTTTAAGAGTTTTGATACTGATAAAGTAATTCTAGA
CAAACGCAAAGGGAGGACACTTCCTCAAGCAAGACGTTGCCGCCTTTGATGCATCGTTCTTCT
CCATCCCTCCCACAGAAGCAAAGAGCATTGACCCTCAACAACGCATTCTTATGGAGACAGTGT
ATGAGGCCATGGAGAGTGGTGGAATGAGTATGGAGGACATGGCTGGCTCCGACACGAGCTGC
TACGTGGGAAGCTTCTCTCGAGACTACTATGAGATTATCGATCGAGATCCGGAAACAGCACCT
CTGTACTCGGTGACGGGTAACGGAGCTGCTATTCTGTCTAACCGTATCAGCTACTTCTATGATCT
CAAGGGCCCCAGTCTGACCCTGGACACAGCCTGTAGTAGTAGTCTAGTTGCCCTACATCTGGC
CTGTCAGAGCTTGCGTTCAGGAGAGAGCCATTGCTCCATCGTCGGTGCCACTAACCTGATTCT
CAACCCAGACATCATCACAGCCATGTCCAACATGCATTTCTTCAGCCCCGACTCACGATGCTAC
AGTTTCGATGCTCGTGCCAACGGCTACTCCCGTGGTGAGGGTATTGCTTGTTTGATTCTCAAGC
CCTTGGCTGATGCCATCCGTGACAATGATACCATTCGTGCTGTGATTCGAGGCACTGCCTGCAA
TCAGGATGGGAAGACCTCAGGTATCATGTTGCCTTCGCGAGAGGCTCAAGAGGAGCTTATCAA
GACGGCGTACAGGGATTCAGGTTGTGATTTTGCTACCACGGCCTATTTCGAGGCTCACGGCAC
TGGAACACAAGCCGGAGATCCCCTTGAGTAAGCCCATTATGCAGAAACAATGTCTAACTCTAC
TAACCTATGTTAGATCTGGAGCCATCGGTAGCACCTTTGGCCCCCACCGCCCTGTGGACGAGA

