Supporting Information

Differences in PLP-Dependent Cysteinyl Processing Lead to Diverse *S*-Functionalization of Lincosamide Antibiotics

Min Wang,[†] Qunfei Zhao,[*,†,‡] Qinglin Zhang,[₸] and Wen Liu[*,†,‡,₸]

[†] State Key Laboratory of Bioorganic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, University of Chinese Academy of Sciences, 345 Lingling Road, Shanghai 200032, China.

[‡] State Key Laboratory of Microbial Metabolism, School of Life Science & Biotechnology, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China.

[₸] Huzhou Center of Bio-Synthetic Innovation, 1366 Hongfeng Road, Huzhou 313000, China.

* To whom correspondence should be addressed: Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 345 Lingling Rd., Shanghai 200032, China. Qunfei Zhao, Email: massfly@sioc.ac.cn; and Wen Liu, Email: wliu@mail.sioc.ac.cn, Tel: 86-21-54925111, Fax: 86-21-64166128

Table of Contents

4. Supplementary Tables

5. Supplementary References

1. Supplementary Methods

1.1 General Materials and Methods

Materials, Bacterial Strains and Plasmids. Biochemicals and media were purchased from Sinopharm Chemical Reagent Co., Ltd. (China), Oxoid Ltd. (U.K.) or Sigma-Aldrich Corporation (USA) unless otherwise stated. Restriction endonucleases were purchased from Thermo Fisher Scientific Co. Ltd. (USA). Chemical reagents were purchased from standard commercial sources. The bacterial strains, plasmids and primers used in this study are summarized in **Tables S1** and **S2**.

DNA Isolation, Manipulation, and Sequencing. DNA isolation and manipulation in *Escherichia coli* or *Streptomyces* strains were carried out according to standard methods[1,2]. PCR amplifications were carried out on an Applied Biosystems Veriti™ Thermal Cycler using either Taq DNA polymerase (Vazyme Biotech Co. Ltd, China) for routine genotype verification or Phanta® Max Super-Fidelity DNA Polymerase (Vazyme Biotech Co. Ltd, China) for high fidelity amplification. Primer synthesis was performed at Shanghai Sangon Biotech Co. Ltd. (China), and DNA sequencing was performed at Shanghai Majorbio Biotech Co. Ltd. or Shenzhen BGI in China.

Analysis of Metabolites and Enzymatic Products. High performance liquid chromatography (HPLC) analysis was carried out on the Agilent 1200 HPLC system (Agilent Technologies Inc., USA). HPLC Electrospray ionization MS (HPLC-ESI-MS) analysis was performed on the Thermo Fisher LTQ Fleet ESI-MS spectrometer (Thermo Fisher Scientific Inc., USA), and the data was analyzed using Thermo Xcalibur software. HPLC ESI-high resolution MS (HPLC-ESI-HRMS) analysis was carried out on the 6530 Accurate-Mass Q-TOF LC/MS System (Agilent Technologies Inc., USA), and the data was analyzed using Agilent MassHunter Qualitative Analysis software. NMR data was recorded on the Bruker DRX400 or Bruker AV500 spectrometers (Bruker Co. Ltd., Germany).

1.2 Gene Inactivation and Complementation

The genomic DNA of the *Streptomyces caelestis* (for celesticetin) or *S. lincolnensis* (for lincomycin A) wild type strain served as the template for PCR amplification unless otherwise stated.

Deletion of *ccbF*. The 2.13-kb fragment obtained using primers ccbF-R-for and ccbF-R-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1101. The 2.01-kb fragment obtained using primers ccbF-L-for and ccbF-L-rev was digested by XbaI and HindIII and cloned into the same site of pLL1101 to yield the recombinant plasmid pLL1102, in which a 813-bp in-frame coding region of *ccbF* was deleted. To transfer pLL1102 into the celesticetin-producing strain *S. caelestis*, conjugation between *E. coli* ET12567-*Streptomyces* was carried out following the standard procedure as previously described[3]. The colonies that were apramycin-resistant at 37 ℃ were identified as integrating mutants, in which a single-crossover homologous recombination event took place. These mutants were cultured for several rounds in the absence of apramycin, and the resulting apramycin-sensitive isolates were subjected to PCR amplification to examine the genotype, as judged by the amplification of the desired 0.44-kb product when using primers ccbF-for and ccbF-rev (**Figure S3A**). Further sequencing of this PCR product confirmed the genotype of LL1101, in which *ccbF* was in-frame deleted.

Deletion of *ccb5*. The 2.11-kb fragment obtained using primers ccb5-L-for and ccb5-L-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1103. The 2.34-kb fragment obtained using primers ccb5-R-for and ccb5-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1103 to yield the recombinant plasmid pLL1104, in which a 538-bp in-frame coding region of *ccb5* was deleted. Following the procedure described above, pLL1104 was introduced into the *S. caelestis* wild type strain for double-crossover recombination. The resulting strain LL1102 was then subjected to PCR amplification to examine the genotype, as judged by the the desired 0.52-kb product when using primers ccb5-for and ccb5-rev (**Figure S3B**). Further sequencing of this PCR product confirmed the genotype of LL1102, in which *ccb5* was in-frame deleted.

Deletion of *ccb4*. The 1.52-kb fragment obtained using primers ccb4-L-for and ccb4-L-rev was

digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1105. The 1.77-kb fragment obtained using primers ccb4-R-for and ccb4-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1105 to yield the recombinant plasmid pLL1106, in which a 474-bp in-frame coding region of *ccb5* was deleted. Following the procedure described above, pLL1106 was introduced into the *S. caelestis* wild type strain for double-crossover recombination. The resulting strain LL1103 was then subjected to PCR amplification to examine the genotype, as judged by the desired 0.64-kb product when using primers ccb4-for and ccb4-rev (**Figure S3C**). Further sequencing of this PCR product confirmed the genotype of LL1103, in which *ccb4* was in-frame deleted.

Deletion of *lmbF*. The 2.5-kb fragment obtained using primers lmbF-L-for and lmbF-L-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1107. The 2.60-kb fragment obtained using primers lmbF-R-for and lmbF-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1107 to yield the recombinant plasmid pLL1108, in which a 504-bp in-frame coding region of *lmbF* was deleted. Following the procedure described above, pLL1108 was introduced into the lincomycin A-producing *S. lincolnensis* wild type strain for double-crossover recombination. The resulting strain LL1104 was then subjected to PCR amplification to examine the genotype, as judged by the desired 1.1-kb product when using primers lmbF-gt-for and lmbF-gt-rev (**Figure S3D**). Further sequencing of this PCR product confirmed the genotype of LL1104, in which *lmbF* was in-frame deleted.

1.3 Production and Analysis of Celcesticetin and Desalicetin, Lincomycin A, and Intermediates.

The *S. caelestis* wild type strain or its derivatives was spread on MS plates that contain the medium composed of 20 g mannitol, 20 g soybean meal, and 20.0 g agar per liter (pH 7.0~7.5), and incubated at 28 ℃ for sporulation and growth. Approximately 1 cm^2 of the sporulated agar of *S. caelestis* was cut, chopped, and inoculated into 25 mL of the seed medium, which is composed of 25 g glucose and 25 g phamamedia (Archer Daniels Midland Co, USA) per liter. After incubation at 28 ℃ and 220 rpm for 36 h, 5 mL of the seed culture broth was transferred into 100 mL of the fermentation

medium, which is composed of 25 g glucose, 2.5 g Brewer's yeast, 5 g $(NH_4)_2SO_4$, 8 g $CaCO_3$, 4 g NaCl, 7 g soybean meal and 0.4 g KH_2PO_4 per liter (pH 7.0). Further incubation was carried out at 28 ℃ and 220 rpm for 7 days.

The *S. lincolnensis* wild type strain or its *ΔlmbF* derivative was spread on the agar plates that contain the medium composed of 19 g starch, 5 g soybean meal, 0.5 g K_2HPO_4, 0.5 g $MgSO_4.7H_2O$, 1.0 g KNO_3, 0.5 g NaCl, 0.01 g $FeSO_4.7H_2O$, and 20.0 g agar per liter (pH 7.0~7.5), and incubated at 28 ℃ for sporulation and growth. Approximately 1 cm^2 of the sporulated agar of *S. lincolnensis* was cut, chopped, and inoculated into 25 mL of the seed medium, which is composed of 20 g starch, 10 g glucose, 10 g soybean meal, 30 g corn steep liquor, 1.5 g $(NH_4)_2SO_4$ and 5 g $CaCO_3$ per liter. After incubation at 28 ℃ and 220 rpm for 36 h, 5 mL of the seed culture broth was transferred into 50 mL of the fermentation medium, which is composed of 100 g glucose, 25 g soybean meal, 2 g corn steep liquor, 8 g $(NH_4)_2SO_4$, 0.2 g KH_2PO_4, 8 g $NaNO_3$, 5 g NaCl and 8 g $CaCO_3$ per liter (pH 7.0). Further incubation was carried out at 28 ℃ and 220 rpm for 7 days.

For product examination, 500 μl of each fermentation broth was mixed with the equal volume of methanol. After centrifugation to remove the residue, the supernatant was subjected to HPLC-ESI-MS analysis on an Agilent Zorbax column (SB-C18, 5 mm, 4.6 × 250 mm, Agilent Technologies Inc., USA) by gradient elution of solvent A (10 mM NH_4Ac) and solvent B (CH_3CN) with a flow rate of 1 mL/min over a 22 min period as follows: T = 0 min, 10% B; T = 9 min, 10% B; T = 17 min, 60% B; T = 18 min, 60% B; T = 20 min, 10 % B; and T = 22 min, 10% B (mAU at 210 nM). The data were analyzed using Thermo Xcalibur software.

1.4 Protein Expression and Purification.

The genes *ccbF*, *ccb5*, *ccb4*, *lmbF* and *lmbG* were amplified by PCR from the genomic DNA of *S. caelestis* or *S. lincolnensis* using primers with engineered NedI and HindIII restriction sites. The sequences of the primers are described in **Table S2**. The PCR-amplified DNA fragments were purified, digested with NdeI and HindIII and ligated into a pET28a(+) vector (Novagen) that had been digested with the same enzymes. The resultant plasmids were used to transform *E. coli* BL21

(DE3) for protein overexpression. The *E. coli* transformants cultures were incubated in Luria-Bertani (LB) medium containing 50 μg/mL kanamycin at 37 ℃ and 250 rpm until the cell density reached 0.6-0.8 at $OD_{600\ nm}$. To induce protein expression, IPTG (0.1 mM) was added to the cultures, which were further incubated at 16 ℃ for 40-48 h. The cells were harvested by centrifugation and stored at -80 ℃ before lysis. The thawed cells were re-suspended in lysis buffer containing 50 mM Tris-HCl (pH 7.6), 300 mM NaCl, 10 mM imidazole and 10% (v/v) glycerol. After disruption by an ultrahigh pressure cells crushed homogenizer (FB-110X, Shanghai Litu Mechanical Equipment Engineering Co.,Ltd, China or JN-02HC, JNBIO, China), the soluble fraction was collected, subjected to purification of each target protein by using a HisTrap FF column (GE Healthcare, USA) and then desalted using a PD-10 Desalting Column (GE Healthcare, USA) according to the manufacturers' protocols. The resulting proteins were concentrated and stored at -80 ℃ for in vitro assays. The purity of the proteins was examined by 12% SDS-PAGE analysis (**Figure S5A**), and the concentration was determined by Bradford assay using bovine serum albumin (BSA) as the standard. The UV-Vis spectra of the recombinant proteins CcbF and LmbF were recorded at a concentration of 1 mg/ml on a Unico 4802 UV/Vis spectrophotometer (UNICO Instruments Co., Ltd., Shanghai) from 250 to 600nm (**Figure S5B**).

1.5 Characterization of CcbF-Catalyzed Conversions *in vitro*

The assays were initially carried out in a 50-μL reaction mixture containing 100 mM Phosphate Buffer (pH 7.5), 2 mM cysteine *S*-conjugate **1**, 2 mM α-keto acid (pyruvate or α-ketoglutarate), 0.2 mM PLP and 20 μM CcbF. To determine the necessity of α-keto acid and exogenous PLP, either of them was omitted from the reaction mixture. The reaction proceeding in the absence of the enzyme was utilized as the negative control. After incubation at 30˚C for 20 min, an equal volume of acetonitrile was added into each mixture to terminate the reaction. After removal of the denatured protein by centrifugation, the supernatants were subjected to HPLC-ESI-MS analysis on an Agilent Zorbax column by gradient elution of solvent A (H_2O containing 0.1% formic acid) and solvent B (CH_3CN containing 0.1% formic acid) with a flow rate of 0.8 mL/min over a 22 min period as follows: T = 0 min, 5% B; T = 5 min, 5% B; T = 15 min, 80% B; T = 17 min, 80% B; T = 18 min, 5% B; and T = 20 min, 5% B (mAU at 210 nm).

To determine whether CcbF has a transamination activity, the assays were carried out at 30˚C for 1 h in a 50-μL reaction mixture containing 100 mM Phosphate Buffer (pH 7.5), 2 mM **1**, 2 mM α-keto acid (pyruvate or α-ketoglutarate) and 20 μM CcbF. After similar reaction termination and centrifugation, 2.5 μL of dansyl chloride (DNSC, 10 mM) was added to the 25 μL of supernatant of the assays, following by further incubation at 50 ℃ for 1 h according the procedure previously described[4]. After removal of precipitate by centrifugation, the mixtures were then subjected to HPLC analysis on an Agilent Zorbax column by gradient elution of solvent A (H_2O containing 0.1% formic acid) and solvent B (CH_3CN containing 0.1% formic acid) with a flow rate of 1 mL/min over a 22 min period as follows: T = 0 min, 10% B; T = 3 min, 10% B; T = 6 min, 30% B; T = 12 min, 30% B; T = 20 min, 80 % B; and T = 22 min, 10% B (mAU at 285 nm). For identifying the production of the amino acid species, 2 mM alanine or glutamic acid was used as an external standard (**Figure S6**).

To determine whether the conversion of PLP to PMP exists in CcbF-catalyzed reaction, the assays were carried out at 30˚C for 1 h in a 50-μL standard reaction mixture containing 100 mM Phosphate Buffer (pH 7.5), 2 mM **1**, 0.2 mM PLP and 20 μM CcbF. The reactions proceeding in the absence of CcbF or PLP were utilized as the negative controls. The absorbance change of 390 nm was monitored at 10-min intervals on a NanoDrop 2000C UV/Vis spectrophotometer (Thermo Fisher Scientific Inc., USA) (**Figure S7**). To determine the necessity of O_2 for CcbF-catalyzed conversion, the reaction carried out under N_2 atmosphere in a type A anaerobic chamber (Coy Laboratory Products Inc., Grass Lake, MI) maintained at 0 ppm O_2 was utilized as the negative control. Buffers containing the substrate (100 mM Phosphate Buffer, 2 mM **1**, 0.2 mM PLP) were vigorously bubbled with N_2 for 2 h, brought into the anaerobic chamber and allowed to equilibrate for at least 12 h before use. Enzyme solutions were transferred in the anaerobic chamber on ice, gently bubbled with the same gas using pipette for 10 min and allowed to equilibrate for at least 1 h before use. The termination of each reaction and HPLC-ESI-MS analysis was carried out according to the MS methods described above. To determine the production of ammonia (NH_3), the resulting assays were then incubated with an equal volume of *o*-phthaldialdehyde (OPA)/mercaptopropionic acid (MPA) agent (2 mg/ml OPA, 0.2% (v/v) MPA in 0.2 M borate buffer, pH 10.2) according to the methods

previously described[5]. After removal of precipitate by centrifugation, the mixtures were then subjected to HPLC analysis on an Agilent Zorbax column by gradient elution of solvent A (H_2O containing 0.1% formic acid) and solvent B (CH_3CN containing 0.1% formic acid) with a flow rate of 1 mL/min over a 22 min period as follows: T = 0 min, 10% B; T = 3 min, 10% B; T = 16 min, 70% B; T = 20 min, 98% B; T = 22 min, 10% B (mAU at 338 nm). For identifying the production of NH_3, 2 mM $(NH_4)_2SO_4$ was used as an external standard (**Figure S8**). To determine the production of hydrogen peroxide (H_2O_2), the reagents (4 mM phenol, 6 mM 4-aminoantipyrine, and 0.2 units of horseradish peroxidase) were added to the standard reaction mixtures (1 h) according to the methods previously described[6]. The absorbance increase at 505 nm was monitored at 10-min intervals on the same UV/Vis spectrophotometer (**Figure S9**).

1.6 *In vitro* Enzymatic Transformation of Cysteine S-Conjugated Intermediate to Desalicetin

The assays were carried out at 30˚C for 2 h in a 50-μL reaction mixture containing 100 mM Phosphate Buffer (pH 7.5), 2 mM cysteine S-conjugate **1**, 0.2 mM PLP, 4 mM SAM, 4 mM NADPH and 20 μM CcbF in the presence of 20 μM Ccb5 and 20 μM Ccb4, 20 μM Ccb5 or 20 μM Ccb4. Reaction termination and HPLC-ESI-MS analysis were carried out according to the methods described above for CcbF-catalyzed conversions.

1.7 Characterization of LmbF-Catalyzed Conversions *in vitro*

The assays were carried out at 30˚C for 2 h in a 50-μL reaction mixture containing 100 mM Phosphate Buffer (pH 7.5), 2 mM cysteine S-conjugate **5**, and 0.2 mM PLP in the presence of 20 μM LmbF. The reaction proceeding in the absence of LmbF was utilized as the negative control. After similar reaction termination and centrifugation, the supernatants of the reaction mixtures were subjected to HPLC-ESI-MS analysis on an Agilent Zorbax column by isocratic elution of 40 % solvent A (H_2O containing 5 mM NH_4Ac) and 60 % solvent B (MeOH) with a flow rate of 0.6 mL/min over a 17 min period (mAU at 210 nm). To determine the production of pyruvate, 2.5 μL of 2,4-dinitrophenylhydrazine (DNPH) solution (0.1 mg/mL in 50% ethanol) was added to the 25 μL of supernatant of the assays, following by similar work-up and HPLC analysis according the

methods described above for amino acid derivation. For identifying the production of pyruvate, 2 mM pyruvate was used as an external standard.

1.8 *In vitro* Enzymatic Transformation for Lincomycin Maturation

The two-enzymes coupled assay was carried out at 30˚C for 2 h in a 50-μL reaction mixture containing 100 mM Phosphate Buffer (pH 7.5), 2 mM **5**, 0.2 mM PLP, 4 mM SAM, and 20 μM CcbF in the presence of 20 μM LmbG. The termination of each reaction and HPLC-ESI-MS analysis was carried out according to the methods described above for LmbF-catalyzed conversion.

1.9 *In vitro* Combinatorial Biosynthesis of Lincosamide Antibiotics

To generate Bu-2545 and related products with a lincomycin-like sulfur appendage, the assays were carried out at 30˚C for 2 h in a 50-μL reaction mixture containing 100 mM Phosphate Buffer (pH 7.5), 2 mM the celesticetin intermediate **1**, 0.2 mM PLP, and 8 mM SAM, in the presence of 20 μM LmbF alone, 20 μM LmbF and 20 μM LmbG, 20 μM LmbF and 20 μM Ccb4 or 20 μM LmbF, 20 μM LmbG and 20 μM Ccb4. HPLC-ESI-MS analysis was carried out according to the methods described above for CcbF-catalyzed conversions.

To generate the products with a celesticetin-like sulfur appendage, the assays were carried out at 30˚C for 2 h in a 50-μL reaction mixture containing 100 mM phosphate buffer (pH 7.5), 2 mM the lincomycin intermediate **5**, 0.2 mM PLP, 4 mM SAM, and 4 mM NADPH, in the presence of 20 μM CcbF alone or 20 μM CcbF and 20 μM Ccb5. HPLC-ESI-MS analysis was carried out according to the methods described above for LmbF-catalyzed conversions.

1.10 Compound Isolation and Purification

For celesticetin and desalicetin (from the *S. caelestis* wild type strain), 100 g of Amberlite XAD-2 resin (Rohm and Haas Co., USA) was incubated with 1 L of the fermentation culture broth overnight to extract the target compound from each mutant strain. After filtrated, the resin was backwashed

with 2 L of water, and then eluted with 3 L of MeOH. According to ESI-MS analysis, the fractions containing the target compounds was combined, evaporated *in vacuo* and re-dissolved in 5 mL of water. The sample was loaded onto a Sephadex LH20 column (3.5 × 200 cm, GE healthcare, USA) for elution of 500 mL of MeOH with a flow rate of 0.5 mL/min. After the similar workup for fraction and concentration, further purification of desalicetin was carried out on an Aglient Zorbax column (SB-C18, 5 μm, 9.4 × 250mm, Agilent Technologies Inc., USA) by gradient elution of solvent A (10 mM NH_4Ac) and solvent B (CH_3CN) with a flow rate of 2 mL/min over a 26 min period as follows: T = 0 min, 5% B; T = 5 min, 5% B; T = 20 min, 25% B; T = 21 min, 40% B; T = 25 min, 40 % B; and T = 26 min, 5% B (mAU at 210 nm); and further purification of celesticetin was carried out on the same Aglient Zorbax column by gradient elution of solvent A (10 mM NH_4Ac) and solvent B (CH_3CN) with a flow rate of 2.5 mL/min over a 26 min period as follows: T = 0 min, 30% B; T = 10 min, 30% B; T = 20 min, 45% B; T = 23 min, 60% B; T = 25 min, 60 % B; and T = 26 min, 30% B (mAU at 210 nm).

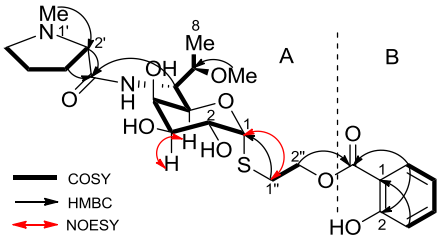
For compound **1** (from the *ΔccbF* mutant strain), 100 g of Amberlite XAD-2 resin was incubated with 1 L of the fermentation culture broth overnight to extract the target compound from each mutant strain. After filtrated, the resin was backwashed with 2 L of water, and then eluted with 3 L of MeOH. According to ESI-MS analysis, the fractions containing the target compound was combined, evaporated *in vacuo* and re-dissolved in 5 mL of water. The sample was loaded onto a Sephadex LH20 column for elution of 500 mL of MeOH with a flow rate of 0.5 mL/min. After the similar workup for fraction and concentration, further purification was carried out on a COSMOSIL HILIC Packed Column (5 μm, 10 × 250 mm, Nacalai Tesque Inc., Japan) by isocratic elution of 21% solvent A (10 mM NH_4Ac) and 79% solvent B (CH_3CN) with a flow rate of 2 mL/min over a 70 min period (mAU at 210 nm).

For Compound **5** (from the *ΔlmbF* mutant strain), 100 g of Amberlite XAD-2 resin was incubated with 1 L of the fermentation culture broth overnight to extract the target compound from each mutant strain. After filtrated, the resin was backwashed with 2 L of water, and then eluted with 3 L of MeOH. According to ESI-MS analysis, the fractions containing the target compound was combined, evaporated *in vacuo* and re-dissolved in 5 mL of water. Further semi-preparative purification was

carried out twice on the same Aglient Zorbax column by isocratic elution of 40% solvent A (5 mM NH_4Ac) and 60% solvent B (MeOH) with a flow rate of 2 mL/min over a 12 min period (mAU at 210 nm).

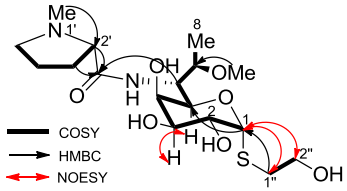
2. Supplementary Results

2.1 Physico-Chemical Properties and Structural Elucidation of Celesticetin



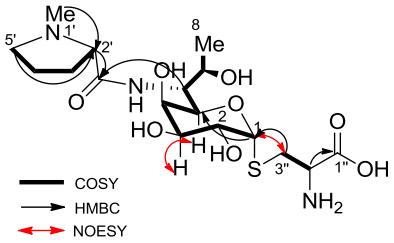
Celesticetin was purified as yellowish white amorphous solid: UV (H_2O) λ_{max} 240 nm; ^1H and ^{13}C NMR (500 and 125 MHz, respectively, D_2O) see **Table S3**; ESI-HR-MS Calcd. for $C_{24}H_{37}N_2O_9S^+$ 529.2220 $[M+H]^+$, found 529.2222 (**Figure S15**). The molecular formula of celesticetin was established to be $C_{24}H_{36}N_2O_9S$ by analysis of its ESI-HR-MS, ^{13}C NMR and DEPT spectra. The planar structure of celesticetin was elucidated by detailed analysis of its 1D and 2D NMR spectra (^1H, ^{13}C, DEPT, HSQC and HMBC) (**Figure S11**). This compound has been previously characterized[7].

2.2 Physico-Chemical Properties and Structural Elucidation of Desalicetin



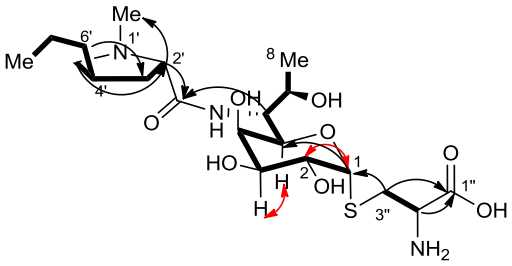
Desalicetin was purified as yellowish white amorphous solid: UV (H_2O) λ_{max} 210 nm; ^1H and ^{13}C NMR (500 and 125 MHz, respectively, D_2O) see **Table S4**; ESI-HR-MS Calcd. for $C_{17}H_{33}N_2O_7S^+$ 409.2008 $[M+H]^+$, found 409.2007 (**Figure S16**). The molecular formula of desalicetin was established to be $C_{17}H_{32}N_2O_7S$ by analysis of its ESI-HR-MS, ^{13}C NMR and DEPT spectra. The planar structure of desalicetin was elucidated by detailed analysis of its 1D and 2D NMR spectra (^1H, ^{13}C, DEPT, HSQC and HMBC) (**Figure S12**). This compound has been previously characterized[8].

2.3 Physico-Chemical Properties and Structural Elucidation of Compound 1



Compound **1** was purified as yellowish white amorphous solid: UV (H_2O) λ_{max} 210 nm; ^1H and ^{13}C NMR (500 and 125 MHz, respectively, D_2O) see **Table S5**; ESI-HR-MS Calcd. for $C_{17}H_{32}N_3O_8S^+$ 438.1910 $[M+H]^+$, found 438.1914 (**Figure S17**). The molecular formula of **1** was established to be $C_{17}H_{31}N_3O_8S$ by analysis of its ESI-HR-MS, ^{13}C NMR and DEPT spectra. The planar structure of **1** was established by detailed analysis of its 1D and 2D NMR spectra (^1H, ^{13}C, DEPT, HSQC and HMBC) (**Figure S13**), and by comparison of the ^1H and ^{13}C NMR spectroscopic data of **1** with those of compound **5** (**Table S5** and **S6**). Compound **1** shared the similar NMR spectra (C8 aminosugar central to an amino acid and a cysteine residue) with that of **5**. The major difference of **1** compared with **5** was the absence of signals for propyl on the amino acid residue, which was consistent with the proline-residue in celesticetin. The anomeric configuration of C1 of **1** was determined to be α-orientation, according to the observed coupling constant ($^3J_{H,H}$ = 5.8 Hz) as same as that in **5**. The absolute configuration of **1** was suggested according to those of **5** and celesticetin[7].

