Discovery of the Leinamycin Family of Natural Products by Mining Actinobacterial Genomes

Guohui Pan[a,1], Zhengren Xu[a,1], Zhikai Guo[a], Hindra[a], Ming Ma[a], Dong Yang[a,b], Hao Zhou[a] Yannick Gansemans[c], Xiangcheng Zhu[d,e], Yong Huang[d], Li-Xing Zhao[f], Yi Jiang[f], Jinhua Cheng[g,h], Filip Van Nieuwerburgh[c], Joo-Won Sun[g,h], Yanwen Duan[d,e], and Ben Shen[a,b,i,*]

[a]Department of Chemistry, The Scripps Research Institute, Jupiter, FL 33458, USA; [b]Natural Products Library Initiatives at The Scripps Research Institute, The Scripps Research Institute, Jupiter, FL 33458, USA; [c]Laboratory of Pharmaceutical Biotechnology, Ghent University, Ghent, Belgium; [d]Xiangya International Academy of Translational Medicine, Central South University, Changsha, Hunan 410013, China; [e]Hunan Engineering Research Center of Combinatorial Biosynthesis and Natural Product Drug Discovery, Changsha, Hunan 410013, China; [f]Yunnan Institute of Microbiology, Yunnan University, Kunming, Yunnan 650091, China; [g]Division of Bioscience and Bioinformatics, College of Natural Science, Myongji University, Yongin, Gyeonggi-Do 17058, Korea; [h]Center for Neutraceutical and Pharmaceutical Materials, Myongji University, Yongin, Gyeonggi-Do 17058, Korea; [i]Department of Molecular Medicine, The Scripps Research Institute, Jupiter, FL 33458, USA;

*Correspondence to E-mail: shenb@scripps.edu, Tel: (561) 228-2456; Fax: (561) 228-2472

Short title: Discovery of the Leinamycin Family of Natural Products

[1]These authors contributed equally

Supporting Online Materials

Table of Contents

Materials and Methods

Bacterial strains, plasmids, chemicals

Strains, plasmids, and PCR primers used in this study are listed in **Tables S2** and **S3**, respectively. PCR primers were purchased from Sigma-Aldrich. Q5 high-fidelity DNA polymerase, restriction endonucleases, and T4 DNA ligase were purchased from New England Biolabs (NEB) and reactions were performed according to the manufacturer's protocols. DNA gel extraction and plasmid preparation kits were purchased from Omega Bio-Tek. DNA sequencing was conducted by Eton Bioscience. Other common chemicals, bio-chemical, and media components were purchased from standard commercial sources.

General procedures

E. coli strains harboring plasmids or cosmids were grown in lysogeny broth (LB) with appropriate antibiotics (1). *E. coli* ET12567/pUZ8002 was used for intergeneric conjugation with *Streptomyces* sp. CB01883 and *Streptomyces* sp. CB02120-2, and the conjugations were carried out following typical procedures (2). *Streptomyces* sp. NRRL F-5630 and *S. aureofaciens* NRRL B-2183 were obtained from NRRL (Agricultural Research Service Culture Collection). *Streptomyces* sp. NRRL F-5630 was cultivated on solid ISP2 medium for sporulation, while all the other *Streptomyces* strains used in this study were cultivated on solid ISP4 medium for sporulation. *Streptomyces* strains were cultivated in liquid tryptic soy broth (TSB) at 28 °C and 250 rpm to prepare the mycelium for genomic DNA (gDNA) isolation. Isolation of gDNA from *Streptomyces* strains were performed using the salting out protocol (2). Cosmid libraries of *Streptomyces* sp. CB01883 and *Streptomyces* sp. CB02120-2 were constructed based on SuperCos 1 cosmid following standard protocols (Agilent Technologies). Screening of the cosmid libraries were performed by PCR with primers shown in **Table S3** using OneTaq® 2X Master Mix with GC buffer (NEB). For Southern analysis, digoxigenin labeled DNA probes preparation, hybridization, and detection were performed according to the manufacturer's protocols (Roche Diagnostics Corp.).

ClustalW was used for multiple sequence alignment (3). All the phylogenetic trees (Figs. 2A and 3A in the main text, Figs. S3A, S3B, S5, and S6 in the SI) in this study were generated by MEGA6 using the maximum likelihood method with a bootstrap test of 500 replicates (4). Numbers next to the branches represent the percentage of replicate trees in which the shown topology was reached. Annotation of the new *lnm*-type gene clusters were done by using antiSMASH 3.0 (5) and 2ndFind (http://biosyn.nih.go.jp/2ndFind/). Construction of GNN of the LNM-type pathways was carried out as described previously (6).

Optical rotation values were measured using an Autopol IV automatic polarimeter. UV spectra were recorded with a NanoDrop 2000C spectrophotometer (Thermo Scientific). IR spectra were collected with a Spectrum One FT-IR spectrometer (PerkinElmer). All ^{1}H, ^{13}C, and 2D-NMR (HSQC, ^{1}H-^{1}H COSY, HMBC, ROESY) spectra were collected at room temperature (25 ^{o}C) with a Bruker Avance III Ultrashield 700 at 700 MHz for ^{1}H and 175 MHz for ^{13}C nuclei. ^{1}H-NMR spectra at higher temperatures were measured using Bruker Avance 400 MHz ULTRAShield instrument. HR-ESI-MS data were acquired on Agilent 6230 TOF LC/MS instrument.