ATGGCTGCAAGCTTCCACTTTATATTGGTAGTGTCAAGACAAATGTCGGCCACACTGAAGGTA
CAAGTGGCCTTGCCGGTGTCATCAAGGCTGTCTTGTCTCTCGAACGTGGCGCTATTGCTCCCA
ACGTGTACTTTGAGACGCCCAATCCGCAGATTGACCTTGAAGGCTGGAACCTGCGAGTCCCCA
CCAAGGTGATAGCCTGGCCTACCAAGGGGCAACGTCGAATTAGTGTCAACTCGTTTGGATTTG
GTGGAACAAACGGACACGTCATTCTTGATGATGCGTTTCATTATCTTAATGGGCGTGGTCTGAA
GGGCGCTCACCGAACGGCCCCTAGGCCACTGCTGGACTTGCAAACTGGTATTGCAATGAGCA
ACGCAATTGCTCCCGTCTTGTCTAATGGCAACGGCGTCAATGGGCATACAAACGGCCATACTAA
TGGCCCTACTAACGGCTATACCAACGGCCACTCCAATGGTCACTCTCTCTCCTCTGCAAAGCGC
CACCGCGTGTTCATCTGGTCTACTCATGAAGAGAACATTGCTGGAAAAAGCGCGGCCGTATAT
GCTGAGCATCTGGCCCACCGCAAGGAGGCAGACGAGGAAGCCTTCCTCGACAACCTCTCGTA
CACACTCTGCAGCCGTCGTTCCATGTTGCCCTGGAAGTCGTTCCTGGTGGCTCAGTCTCTACCC
GACTTAGCCGAGAAGGTGGCCGCGACGCGCCAGAAGCCAGTGCGCGCATCCATCAGCCCGCT
CAAGATCGGCTTTGTCTTTACAGGCCAGGGCGCGCAGTGGTTCGCCATGGGTCGAGAGTTGAT
ACAGACGTACCCCCTCTTCCAACAGCGGCTGGAATATTCTGACCGGTTTGTCAAGAGTCTTGG
AGCCAAATGGTCTCTCATTGGTAAGTAACTCTGGGTCGTGGATCTTTGGGGCCGATACTAACTG
GCATAGATGAACTGAACAAGGATGAGGAGACATCTCGCATTAACGAGTCCCTCGTTAGCCAGA
CCGCTTGCACGGCCCTCCAGCTTGCCTTGGTTGACCTGCTCGCCAGCTGGGGTATTCGTCCCC
AGAAGGTCATAGGCCACTCGAGTGGTGAGATCGCCGCCGCTTATGCTTCAGGTATCTTGCCAA
CGGAATCAGCCCTCAAGGCGGCCTACTTCCGCGGTCTGCACTCTTCTCGTGTGCAAGAAAAGC
TGCCGTCAGCTGATGGCGCCATGATGGCTGTTGGTCTTGCTGAGGAAGATGCCAACAATTACA
TCAGTGCCCTTGCTCCGGAGCTCGGGAAGGCATCCGTCGCCTGCATCAACAGCCCGACCAATG
TCACTGTCTCGGGTGATCGACCTGCCCTAGCCGCTCTGGCAGAAGCTCTCCAGAAGGATTCCG
TCTTTGCGCGGTTACTCAAGGTCTCGACGGCGTATCATTCCCACCACATGGATGCTGTTGCTAG
CGATTACGAGCACTCTCTGGCTGGTATGGAGGTCAATACGCCCTCGGAAACTATTGAGATGGTC
TCGGCCGTTACTGGAGAGCTATTGAAGTCCTCTGACATCCTTGGCCCTAAGTACTGGGGGGAG
AACATGGTGTCCCCCGTCCGCTTCAACACTGGTCTGCAGACTCTCTGTGCTCCCCGTGCCACT
GCCAAGCGTACTCGTCGTCTTGCCGCCAATCAAGCATCTGTAGACGTCTTGGTAGAAGTCGGC
CCTCACGCTGCCCTGGCAGGTCCAATCAAGCAAATTCTGGGCGTTCCCGTCTTGGAGAAGAGC
GGCATTGTATACAAGTCTGTGCTATCTAGGGGACAGGATGCTTGCGCGACTGCTCTTGAGGCTG
CGGCATTCCTCTTTGCCAAGGGATGTACGACTATTGACCTCTACAAGGTCAACTCCCCCGCTGG
GTCCCATACAGCACAGCCGCGTGTGCTCGTTGATCTGCCTCCATACTTCTGGAACCACACCAA
GAGCCACTGGATGGAGTCGCGGTTGAGTATCGAGCATCGCTTCCGCCGTCATGCTCGCACTGA
CTTCTTGGGATATCCTGTCAATGACTGGAACCCTATGGATCCACGCTGGCGCAATCTCATTCGT
CTCCGCGAGCAGCCCTGGCTGAAGGGCCATATCGTGCAGGGCTCATATGTCTACCCTGGTACA
GGCTACTTGTGCATGGCTCTCGAGGCGATGTATCATCTGCGACAAATGCCTGAGTATACAGCAC
CTCAGGGTGACTTTGTGGGATACCGGCTGAAAGATGTCAAGCTGATTCGCGCTCTGATGGTTC
CAGCAAGTGAAGAGGGAGTTGAGACGCTGTTCTCTCTGCGTCACTACAAGGAAAGTACAGCT
TCTTACTCCAACAATTGGTATGAGTTCCGTGTCTTTTCTTGGTCTGCCGCTGATGGGTGGGCTG
AACACGCTCACGGTATGATTACTGCCGTCTATGATGCCCCGAACACGGCTCATTTAAACCGTAT
TCCCTTCAACCCCACGATGGATATTGCAGTGGACTTGGAGAACTTTGCGGGAACCCGCACAGC
TGAGAACTTGTATGACATCGTCGATGCTGTGGGTTTGGTGTACGAAGAACCATTCCGGAACTT
GACTGGAGACCTCGTGTCCAACGAGGGCACAGCAAAGGGGGTCGTCACTGTGCCTGATACCA