2.4 Physico-Chemical Properties and Structural Elucidation of Compound 5



Compound **5** was purified as yellowish white amorphous solid: UV (H_2O) λ_{max} 210 nm; ^1H and ^{13}C NMR (500 and 125 MHz, respectively, D_2O) see **Table S6**; ESI-HR-MS Calcd. for $C_{20}H_{38}N_3O_8S^+$ 480.2380 $[M+H]^+$, found 480.2381 (**Figure S21**). The molecular formula of **5** was established to be $C_{20}H_{37}N_3O_8S$ by analysis of its ESI-HR-MS, ^{13}C NMR and DEPT spectra. The ^{13}C NMR and DEPT spectra of **5** indicated that the molecule has two carbonyl carbon atoms, ten methine carbon atoms, five methylene carbon atoms and three methyl carbon atoms. The planar structure of **5** was established by detailed analysis of its 1D and 2D NMR spectra (^1H, ^{13}C, DEPT, HSQC and HMBC)

(**Figure S14**), and by comparison of the ^{1}H and ^{13}C NMR spectroscopic data of **5** with those of lincomycin A[9,10] (**Table S6**). Compound **5** shared the similar NMR spectra with that of lincomycin A. The major difference of **5** compared with lincomycin was the absence of signals for the *S*-methyl and the presence of signals for alanine. The absence of the *S*-methyl group signals and the obvious HMBC couplings of H3" to C1 suggested the attachment of alanine moiety onto C8 aminosugar moiety via a C-S bond linkage between C3" and the sulfur atom. The anomeric configuration of C1 of **5** was determined to be α-orientation, according to the observed coupling constant ($^{3}J_{H,H}$ = 5.8 Hz) as same as that in lincomycin A. The absolute configuration of **5** was suggested according to those of lincomycin A[9,10].

1. Supplementary Figures

Fig. S1. Biosynthetic gene clusters of celesticetin (top, *ccb*) and lincomycin A (bottom, *lmb*)[11,12]. The tailoring genes relevant to this study are indicated by color. Red, the PLP-dependent proteins CcbF and LmbF; blue, S-adenosyl-L-methionine (SAM)-dependent *O*-methyltransferase Ccb4; yellow, NAD(P)H-dependent oxidoreductase Ccb5; and purple, *S*-methyltransferase LmbG. ID, identity.

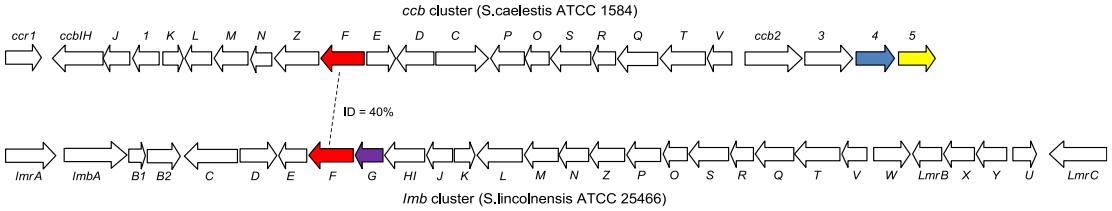


Fig. S2. Alignment of the amino acid sequences of CcbF and LmbF with those of ten homologous but functionally unknown proteins from the NCBI database. The conserved residue lysine at the PLP-binding site is indicated by a vertical arrow.

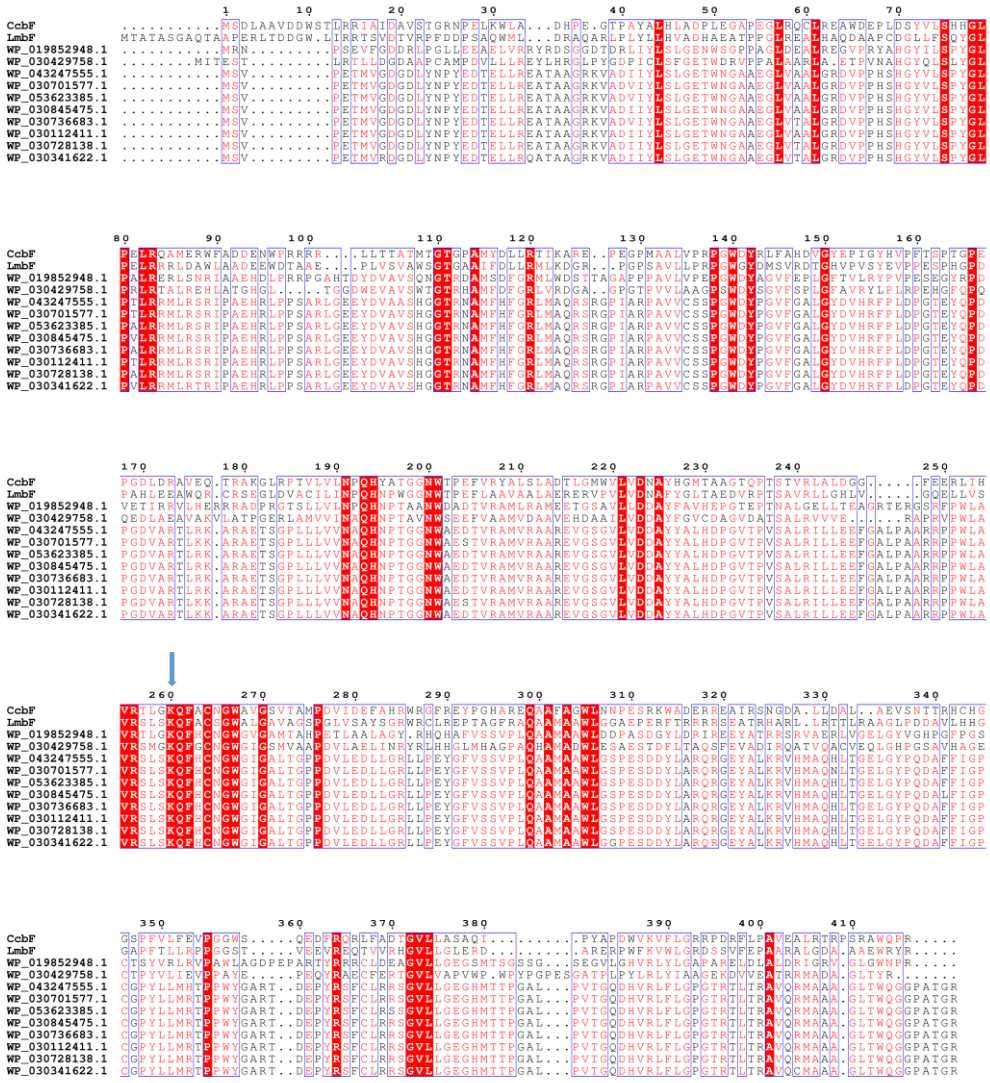
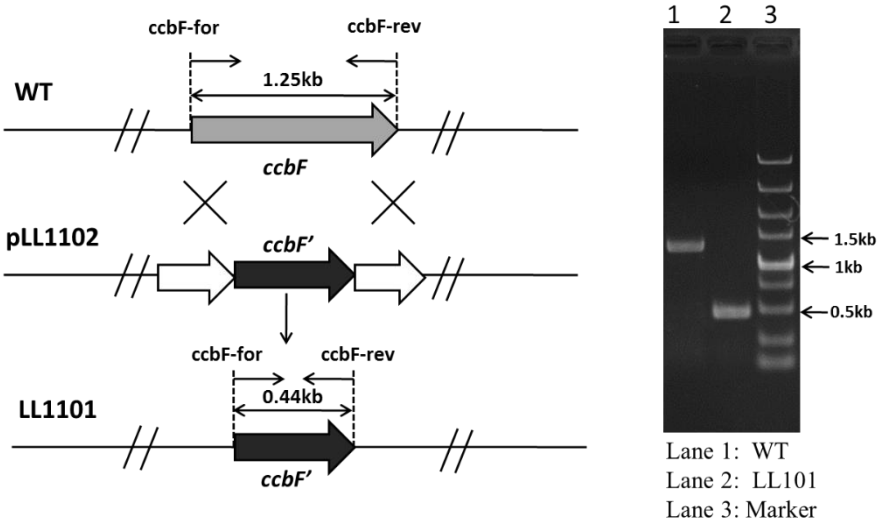
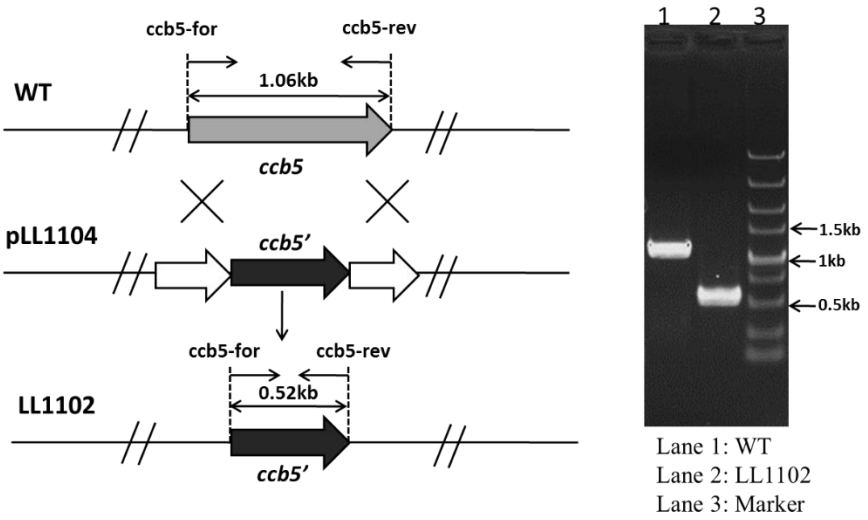


Fig. S3. Generation of the *S. caelestis* or *S. lincolnensis* mutant strains. (**A**) LL1001 (*ΔccbF*); (**B**) LL1002 (*Δccb5*); (**C**) LL1003 (*Δccb4*); (**D**) LL1004 (*ΔlmbF*). Validation of each genotype was performed by PCR amplification using genomic DNA of the mutant or wild type strain as the template. The PCR primers are labeled with the predicted sizes of their resulting products.

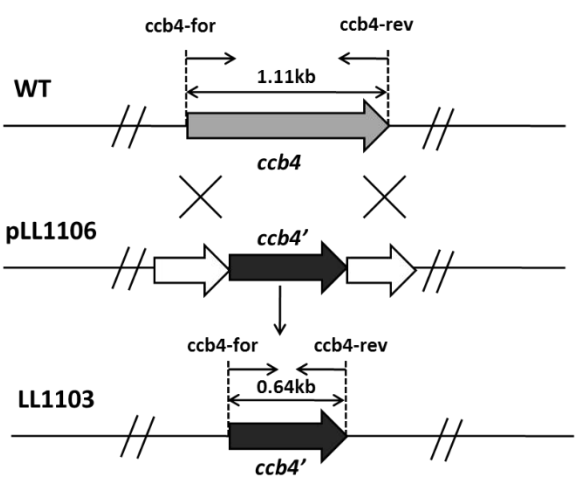
A

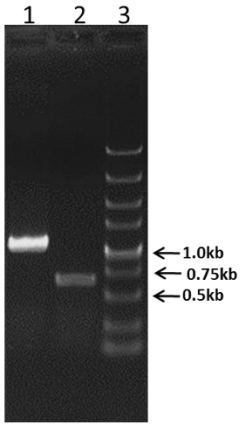


B



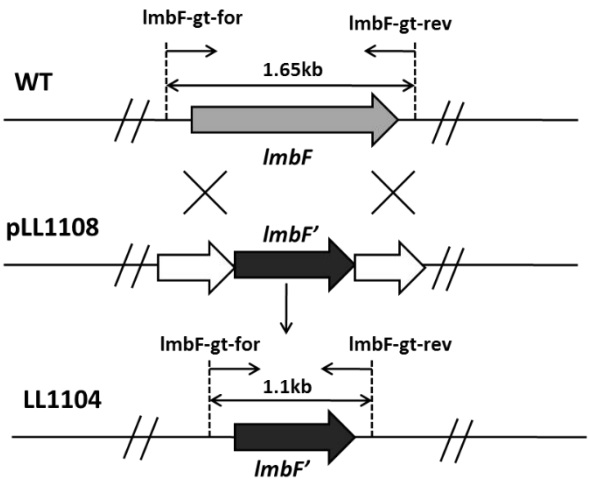
C

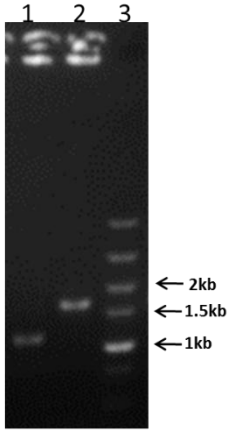




Lane 1: WT
Lane 2: LL1103
Lane 3: Marker

D





Lane 1: LL1104
Lane 2: WT
Lane 3: Marker

Fig. S4. Production analysis of lincosamide antibiotics *in vivo*. The ESI m/z $[M + H]^+$ modes of target compounds are indicated (right) and colored in the HPLC-MS traces (left). (**A**) Product profiles of the *S. caelestis* strains. i), *ΔccbF*; ii), *Δccb5*; iii), *Δccb4*; and iv), wild-type. (**B**) Product profiles of the *S. lincolnensis* strains. i), Δ*lmbF*; and ii), wild-type.

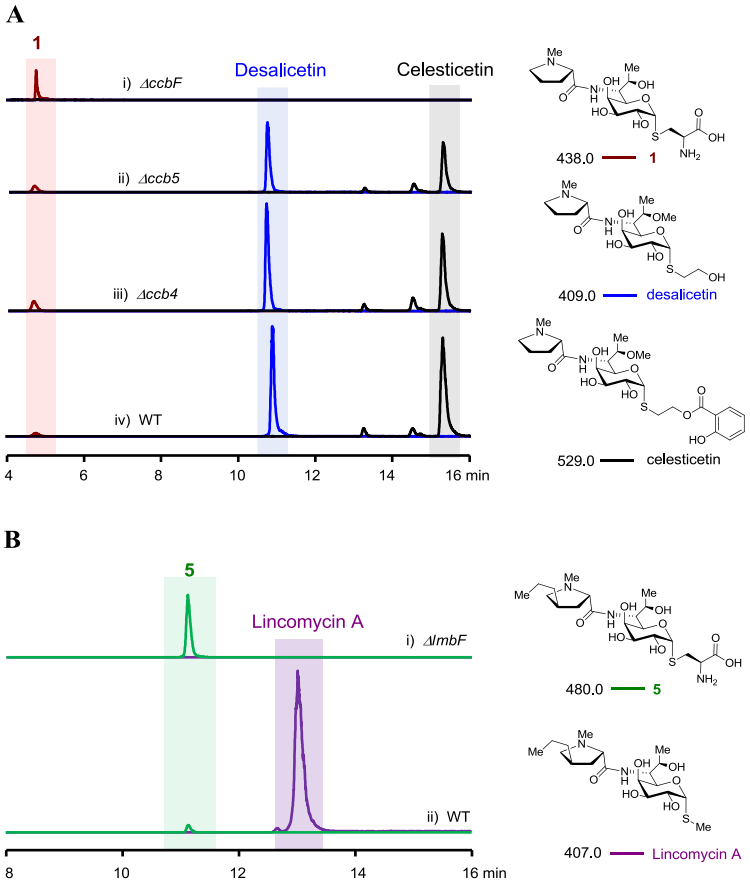
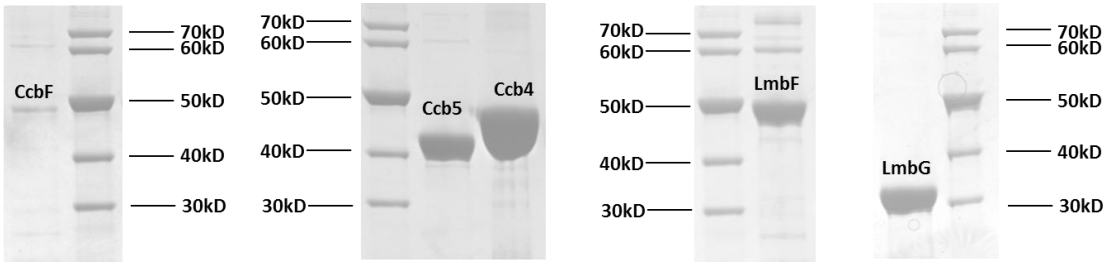


Fig. S5. Characterization of purified recombinant proteins. (**A**) Coomassie-stained SDS-PAGE analysis of purified His_6-CcbF (48.8 kD), His_6-Ccb5 (42.5 kD), His_6-Ccb4 (42.8kD), His_6-LmbF (48.9 kD) and His_6-LmbG (31.4 kD). **(B**) UV-vis absorption spectrum of the purified recombinant proteins His_6-CcbF (left) and His_6-LmbF (right).

A



B

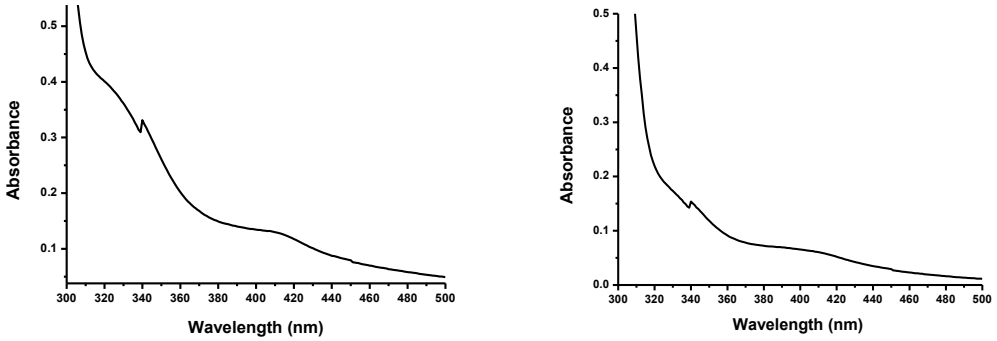


Fig. S6. Determination of possible transamination activity of CcbF in the presence of α-keto acids. (**A**) Derivatization of L-alanine (L-Ala) and L-glutamic acid (L-Glu) with dansyl chloride (DNSC). (**B**) HPLC traces of CcbF-catalyzed reactions, which were conducted by incubating substrate **1** with CcbF and pyruvate (i), pyruvate alone (ii), CcbF and α-ketoglutarate (iii), or α-ketoglutarate alone (iv). L-Ala (v) and L-Glu (vi) were used as external standards. The reactions were incubated under aerobic conditions for 1-h and then treated with DNSC. Neither L-Ala nor L-Glu was detected, showing that neither pyruvate nor α-ketoglutarate could accept ammonia from the cysteinyl group of **1** and CcbF had no transamination activity.

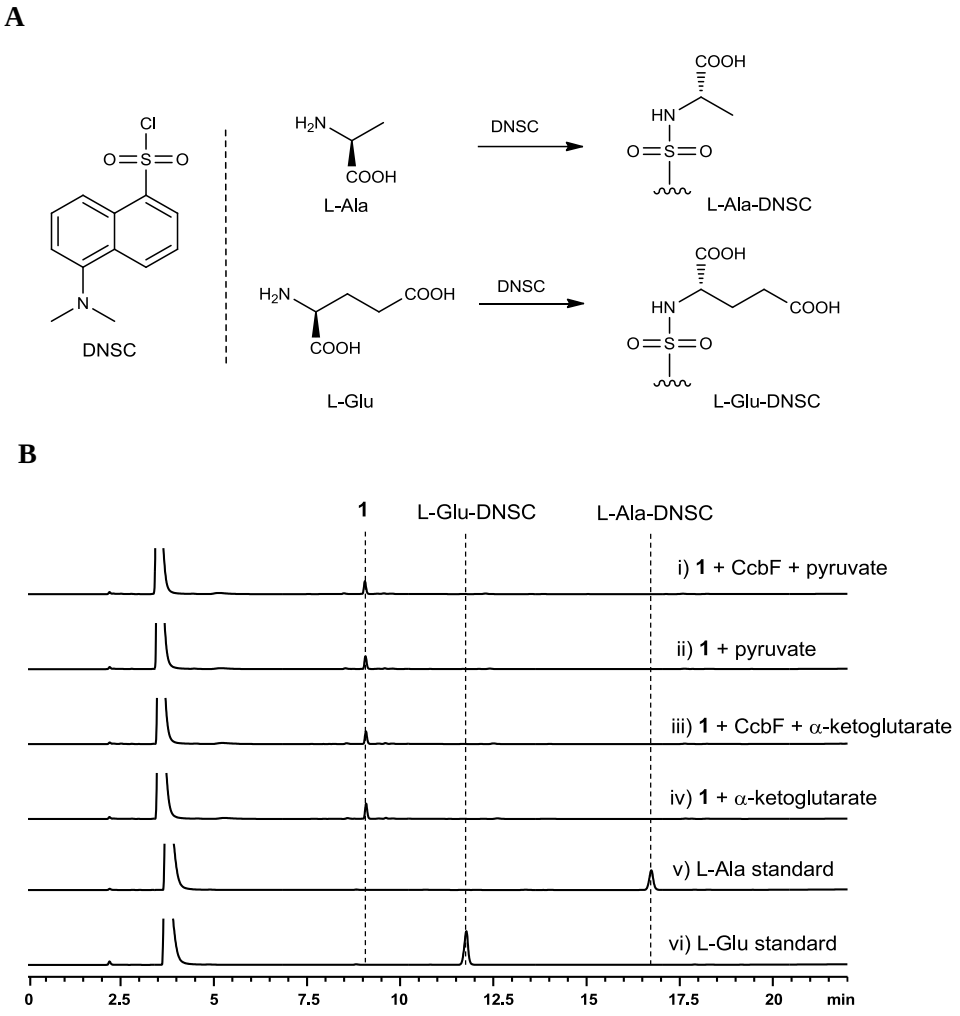


Fig. S7. Determination of the possible conversion of PLP to PMP in CcbF-catalyzed reaction. (**A**) Conversion of substrate **1** to product **2** *in vitro*. The ESI m/z $[M + H]^+$ modes of target compounds are indicated (left) and colored in the HPLC-MS traces (right). The transformations proceeding here derived from the reaction in which CcbF, **1**, and exogenous PLP were incubated under aerobic conditions during a 1-h incubation period (i, full assay), and included those in which PLP (ii) or CcbF (iii, control) was selectively omitted. (**B**) Time course analysis of UV-vis absorption spectrum of the full assay during a 1-h incubation period. Both results showed that PLP was not consumed as co-substrate and no PMP was produced during the reaction course of CcbF.

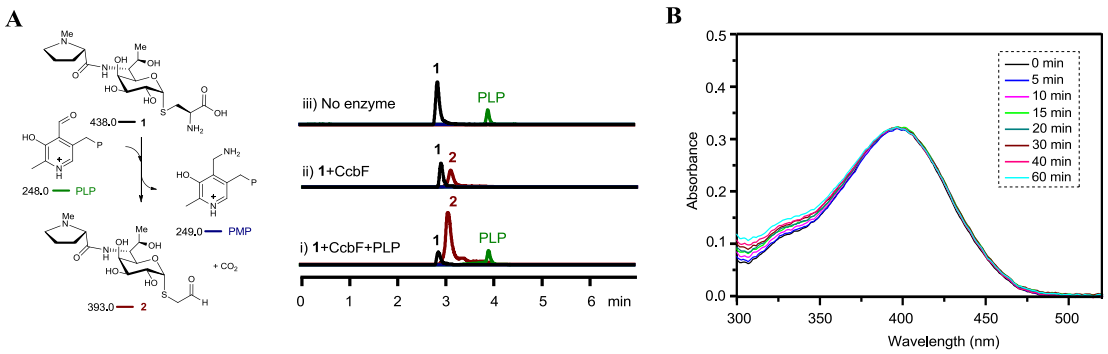


Fig. S8. Characterization of NH_3 production in CcbF-catalyzed reaction. (**A**) Derivatization of NH_3 with *o*-phthaldialdehyde (OPA)/mercaptopropionic acid (MPA). The substrate **1** and Tris component of the enzyme solution can also react with OPA/MPA. (**B**) HPLC traces of CcbF-catalyzed reactions, which were conducted by incubating substrate **1**, exogenous PLP and CcbF (i), substrate **1** and exogenous PLP (ii), or CcbF and exogenous PLP (iii). 2 mM $(NH_4)_2SO_4$ (iv) was used to generate external standard NH_3. The reactions were incubated under aerobic conditions for 1-h and then treated with OPA/MPA reagent. (**C**) The ESI-MS spectra of NH_3-OPA/MP.A (left, calculated for $C_{19}H_{16}NO_3S^+$ $[M+H]^+$ 338.0845), Tris-OPA/MPA (middle, calculated for $C_{15}H_{20}NO_5S^+$ $[M+H]^+$ 326.1062) and **1**-OPA/MPA (right, calculated for $C_{28}H_{40}N_3O_{10}S_2^+$ $[M+H]^+$ 642.2155).

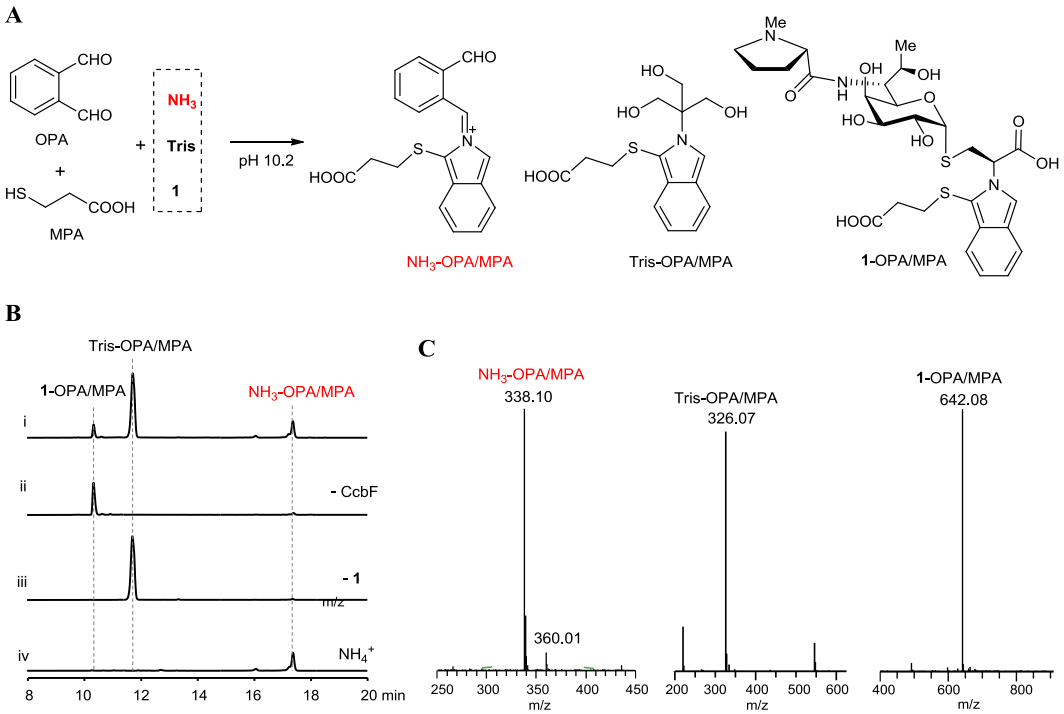


Fig. S9. Characterization of H_2O_2 production in CcbF-catalyzed reaction. (**A**) Derivatization of H_2O_2 with 4-aminoantipyrine and phenol under the catalysis of horseradish peroxidase. (**B**) Time course analysis of UV-vis absorption spectrum of the coupled assay, which was conducted by incubating substrate **1** with CcbF in the presence of exogenous PLP and the derivation regents (4-aminoantipyrine, phenol and horseradish peroxidase) during a 1-h incubation period under aerobic conditions. (**C**) Detection of H_2O_2 in the presence (i) and absence (ii) of CcbF for the above assay.