Genome survey of 5,000 strains from the actinomycetes collection at TSRI for producers of the leinamycin (LNM) family of natural products

Real-time PCR was performed using an Applied Biosystems 7900HT Fast Real-Time PCR system. To identify strains harboring *lnm*-type gene clusters, the conserved DUF-SH didomain was chosen as the screening target. Degenerate primers DUF-F (derived from the conserved sequences within DUF domains) and SH-sub-R (derived from the conserved sequences

predicted to be involved in substrate and potassium binding within SH domains), aiming to amplify a 500 bp DNA fragment from DUF-SH didomain, were used in the screening with final concentrations of 1 μM. Preparation of gDNA arrays and application of the real-time PCR screening followed a published protocol (7). Specifically, gDNA samples were arrayed in 384-well plates and subjected to real-time PCR. Each plate contained a positive control, with the gDNA of *S. atroolivaceus* S-140 (LNM producer) as the template which will give PCR products with a theoretical T_m of 93.1 °C, and a negative control without template. The PCR conditions consisted of a background check at 50 °C for 2 min; initial denaturation at 95 °C for 7 min; 35 cycles of denaturation at 95 °C for 30 s, primer annealing at 52 °C for 15 s, extension at 68 °C for 45 s; and melting steps with a ramp rate of 2%. Samples with T_m between 91.6 to 93.5 °C were considered as potential hits (**Fig. S2**).

To verify the identity of hit strains obtained from the real-time PCR screening, another round of PCR was performed with degenerate primers Seq-DUF-F and Seq-SH-R/Seq-SH-2120-R, aiming to amplify a 1.2 kb DNA fragment from DUF-SH didomain, using the gDNAs of the hit strains as templates. The resultant PCR products were subjected to DNA sequencing and the sequence results have been deposited into GenBank with their accession numbers summarized in **Table S1**. For taxonomic analysis, three housekeeping genes (16S rRNA, *recA*, and *trpB*) were amplified, sequenced, and analyzed. All the DNA sequences have been deposited into GenBank (**Table S1**).

DNA Sequencing and Genome Assembly

Genome sequencing of the representative hit strains was performed using an Illumina MiSeq sequencer (2 x 300 paired end sequencing) at the Next Generation Sequencing and Microarray Core Facility, The Scripps Research Institute. Read quality filtering was performed using an in-house developed tool. Adapter trimming and de novo assembly were done with CLC Genomics Workbench version 7.5.1 (CLC Bio.) using default settings. The resulting contigs were further extended and joined into a final scaffold by SSPACE version 2.0 using all quality filtered reads (8). The remaining gaps within the final scaffold were partially or completely filled using the quality filtered reads by GapFiller version 1.10 (9). Genome sequences (under BioProject PRJNA393980) have been deposited into GenBank with accession numbers summarized in **Table S1**.

Inactivation and complementation of *gnmB* (encoding hybrid NRPS/PKS) and *gnmO* (encoding N-acetylglucosamine deacetylase) in Strain *Streptomyces* sp. CB01883

To construct a plasmid for inactivation of *gnmB*, a 1555 bp DNA fragment upstream of *gnmB* was amplified with primers 1883orf155-I5 and 1883orf155-I3, and a 1483 bp DNA fragment downstream of *gnmB* was amplified with primers 1883orf155-II5 and 1883orf155-II3, using cosmid pBS21001 (1C4) of strain CB01883 as template. The two DNA fragments were trimmed with appropriate enzymes and cloned into the *Hin*dIII and *Eco*RI sites of pOJ260 to obtain pBS21004 (10). Then pBS21004 was transformed into *E. coli* ET12567(pUZ8002) and introduced into strain CB01883 by intergeneric conjugation. After several rounds of passaging the exconjugants on solid ISP4 medium, apramycin sensitive double-crossover mutants were obtained. The genotype of the in-frame deletion mutant strain SB21001 (i.e. Δ*gnmB*) was verified by Southern analysis (**Fig. S8**).

For complementation of SB21001, a 7687 bp DNA fragment containing *gnmB* was amplified with primers orf155-kasO-5 and orf155-kasO-3 using cosmid pBS21001 as template, and then cloned into the *Afl*II and *Spe*I sites of pBS21003 (pSET-KasO*: constructed by digestion of PCR products obtained with primers PSET-kasO-F and PSET-kasO-R using pSETTurdR as template with *Nsi*I followed by self-ligation) (11) to obtain pBS21005, in which *gnmB* was under the

control of the strong promoter *kasO** (12). Then pBS21005 was introduced into SB21001 by intergeneric conjugation to afford the complementation strain SB21002 (i.e. Δ*gnmB*/*gnmB*).
Similar strategy was used to generate the in-frame deletion mutant strain SB21003 (i.e. Δ*gnmO*). An upstream 3085 bp DNA fragment was amplified with primers 1883orf132-I5 and 1883orf132-I3, and a downstream 3070 bp DNA fragment was amplified with primers 1883orf132-II5 and 1883orf132-II3, using cosmid pBS21001 as template. The two DNA fragments were digested with appropriate enzymes and cloned into the *Hin*dIII and *Eco*RI sites of pOJ260. The resultant plasmid pBS21006 was introduced into strain CB01883 by intergeneric conjugation. After several rounds of screening, apramycin sensitive double-crossover mutants were obtained. The genotype of the in-frame deletion mutant strain SB21003 was verified by Southern analysis (**Fig. S30**).