AGTCCCTGATGCCTTTTGAGTTTGAGTATCCTTATCTCGTCCACCCTGCCACTATGGACGGGTTT
GTCCAGATGGTGTTCCCCGCGTTGCTACATGCCCAGACCTCCATCGCGTCTCCGTATCTACCAG
TCTTCTTCGCCGACACTTTTGTACGAGGCGATATTGCTCAAGCTGCTGGCCATCGGTTTGACTG
CATCGCAACTGCGGGATACACCGGCTTCCGAGAGGTCACGACCAATGTTCTTGTCCGCGACGA
GGGCACCGGCGACGCGGTGGTCAGCTTTAAGGACGTCAAGTGCAGCGGAGTTGACTCTGGAG
AGAATGCCAATGACGGTCTATCGGGTCGAGAAGCTATCAAGAAGCTGTGCTTCCACTCCACGT
GGCACCCGGACCCAGAGCTCCTATCCCAGGACAAGGGTGACGAGTTAATGCGTTCCTTTATTC
CTGTCCCTGAAGACCCTCAGCGAGTGGCTAACCTCGAGTCGATCGCCTACTACTACTACTACC
GGGTTCTGCAGACTGTTACTGAAGACCAGGTTCCCAGCATGAAGCCTCACCACCAGAAGTTCT
TCCGATTCATGCAGTACCAGAGTGACCTCGTTTTGGCTCACAAGAGCCCTCACCAGACACCCG
AGTGGGAGCAACTTGATGACCCAGTCATCAGCCAGAAGATGGAGAATCTGGCTGTCCGTCTG
GAACCCTCTGGTACCGACGCTCAGCTCATCTGCCGAGTGGGCCGAAAGCTCGACAAGATCCTC
ACCGGCGTCGTCGACCCGCTTGCGCTGATGCTGGAAGACGAGCTTCTGTACAAGTACTACGAG
TCTGTCATGGGCTACGACATCATTTACCACTACGCCGAGATCCTTGCCAACAAGAACCCTAATA
TGCAGGTTCTGGAAATTGGTGGCGGTACTGGTGGTGCTACGGCCCCAATCTTGTCAGCTTTCG
GCGGAAATAACGGGAAGTACCCCAAGTTTACCTCGTACACTTTCACCGATATCTCTTCGGGCTT
CTTCGAGGAAGCCGAGACCAAGTTCAAGGACTGGGAGGGCCTTGTTGAGTATAGACGACTCA
ACATTGAAGAAGACCCTGTTGATCAGGGTTTCGAGGCTGGAAAATATGATCTTGTCGTTGCAG
CTTCGGTGTTACACGCTACGGCCAACATCACTCGGACTCTGGAGAACACGCGAAAGCTGCTTA
AGCCGGGTGGTCGTCTCATCTTGGTTGAGATCTCCAATATACTCAACCAGGCCTTCCTCTTGTT
TGGCTGCCTGCCCGGCTGGTGGATGTCTGAGGAGAGCTACCGCCCTTGGGGTCCTACTATGGA
TCAGGATATGTGGGCAGAGCACCTCAAGCGTGTTGGGTTCAGTGACCTGACTCTCGCTGCACC
AGATAGCGAGGACCCCAAGGACGAGCTGGGCCGGGTCTTCTCTTGTGTTGCTGTTGAGCAGC
CTGCCCCGCCATCCGTCGACCCATCTATTTCCTCGGTAGTACTGATCGTCGATGACGGCGCACC
TAGTGATTTGGAGCAGTGTATTGACGAGAAATTTTCGAAGCTGGGTGTTCCAGTTCGTAGGAC
TCGACTTGCTGAAGCGGCAGACACACCCCTAGAGAACACTTGTTGCGTGTCTCTGGCTGAGG
TCAACCGCTCGGTCATCAAGGATATGACATTGGCCGAGTTTGAAGGCATCAAGACTATCTTCA
AGTCTAGCCGGGCTTTGCTCTGGGTGACCGAAGGAGCATTCAACAACGCAAGCCGTCCAGAG
TCGGCCCTCTTCCACGGTCTTGCTCGCTCACTCCGGGCTGAGAACGAGACCTTCCCGTTGACC
ACGGCGGACTTTGCGTCCATCAATCGTAAGAACCCTTCAGAAACTGCACAACAATTGCTTGAG
CTGTTGAAGCAAGTCCTGCGGAACCCCGGAACTCAGGAGCCCGAGTATCACTATGAAGATGG
CTCTTGGAAAGTCTGCCGCTGCCTCGAGTCTGTTGATGCCAGCGCTCAGATCCATGGCTGCCT
GCACCATGAGTCAGCTGTCAGCGATAAGACAGAGTTGCAGCCCTTCTACCAGCCCGGTAGGCC
CTTGAAGTTGAGCATCAAGACGCCTGGTCTTCTGGATACGCTGCGCTTCGAAGACGATCCTCT
ACCAGCCGAACCTCTTGGCACTGACGAGGTTGAAGTTGAAGTCAAGGCATCCGGCGTCAACT
TCCGCGATATCATGATCTGCACGGGACAGATGTCTGACCCTTCGCTTGGACTTGAGTGTGCTGG
CATTGTCCACCGCGTCGGAGACAATGTCTCTAATGTCAAGGTCGGTCAGCGTGTGGTAGCGTG
GACTCGATATAACTACTCCAACTATGCTCGTACTCCCGCTTGCGTTGTCCAACCAATCCCTGAC
GACATGAGCTACGCCACCGCGGCCTCCTTGCCTATCGTATACAACACCGTCATCTATGGGTTAC
AGCACATGGCTCGTCTCCGCAAGGGCGAGTCGATCCTCATCCACGCAGCTGCTGGAGGTGTTG
GTCAAGCGGCTATAACTCTAGCCCAACGTCTTGGGGCGGATATCTTTGTAACGGTTGGAACTCA
AGACAAGCGAGACCTCATGAAGAGCGAGTTTGGAATCCCTGACGAGCGTATCTTTTCTAGTCG