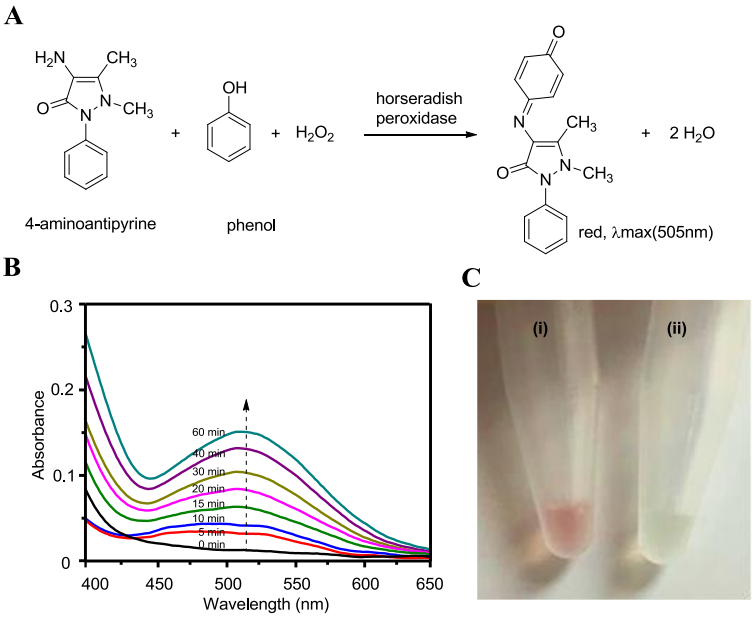


Fig. S10. Characterization of pyruvate production in LmbF-catalyzed reaction. (**A**) Derivatization of pyruvate with 2,4-dinitrophenylhydrazine (DNPH). (**B**) HPLC traces of LmbF-catalyzed reactions, which were conducted by incubating substrate **5** and exogenous PLP in the absence (i) or presence (ii) of LmbF. Pyruvate was used as external standard (iii). The reactions were incubated for 2-h and then treated with DNPH.

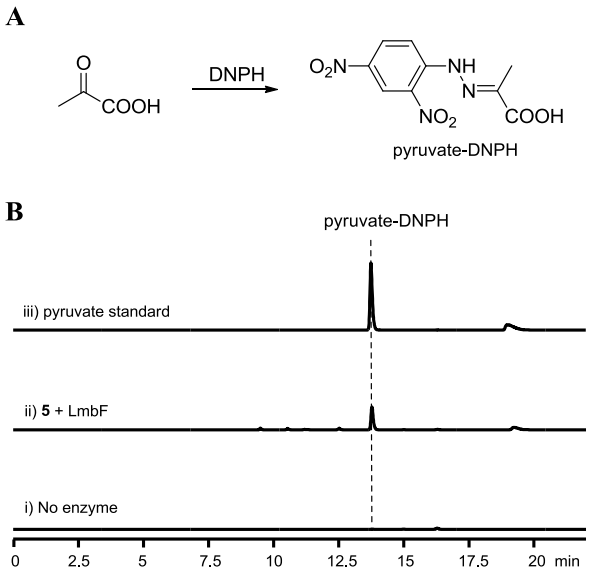
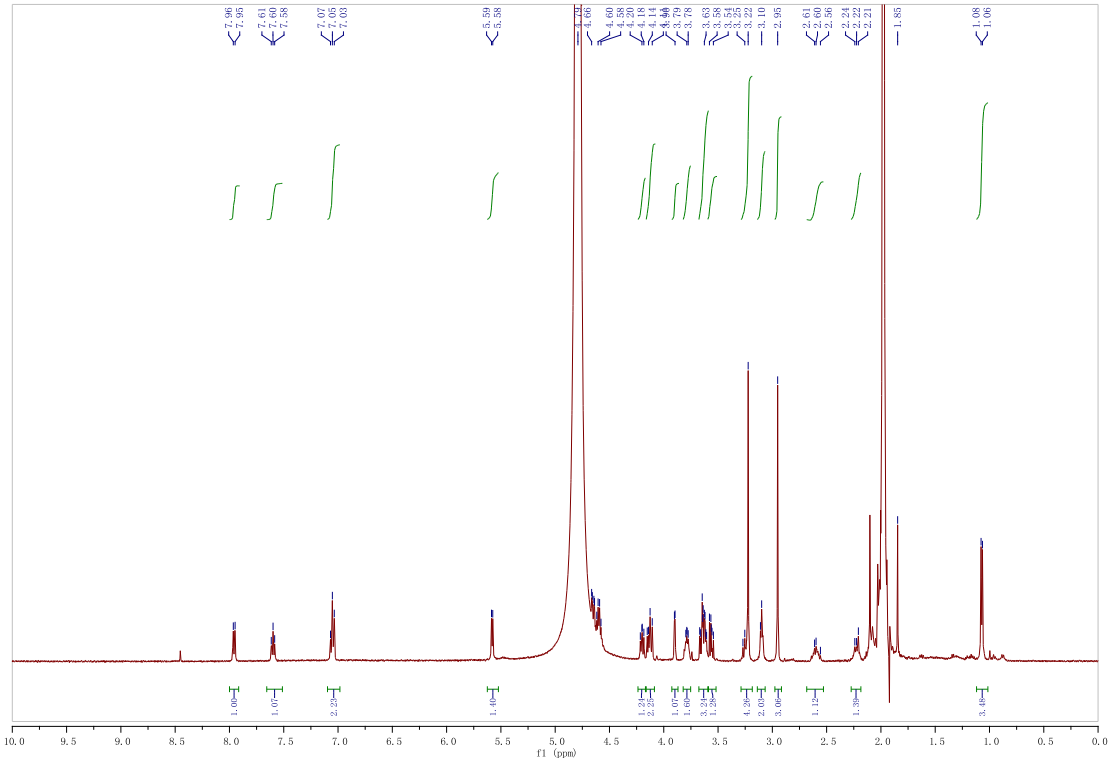
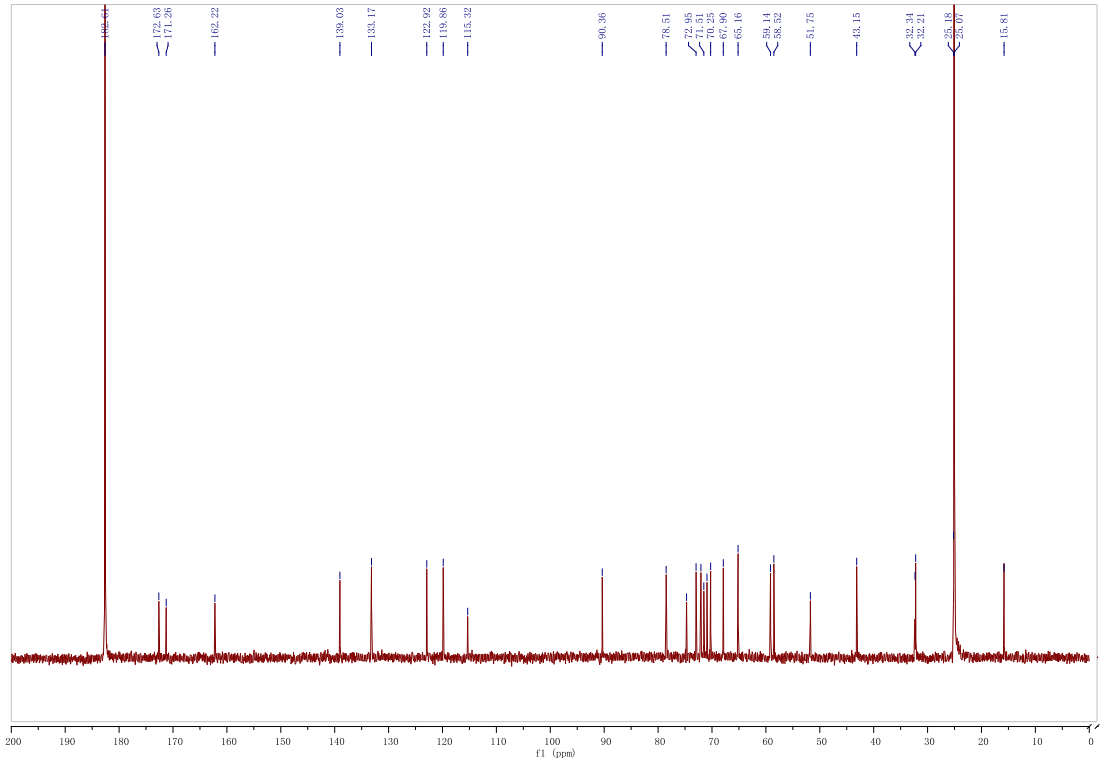


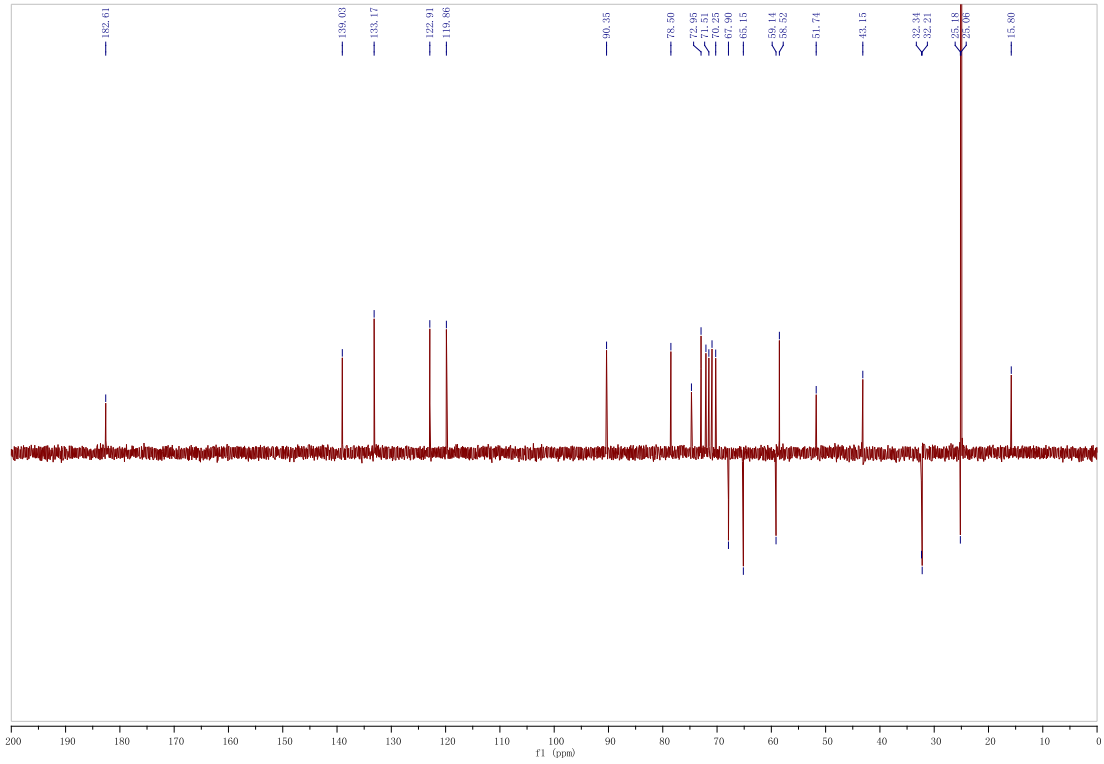
Fig. S11. NMR spectra of celesticetin. (**A**) ^{1}H NMR spectrum. (**B**) ^{13}C and DEPT 135 spectrum. (**C**) ^{1}H-^{1}H COSY spectrum. (**D**) HSQC spectrum. (**E**) HMBC spectrum. (**F**) NOESY spectrum.

A

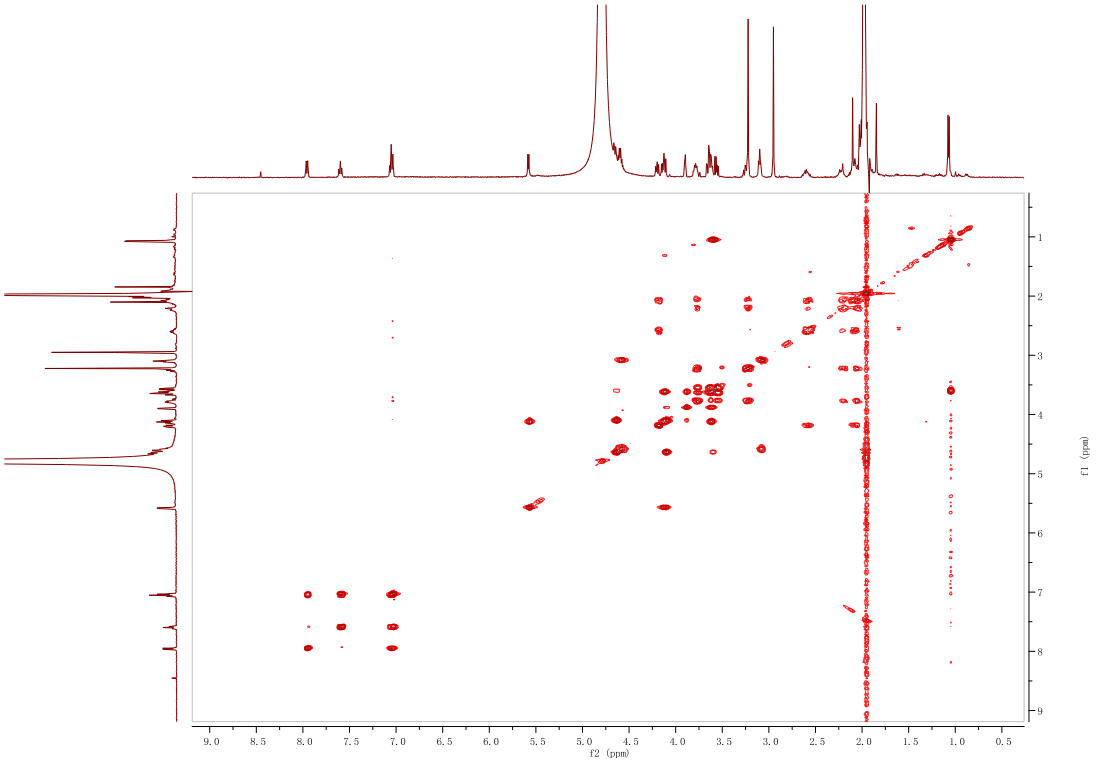


B

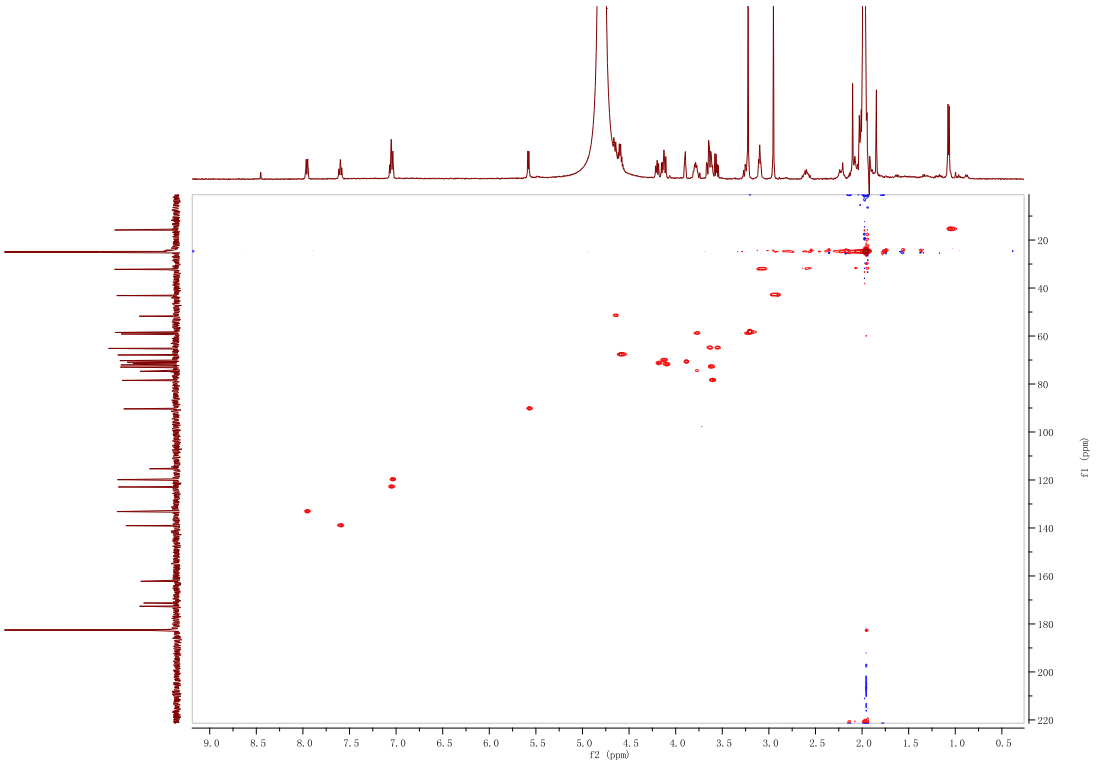

182.61
172.63
171.26
162.22
139.03
133.17
122.92
119.86
115.32
90.36
78.51
72.95
71.51
70.25
67.90
65.16
59.14
58.52
51.75
43.15
32.34
32.21
25.18
25.07
15.81
f1 (ppm)


182.61
139.03
133.17
122.91
119.86
90.35
78.50
72.95
71.51
70.25
67.90
65.15
59.14
58.52
51.74
43.15
32.34
32.21
25.18
25.06
15.80
f1 (ppm)

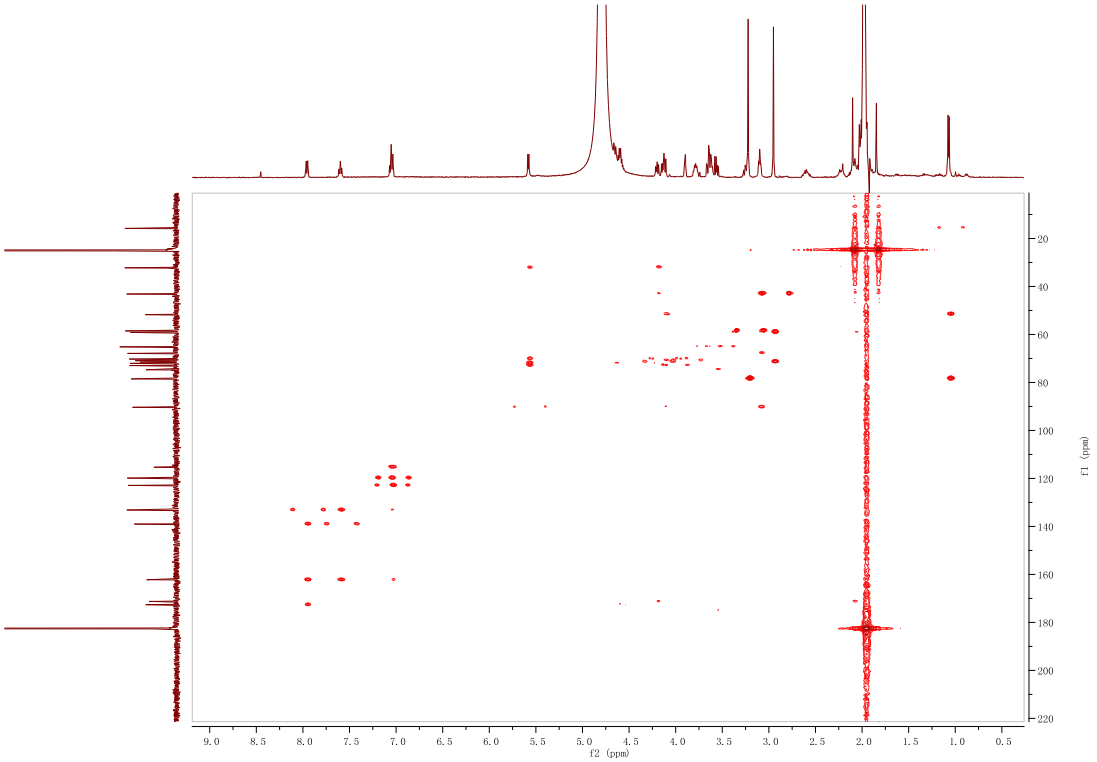
C



D



E



F

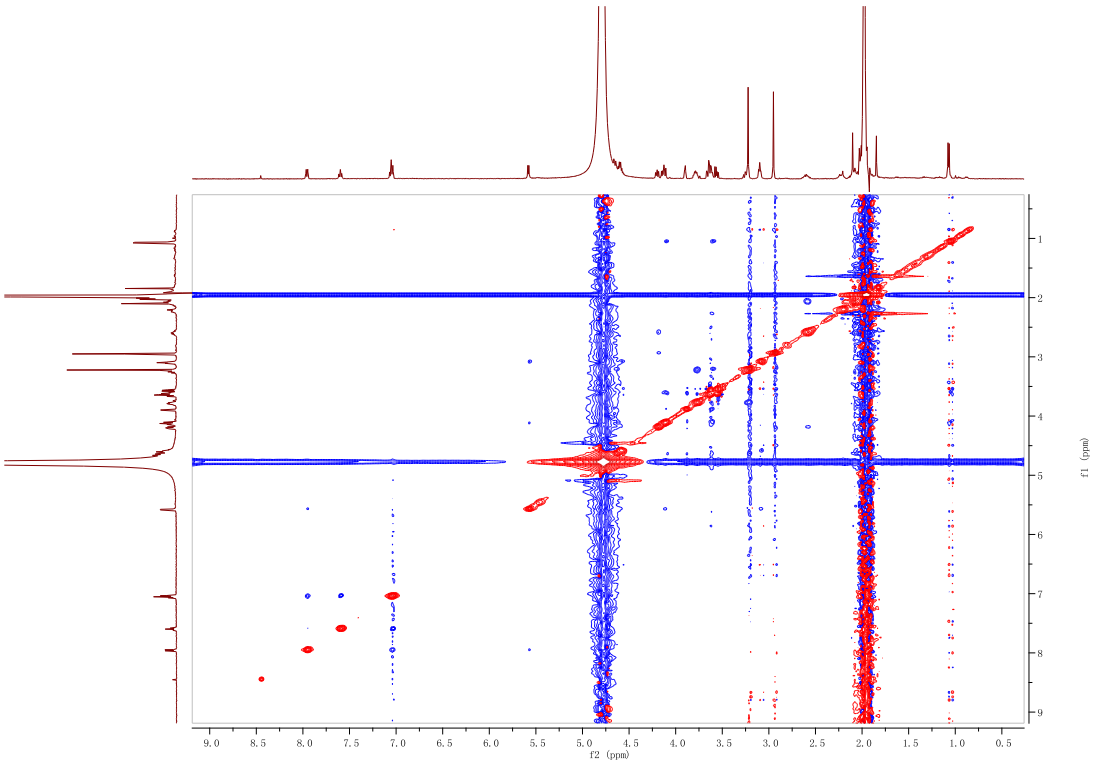
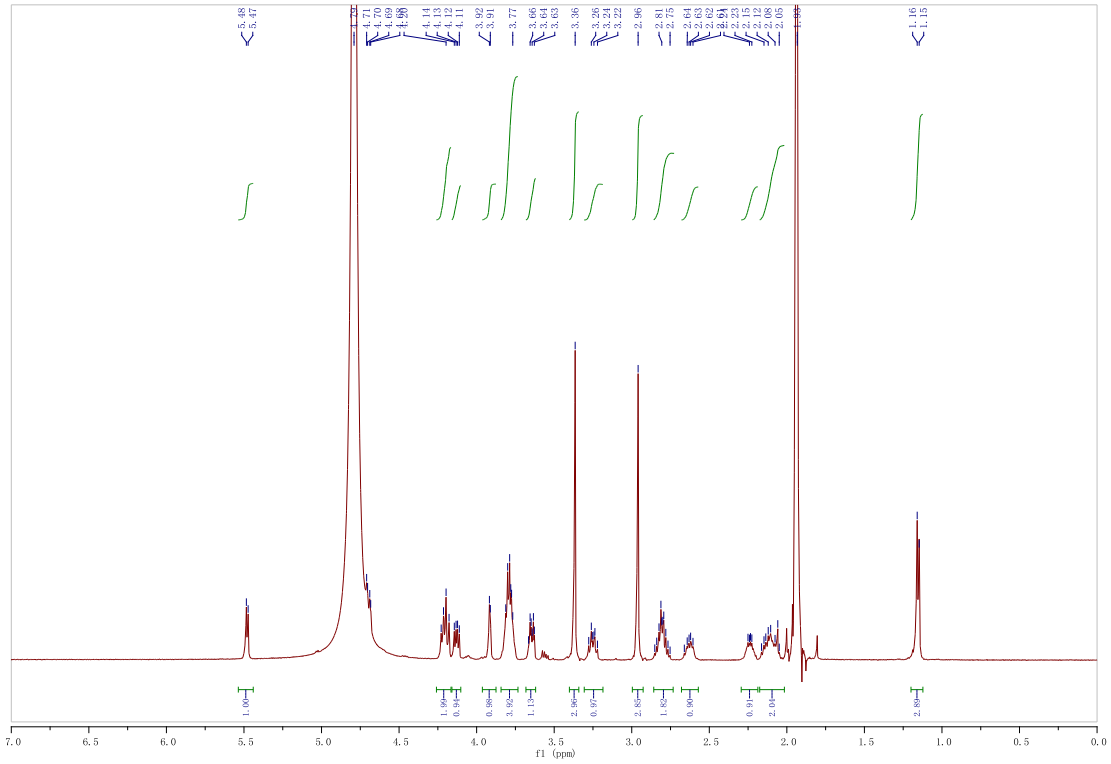
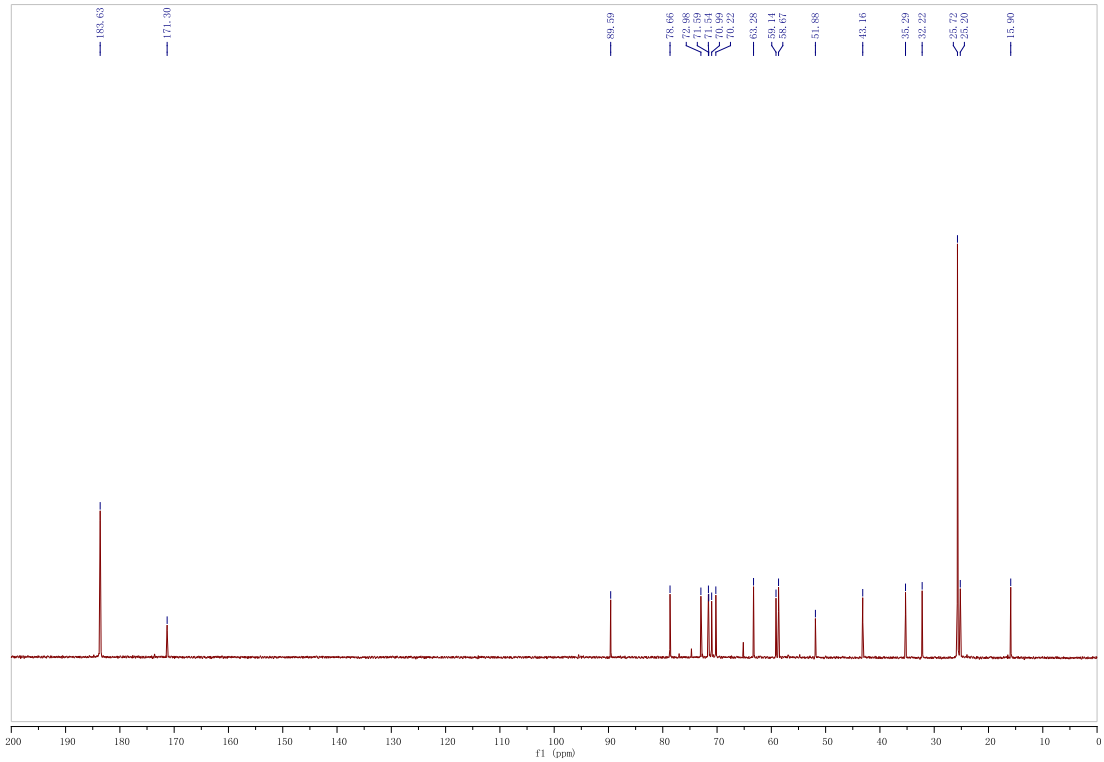