For complementation of SB21003, a 1007 bp DNA fragment containing *gnmO* obtained by PCR with primers orf132-com-F and orf132-com-R using cosmid pBS21001 as template, was cloned into the *Afl*II and *Spe*I sites of pBS21003 to generate pBS21007, which was then introduced into SB21003 by intergeneric conjugation to afford the complementation strain SB21004 (i.e. Δ*gnmO*/*gnmO*).

Inactivation of *wsmW* (encoding hybrid NRPS/PKS), *wsmZ3* (encoding 3-oxyacyl-ACP synthase, KASIII) and *wsmZ4* (encoding 3-hydroxylacyl-CoA dehydrogenase, HBDH) in *Streptomyces* sp. CB02120-2
To construct a plasmid for inactivation of *wsmW*, an upstream 3038 bp DNA fragment was amplified with primers 2120orf25-I5 and 2120orf25-I3, and a downstream 3033 bp DNA fragment was amplified with primers 2120orf25-II5 and 2120orf25-II3 using cosmid pBS22001 (8E4) of strain CB02120-2 as template. The two DNA fragments were then cloned into the *Hin*dIII and *Eco*RI sites of pOJ260. The resultant plasmid pBS22004 was introduced into strain CB02120-2 by intergeneric conjugation. After several rounds of passaging the exconjugants on solid ISP4 medium, apramycin sensitive double-crossover mutants were obtained. The genotype of the in-frame deletion mutant strain SB22001 (i.e. Δ*wsmW*) was verified by Southern analysis (**Fig. S18**).

Similar strategy was used to generate the in-frame deletion mutant strains SB22002 (i.e. Δ*wsmZ3*) and SB22004 (i.e. Δ*wsmZ4*). For knock out of *wsmZ3*, an upstream 3165 bp DNA fragment was amplified with primers 2120orfZ3-I5 and 2120orfZ3-I3, and a downstream 3182 bp DNA fragment was amplified with primers 2120orfZ3-II5 and 2120orfZ3-II3, using cosmid pBS22001 as template. The two DNA fragments were then cloned into the *Hin*dIII and *EcoR*I sites of pOJ260 to obtain pBS22005. For knock out of *wsmZ4*, an upstream 3155 bp DNA fragment was amplified with primers 2120orfZ4-I5 and 2120orfZ4-I3, and a downstream 3170 bp DNA fragment was amplified with primers 2120orfZ4-II5 and 2120orfZ4-II3, using cosmid pBS22001 as template. And then the two DNA fragments were cloned into the *Hin*dIII and *Eco*RI sites of pOJ260 to obtain pBS22007. The plasmids pBS22005 and pBS22007 were introduced into strain CB02120-2, respectively, through intergeneric conjugation. After several rounds of screening, apramycin sensitive double-crossover mutant strains SB22002 and SB22004 were obtained, and their genotypes were verified by Southern analysis (**Fig. S26**).

For complementation of SB22002, a 1283 bp DNA fragment containing *wsmZ3* was amplified with primers orfZ3-kasO-5 and orfZ3-kasO-3 using cosmid pBS22001 as template, and then cloned into the *Afl*II and *Spe*I sites of pBS21003 to obtain pBS22006. For complementation of SB22004, a 1264 bp DNA fragment containing *wsmZ4* was amplified with primers orfZ4-kasO-5 and orfZ4-kasO-3 using cosmid pBS22001 as template, and then cloned into the *Afl*II and *Spe*I sites of pBS21003 to generate pBS22008. Then pBS22006 and pBS22008 were introduced into

SB22002 and SB22004, respectively, to afford the complementation strains SB22003 (i.e. Δ*wsmZ3*/*wsmZ3*) and SB22005 (i.e. Δ*wsmZ4*/*wsmZ4*), by intergeneric conjugation.

Fermentation, HPLC analysis and isolation of guangnanmycin (GNM) A, GNM B, GNM B1, GNM B2 and GNM B3

Fresh spores of strain CB01883, its derivative mutant and complementation strains were individually inoculated into 250-mL baffled flasks containing 50 mL TSB seed medium and cultured for 36 h at 28 °C and 250 rpm. For small-scale fermentations, seed culture was inoculated (10%, v/v) into 250-mL baffled flasks each containing 50 mL production medium LM1 (soluble starch 3%, soy flour 1%, $CaCO_3$ 0.5%, KH_2PO_4 0.05%, $MgSO_4$ 0.025%, $ZnSO_4{\cdot}7H_2O$ 0.004 %, L-methionine 0.01%, vitamin B_{12} 0.0001%, pH 7.2), a medium modified from the one for the production of LNM by *S. atroolivaceus* S-140 (13). Amberlite XAD-16 resin (Sigma) was added to each flask (4%, w/v) at 24 h after inoculation, and the fermentation was allowed for another 72 h. The resin was harvested from the fermentation broth, washed with water and allowed to air dry. The dry resin was extracted with 6 mL methanol, and the methanol extract was directly used for HPLC and LC-MS analysis. HPLC analysis was carried out on an Agilent 1260 Infinity LC System with an Agilent SB-C18 column (5 μm, 250 × 4.6 mm) at a flow rate of 1 mL/min. The column was equilibrated with 100% solvent A (20% acetonitrile with acetic acid, pH 3.6), and developed with a 56 min analytical program consisting of a 40 min linear gradient from 0% to 68% of solvent B (80% acetonitrile with acetic acid, pH 3.6); 68–100% of solvent B from 40 to 42 min; 100% of solvent B from 42 to 50 min; 100–0% of solvent B from 50 to 51 min; and 0% of solvent B for further 5 min. HPLC-MS was performed on an Agilent 1260 Infinity LC coupled to a 6230 TOF (HRESI) equipped with an Agilent Po-roshell 120 EC-C18 column (2.7 μm, 250 × 4.6 mm). Chromatography for LC-MS was conducted using an 18 min solvent gradient from 5–100% acetonitrile containing 0.1% formic acid in H_2O containing 0.1% formic acid at a flow rate of 0.4 mL/min.