TGGCATTAGCTTTGTGCAAGACATCAAGAAGGCAACCAACGGCCGCGGCGTTGATGTCGTGCT
CAACAGCTTGGCGGGAGAGATGCTGAGTGCTACATGGGACTGCATTGCTACATTTGGCCGCTT
CATCGAGATAGGCAAGAAGGATATGATAGACAACCGTCGTCTCGATATGGCGCCGTTCCTGCGT
AATGTCATGTTTGCCTCTATCGATCTCATCACCGTCTATGAGCAGAACATCTCGTTGGCAGGTG
AGCTTCTCCATGAGACGATGGACCTCATCCGAACCAAGGCCATCAAACCTATTCCCCGCATCA
AGATCTTCTCATTCTCCCAATTTGAGGAGGCCTTTCGTTTCATGCAGCAAGGCAAGCATGTCGG
TAAGATTGTTATGGTGCCTGAAGAGAACGACATGGTCCCAGTAAGTTTCACTTTGCAATGCCAT
TCCTGCTTGTCTAATAATTTCTACTAGGCTATCCCAGCATCATCGGGTCCCTTCACCTTCCAGGA
GGACGCGTCATATCTCATTACTGGATTTGGCGGTCTCGGTCGCAGTATGGCACGCTGGATGGCT
TCGCGAGGTGCCAAGAACCTCATCTTTGCTTCGCGTTCTGGAGACAGCCGCCCTGAAGTGCGC
GAGCTCATTGATGAACTTGCTGAGGTTGGCACGCGAGTTAAGGCTCTGCCAGTTGACATCACC
AATGGACCAGCACTAGATGCCGCCATAAGAGGCATTGCCGATAGTTTCCCACCTCTCCGTGGT
GTTGTTCAAGCTGCCATGGTTCTGGACGATGCCATCTTCGACAACATGTCTCTCAAGTCGTTCA
ACAACGGCATTAGGCCCAAGGTCCAAGGAACTTGGAATCTGCACCAGTCTACGTTGAACCAA
CCGCTCGACTTCTTCGTCATCATGGCGTCTTCAGTCGGTGTTCTTGGTAACTCCGGTCAGGCCA
ACTACTCAGCCGGTAACGCGTACGAGGATGCTCTCGCACAGTACCGCCGCTCACAAGGACTCC
CTGCCATCTCGGTTGATCTGGGTATGATCTTGTCTGTGGGAGCCGTGGCTCAGGATTCTACGGG
AATGATCCGCAACAATCTAGAGAGCAAAGGATTCGTGGGTATTGAGGAGGATGAGTTCCTTGC
CATCCTCGAGACTTCCATTCGAGAGTCATCCTCTTCATCAGCTGCCAGTCAGATGATTACTGGC
ATCCAGACACAATCAACCGTTCCAGGTGACGACGGCGTTCCCGTTGACGAACCCTTCTGGAAA
TCTAGCCCAGTCTTCTCTCACCTCCCGAAATTGGGCGCCCGCTCTTCCGGTCAGTCCAACAATG
CGGGCGAGCAGTCCATTCAGTCTCTCCTCAAGGGAGCCCTCTTACTAACTGATGCTGTGACGG
TCATCATCGACGCAATGACAATCAAGCTCAGCCGCAGTCTGATGATGGACGTCGCAGAGCTGG
ACCCGACCCGCCCAACCAGTGCATTTGGAATAGACAGCTTGGTGGCTGTCGAGCTCCGAAACT
GGTTCCAGAAACACATGAAGGCTGATATTGCCGTATTTGAGATTTTGCAGACCAATAGTCTGCA
GACGCTTGCCTTTCGAGTTGCAGAGAAGAGTACTTTGGTGGAAGGCTCCCTGGCAGATGGAC
AGTGA