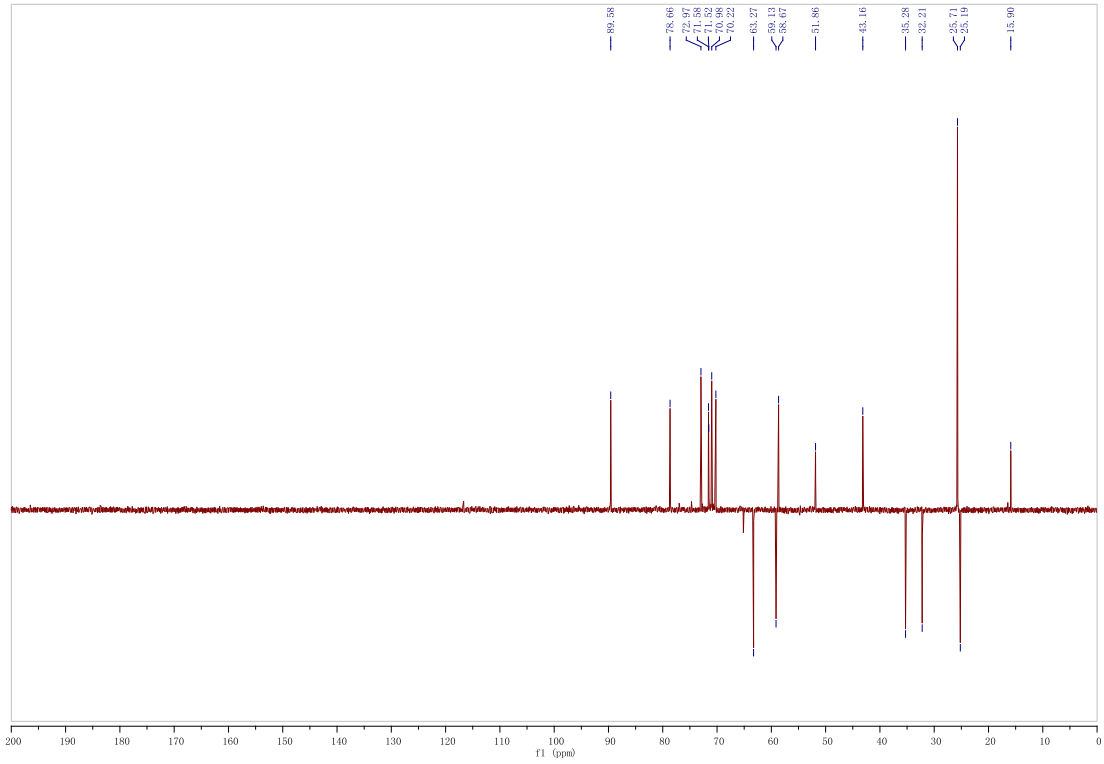
Fig. S12. NMR spectra of desalicetin. (**A**) ^{1}H NMR spectrum. (**B**) ^{13}C and DEPT 135 spectrum. (**C**) ^{1}H-^{1}H COSY spectrum. (**D**) HSQC spectrum. (**E**) HMBC spectrum. (**F**) NOESY spectrum.

A

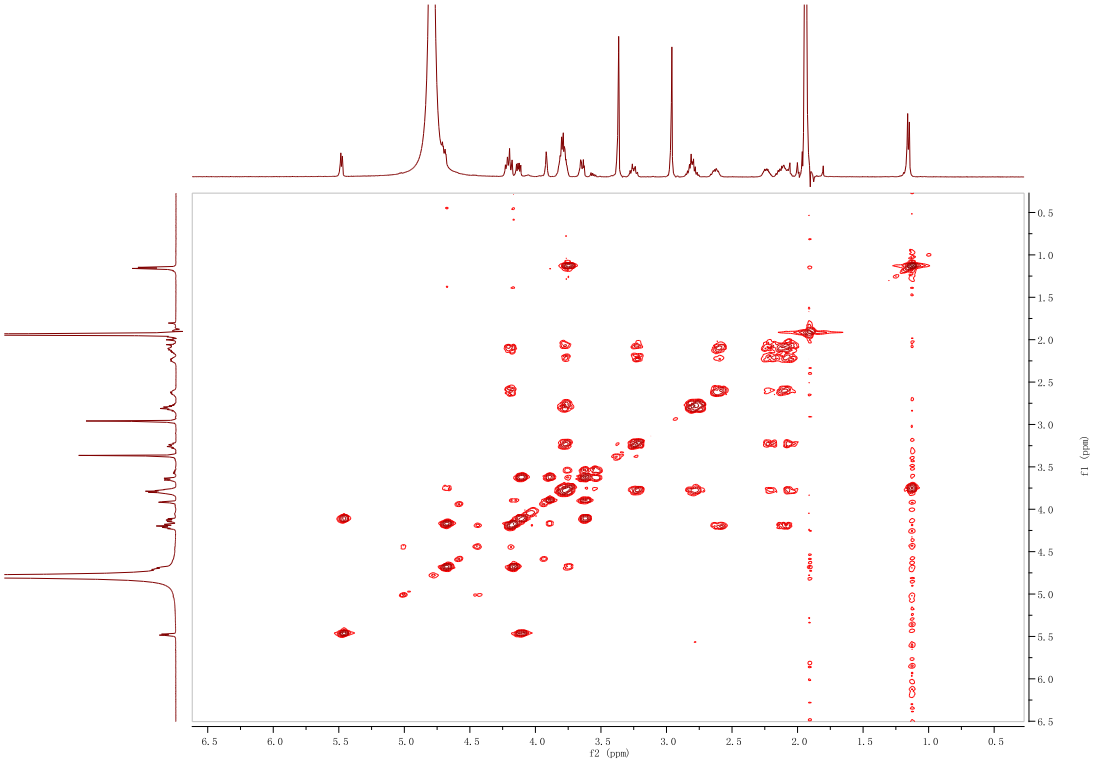


B

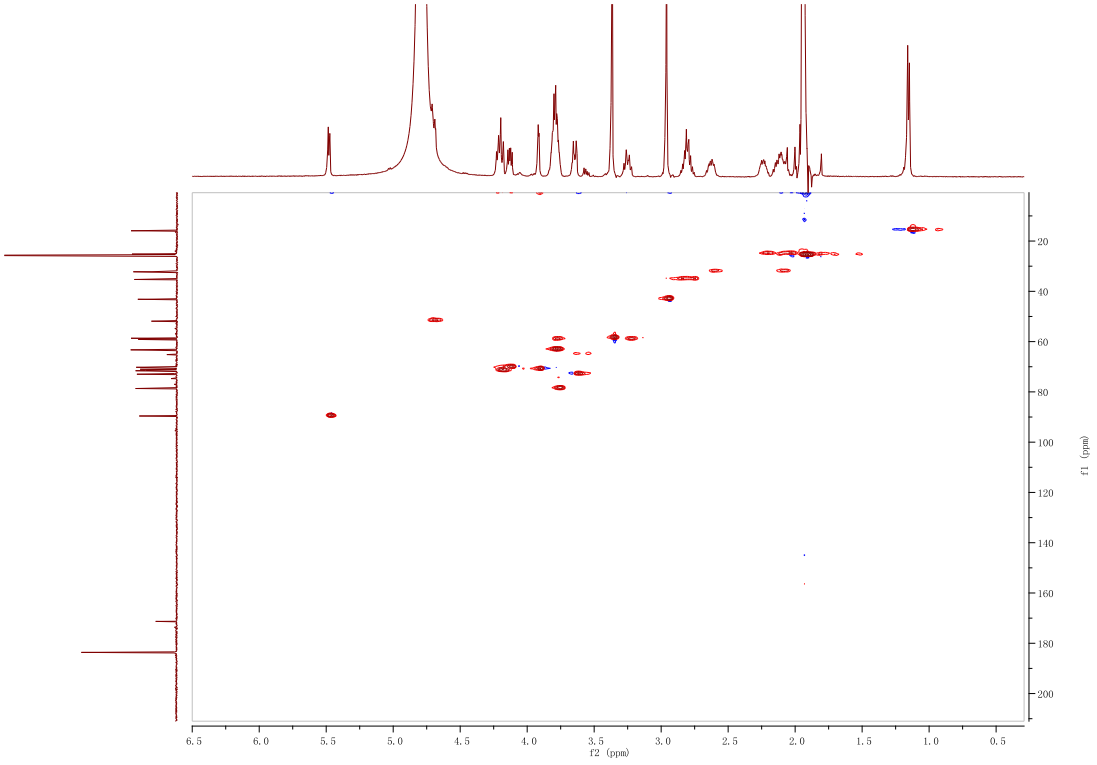

183.63
171.30
89.59
78.66
72.98
71.58
71.54
70.99
70.22
63.28
59.14
58.67
51.88
43.16
35.29
32.22
25.72
25.20
15.90
f1 (ppm)


89.58
78.66
72.97
71.98
71.52
70.98
70.22
63.27
59.13
58.67
51.86
43.16
35.28
32.21
25.71
25.19
15.90
f1 (ppm)

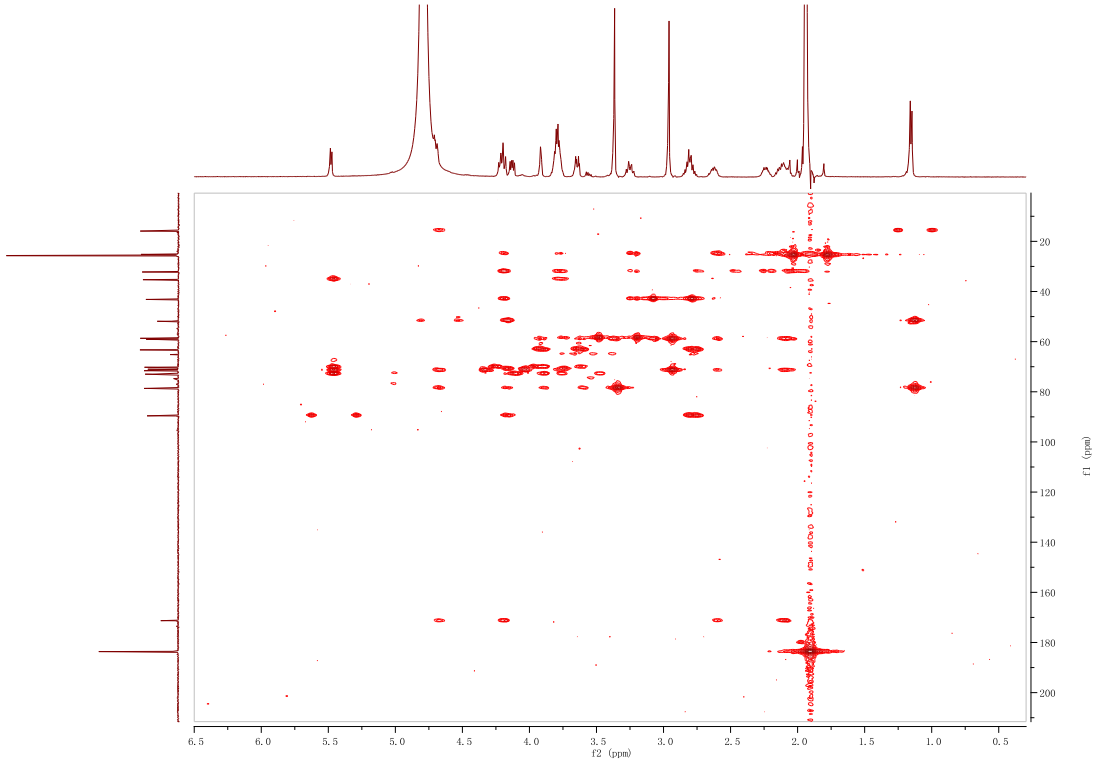
C



D



E



F

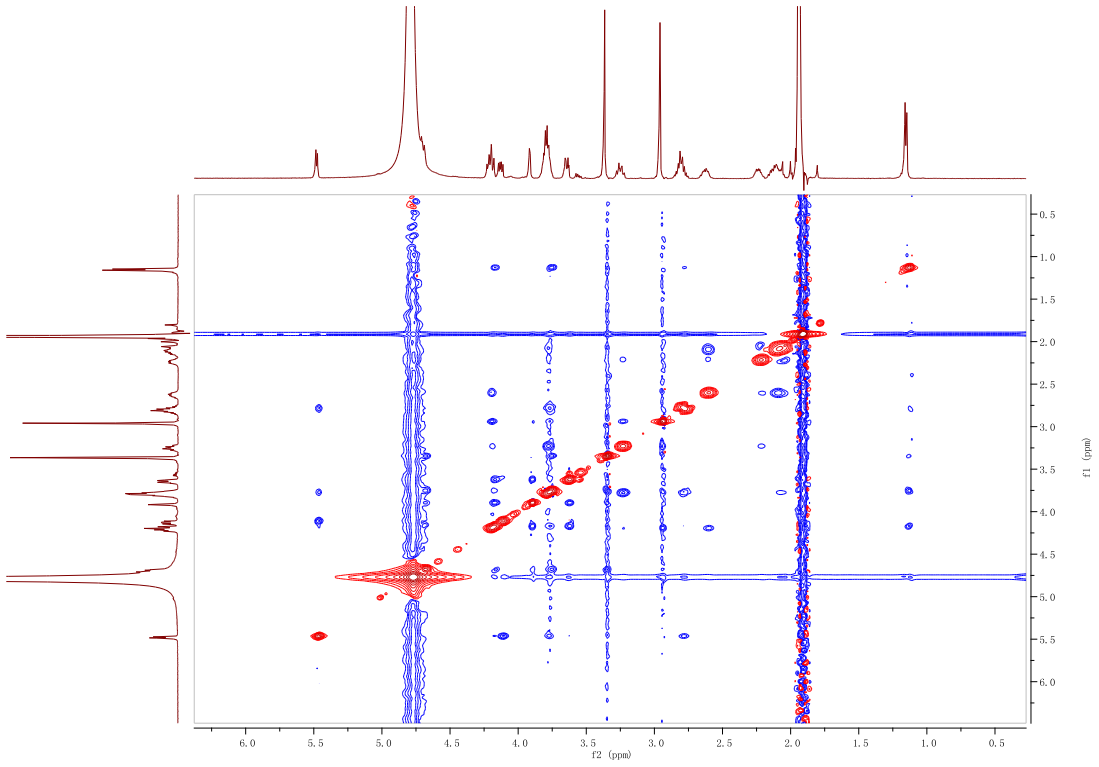
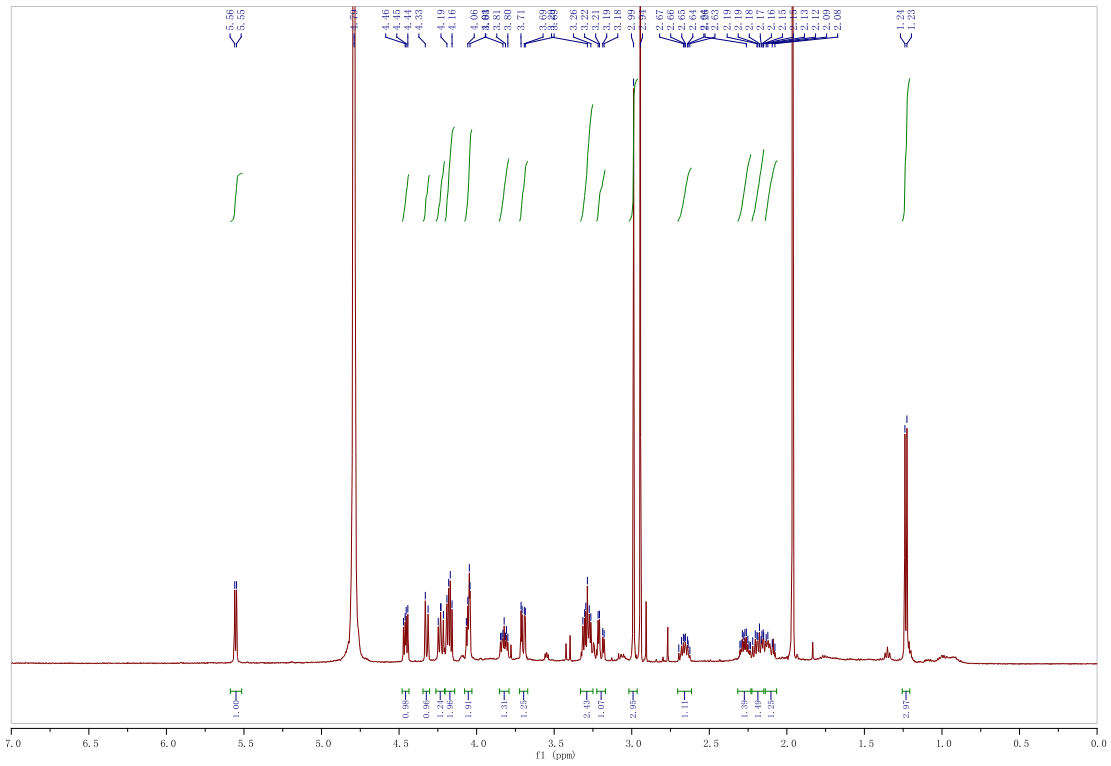
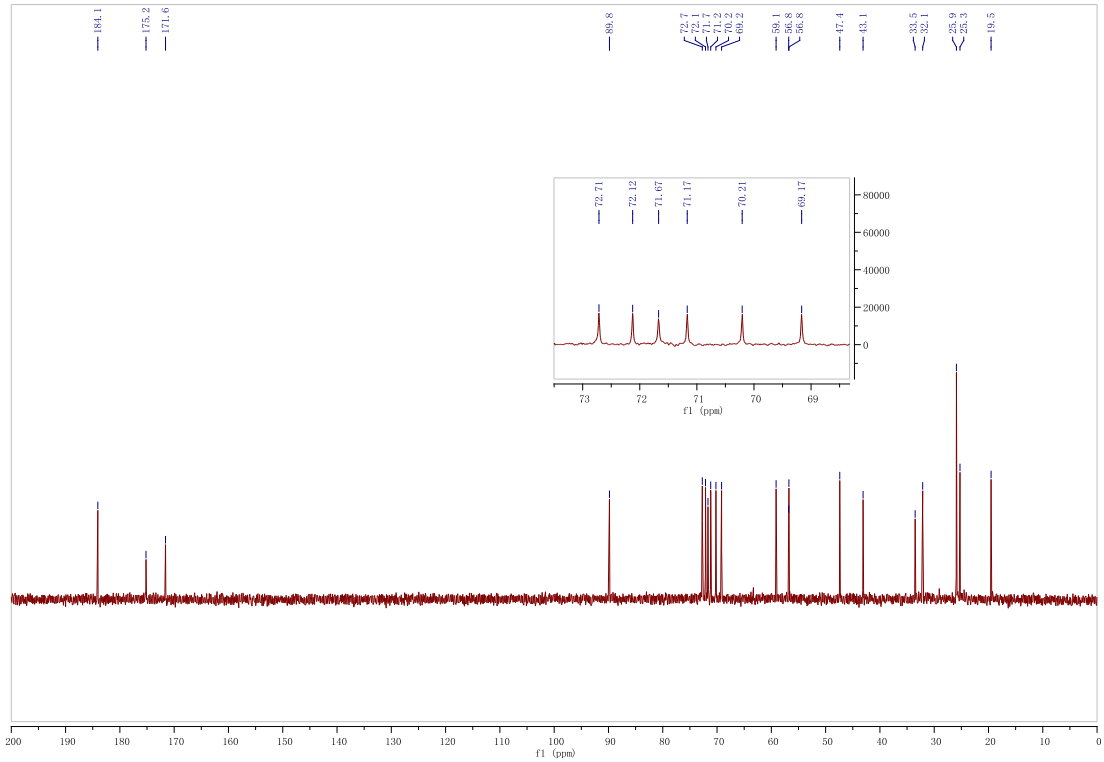


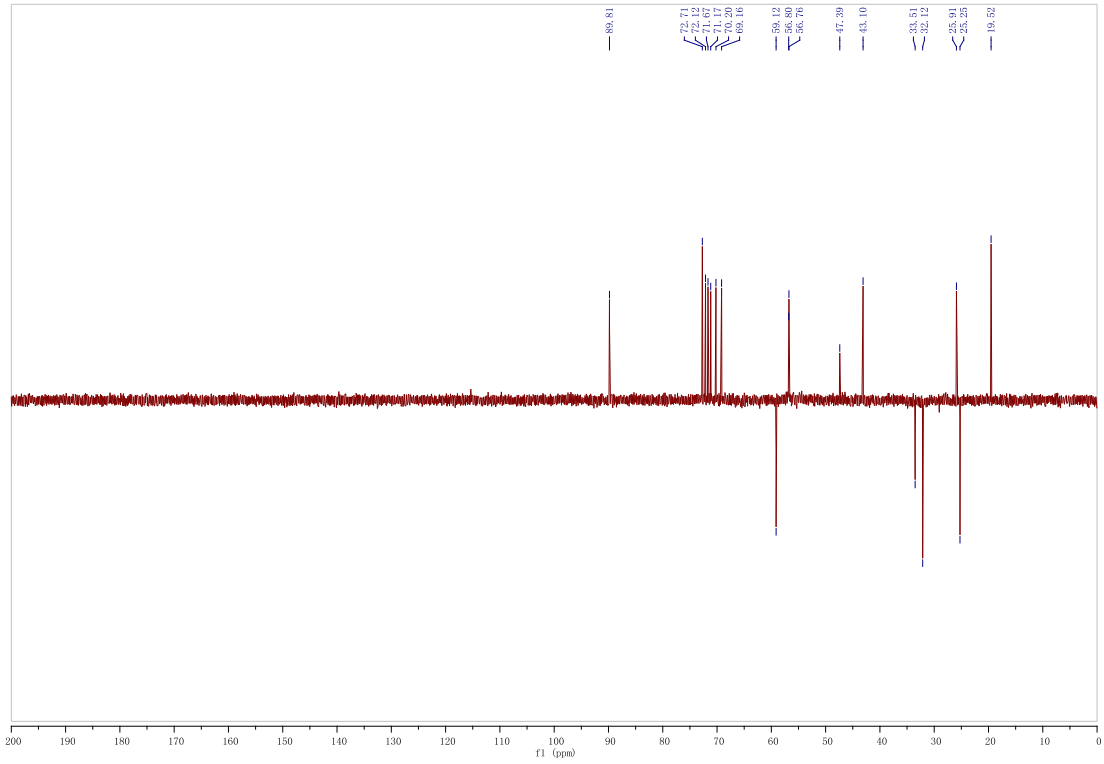
Fig. S13. NMR spectra of compound **1**. (**A**) ^{1}H NMR spectrum. (**B**) ^{13}C and DEPT 135 spectrum. (**C**) ^{1}H-^{1}H COSY spectrum. (**D**) HSQC spectrum. (**E**) HMBC spectrum. (**F**) NOESY spectrum.