For isolation of GNM A and GNM B1, a large scale fermentation (12 L) of *Streptomyces* sp. CB01883 was performed with medium LM1. The resins were harvest, and extracted with methanol (2 L × 3). The solvent was evaporated to give the crude extract, which was suspended with water (500 mL), and extracted with ethyl acetate (1 L × 3). The combined organic phases were evaporated to dryness, and the residue was fractionated with Agilent 1260 Infinity LC System using an Eclipse XDB-C18 column (7 μm, 212 × 250 mm). The mobile phase consisted of solvent A (20% acetonitrile with acetic acid, pH 3.6) and B (80% acetonitrile with acetic acid, pH 3.6). The crude extract was eluted with a linear gradient of 0% B to 100% B for 20 min, followed with 100% B for 5 min at a flow rate of 17 mL/min. The fractions containing compounds of interest were combined and concentrated. The residue was subjected to a Sephadex LH-20 column, and eluted with methanol. The fractions containing compounds of interest were combined and finally purified by Varian LC system with a Zorbax SB-C13 column (5 μm, 9.4 × 250 mm) to give GNM A (5.9 mg) as a white powder, and GNM B1 (1.2 mg) as a yellowish oil.

For isolation of GNM B, GNM B2 and GNM B3, a large scale fermentation (20 L) of the mutant strain SB21003 was performed with medium LM3 (soluble starch 0.5%, malt extract 1.5%, glucose 1%, tryptone 0.4%, K_2HPO_4 0.05%, $MgSO_4{\cdot}7H_2O$ 0.02%, methionine 0.01%, vitamin B12 0.0001%, pH 7.2). The fermentation was carried out for 6 days at 28 °C and 250 rpm, with 3% (w/v) Amberlite XAD-16 resin added after 1 day of fermentation. The mixture of the resin and mycelia were harvested by centrifugation and extracted by acetone. The solvent was removed under reduced pressure. The crude extract was suspended with water, and extracted with ethyl acetate. The combined organic phases were concentrated in vacuum, and the residue was subjected to Sephadex LH-20 column chromatography and eluted with methanol. The fractions with compounds of interest were combined, and further purified with Varian LC system

with a Zorbax SB-C13 column (5 μm, 9.4 × 250 mm) to afford GNM B (53.2 mg), GNM B2 (1.1 mg), and GNM B3 (0.8 mg).

Fermentation of *Streptomyces* sp. CB01635, *Streptomyces* sp. NRRL F-5630 and *S. aureofaciens* NRRL B-2183, and analysis of their metabolite profiles

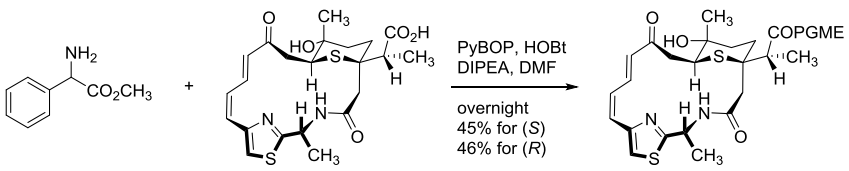
Strain CB01635 and LNM original producing strain *S. atroolivaceus* S-140 were cultivated using a two-stage fermentation similar to the previously published procedure (13). Briefly, fresh spores were inoculated into 250-mL baffled flasks containing 50 mL seed medium (glucose 1%, soluble starch 1%, beef extract 0.3%, yeast extract 0.5%, tryptone 0.5%, $CaCO_3$ 0.2%, pH 7.2), and cultured for 48 h at 28 °C and 250 rpm. Seed culture (5 mL) was then inoculated into 250-mL baffled flasks containing 50 mL production medium LM1. Amberlite XAD-16 resin (Sigma) were added to each flask (4%, w/v) at 16 h after inoculation, and the fermentation was continued for another 72 h. Crude extracts were analyzed with HPLC and LC-MS by the same procedures as those used for strain CB01883.

Streptomyces sp. NRRL F-5630 and *S. aureofaciens* NRRL B-2183 were cultivated and fermented in a similar way to that of strain CB01883 except for the production medium, in which medium LM3 was used. The HPLC and LC-MS analysis were carried out using the same procedures described previously for strain CB01883.