ProG gDNA nucleotide sequence

ATGAGGTTTCTTTGCTTGCATGGTGCCGGCACTAACTCAGATGTACGTCGTATTCTGAATTCTCC
CCATGTCCCTCACTGACTGGTCTATTTGTAAATCAGGTCCTTGACATTCAGACTGGTAATTTAAA
TACTCTGTATTAACGGGACATAATGTAACACAAGATCCAGGACCAATCCGTGATTCGCTCGACA
GCAACGCCACGTTTAAGTTCTATGATGGGTTTTGGGATGTAGAACCAGTGGAAGGTAACTCGC
CTTCTCAACTCCATAATGCCCCGTCATGCTTAACTTGGTGGATCAGAGATCAAGAACATATTTG
CCGGTCCCTTCTACACTTGGTACTCTCCTGGCCTCGGCGGTCGCACTCTCACAGAAGCCAAAG
CCGAATTGCTCGATCTCATAGCGACTGAGGGACCATTCGACGCCTGCATGGGCTTCTCACAAG
GCGCAGCCCTCCTTGCAGCTGTCATCATCGATCATCAGGTACAGAACCCATTCGGGCCCAACCT
ATTCAAATTTGCCGCGTTCATCTGTGGTGGGAGTCCCCTGCTTGTTACCAAGGCGCTTGAGCA
GGACCATCTTGACTACCAGCCAACTGTAGATAGAATGGCCCCATTGACGGAGCCTTGGCTGGG
TCCATATGTACCGGGCCACGAGCCACATCCTGACGAACAGTGGAACATGCTCGTGGCCCACAG
AGTAAGGGAGGCTGGCTTAACCATTAGGATTCCCACAGCGCACATCTATGGGGCTAAGGATCA
CACCGTCAAGGAAAGCCTGAACTTGAGGGACATGTGTGATCCCAGGAGGAGAGTTGAGTTTG

ATCATGGAGGAAGACACGAGGTACCCAGAGCCACTAGAGTTGTTCAGCAGATGGCAATGACG
TTGAGGAGGGGAATTGATAAGGCCATGACCGCTGTATAG

ProI gDNA nucleotide sequence

ATGACAGACAAGAAACCCATTCGTGTTGCCATCATTGGCGCCGGCCCAGGTGGTCTCGTACTC
GCGCAGACACTTCGACAAGACCCTCGCTTTAGTGTGACTGTCTACGAGCGCGGCGTGCGTGAT
GGTAGCGGCGTCAGTTCGTTGGTTGGTTTCCGTATCCTTCTTCCACCCTCTATCCTCGATAATTT
GCGCAGCCAGCTTCCTGCATCTGTAGCGACTCTTATCGATGACGCCATTGGCGTTCCACAAGCT
CAGGGAAATCGCGTGGCGTTTATGGATGAGCAATGCGGGATCATCTGTCGCCTGGATGTGCAG
CAGTCCAGGGACATGTGCTCTGTATCAAGATGGAAGTTGAGGGAGGCTCTGCTGCACGACGC
AGAAGAGATTGTCCAATTCGGCAAACAGTTTTCATCTTATGAGCAACTAGGTGGAGAGTCGGG
CGACGTAAAGGTGCGCTTTGCAGATGGGGATGAGATAGAATGCGATGTCCTAGTCGGAGCCGA
CGGTGCCGGTTCAAAAGTGCGGAAACTGTTACTACCAAACAGCCAAAGGAGTGCCAGTGGCC
TTACTGTTGTCTACTTCAAAGCACCATTTACTCCCGAGACTGAAGCCATGATTCCTTGGAAGTC
TGGATGTGTGGTAGGTTTACACTATTCGCTTGTAGCTGAATGCGACTGCTAACATCCAGCTAGG
CGATTACACCCAGGCGTTCCATGGTGGTGGCTTACTACAAAGATCGCCGGAGACCATATGGAC
CCTATGACTTGGAAAAGATCGACCCTGCCGACTCGTTTCTCATGTTTGGGCTGGGCTGCTACAC
CAACGAGTTCGTGAACCAGAGCAAGCATCCCGACGAGATGACTCCCGAGGAGCTGAAGGATG
AGTGTTTGGCCAGAGCCAAGGATTGGAGTCCGCTTCTTCGGGCTCTTATTGCGCTGAGTGTGC
CCAGCTCAGTGTTTGTGTCGCATGTCAAGACACAAGATCCGATCAAGCCATGGGAGTCGGGA
AGGGTCACTCTTTTGGGGGACGCCGCTCATAGGTGAGTCTGTCACTCTTAGATTGCCGGTGAG
AATGAATCGGGTGTTAACCAAGTCGACAGTATGACCCCATACCTAGGCAAGGGTGCCAGCAGT
GCCATGATAGACGCCATGTCCCTTGCCAAAGCTCTCAAATCCGAGCCAAAGCAAGGCCAAGGT
GACTTTCTCAAAGCGCAGCTATCCATCTATGAGGAAGCAATGCTGAAGCATGGCTTTGAAGCA
GCTAGACAGAGCATGACTGCGCAAAAGTTCACTTTCAATGCTGGAGACACCCCGTGGAAGTG
TTGGTGGCGGAACCTGGCGCTCAAGGCGTGGGACTGGTGGATGTCACATCCGCCAGCGATGG
AGGAGAATTTCCCAGGTATGTGTCTTTGGCTCAAGTATATCGCACGAATTACTGACTGGTGGCT
AGTGAGCTATAGTGAGATGAAGAAGGACGTATGA