A

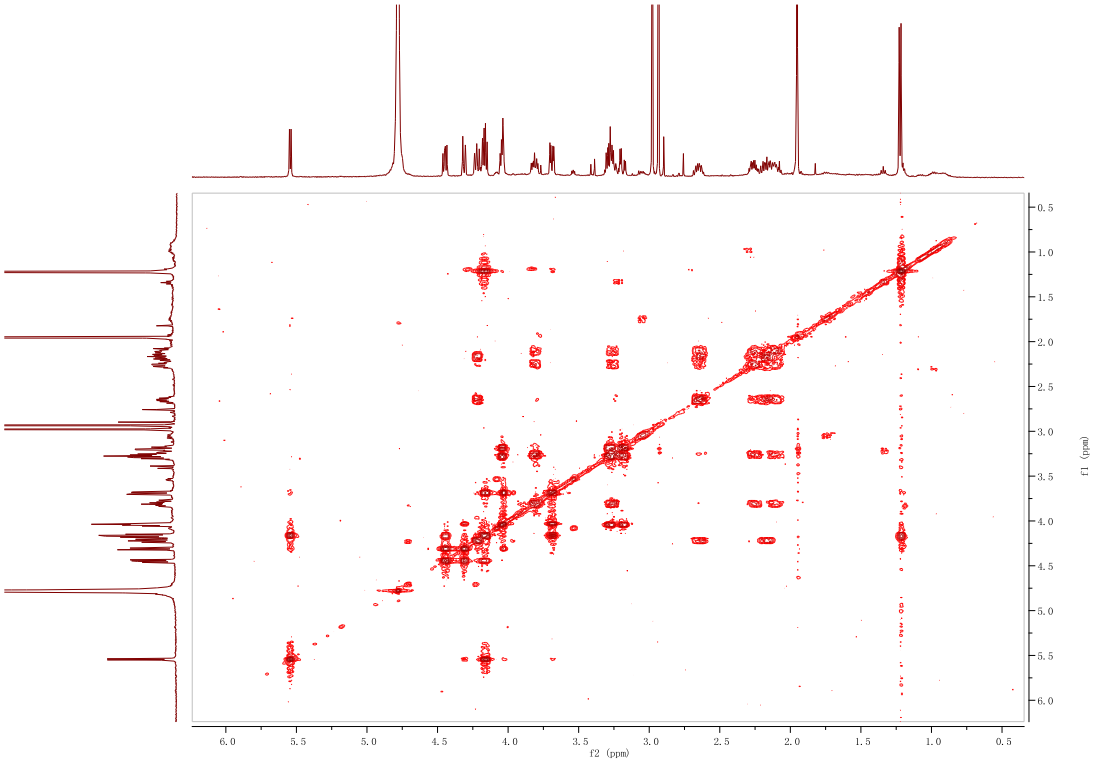


B

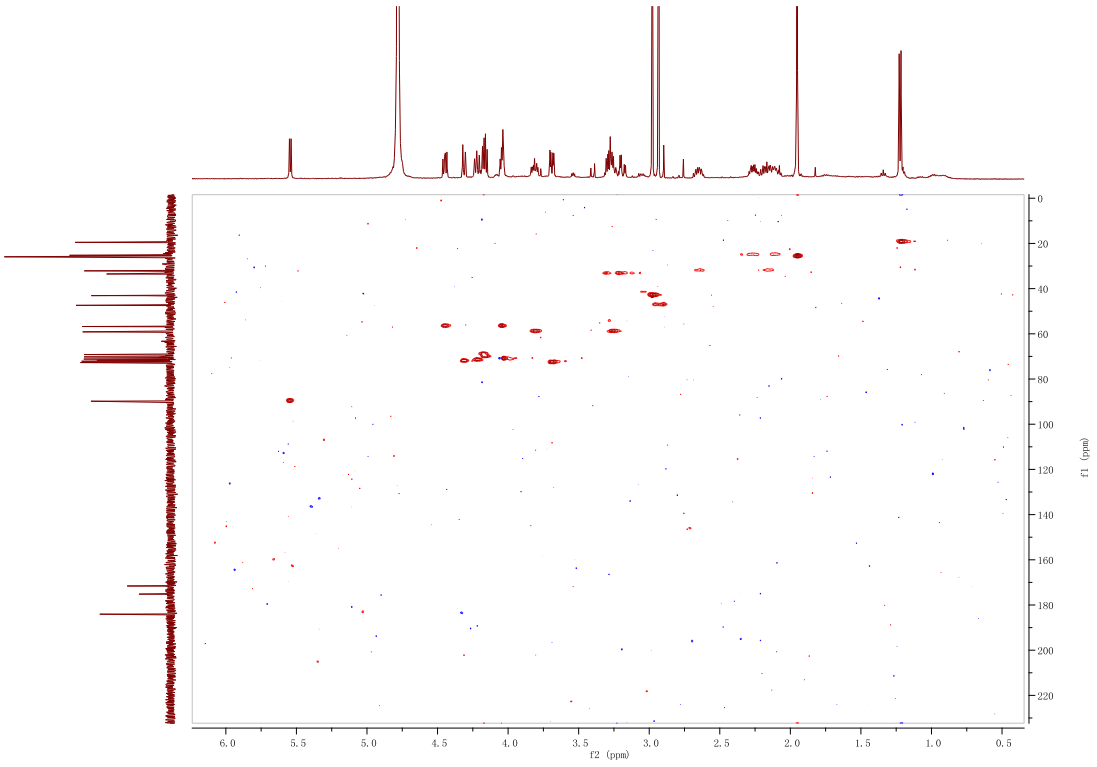

184.1
175.2
171.6
89.8
72.7
72.1
71.7
71.2
70.2
69.2
59.1
56.8
56.8
47.4
43.1
33.5
32.1
25.9
25.3
19.5
72.71
72.12
71.67
71.17
70.21
69.17
80000
60000
40000
20000
0
73
72
71
70
69
f1 (ppm)
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


89.81
72.71
72.12
71.67
71.17
70.20
69.16
59.12
56.80
56.76
47.39
43.10
33.51
32.12
25.91
25.25
19.52
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

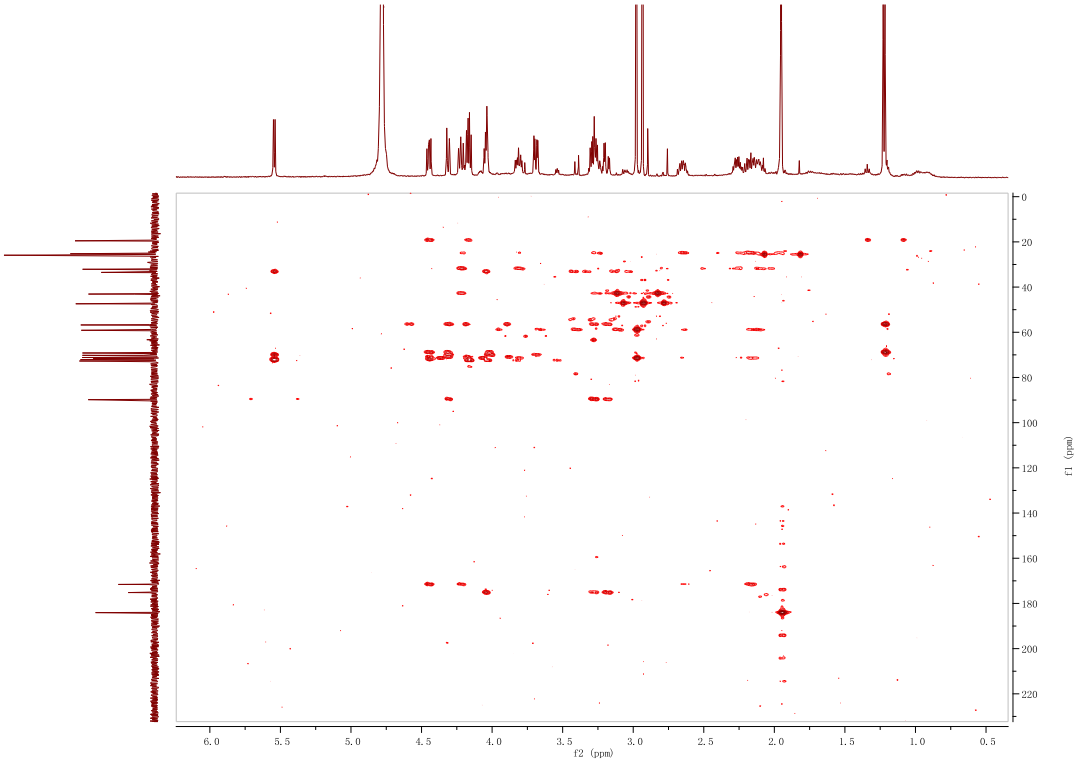
C



D



E



F

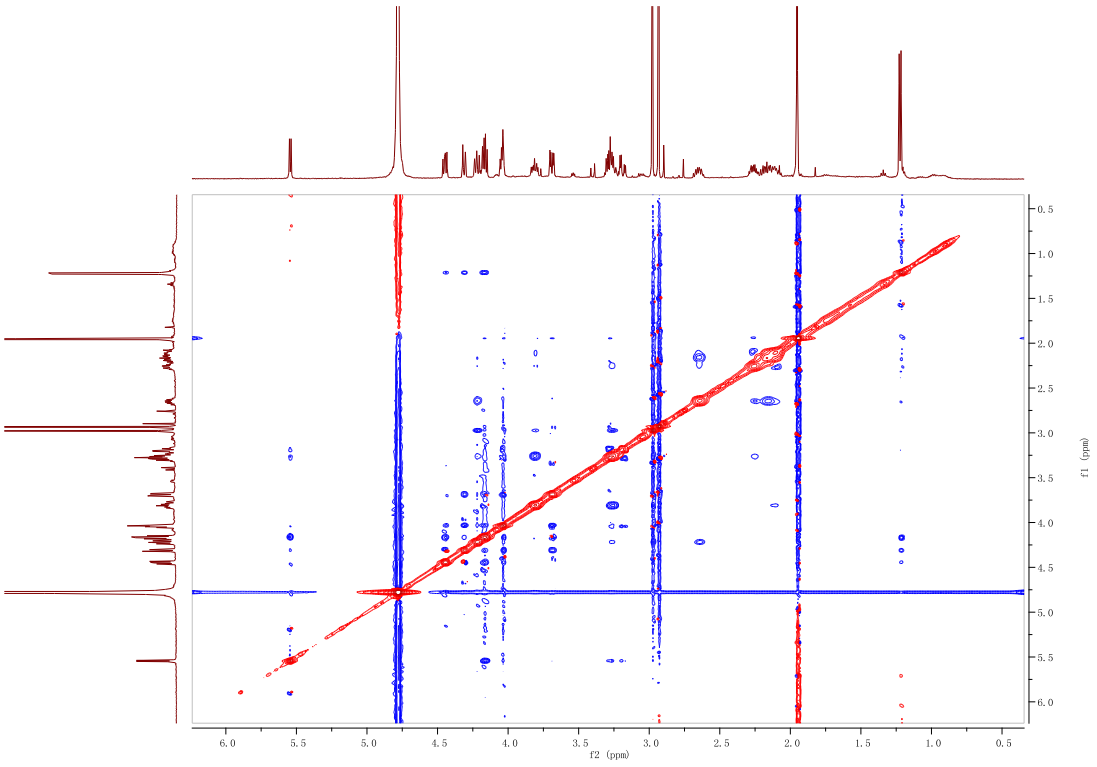
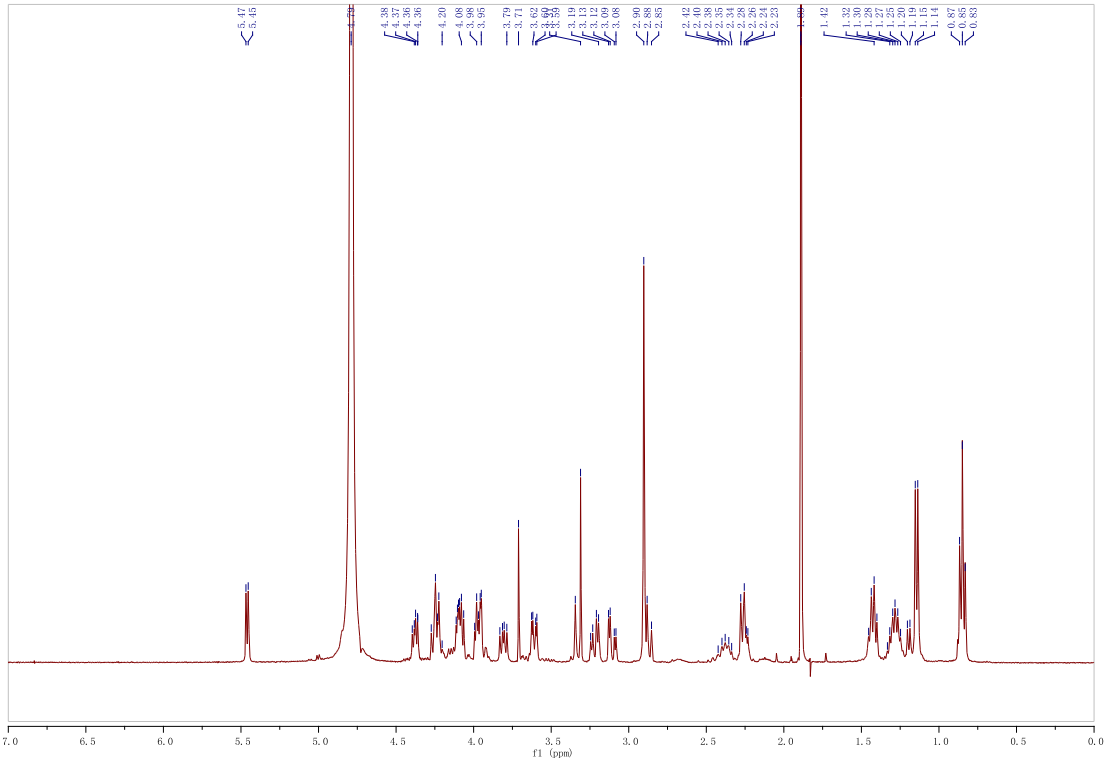
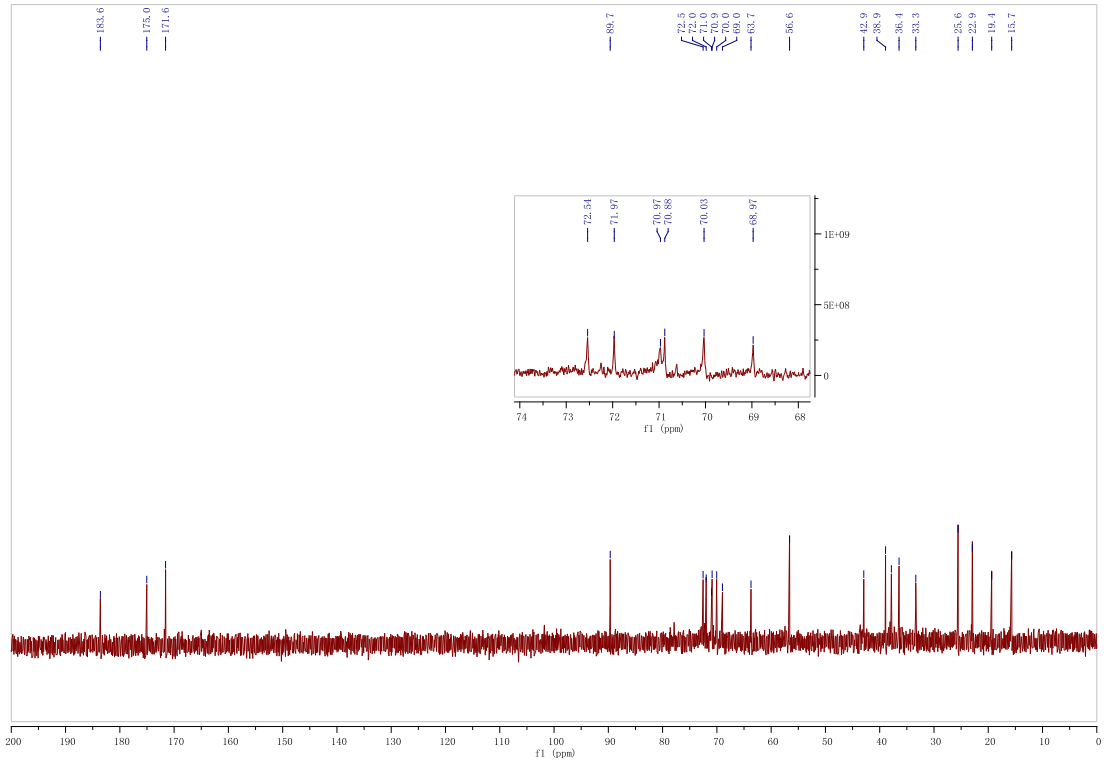


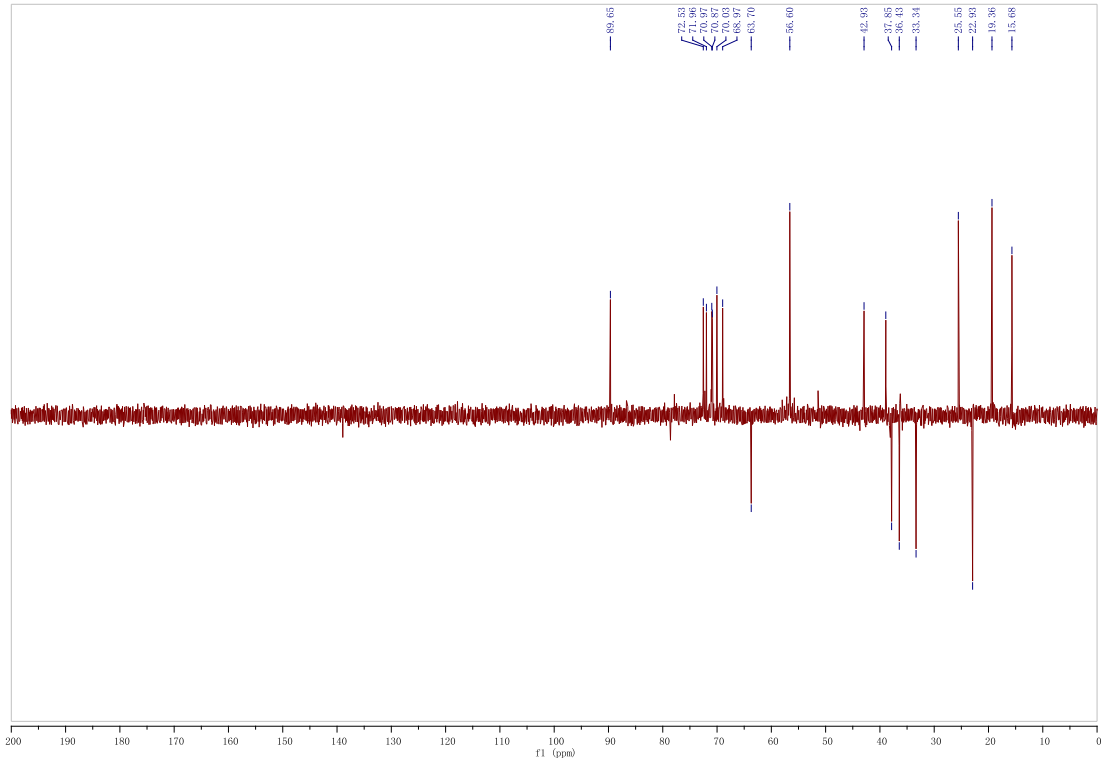
Fig. S14. NMR spectra of compound **5**. (**A**) ^{1}H NMR spectrum. (**B**) ^{13}C and DEPT 135 spectrum. (**C**) ^{1}H-^{1}H COSY spectrum. (**D**) HSQC spectrum. (**E**) HMBC spectrum. (**F**) NOESY spectrum.

A

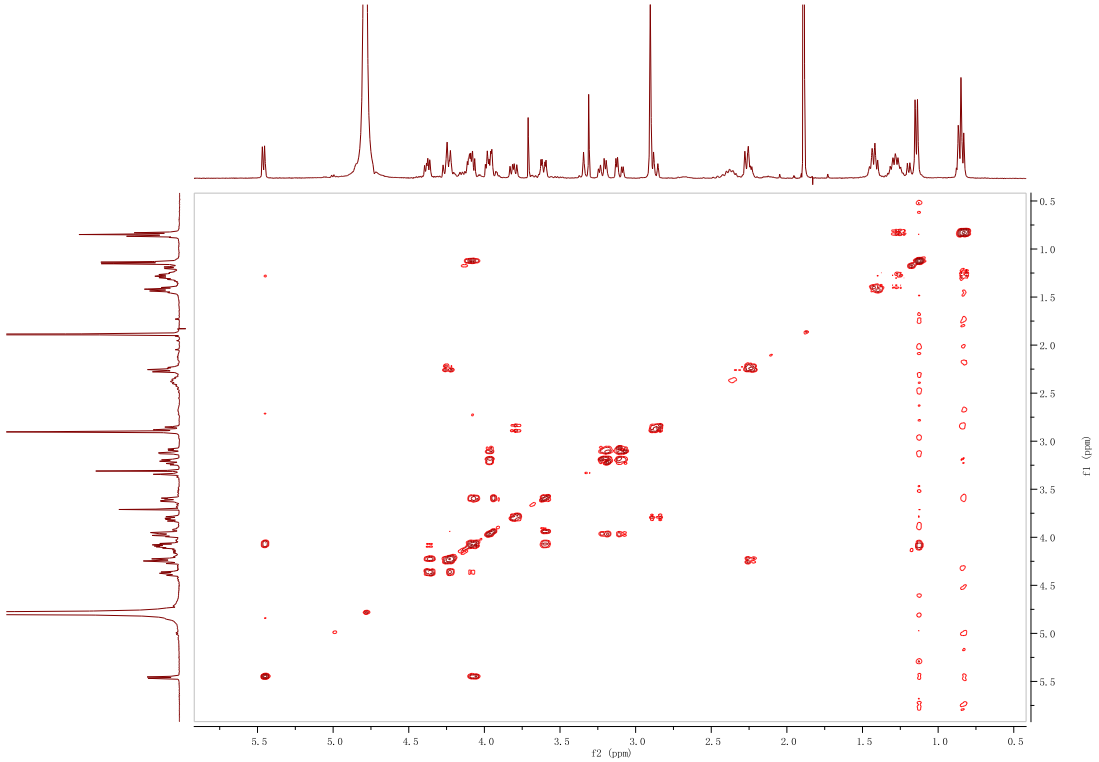


B

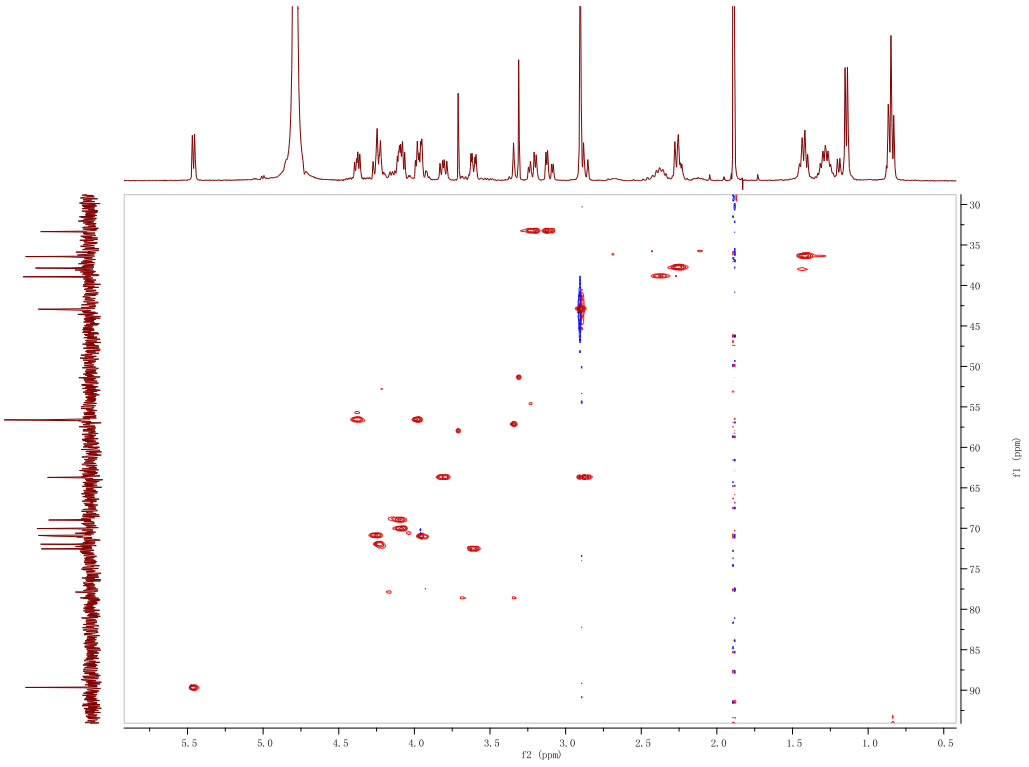

183.6
175.0
171.6
89.7
72.5
72.0
71.0
70.9
70.0
69.0
63.7
56.6
42.9
38.9
36.4
33.3
25.6
22.9
19.4
15.7
72.54
71.97
70.97
70.88
70.03
68.97
1E+09
5E+08
0
74
73
72
71
70
69
68
f1 (ppm)
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


89.65
72.53
71.96
70.97
70.87
70.03
68.97
63.70
56.60
42.93
37.85
36.43
33.34
25.55
22.93
19.36
15.68
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

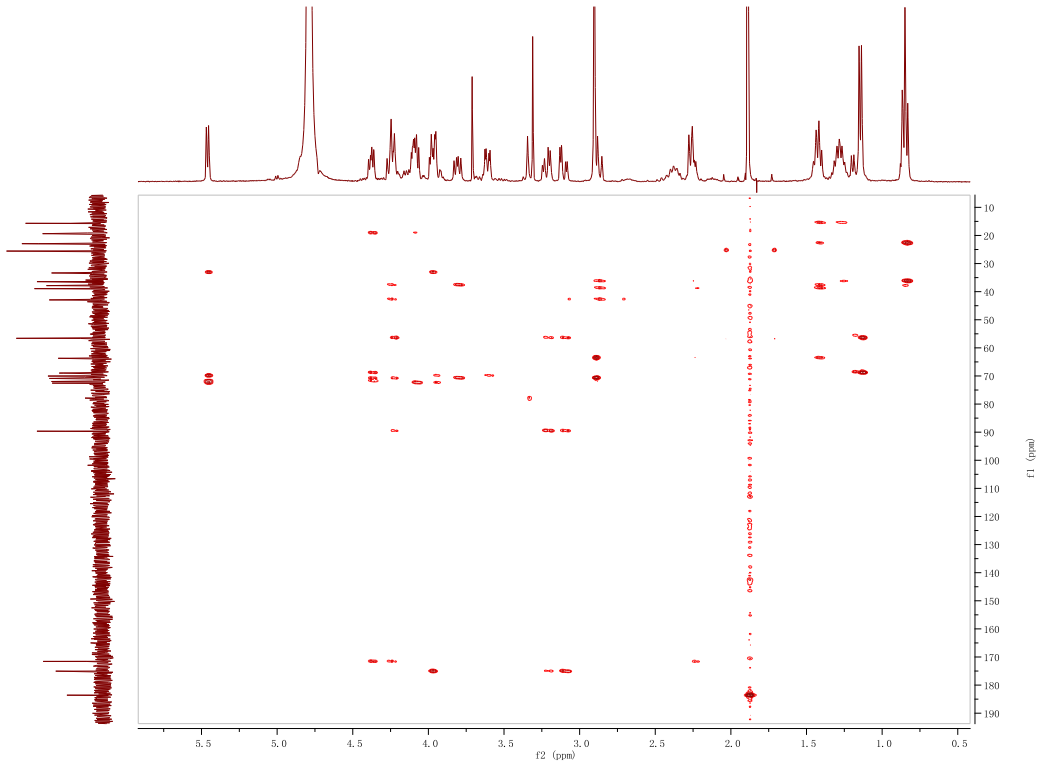
C



D



E



F

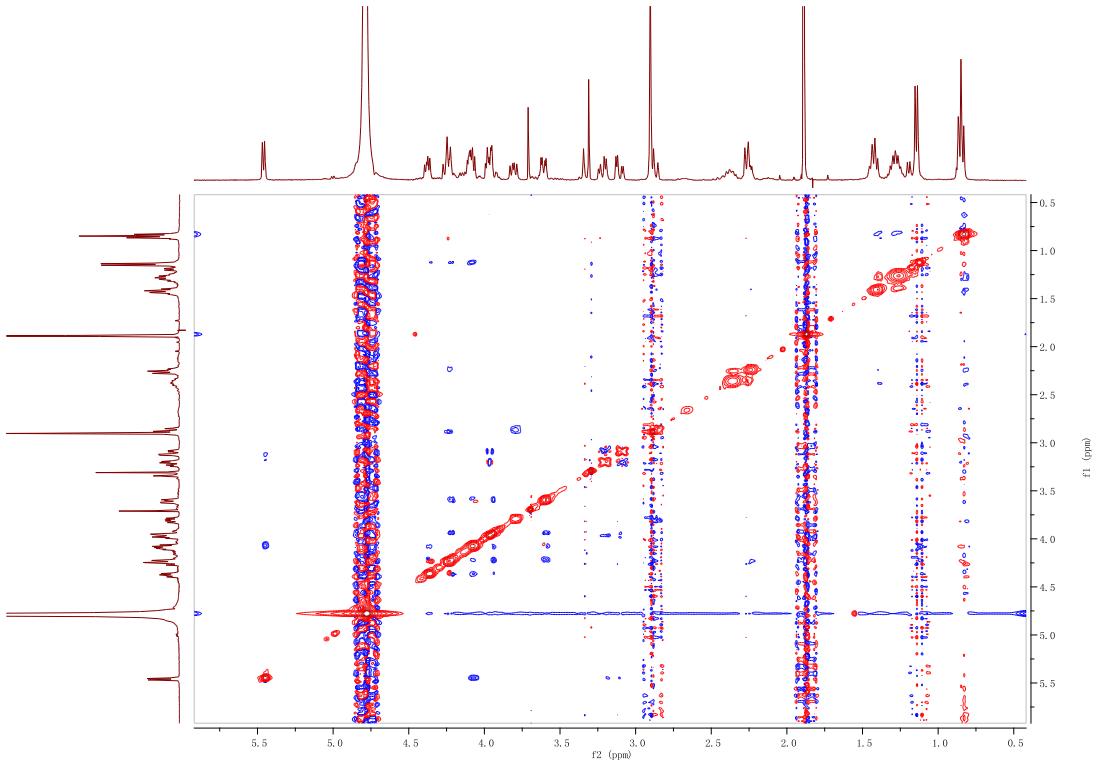
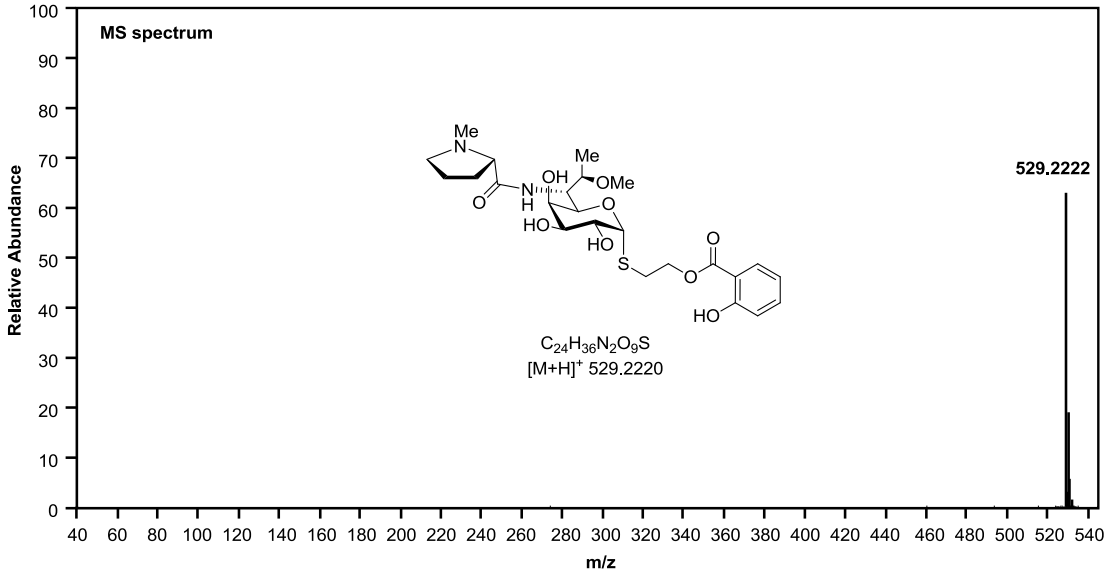


Fig. S15. MS and MS/MS spectra of celesticetin. Selective ion monitoring on $[M+H]^+$ (m/z 529.2220) ion of celesticetin.