Fermentation of *Streptomyces* sp. CB02120-2, HPLC analysis of its metabolite profiles and isolation of weishanmycin (WSM) A1-A3

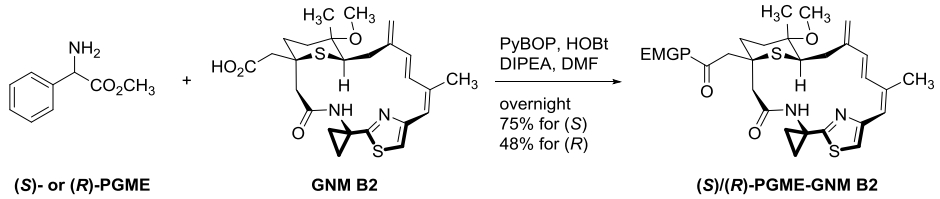
Small-scale fermentations and HPLC-MS analysis of strain CB02120-2, its derivative mutant and complementation strains were carried out using the same procedure for the production of guangnanmycins by strain CB01883. For isolation of WSM A1-A3, a large-scale (12 L) fermentation of strain CB02120-2 was performed with medium LM1. The resins were harvest, and extracted with methanol (2 L × 3). The solvent was evaporated to give the crude extract, which was suspended with water (500 mL), and extracted with ethyl acetate (1 L × 3). The combined organic phases were evaporated to dryness, and the residue was fractionated with Agilent 1260 Infinity LC System using an Eclipse XDB-C18 column (7 μm, 212 × 250 mm). The mobile phase consisted of solvent A (20% acetonitrile with acetic acid, pH 3.6) and B (80% acetonitrile with acetic acid, pH 3.6). The crude extract was eluted with a linear gradient of 0% B to 100% B for 20 min, followed with 100% B for 15 min at a flow rate of 17 mL/min. The fractions containing compounds of interest were combined and concentrated. The residue was subjected to a Sephadex LH-20 column, and eluted with methanol. The fractions containing compounds of interest were combined and finally purified by Varian LC system with a Zorbax SB-C13 column (5 μm, 9.4 × 250 mm) to give WSM A1 (18.1 mg) as a yellowish oil, WSM A2 (16.8 mg) as a yellowish oil, and WSM A3 (2.4 mg) as a yellowish oil.

Chemical transformations of leinamycin E2 (LNM E2), GNM B2 and WSM A2 to their phenylglycine methyl ester (PGME) derivatives for absolute configuration determination (14)

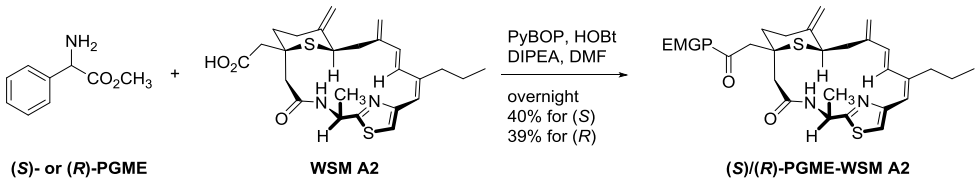


(*S*)- or (*R*)-PGME LNME2 (*S*)/(*R*)-PGME-LNME2

(*S*)/(*R*)-PGME-LNM E2. To a stirring solution of LNM E2 (1.0 equiv) in DMF (0.01 M) was added (*S*)/(*R*)-PGME (1.3 equiv), PyBOP (1.3 equiv), HOBt (1.3 equiv) and DIPEA (20 equiv), and the reaction mixture was allowed to stir at room temperature overnight. The desired product was purified with Agilent 1260 Infinity LC System using an Eclipse XDB-C18 column (7 μm, 212 × 250 mm) with the mobile phase containing A (water) and B (acetonitrile), and eluted with a linear gradient of 5% B to 80% B for 20 min, followed with 80% B for 5 min at a flow rate of 17 mL/min. (*S*)-PGME-LNM E2 (45%) and (*R*)-PGME-LNM E2 (46%) were obtained as white powders.



(*S*)/(*R*)-PGME-GNM B2. To a stirring solution of GNM B2 (1.0 equiv) in DMF (0.01 M) was added (*S*)/(*R*)-PGME (1.3 equiv), PyBOP (1.3 equiv), HOBt (1.3 equiv) and DIPEA (20 equiv), and the reaction mixture was allowed to stir at room temperature overnight. The desired product was purified with Agilent 1260 Infinity LC System using an Eclipse XDB-C18 column (7 μm, 212 × 250 mm) with the mobile phase containing A (H_2O) and B (CH_3CN), and eluted with a linear gradient of 5% B to 100% B for 20 min, followed with 100% B for 7 min at a flow rate of 17 mL/min. (*S*)-PGME-GNM B2 (75%) and (*R*)-PGME-GNM B2 (48%) were obtained as white powders.



(*S*)/(*R*)-PGME-WSM A2. To a stirring solution of WSM A2 (1.0 equiv) in DMF (0.01 M) was added (*S*)/(*R*)-PGME (1.3 equiv), PyBOP (1.3 equiv), HOBt (1.3 equiv) and DIPEA (20 equiv), and the reaction mixture was allowed to stir at room temperature overnight. The desired product was purified with Agilent 1260 Infinity LC System using an Eclipse XDB-C18 column (7 μm, 212 × 250 mm) with the mobile phase containing A (H_2O) and B (CH_3CN), and eluted with a linear gradient of 5% B to 100% B for 20 min, followed with 100% B for 7 min at a flow rate of 17 mL/min. (*S*)-PGME-WSM A2 (40%) and (*R*)-PGME-WSM A2 (39%) were obtained as white powder.

Cloning and production of adenylation domains from *Streptomyces* sp. CB01883 (GnmS), *Micromonospora aurantiaca* (Maur_X), *Streptomyces* sp. CB02120-2 (WsmQ), *Streptomyces* sp. CB01373 (CB01373_Q) and *Streptomyces* sp. CB02613 (CB02613_Z17 and CB02613_M)

To construct a plasmid for production of GnmS, the region coding *gnmS* was amplified by PCR from the cosmid pBS21001 of strain CB01883 using primers 1883orf127-F and 1883orf127-R, and then cloned into the *Nde*I and *Hin*dIII sites of pET28a(+) (Novagen) to afford pBS21008.