ProJ gDNA nucleotide sequence

ATGCCTAGTTCCGAGTTGGCGCACCCGTCCCGCAAGATCCTGCTTGTGGTGACGACGGGCGGA
TTCACCCATGCCAGTACGTTAGACTAGCCCCTCTCCATATCACGCCACTCACTTCCAAGTTCTA
ACCCTCGCCCAGGTCCCGTATTCGAGATTGGCAAAATTCTGGCCGTAAGAGGCCACACCATTG
AGTTTGCTACGCTCGAGGGCCAGGAAGCCTGGACTAAGGGCTATGAGTTCATCACCAAGGTCC
ACACCCTGGGCCCGGGCCCGACTCACGACCAGATGAATGCCCACTACCTGCGCATGCGAACG
TGGGACATCTCAAAGGGCATCTCGGGGACTATGCCCAGCAAGTATCTATGGGACTCGTTCTGG
CCACAGACATATCGGGGACTCAAGGCCATCATGGACGACCCCAAGACAAGGCCAAGCATGAT
GATTGCCGACTTCTTTGTTGATGCCGTCAAGGATATTCACGTCGAGTATAATCTGCCAATCACTC
AAGTCTGGCCCCAGATGCCCTTCTTGATGATGCCATGCTCGTACATTCCTGGCCAACCGGGCTT
TCAGCTTGAGGGAACACTTACCTCTGAGCATGCTAGCCTCTGGCATCGGATCAAGAACGAACT
CGTCATCTTCTTTGACCTCCCCGTTATCGTGAAATGGATGAAGTGGACCAAGAAGATGCGCCTC
GAAAATGGTGTCAAATATCCTCCTCACAAGATCCAGAAGCCTGACTACCTTATTTTTGTTAACT

CCTTCCTTGGCCTGGAGATTCCGCGTGATCTACCTCCTACCTGTGCCCCGGTTGGACCCTTGAT
CAGTGACACTTATCCACCACTCAACGAGGAGTGCAAGCAGTTCTTGACCAAGCACACCAAGG
TCATCTATATCGCCCTGGGCACACACATTATCCTCACCAATGCCGACGCAGCCAAGATCATCAA
CGGTCTGCTGCTTCTTCTTGAAGGGAGCTTATTGGACGGCGTCATTTGGTCTATTCCCAAGAGT
GGCCGTCAAGATCTCGATGTAAACACGACCTATAAGACTGGTAACAAAACGCTCCGCCTGGGG
GATATTCTGGATGGAAAGAACCCTGACTGGCTGTGTTCCACCTTTGTCCCCCAGCGAGCTATCC
TCGACCACCCCTCGACAAAACTCTACTACACACATGGTGGTGGTTCTAGCGCCAACGAGGGTC
TCTTCCACGGGGTACCCATGCTGTCCATGGGTGTCTTCATGGACCAAATCTCCAACACAGCTCG
ACTCGTCGATGGAGGCGTGGCTGAGCCGTTAAACAAGTTCCGCTTTACGTCCCAAGAGATCTA
CATAAAGGCAAAGAAGATTCTCCTTGATGGAGATGGAAGCTATAAGCGCAATAGCCTGCGGTT
GATGCGCATCGCTCATGTGGCTGCCCGACGGAAAAAACATGCCGCGGACCTGGTCGAGGAGC
TGATTTACGACACTGAGCTGAGGTTCAAGGATGGAAAGGAACTCAGGCCCATGCATCTGCAG
ACGGCGGATATGCGTATGCCAGTGTGGAAAGCCAAGAACTGGGATATTTGGGCTGTCAGTCTT
CTGGGCATTGGTGCTGTCTTTGGAGGACTGGGGATTGGCGGAAGAATGTTGTGGCTGCACCGG
GTTTGGCTGACTGGATCAGTGAAGGGATTCGTGGGAAGCCAAGAGTGGTTGAGGAACCTGAT
CAAGTAG