A



B

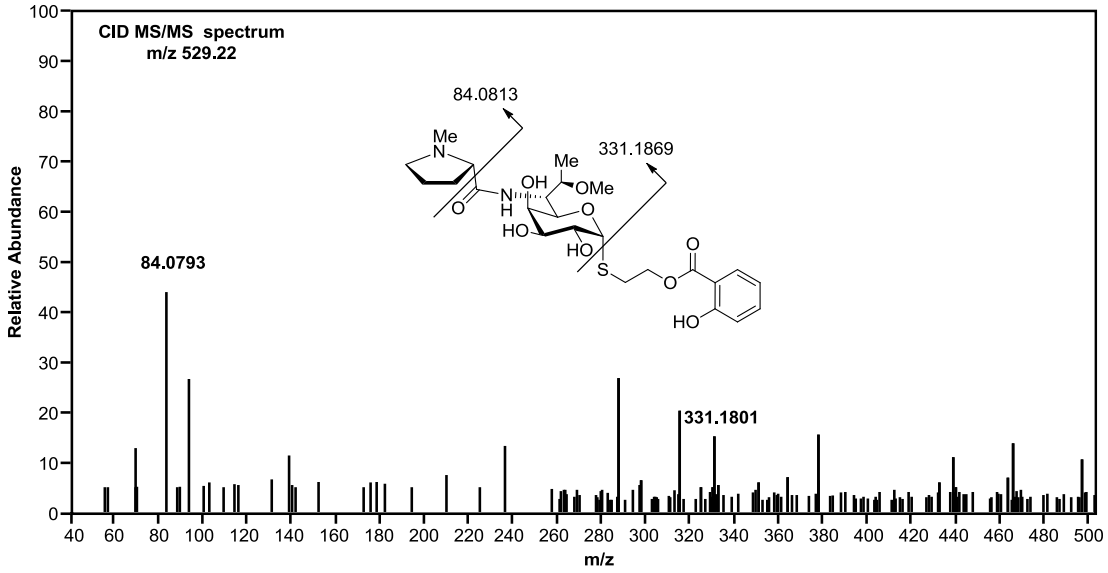
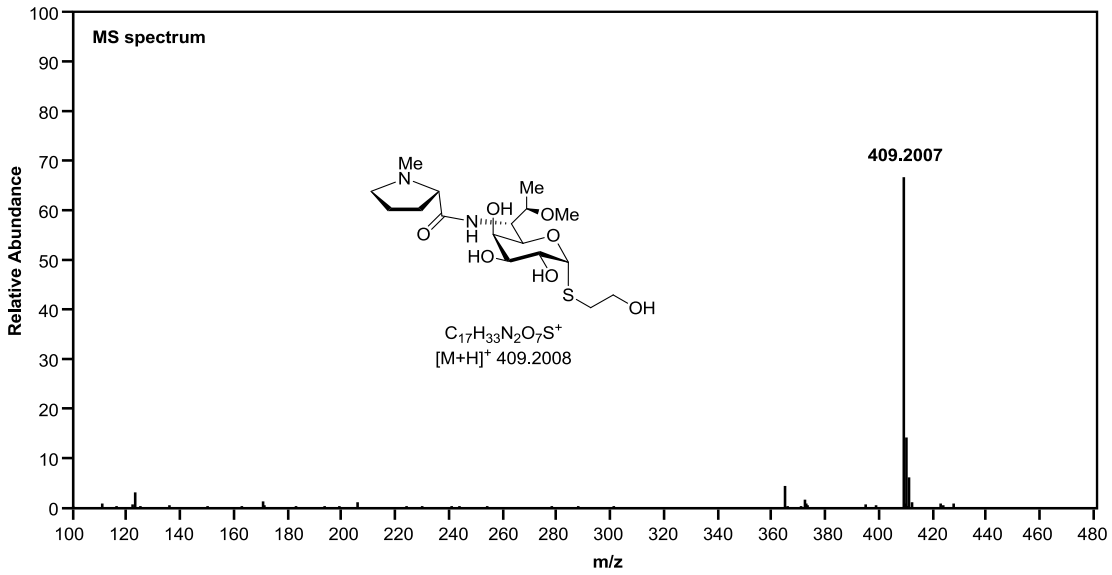


Fig. S16. MS and MS/MS spectra of desalicetin. Selective ion monitoring on $[M+H]^+$ (m/z 409.2008) ion of desalicetin.

A



B

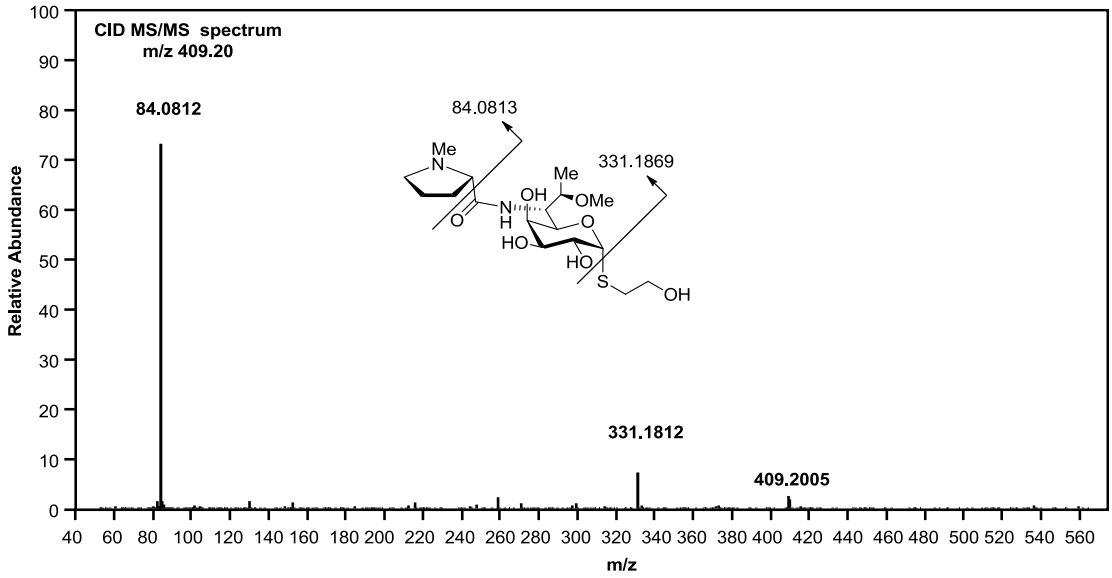
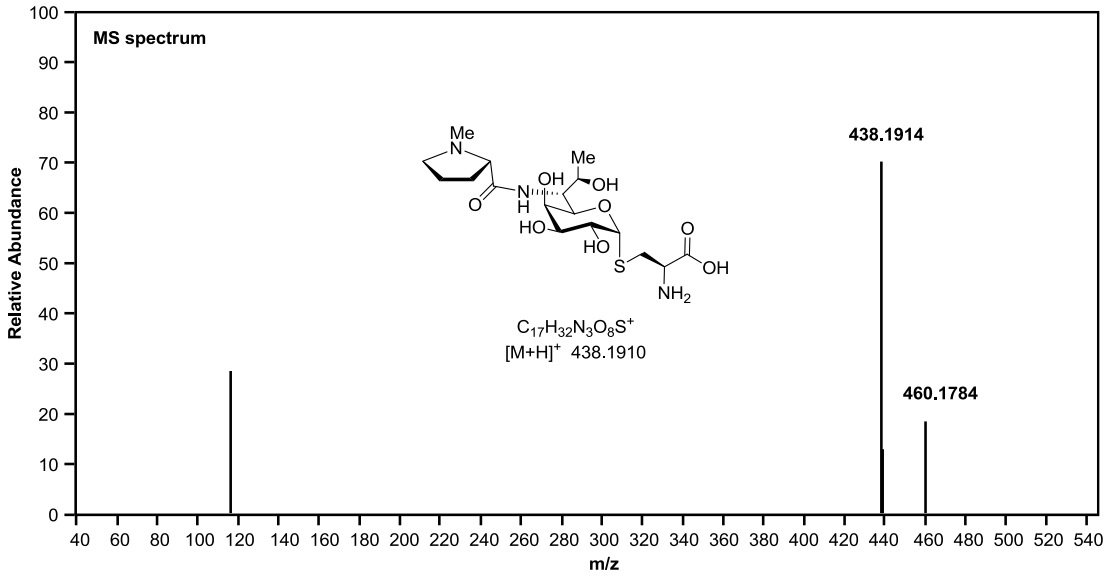


Fig. S17. MS and MS/MS spectra of compound **1**. Selective ion monitoring on $[M+H]^+$ (m/z 438.1910) ion of compound **1**.

A



B

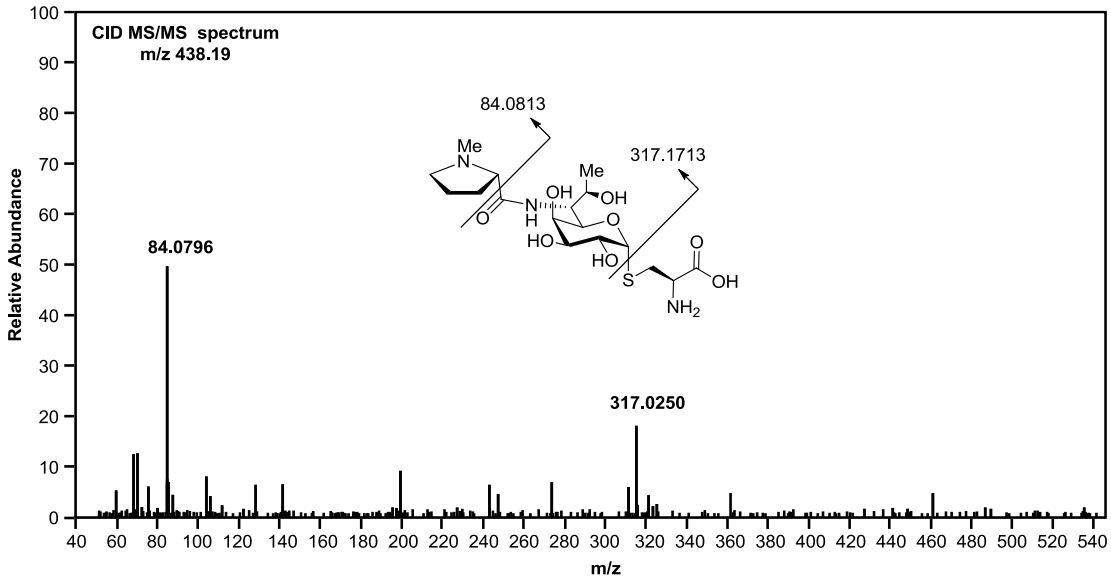
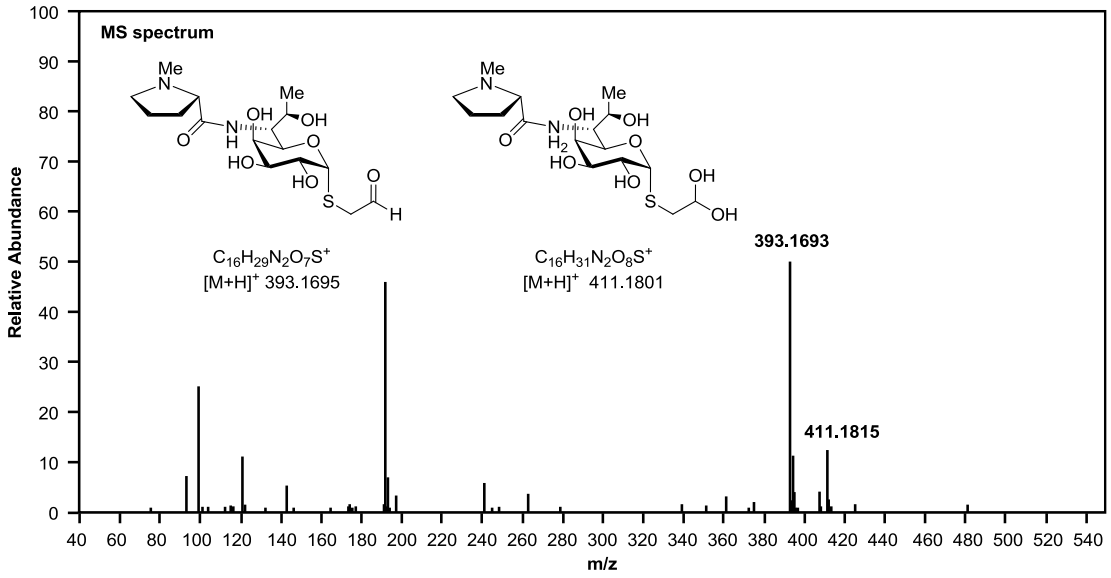


Fig. S18. MS and MS/MS spectra of compound **2**. Selective ion monitoring on $[M+H]^+$ (m/z 393.1695 and 411.1801) ion of compound **2** and its hydrate.

A



B

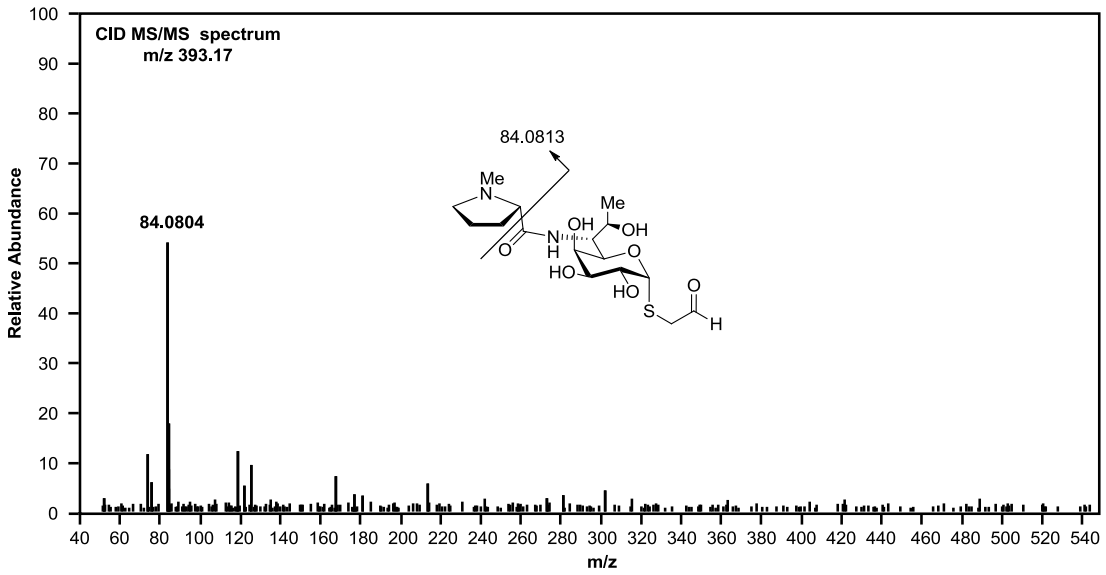
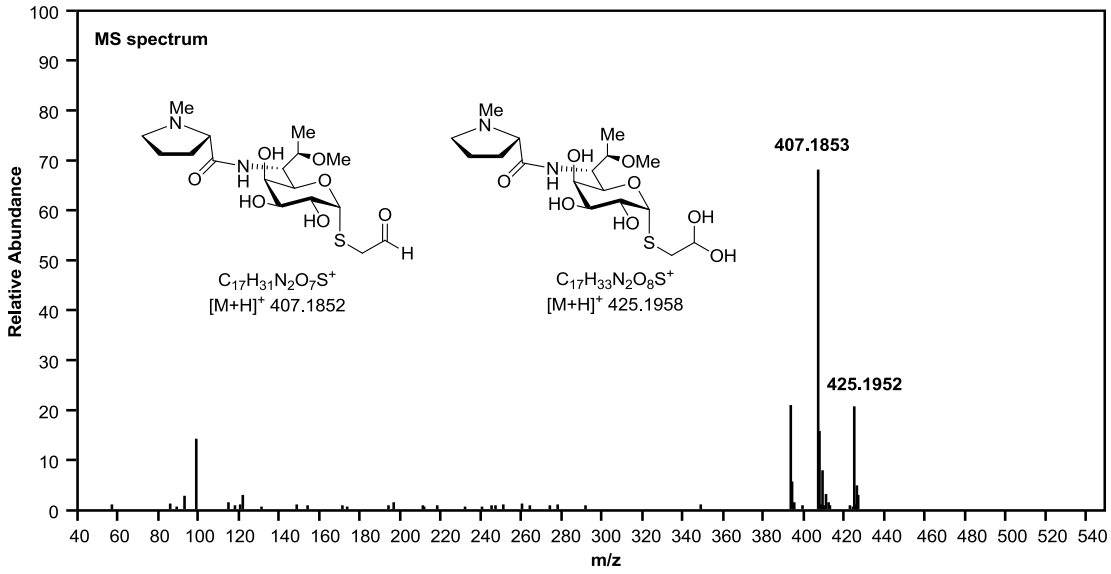


Fig. S19. MS and MS/MS spectra of compound **3**. Selective ion monitoring on $[M+H]^+$ (m/z 407.1852 and 425.1958) ion of compound **3** and its hydrate.

A



B

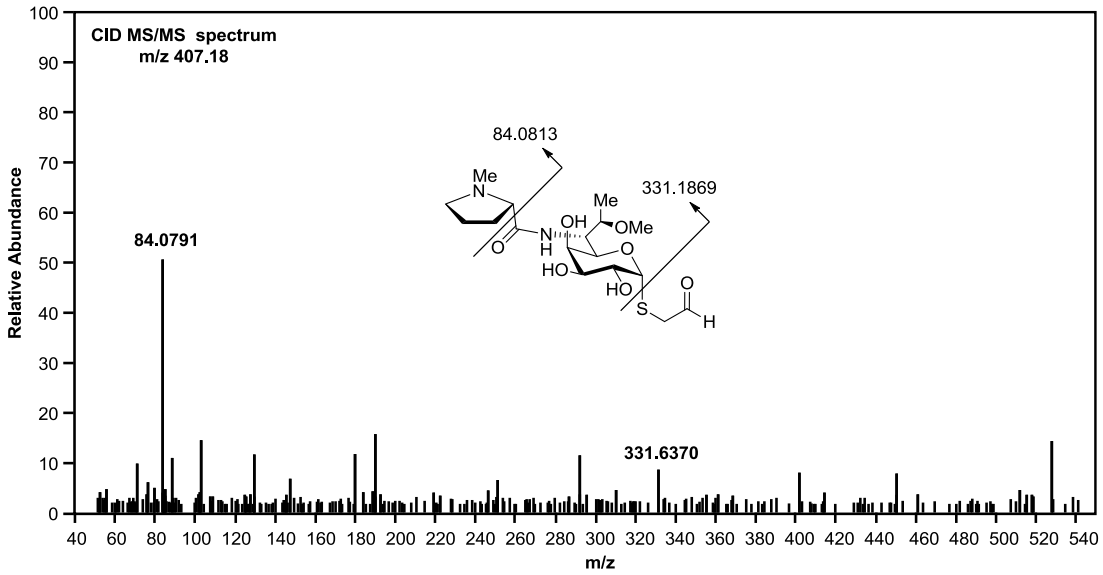
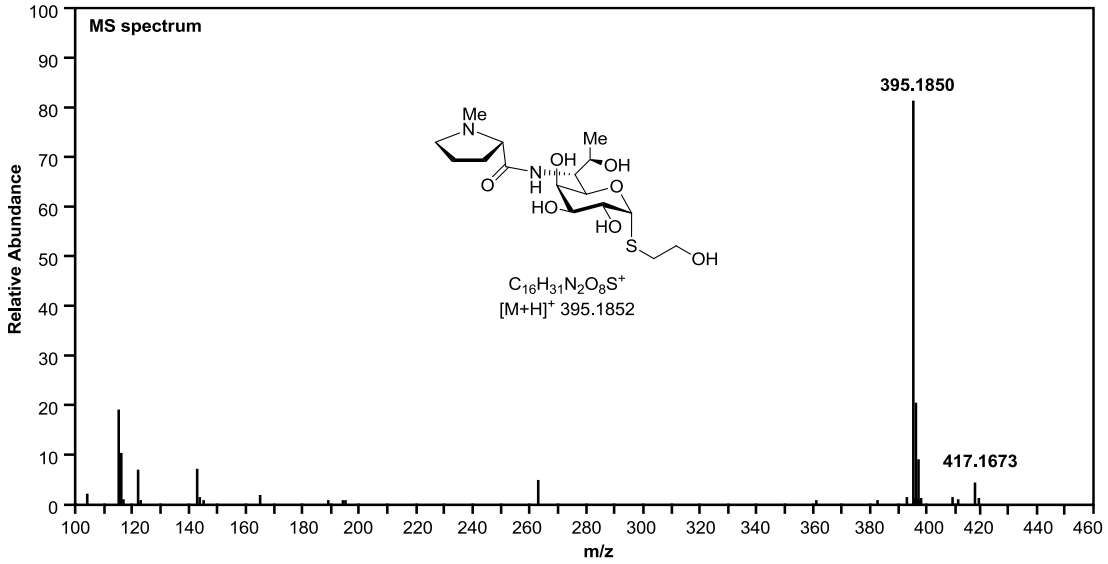


Fig. S20. MS and MS/MS spectra of compound **4**. Selective ion monitoring on $[M+H]^+$ (m/z 395.1852) ion of compound **4**.