To construct a plasmid for production of WsmQ, the region coding *wsmQ* was amplified by PCR from the cosmid pBS22002 (7F6) of strain CB02120-2 using primers 2120orf32-N-5 and

2120orf32-N-3, and then cloned into pBS3080 (15) according to the ligation independent procedures (16) to afford pBS22009.

The DNA region encoding the adenylation domain of Maur_X was amplified by PCR from gDNA of *Micromonospora aurantiaca* with primers MaurX-Ad-N5 and MaurX-Ad-N3, and cloned into the *Bam*HI and *Hin*dIII sites of pBS3082 (15) to obtain pBS21009.
The DNA region encoding the adenylation domain of CB01373_Q was obtained by PCR from gDNA of strain CB01373 with primers 1373Q-Ad-N5 and 1373Q-Ad-N3, and cloned into pBS3080 according to the ligation independent procedures to give pBS21010.

The DNA regions encoding the adenylation domains CB02613_Z17 and CB02613_M were amplified by PCR from gDNA of strain CB02613 with primers 2613Z17-Ad-N5/2613Z17-Ad-N3, and 2613M-Ad-N5/2613M-Ad-N3, respectively. The two PCR products were cloned into pBS3080 according to the ligation independent procedures to give pBS21011 and pBS21012, respectively.

For protein production, plasmids containing the desired gene fragments were transformed into *E. coli* BL21(DE3) (Life Technologies). The resultant recombinant strains were cultivated in LB at 37 °C and 250 rpm until an OD_{600} of 0.6 was reached, and then induced by 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 20 h at 18 °C. The cells were harvested by centrifugation, resuspended in lysis buffer (100 mM Tris, pH 8.0, containing 300 mM NaCl, 15 mM imidazole, and 10% glycerol) and lysed by sonication. After centrifugation at 15,000g for 20 min at 4 °C, the supernatants containing the corresponding proteins were first purified by nickel-affinity chromatography using an ÄKTA FPLC system (GE Healthcare Biosciences) equipped with a HisTrap HP column. The fractions containing the target protein (with an N-terminal His_6-tag) were combined and further purified by ion exchange chromatography using a HiTrap Q HP column (GE Healthcare Biosciences). The purified proteins were desalted using a HiPrep desalting column (GE Healthcare Biosciences) and concentrated using an Amicon Ultra-15 concentrator (Millipore) in 25 mM Tris, pH 7.5, containing 25 mM NaCl, and 10% glycerol. The resultant proteins were stored at −80 °C until use.

Hydroxylamine-trapping assay of adenylation proteins (17)
A typical 150 µl reaction mixture contained 25 mM Tris-HCl (pH 8.0), 15 mM $MgCl_2$, 2.25 mM ATP, 150 mM hydroxylamine (stock solution was prepared as 2 M hydroxylamine in 3.5 M NaOH, and pH adjusted to 8 with concentrated HCl), 3 mM amino acid, and purified A protein (final concentrations of GnmS and WsmQ were 2.5 µM, and the final concentrations of the other adenylation proteins were 7.5 µM). The reaction with boiled enzyme was used as the negative control. The reactions were performed in 96-well plates at 30 °C for 1 h, and stopped by addition of 150 µl of stopping solution containing 10% (w/v) $FeCl_3$ $6H_2O$, 3.3% (v/v) trichloracetic acid, and 0.7 M HCl. The absorbance at 540 nm of each reaction was recorded on a SpectraMax M5 Multi-Mode Microplate Reader (Molecular Devices).

Physicochemical properties of new compounds

Guangnanmycin A. White powder; $[\alpha]_D^{25}$ +15.3 (*c* 0.32, $CHCl_3$/MeOH, 1:1); IR (neat) ν_{max} 3277, 2921, 1669, 1515, 1422, 1201, 1025 cm^{-1}; UV (DMSO) λ_{max} (log ε) 259 (4.09) nm; ^1H and ^{13}C NMR data see **Table S38**; HR-ESI-MS *m/z* 505.1647 $[M + H]^+$ (calcd for $C_{25}H_{33}N_2O_3S_3$, 505.1648).

Guangnanmycin B. White powder; $[\alpha]_D^{25}$ +66.3 (*c* 0.21, $CHCl_3$); IR (neat) ν_{max} 3303, 2925, 1660, 1504, 1445, 1260, 1082 cm^{-1}; UV (DMSO) λ_{max} (log ε) 300 (3.97) nm; ^1H and ^{13}C NMR data see **Table S39**; HR-ESI-MS *m/z* 459.1772 $[M + H]^+$ (calcd for $C_{24}H_{31}N_2O_3S_2$, 459.1771).

Guangnanmycin B1. Yellowish oil; $[\alpha]_D^{25}$ +54.2 (*c* 0.16, $CHCl_3$/MeOH, 1:1); IR (neat) ν_{max} 3264, 2927, 1665, 1535, 1379, 1027 cm^{-1}; UV (DMSO) λ_{max} (log ε) 258 (4.02) nm; 304 (4.06) nm; ^1H and ^{13}C NMR data see **Table S44**; HR-ESI-MS *m/z* 475.1719 $[M + H]^+$ (calcd for $C_{24}H_{31}N_2O_4S_2$, 475.1720).