ProK gDNA nucleotide sequence

ATGGCTAGGGAGTATTCCGGTACTTGCATCACAGCTGTTGATACGCCGTATTTGCAAGCGTCTA
AAAGTCTGAGCGAAGTGGCATTCGCAGTTCGTGAAAGGAGTCGCTGGGCTAACGCCAGAGAA
ACTGGCCAAGGGCGATGAGTGGATTCAGTCAATTCAAGCTATCAACGGGCTCATGGACTTGGC
CCCCGCATTTGACACTGATACTTTGTATGTAGATGATTGGACCAAAGTTTCTATGGAGGAGATG
AATTTCGGACAGGAGCAGCATATGTTCTGCCAGCCGTTAGAAGTGGAGGCTCTGCCAGTCCGA
AACTGGTGCATGATTTATTCGGCAAAGGCTGCTAATGCAGGAGACGCGAATAGCCTGGGATAC
CAGGTACAAGTCGTGGTGCTAAAGGGACGCGCAGCTGAGCTCATTAATGGGATAGTCACGGAT
CTTCAGGAGGGATTTGACGAGTACTTTTGGTAG

ProL gDNA nucleotide sequence

ATGTATCGACTCAGTTTGGCATCTTTACTATCACTGGCCTTGCCCGTTGTAGCAGGATCTCAATT
ATCATCAGGTTAGATCGTCCTTCGCACCTCGGCCGAGGGTTACTGACGTGGTACAGCAATCGA
CCCCTCGCAATGGCTCAACGCTCGTGGTAAAGCAGACTCACTGACCTCCAAGATGACACTGG
AGGAAAAGTCCTCGATGGTGACAGGCACCTTTGACGGCACATGCATCGAGCACATCGCACCC
CTCAAACGTCTTGGCTTCGGCGGTCTTTGCATCCAAGACGGACCAATTGGTCTTCGCCTTGGC
GATCTAGTCAGTGTCTTTCCATCGGGCGTTACTACCGCTGCAACGTGGGACAGGCAGCTGATG
GCGCTGAGGGGAGAGGCAATGGCGGAAGAATTCAAAGCAAAGGGTGCCCACGTGATTCTTGG
GTCAGATCTGCCACTCCTCCTCGTGTTGGAATCACTGACGATAGATTAGACCCGTTGCAGGAG
CCCTGGGAAGAAGTCCATACGGCGGTCGCAACTGGGAGGGCTTCTCGCCAGATCCATACCTAA
CCGGCATCGCCATGGGTGAGACAATTCGAGCAATCCAAGACACCGGTGTCCAGGCTACAGCC
AAACATCTCGTCGGCAACGAGCAAGAAACGCAGCGCAAGCCCACACTCATCAACGGAAAGAT
GGTGGACGCAGTATCCTCCAACATTGACGATCGCACCATGCACGAGCTGTACATGTGGCCCTT
TGCAGACGCCGTTCACGCTGGTGTATCATCGGTCATGTGTGGCTACAACCGAGTCAACGGCAC
ATATAGCTGCGAGAACAAGCATCTGATCAATAACCTCTTGAAGAAAGAGCTTGGTTTTGAGGG