A



B

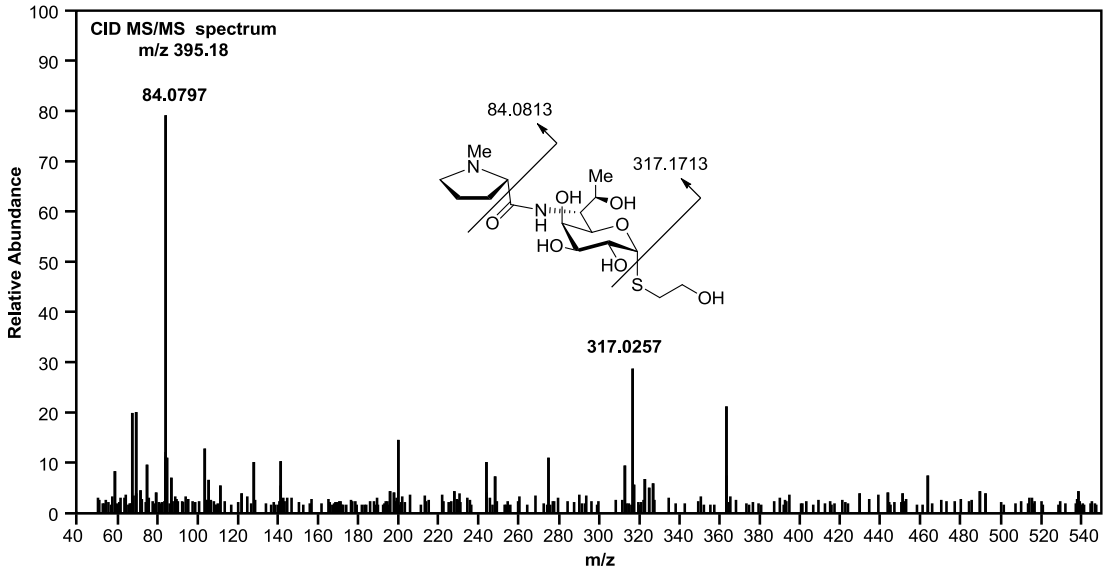
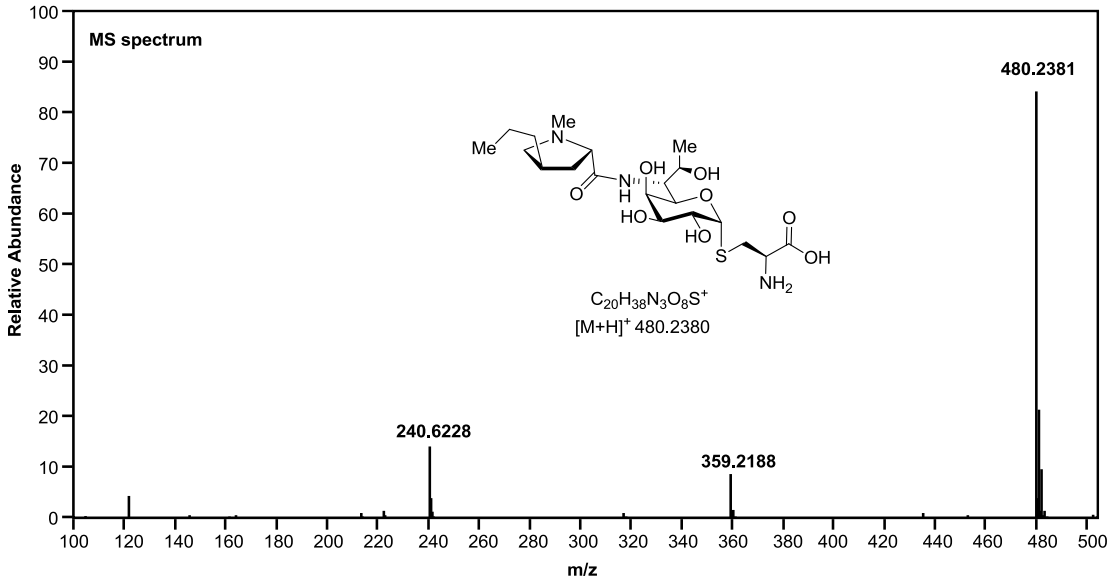


Fig. S21. MS and MS/MS spectra of compound **5**. Selective ion monitoring on $[M+H]^+$ (m/z 480.2380) ion of compound **5**.

A



B

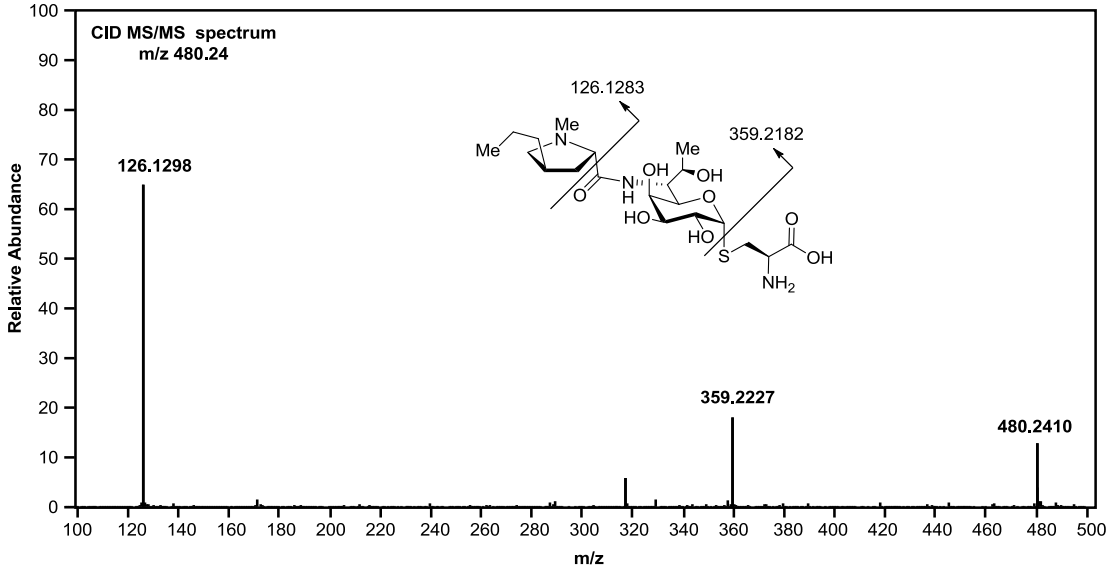
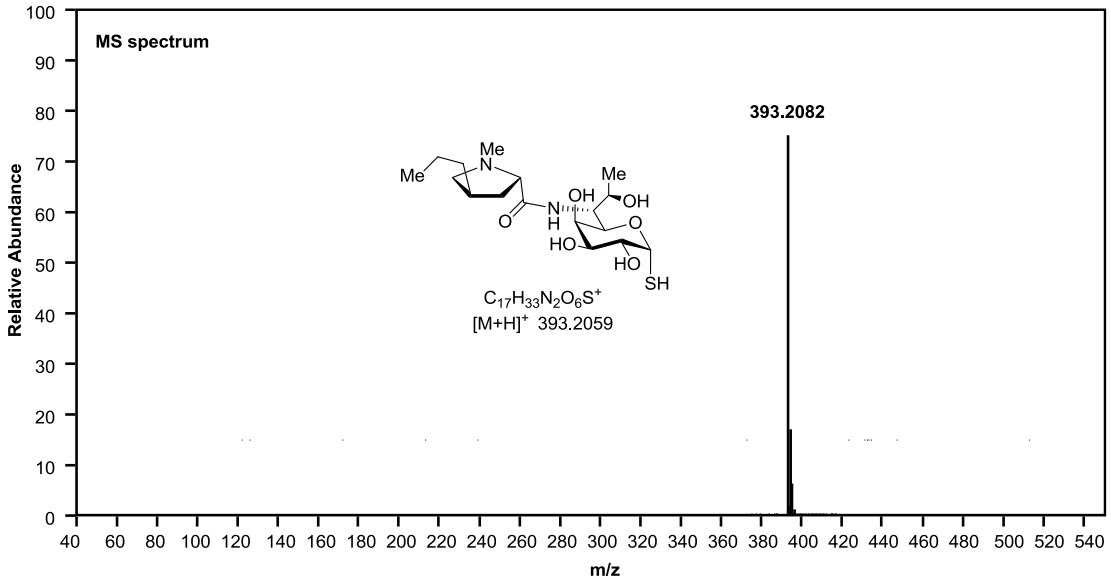


Fig. S22. MS and MS/MS spectra of compound **6**. Selective ion monitoring on $[M+H]^+$ (m/z 393.2082) ion of compound **6**.

A



B

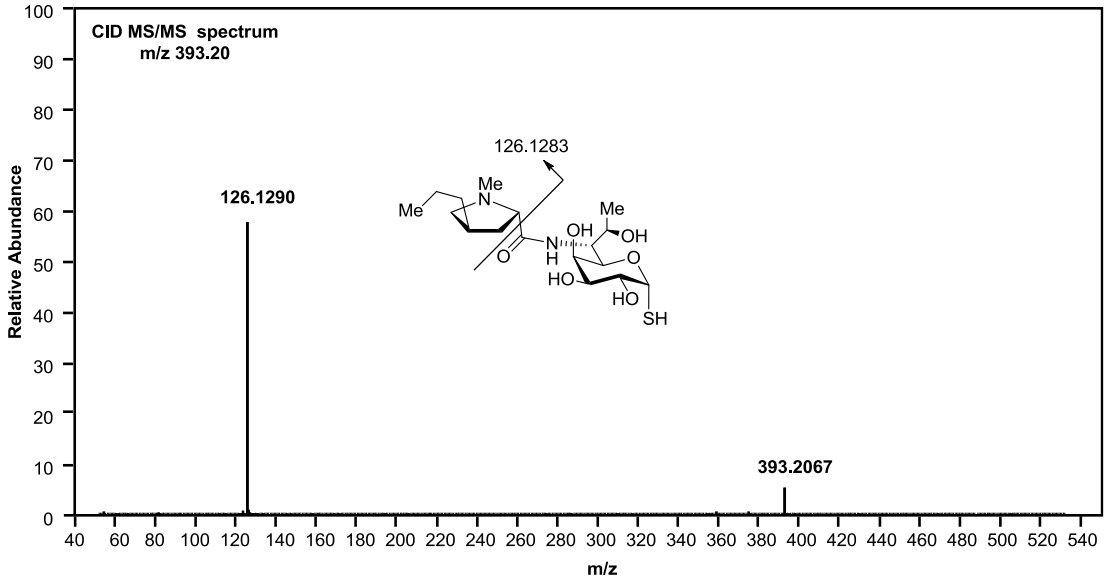
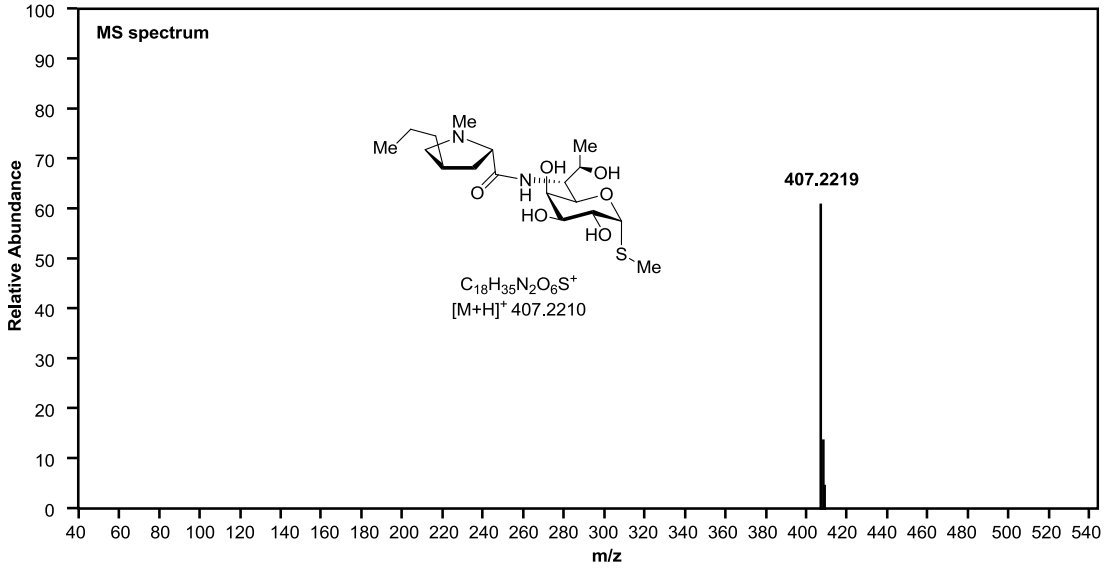


Fig. S23. MS and MS/MS spectra of lincomycin A. Selective ion monitoring on $[M+H]^+$ (m/z 407.2210) ion of lincomycin A.

A



B

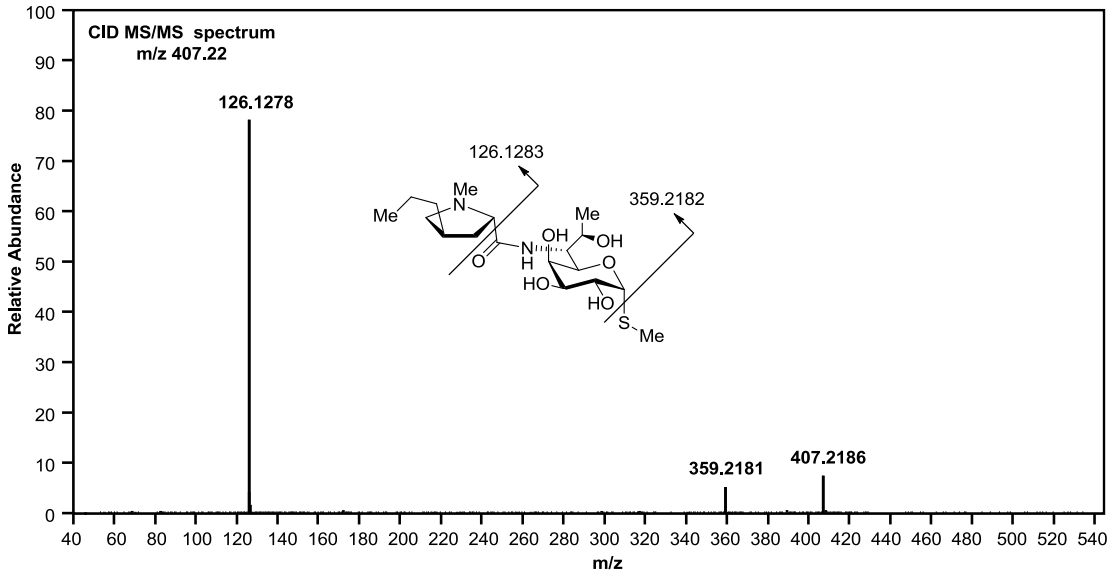
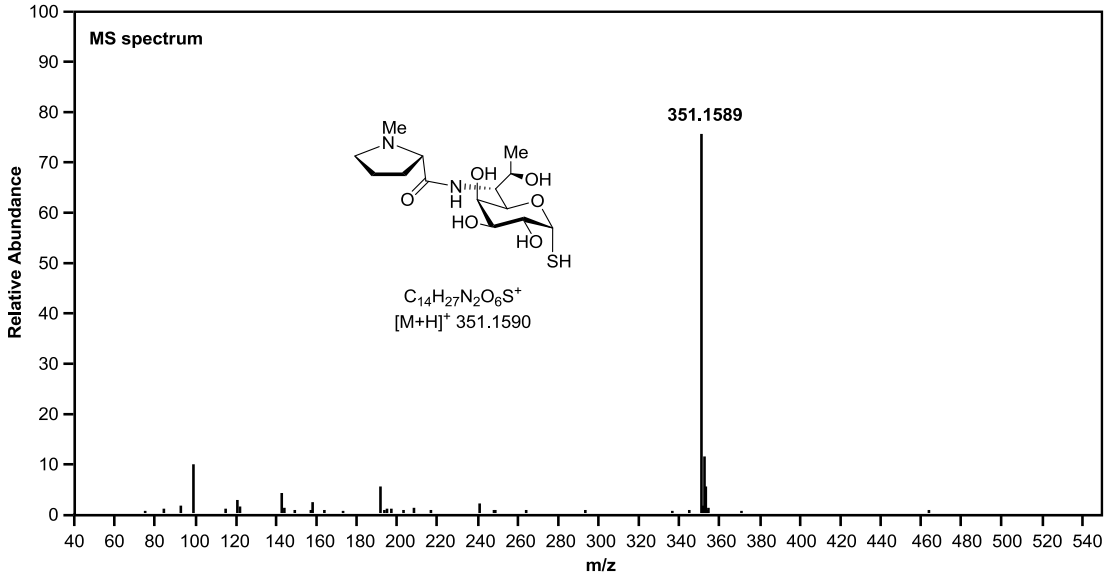


Fig. S24. MS and MS/MS spectra of compound **7**. Selective ion monitoring on $[M+H]^+$ (m/z 351.1590) ion of compound **7**.

A



B

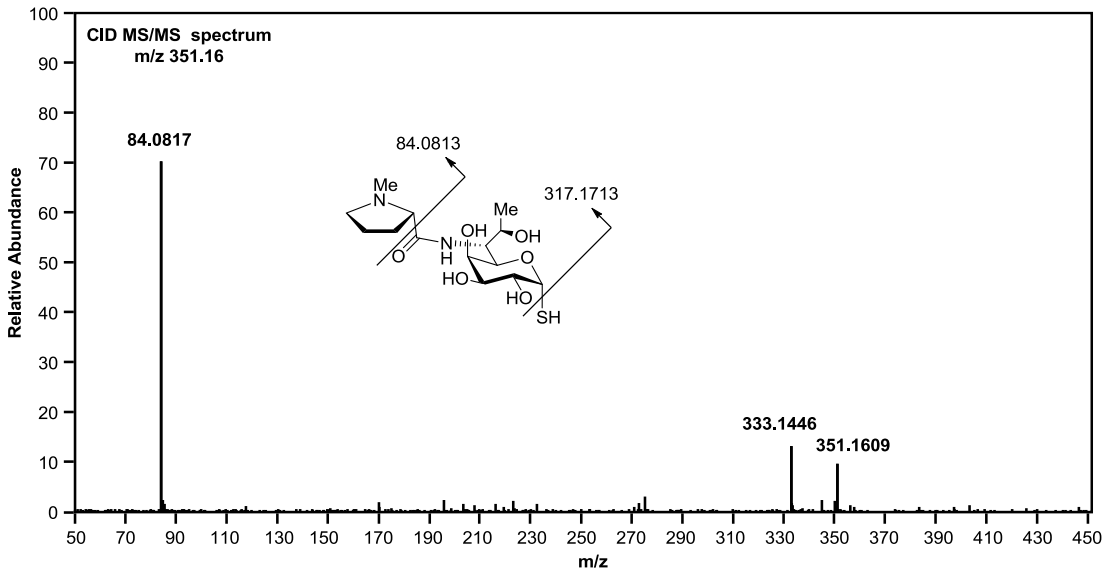
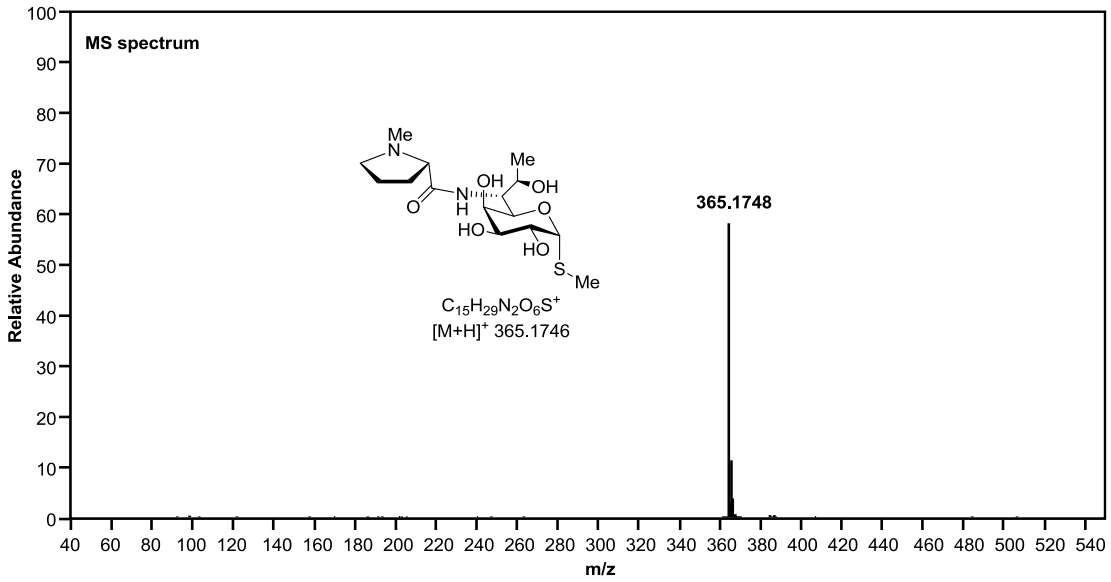


Fig. S25. MS and MS/MS spectra of compound **8**. Selective ion monitoring on $[M+H]^+$ (m/z 365.1746) ion of compound **8**.

A



B

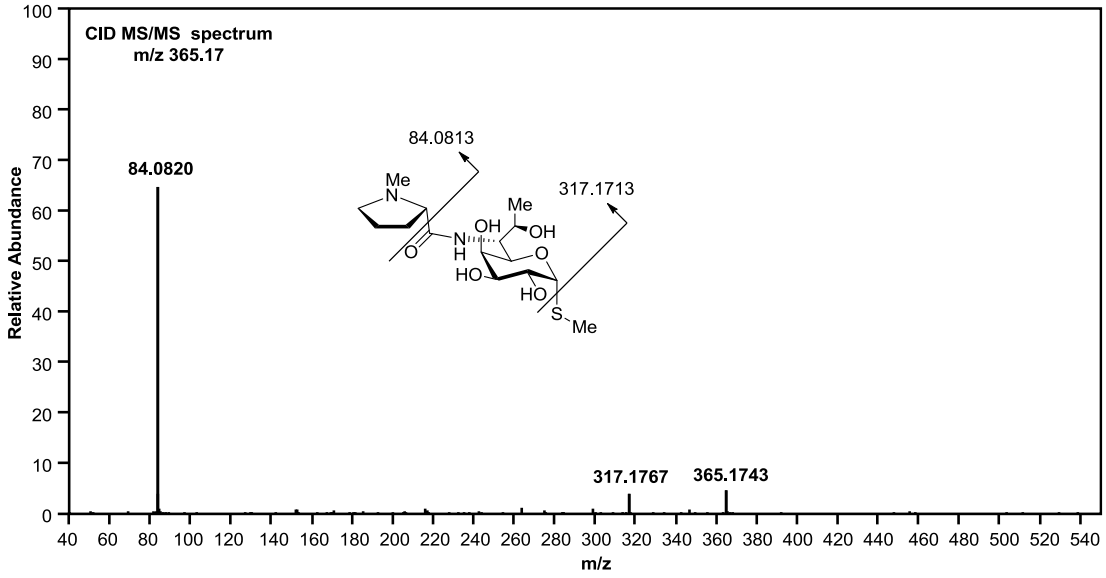
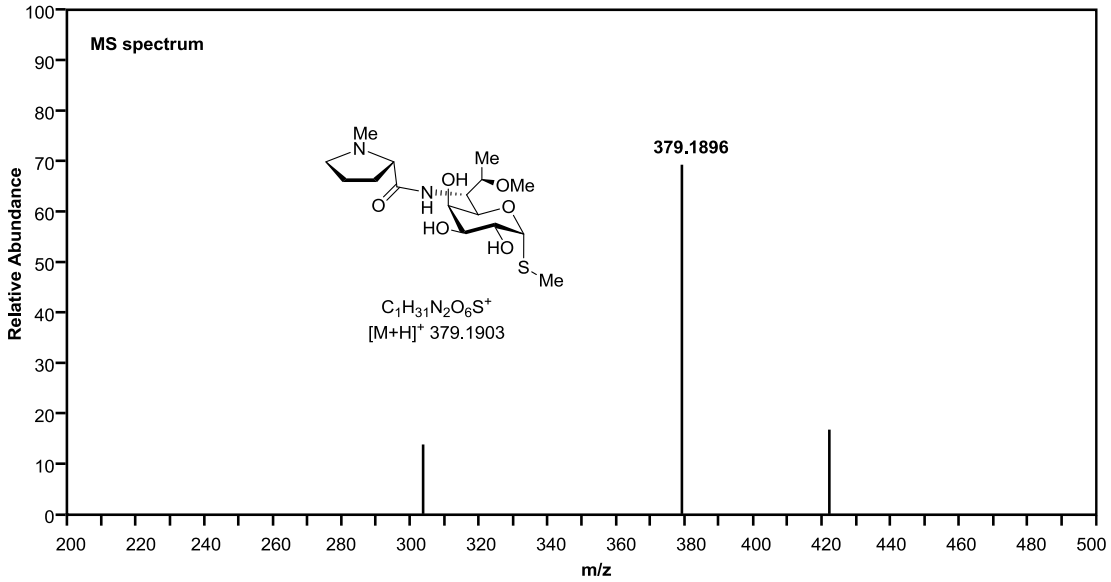


Fig. S26. MS and MS/MS spectra of Bu-2545[13]. Selective ion monitoring on $[M+H]^+$ (m/z 379.1896) ion of Bu-2545.

A



B

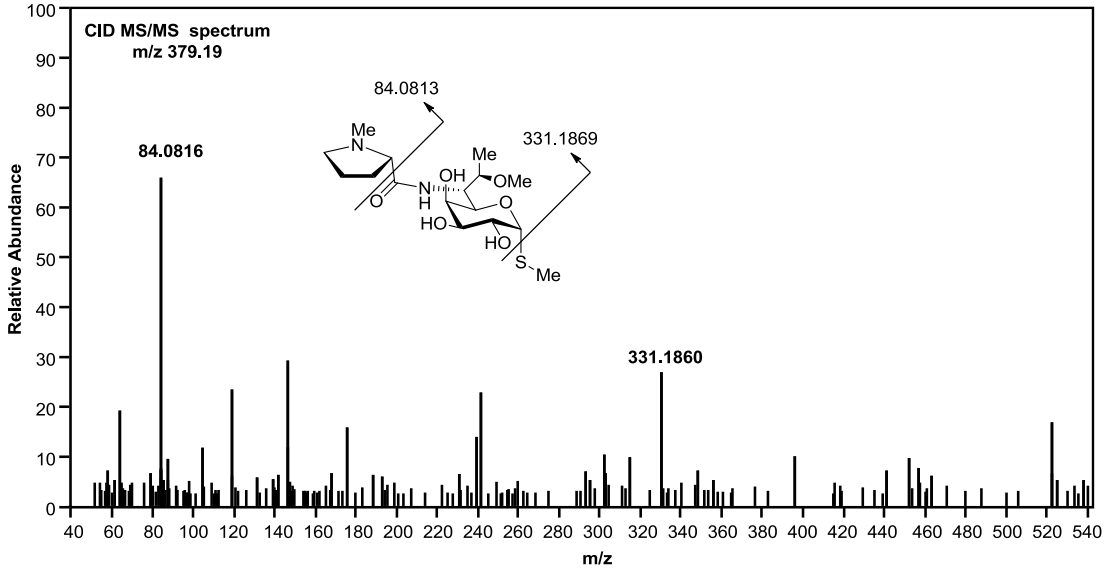
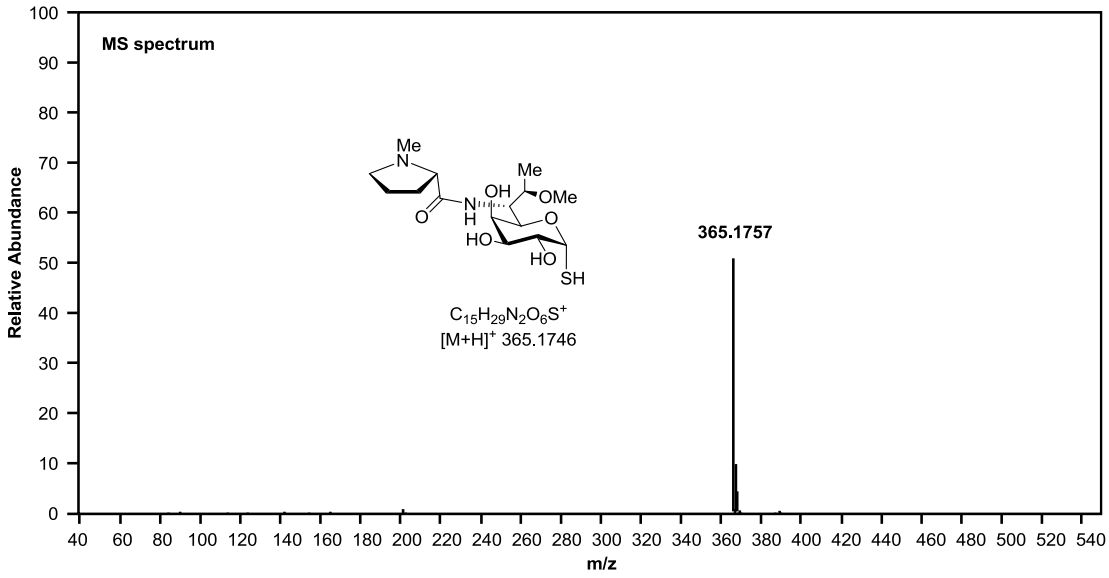


Fig. S27. MS and MS/MS spectra of compound **9**. Selective ion monitoring on $[M+H]^+$ (m/z 365.1746) ion of compound **9**.