Guangnanmycin B2. Yellowish oil; $[\alpha]_D^{25}$ +18.2 (*c* 0.06, $CHCl_3$/MeOH, 1:1); IR (neat) ν_{max} 3257, 2925, 1663, 1446, 1377, 1259, 1092 cm^{-1}; UV (DMSO) λ_{max} (log ε) 313 (4.08) nm; ^1H and ^{13}C NMR data see **Table S44**; HR-ESI-MS *m/z* 489.1875 $[M + H]^+$ (calcd for $C_{25}H_{33}N_2O_4S_2$, 489.1876).

Guangnanmycin B3. Yellowish oil; $[\alpha]_D^{25}$ -0.69 (*c* 0.15, $CHCl_3$); IR (neat) ν_{max} 3365, 2925, 1708, 1661, 1451,1260, 1023, 799 cm^{-1}; UV (DMSO) λ_{max} (log ε) 298 (4.13) nm; ^1H and ^{13}C NMR data see **Table S39**; HR-ESI-MS *m/z* 915.3314 $[M + H]^+$ (calcd for $C_{48}H_{59}N_4O_6S_4$, 915.3312).

Weishanmycin A1. Colorless oil; $[\alpha]_D^{25}$ +35.6 (*c* 0.93, $CHCl_3$/MeOH, 1:1); IR (neat) ν_{max} 3378, 2929, 1713, 1647, 1524, 1444, 1023 cm^{-1}; UV (DMSO) λ_{max} (log ε) 279 (4.09) nm; ^1H and ^{13}C NMR data see **Table S42**; HR-ESI-MS *m/z* 491.2029 $[M + H]^+$ (calcd for $C_{25}H_{35}N_2O_4S_2$, 491.2033).

Weishanmycin A2. Yellowish oil; $[\alpha]_D^{25}$ +206 (*c* 0.84, $CHCl_3$/MeOH, 1:1); IR (neat) ν_{max} 3336, 2932, 1712, 1647, 1021 cm^{-1}; UV (DMSO) λ_{max} (log ε) 267 (4.21)nm, 303 (3.89) nm; ^1H and ^{13}C NMR data see **Table S42**; HR-ESI-MS *m/z* 473.1928 $[M + H]^+$ (calcd for $C_{25}H_{33}N_2O_3S_2$, 473.1927).

Weishanmycin A3. Yellowish oil; $[\alpha]_D^{25}$ -140.8 (*c* 0.12, $CHCl_3$/MeOH, 1:1); IR (neat) ν_{max} 3367, 2941, 1653, 1017 cm^{-1}; UV (DMSO) λ_{max} (log ε) 287 (4.49) nm; ^1H and ^{13}C NMR data see **Table S42**; HR-ESI-MS *m/z* 507.1984 $[M + H]^+$ (calcd for $C_{25}H_{35}N_2O_5S_2$, 507.1982).

(*S*)-PGME-LNM E2. White powder; $[\alpha]_D^{25}$ +89.8 (*c* 0.32, $CHCl_3$/MeOH, 1:1); IR (neat) ν_{max} 3288, 2926, 1742, 1645, 1525, 1023, 1003, 699 cm^{-1}; ^1H and ^{13}C NMR data see **Table S41**; HR-ESI-MS *m/z* 612.2194 $[M + H]^+$ (calcd for $C_{31}H_{38}N_3O_6S_2$, 612.2197).

(*R*)-PGME-LNM E2. White powder; $[\alpha]_D^{25}$ +0.7 (*c* 0.31, $CHCl_3$/MeOH, 1:1); IR (neat) ν_{max} 3295, 2968, 1739, 1644, 1525, 1269, 1023, 1003, 699 cm^{-1}; ^1H and ^{13}C NMR data see **Table S41**; HR-ESI-MS *m/z* 612.2199 $[M + H]^+$ (calcd for $C_{31}H_{38}N_3O_6S_2$, 612.2197).

(*S*)-PGME-GNM B2. White powder; $[\alpha]_D^{25}$ +87.0 (*c* 0.20, $CHCl_3$/MeOH, 1:1); IR (neat) ν_{max} 3263, 2937, 1743, 1646, 1531, 1436, 1051, 1025, 1005, 697 cm^{-1}; ^1H and ^{13}C NMR data see **Table S40**; HR-ESI-MS *m/z* 636.2559 $[M + H]^+$ (calcd for $C_{34}H_{42}N_3O_5S_2$, 636.2560).

(*R*)-PGME-GNM B2. White powder; $[\alpha]_D^{25}$ +19.3 (*c* 0.14, $CHCl_3$/MeOH, 1:1); IR (neat) ν_{max} 3253, 2940, 1742, 1645, 1532, 1436, 1211, 1052, 1024, 1005, 697 cm^{-1}; ^1H and ^{13}C NMR data see **Table S40**; HR-ESI-MS *m/z* 636.2562 $[M + H]^+$ (calcd for $C_{34}H_{42}N_3O_5S_2$, 636.2560).

(*S*)-PGME-WSM A2. White powder; $[\alpha]_D^{25}$ +51.0 (*c* 0.54, $CHCl_3$/MeOH, 1:1); IR (neat) ν_{max} 3253, 2929, 1744, 1669, 1644, 1521, 1439, 1209, 1169, 1053, 1026, 977, 891, 731, 697 cm^{-1}; ^1H and ^{13}C NMR data see **Table S43**; HR-ESI-MS *m/z* 620.2612 $[M + H]^+$ (calcd for $C_{34}H_{42}N_3O_4S_2$, 620.2611).