CTATGTCATGTCGGACTTTCTGGCTACGCCGCCTGGACTTGGTCCCGTCAAGGCAGGGCTTGAT
ATGAACCAGCCCGGACCTGTAAATCCACTTTCCCCTGTTGACACGTACTGGGGCGACAATCTC
GTTGAGTGTGTCAAGAACAAGACACTCTCCGAGTCGGATCTCAATGGAATGGTGCGCCGCATC
TTGACGCCTTACTTCTATCTTGGTCAGGACAAAGACTACCCGTCTAAGGATCCATCTTCTCAGC
CTCTCGTCTACAACGGCTTCGGGTATCCTTATCCAGGCCCTTCTCCAGTTGGCCGCGATGTCCG
CGGCAACCACTCAGCGCTCATCCGTGAGATCGCCGCTGCCGGAACAGTCTTGCTCAAGAATCA
AGGTTCCATTCTTCCTCTTAACAAGTCCCTCACCAACATCGGTCTCTTTGGTAACGATGCTGCT
GATCCCTCCGTCGGTACCCTCTTCAGCGACCACGATGGCATCGACATCGGTACTCTTATCTCCG
GTGGCGGTTCAGGAAGCGGCCGACCAAGCTACGTTATCAGTCCCTTGGACACGTTCAAGTCTT
ACGCGAAAGCCAACGGCAAGCGGCTCCAGTACGTCACCAACAATACTGCTATTCTCAGCATTA
TGCCTGGTCTGTACCCTTGGCCAGCGGTCTGTATAGTCTTCCTCAAATCTTTCGCCACCGAGGG
TTTTGATCGCAAGACTCTTGTGGCAGATGACAACTCAGTTCAAGTCGTCAATAGTATTGCATCG
CGCTGCCCTCGCCGCACCGTCGTAGTCACTCACTCTGGAGGGCCAGACGTGATGCCGTGGGC
GACGAACCCTAATGTTAGCGCTATTGTCGCTGCGCATTACCCAGGCCAGGAATCAGGCAATTCT
ATCCTGGACGTCTTGATCGGAAAGGTCAATCCCTCTGGAAAGCTCCCATATACAATTGCCAAG
AAAGAGGAGGACTACAACGGAAAGATCACCAACATCACTGGCTCCGCAGCTGAAGACTCTTC
GAATTGGCAAAGTGAATTTAGTGAGGGCCTGTTCATTGACTACCGACACTTCGAAAACAAGGG
CCTAGAGCCTCTCTATGAGTTTGGCTACGGTCTCAGCTACACCACCTTTAAACTATCCTCCACC
CTCGTCGTCTCCTCAGCAAACAAGATCTCCGCCCGCGCATCCCCCTCCAACGCTACACTCACC
CTAGGCGGTAACCCTCACCTCTGGGAAACAGTCATCAAGTGCCACGTCGAGGTATCCAACACT
GGTCGGGTTGCAGGCGCCACAGTCATCCAGCTGTATGCGTCTCTGCCCAAGAACAATATACCA
GCCAACAGCCCAGTACGGATGCTTCGTGGGTTCGAGAAGGTCTATCTTGACCCTGGGGAGGCC
AAGAGAGTGTCATTCGCCCTAAGGCGTCGGGATCTGAGTTATTGGGATGTTACTATCCAGGACT
GGATGGTACCAAAGGGTGGCATAAAGCTCAGTGTTGGGTTTAGCTCGAAGGATCTGAGGTCTT
CGACAACCGTTCAACTGGTAAAATAA

5. Supplementary references

1. T. R. Hoye, C. S. Jeffrey and F. Shao, Mosher ester analysis for the determination of absolute configuration of stereogenic (chiral) carbinol carbons, *Nature Protoc.*, 2007, **2**, 2451-2458.

2. Y.-M. Chiang, M. Ahuja, C. E. Oakley, R. Entwistle, A. Asokan, C. Zutz, C. C. C. Wang and B. R. Oakley, Development of genetic dereplication strains in *Aspergillus nidulans* results in the discovery of aspercryptin, *Angew. Chem. Int. Ed.*, 2016, **55**, 1662-1665.

3. W. Xu, X. Cai, M. E. Jung and Y. Tang, Analysis of intact and dissected fungal polyketide synthase-nonribosomal peptide synthetase *in vitro* and in *Saccharomyces cerevisiae*, *J. Am. Chem. Soc.*, 2010, **132**, 13604-13607.

4. J. Fang, Y. Zhang, L. Huang, X. Jia, Q. Zhang, X. Zhang, G. Tang and W. Liu, Cloning and characterization of the tetrocarcin A gene cluster from micromonospora chalcea NRRL 11289 reveals a highly conserved strategy for tetronate biosynthesis in spirotetronate antibiotics, *J. Bacteriol.*, 2008, **190**, 6014-6025.

5. H. Zeng, G. Yin, Q. Wei, D. Li, Y. Wang, Y. Hu, C. Hu and Y. Zou, Unprecedented [5.5.5.6]dioxafenestrane ring construction in fungal insecticidal sesquiterpene biosynthesis, *Angew. Chem. Int. Edit.*, 2019, **58**, 6569-6573.