A



B

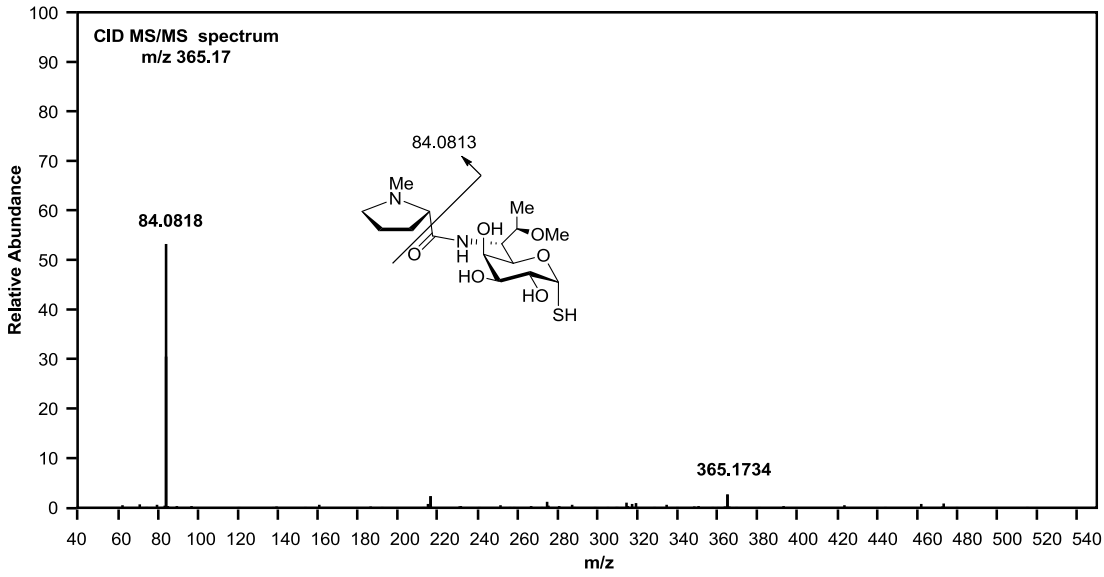
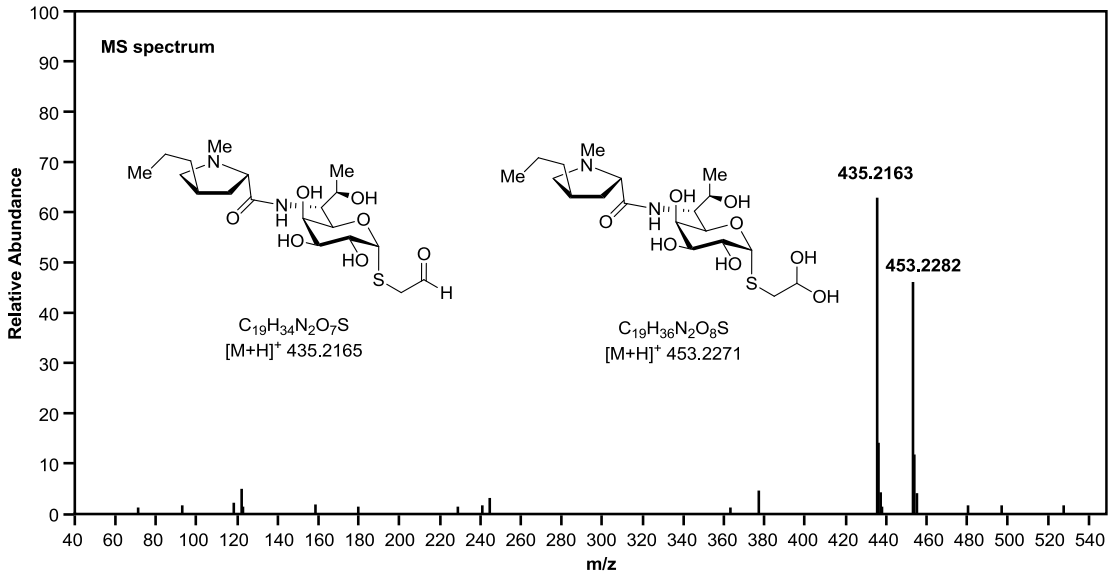


Fig. S28. MS and MS/MS spectra of compound **10**. Selective ion monitoring on $[M+H]^+$ (m/z 435.2165 and 453.2271) ion of compound **10** and its hydrate.

A



B

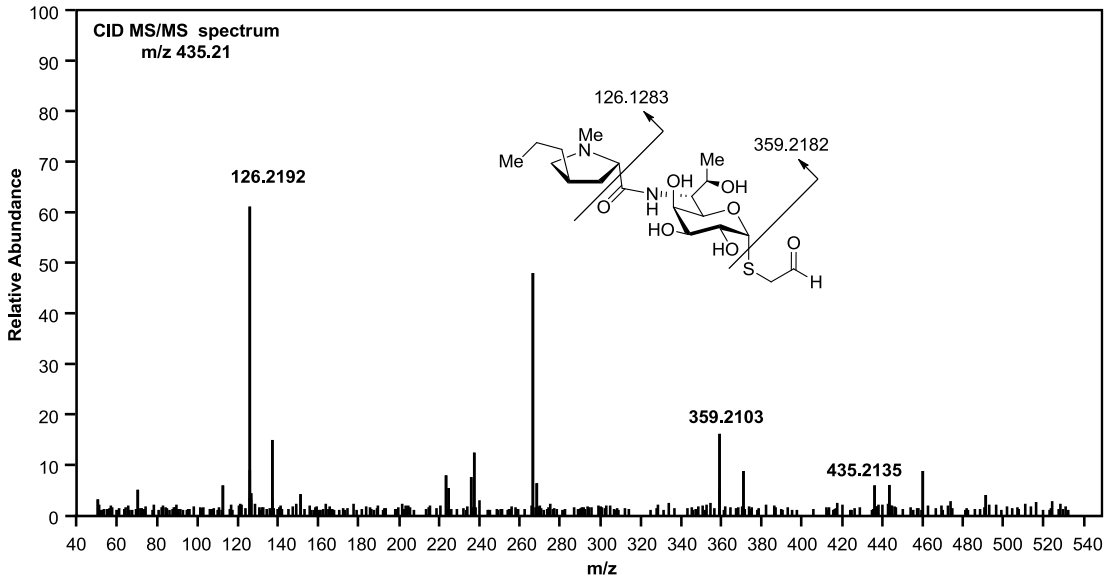
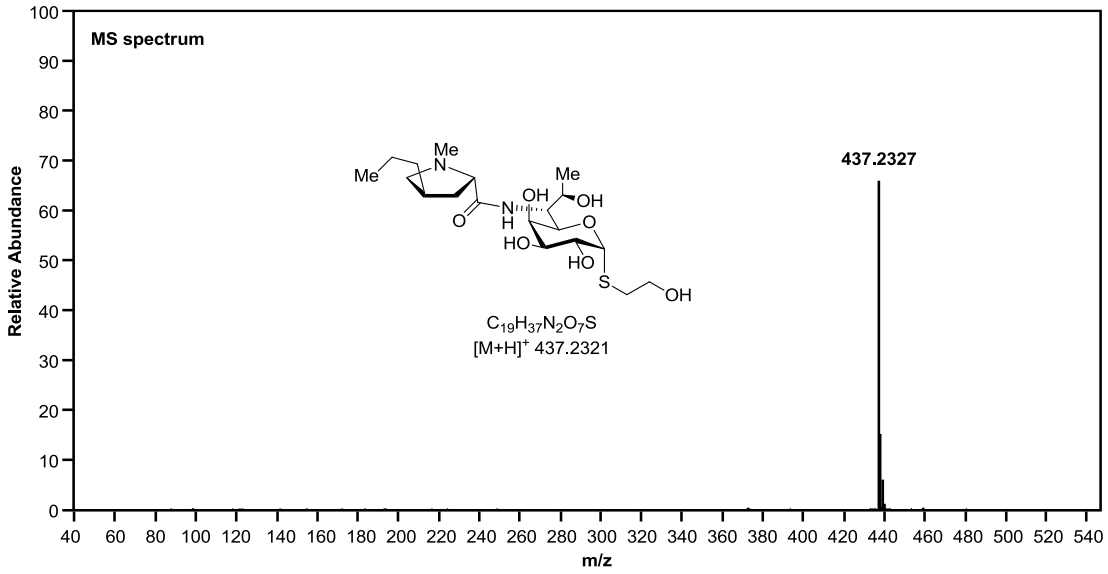
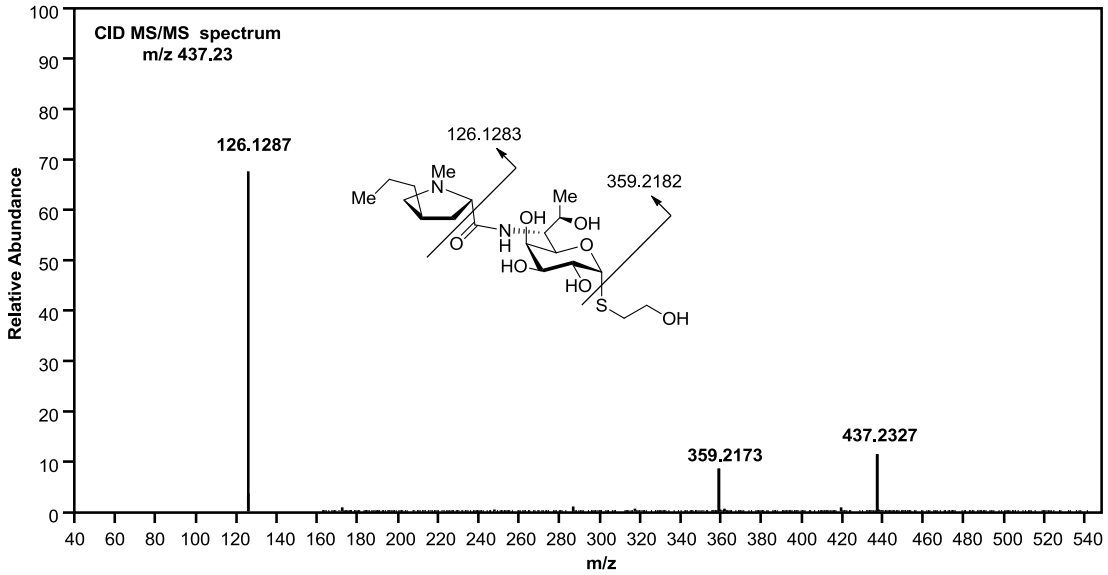


Fig. S29. MS and MS/MS spectra of compound **11**. Selective ion monitoring on $[M+H]^+$ (m/z 437.2321) ion of compound **11**.

A



B



2. Supplementary Tables

Table S1. Bacterial strains and plasmids.

Strain/Plasmid	**Characteristic(s)**	**Sources/ Reference**
S.caelestis		
NRRL2418	Celesticetin-producing strain	NRRL
S. lincolnensis		
NRRL ISP-5355	Wild type strain, lincomycin-producing, identical to the ATCC 25466 strain	NRRL
LL1101	The *ccbF* in-frame deletion mutant of *S.caelestis* NRRL 2418	This study
LL1102	The *ccb5* in-frame deletion mutant of *S.caelestis* NRRL 2418	This study
LL1103	The *ccb4* in-frame deletion mutant of *S.caelestis* NRRL 2418	This study
LL1104	The *lmbF* in-frame deletion mutant of S. lincolnensis NRRL ISP-5355	This study
E. coli		
DH5α	Host for general cloning	Invitrogen
ET12567/pUZ8002	Donor strain for conjugation between *E. coli* and *Streptomyces*	14
BL21 (DE3)	Host for protein expression	NEB
LL1105	BL21 (ED3) derivative, containing pLL1030 for producing CcbF	This study
LL1106	BL21 (ED3) derivative, containing pLL1030 for producing Ccb5	This study
LL1107	BL21 (ED3) derivative, containing pLL1030 for producing Ccb4	This study

LL1108	BL21 (ED3) derivative, containing pLL1026 for producing LmbF	This study
LL1109	BL21 (ED3) derivative, containing pLL1028 for producing LmbG	This study
Plasmids		
pMD19-T	*E. coli* subcloning vector	Takara
pKC1139	*E.coli-Streptomyces* shuttle vector for gene inactivation and complementary, temperature sensitive replication in *Streptomyces*	1
pET-28a(+)	Protein expression vector used in *E.coli*, encoding N-terminal His-tag, kanamycin resistance	Novagen
pWHM79	pGEM-3zf derivative carrying a 0.45 kb fragment containing the *PermE** promoter	15
pLL1101	pKC1139 derivative containing partial *ccbF* fragment	This study
pLL1102	pKC1139 derivative for *ccbF* in-frame deletion	This study
pLL1103	pKC1139 derivative containing partial *ccb5* fragment	This study
pLL1104	pKC1139 derivative for *ccb5* in-frame deletion	This study
pLL1105	pKC1139 derivative containing partial *ccb4*fragment	This study
pLL1106	pKC1139 derivative for *ccb4* in-frame deletion	This study
pLL1107	pKC1139 derivative containing partial *lmbF* fragment	This study
pLL1108	pKC1139 derivative for *lmbF* in-frame deletion	This study
pLL1109	pMD19-T derivative, containing a 1251 bp PCR product that encodes CcbF	This study
pLL1110	pET-28a(+) derivative, containing a 1251 bp PCR product that encodes CcbF	This study
pLL1111	pMD19-T derivative, containing a 1056 bp PCR product that encodes Ccb5	This study
pLL1112	pET-28a(+) derivative, containing a 1056 bp PCR product that encodes Ccb5	This study
pLL1113	pMD19-T derivative, containing a 1110 bp PCR product that encodes Ccb4	This study
pLL1114	pET-28a(+) derivative, containing a 1110 bp PCR product that encodes Ccb4	This study
pLL1115	pMD19-T derivative, containing a1284 bp PCR product that	This study

	encodes LmbF	
pLL1116	pET-28a(+) derivative, containing a 1284 bp PCR product that encodes LmbF	This study
pLL1117	pMD19-T derivative, containing a 801 bp PCR product that encodes LmbG	This study
pLL1118	pET-28a(+) derivative, containing a 801 bp PCR product that encodes LmbG	This study

Table S2. Primers used in this study.

Primers	**Sequence (restriction sites are underlined)**
ccbF-L-for	5’-TATGAATTCGGTGAAGTGCGGCACGTTGATG-3’ HindIII
ccbF-L-rev	5’-TATTCTAGAGACTTCCGGCAACGCCTGTTTG-3’ XbaI
ccbF-R-for	5’-TATTCTAGAGAACCACCGTTCCATCGCTTGG-3’ XbaI
ccbF-R-rev	5’-TATAAGCTTCTTCCCTCAGCCATTACTGGAC-3’ EcoRI
ccbF-gt-for	5’-ATGTCCGACTTAGCTGCCGTTG -3’
ccbF-gt-rev	5’-GCGGGGCTGCCAGGCGCGTGAG -3’
Ccb5-L-for	5’-TATGAATTCCACGTACATCGAGGGACGCAAG-3’ EcoRI
Ccb5-L-rev	5’-TATTCTAGACCACACCCGCTATCTCATGTC-3’ XbaI
Ccb5-R-for	5’-TATTCTAGACGTTGGAGCTGGACTCCTATC-3’ XbaI
Ccb5-R-rev	5’-TATAAGCTTCGACGTTGTCGAGGAAGTACC-3’ HindIII
Ccb5-gt-for	5’-ATGGCGACCGTCCCCGCCCTG-3’
Ccb5-gt-rev	5’-TGATGAGTCCGCGCGCCCCAACGTC-3’
Ccb4-L-for	5’-TATGAATTCGAACTTCCTGTGGAGGACTGGC-3’ EcoRI
Ccb4-L-rev	5’-TATTCTAGACGTTGTAGCACAGGACGAGGTC-3’ XbaI
Ccb4-R-for	5’-TATTCTAGAGCAAGGTCATCAAGGCGGAAGG-3’ XbaI
Ccb4-R-rev	5’-TATAAGCTTCTCGCGCAGAACCTTCAGCAGAC-3’ HindIII
Ccb4-gt-for	5’-ATGAAGACGCCCGGTACATCGCAC-3’
Ccb4-gt-rev	5’-TCAGCACGGAGTGGCCTCAAGCAC-3’
lmbF-L-for	5’-TATGAATTCGGAGAGTCGAACGGAGGACGAC-3’ EcoRI
lmbF-L-rev	5’-TATTCTAGACAGGAACTCCTGGTGTCGGTGC-3’ XbaI
lmbF-R-for	5’-TATTCTAGACAGGCCGTACTGGGAGAACAGC-3’ XbaI
lmbF-R-rev	5’-TATAAGCTTCCGCTACCACGTGTCCGTCTAC-3’ HindIII
lmbF-gt-for	5’-CGAACCGTTCGGGGGTGAAGTC-3’
lmbF-gt-rev	5’-GAACGGTACGTGCAACTCCTCG-3’
CcbF-for	5’-TATCATATGTCCGACTTAGCTGCCGTTG-3’ NdeI

CcbF-rev	5’-TATAAGCTTTCAGCGGGGCTGCCAGGCGCGTGAG-3’ HindIII
Ccb5-for	5’-TATCATATGGCGACCGTCCCCGCCCTG-3’ NdeI
Ccb5-rev	5’-TATAAGCTTTCATGAGTCCGCGCGCCCCAAC-3’ HindIII
Ccb4-for	5’-TATCATATGAAGACGCCCGGTACATCGCAC-3’ NdeI
Ccb4-rev	5’-TATAAGCTTTCAGCACGGAGTGGCCTCAAGCAC-3’ HindIII
lmbF-for	5’-TATGAATTCCATATGACCGCCACGGCGAGCGGC-3’ NdeI
lmbF-rev	5’-TATAAGCTTACTCGAGCCGGTACCGCCACTCGGCCGC-3’ HindIII
lmbG-for	5’-TATGAATTCCATATGCGGGACTACCGTCTCTGG-3’ NdeI
lmbG-rev	5’-TATAAGCTTACTCGAGTCGGGTCGCCCCCGCCGTGCC-3’ HindIII

Table S3. NMR spectroscopic data for celesticetin[a].

position	δ_H(mult., *J* in Hz)	δ_C	mutl.
Moiety A			
1	5.58 (1H, d, 5.8)	90.4	CH
2	4.14 (1H, m)	70.3	CH
3	3.63 (1H, m)	73.0	CH
4	3.90 (1H, d, 2.8)	70.9	CH
5	4.11 (1H, m)	72.1	CH
6	4.65 (1H, d, 3.3, 10.3)	51.8	CH
7	3.62 (1H, m)	78.5	CH
7-OMe	3.22 (3H, s)	58.5	CH_3
8	1.07 (3H, d, 6.4)	15.8	CH_3
1'-NMe	2.95 (3H, s)	43.2	CH_3
2'	4.20 (1H, dd, 7.3, 8.7)	71.5	CH
2'-C=O		171.2	C
3'a	2.60 (1H, m)	32.2	CH_2
3'b	2.09 (1H, m)		
4'a	2.23 (1H, m)	25.2	CH_2
4'b	2.08 (1H, m)		
5'a	3.79 (1H, m)	59.1	CH_2
5'b	3.25 (1H, m)		
1"	3.10 (2H, t, 5.6)	32.3	CH_2
2"	4.60 (2H, m)	67.9	CH_2
Moiety B			
1		115.3	C
1-C=O		172.6	C
2		162.2	C
3	7.03 (1H, d)	119.9	CH
4	7.60 (1H, t, 8.2)	139.0	CH
5	7.05 (1H, t, 8.6)	122.9	CH
6	7.96 (1H, d, 8.2)	133.2	CH

[a] In D_2O, 500MHz for ^1H and 125MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^1H -^1H COSY, HSQC, HMBC and NOESY correlation.

Table S4. NMR spectroscopic data for desalicetin[a].

position	δ_H(mult., *J* in Hz)	δ_C	mutl.
1	5.48 (1H, d, 5.7)	89.6	CH
2	4.13 (1H, dd, 5.7, 10.5)	70.2	CH
3	3.64 (1H, m)	73.0	CH
4	3.91 (1H, d, 2.9)	71.0	CH
5	4.18 (1H, m)	71.5	CH
6	4.70 (1H, dd, 3.1, 10.2)	51.9	CH
7	3.77 (1H, m)	78.7	CH
7-OMe	3.36 (3H, s)	58.7	CH_3
8	1.16 (3H, d, 6.4)	15.9	CH_3
1'-NMe	2.96 (3H, s)	43.2	CH_3
2'	4.21 (1H, m)	71.6	CH
2'-C=O		171.3	C
3'a	2.62 (1H, m)	32.2	CH_2
3'b	2.11 (1H, m)		
4'a	2.23 (1H, m)	25.2	CH_2
4'b	2.08 (1H, m)		
5'a	3.79 (1H, m)	59.1	CH_2
5'b	3.25 (1H, m)		
1''	2.80 (2H, m)	35.3	CH_2
2''	3.80 (2H, m)	63.3	CH_2

[a] In D_2O, 500MHz for ^1H and 125MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^1H -^1H COSY, HSQC, HMBC and NOESY correlation.

Table S5. NMR spectroscopic data for compound **1**[a].

position	δ_H(mult., *J* in Hz)	δ_C	mutl.
1	5.55 (1H, d, 5.8)	69.8	CH
2	4.17 (1H, m)	70.2	CH
3	3.70 (1H, dd, 3.3, 10.6)	72.7	CH
4	4.05 (1H, m)	71.2	CH
5	4.32 (1H, d, 9.0)	72.1	CH
6	4.46 (1H, dd, 5.4, 8.7)	56.8	CH
7	4.18 (1H, m)	69.2	CH
8	1.23 (1H, d, 6.5)	19.5	CH_3
1'-NMe	2.99 (3H, s)	43.1	CH_3
2'	4.23 (1H, dd, 7.6, 8.8)	71.7	CH
2'-C=O		171.6	C
3'a	2.65 (1H, m)	32.1	CH_2
3'b	2.16 (1H, m)		
4'a	2.27 (1H, m)	25.3	CH_2
4'b	2.12 (1H, m)		
5'a	3.82 (1H, m)	59.1	CH_2
5'b	3.26 (1H, m)		
1''	175.2		C
2''	4.05 (1H, m)	56.8	CH
3''a	3.31 (1H, m)	33.5	CH_2
3''b	3.20 (1H, dd, 4.4, 14.9)		

[a] In D_2O, 500MHz for ^1H and 125MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^1H -^1H COSY, HSQC, HMBC and NOESY correlation.

Table S6. NMR spectroscopic data for compound **5**[a].

position	δ_H(mult., *J* in Hz)	δc	mutl.
1	5.46 (1H, d, 5.8)	89.7	CH
2	4.09 (1H, m)	70.0	CH
3	3.61 (1H, dd, 3.1, 10.6)	72.5	CH
4	3.96 (1H, m)	71.0	CH
5	4.23 (1H, m)	72.0	CH
6	4.38 (1H, m)	56.6	CH
7	4.10 (1H, m)	69.0	CH
8	1.14 (3H, d, 6.4)	19.4	CH_3
1'-N-CH_3	2.90 (3H, s)	42.9	CH_3
2'	4.25 (1H, m)	70.9	CH
2'-C=O		171.6	C
3'	2.25 (2H, m)	37.8	CH_2
4'	2.38 (1H, m)	38.9	CH
5'a	3.81 (1H, dd, 6.8, 11.1)	63.7	CH_2
5'b	2.87 (1H, d, 11.3)		
6'	1.41 (2H, m)	36.4	CH_2
7'	1.28 (2H, m)	22.9	CH_2
8'	0.85 (3H, t, 7.2)	15.7	CH_3
1"		175.0	C
2"	3.98 (1H, m)	56.6	CH
3"a	3.22 (1H, dd, 5.4, 14.9);	33.3	CH_2
3"b	3.10 (1H, dd, 4.2, 14.8)		

[a] In D_2O, 400MHz for ^1H and 100MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^1H -^1H COSY, HSQC, HMBC and NOESY correlation.

6. Supplementary References

1. Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F. Hopwood, D. A. *Practical Streptomyces Genetics*; John Innes Foundation: Norwich, 2000.
2. Green, M. R; Sambrook, J. *Molecular Cloning: A Laboratory Manual*, 4th ed.; Cold Spring Harbor Laboratory Press: New York, 2012.
3. Zhao, Q.; Wang, M.; Xu, D.; Zhang, Q.; Liu, W. *Nature* **2015**, *518*, 115.
4. Zhang, Q.; Li, Y.; Chen, D.; Yu, Y.; Duan, L.; Shen, B.; Liu, W. *Nat. Chem. Biol.* **2011**, *7*, 154.
5. Du, Y.-L.; Singh, R.; Alkhalaf, L. M.; Kuatsjah, E.; He, H.-Y.; Eltis, L. D.; Ryan, K. S. *Nat. Chem. Biol.* **2016**, *12*, 194.
6. Kaminaga, Y.; Schnepp, J.; Peel, G.; Kish, C. M.; Ben-Nissan, G.; Weiss, D.; Orlova, I.; Lavie, O.; Rhodes, D.; Wood, K.; Porterfield, D. M.; Cooper, A. J.; Schloss, J. V.; Pichersky, E.; Vainstein, A.; Dudareva, N. *J. Biol. Chem.* **2006**, *281*, 23357.
7. Hoeksema, H. *J. Am. Chem. Soc*. **1968**, *90*, 755-757.
8. Argoudelis, A. D.; Brodasky, T. F. *J. Antibiot.* **1972**, *25*, 194.
9. Moloney, G.; Craik, D.; Iskander, M.; Munro, S. *Aust. J. Chem.* **1990**, *43*, 535.
10. Verdier, L.; Bertho, G.; Gharbi-Benarous, J.; Girault, J.-P. *Biorg. Med. Chem.* **2000**, *8*, 1225.
11. Čermák, L.; Novotná, J.; Ságová-Marečková, M.; Kopecký, J.; Najmanová, L.; Janata, J. *Folia Microbiol.* **2007**, *52*, 457.
12. Kobersk á, M.; Kopeck ý, J.; Olsovsk á, J.; Jel ínkov á, M.; Ulanova, D.; Man, P.; Flieger, M.; Janata, J. *Folia Microbiol.* **2008**, *53*, 395.
13. Hanada, M.; Tsunakawa, M.; Tomita, K.; Tsukiura, H.; Kawaguchi, H. *J. Antibiot* **1980**, *33*, 751
14. MacNeil, D. J.; Gewain, K. M.; Ruby, C. L.; Dezeny, G.; Gibbons, P. H.; MacNeil, T. *Gene* **1992**, *111*, 61.
15. Shen, B.; Hutchinson, C. R. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 6600.