(*R*)-PGME-WSM A2. White powder; $[\alpha]_D^{25}$ -40.7 (*c* 0.29, $CHCl_3$/MeOH, 1:1); IR (neat) ν_{max} 3247, 2927, 1743, 1671, 1642, 1521, 1496, 1440, 1212, 1170, 1053, 1026, 976, 890, 731, 696 cm^{-1}; ^1H and ^{13}C NMR data see **Table S43**; HR-ESI-MS *m/z* 620.2610 $[M + H]^+$ (calcd for $C_{34}H_{42}N_3O_4S_2$, 620.2611).

Table S1*A*. Taxonomy, geography, and genome sequence information of the 19 hit strains from public databases.

Strain name[a]	location[b]	clade[c]	NCBI Genome accession number
Micromonospora tulbaghiae DSM 45142	South Africa	IV	NZ_FMCQ00000000
Streptomyces leeuwenhoekii DSM 42122	Chile	V	NZ_AZSD01000122
Streptomyces hygroscopicus NBRC 16556	Japan	V	NZ_BBOU00000000
Streptomyces sp. NBRC 109436	Japan	V	NZ_BBON00000000
Salinispora arenicola CNH964	Gulf of Mexico	XVII	2515154125 (IMG genome ID)[d]
Streptomyces canus ATCC 12647	USA	VII	NZ_LQCG00000000
Catenulispora acidiphila DSM 44928	Italy	VIII	NC_013131
Streptomyces sp. NRRL F-5630	unknown	IX	NZ_JOGO00000000
Streptomyces aureofaciens NRRL B-2183	USA	IX	NZ_JODU00000000
Saccharothrix sp. ST-888	Japan	XI	NZ_JYJF01000000
Streptomyces sp. NBRC 110035	Japan	XII	NZ_BBNN00000000
Streptomyces novaecaesareae NRRL B-1267	USA	XIII	NZ_JNWQ01000000
Saccharothrix espanaensis DSM 44229	Spain	XIV	NC_019673
Micromonospora sp. L5	Mexico	XV	NC_014815
Micromonospora marina DSM 45555	Thailand	XV	NZ_FMCV00000000
Micromonospora sp. CNB394	unknown	XV	NZ_ARGW00000000
Micromonospora globosa NRRL B-2673	unknown	XV	NZ_JNZR00000000
Micromonospora aurantiaca ATCC 27029	Russia	XV	NC_014391
Burkholderia ubonensis RF25-BP1	Thailand	XVIII	NZ_LOTL01000000

[a]*Burkholderia ubonensis* RF25-BP1 is Gram-negative bacteria, while all the others are actinomycetes (Gram-positive).

[b]Geographic distribution of the 19 hits.

[c]See **Fig. 2** and **Fig. S3B** for detailed clade information.

[d]The IMG genome ID of *Salinispora arenicola* CNH964 is provided, as the NCBI genome accession number of this strain is not available. IMG: Integrated Microbial Genomes (https://jgi.doe.gov/data-and-tools/img/).

Table S1*B*. Taxonomy, geography, accession numbers of partial DUF-SH didomain sequences and selected housekeeping genes of 30 hit strains from the actinomycetes collections at TSRI, and genome sequence information of the nine selected hits.

				NCBI accession number				
strain	location[a]	clade[b]	Taxonomy[c]	*DUF-SH*[d]	*16S rRNA*	*recA*	*trpB*	genome
CB01635	China	I	*Streptomyces* sp.	MF455337	MF455323	MF455352	MF455367	NNBL00000000
MJM2585[f]	Korea	I	*Streptomyces* sp.	MF925449	MF925434	MF925464	MF925479	--
CB02930	China	II	*Streptomyces* sp.	MF455335	MF455321	MF455350	MF455365	--
CB02959	China	II	*Streptomyces* sp.	MF455336	MF455322	MF455351	MF455366	NNBP00000000
CB02891	China	III	*Kitasatospora* sp.	MF455338	MF455324	MF455353	MF455368	NNBO00000000
TSRI0384-2	TSRI	V	*Streptomyces* sp.	MF455339	MF455325	MF455354	MF455369	NOWW00000000
MJM2046[f]	Korea	V	*Streptomyces* sp.	MF925446	MF925431	MF925461	MF925476	--
MJM2478[f]	Korea	V	*Streptomyces* sp.	MF925447	MF925432	MF925462	MF925477	--
MJM2479[f]	Korea	V	*Streptomyces* sp.	MF925448	MF925433	MF925463	MF925478	--
CB02613	UAE	XVI	*Streptomyces* sp.	MF455340	MF455326	MF455355	MF455370	NNBN00000000
CB04103	USA	VI	*Streptomyces* sp.	MF925441	MF925426	MF925456	MF925471	--
CB04055	USA	VII	*Streptomyces* sp.	MF925436	MF925421	MF925451	MF925466	--
I-43184[g]	Italy	VII	*Streptomyces* sp.	MF925445	MF925430	MF925460	MF925475	--
CB04056	USA	VII	*Streptomyces* sp.	MF925437	MF925422	MF925452	MF925467	--
CB04069	USA	VII	*Streptomyces* sp.	MF925438	MF925423	MF925453	MF925468	--
CB04083	USA	VII	*Streptomyces* sp.	MF925439	MF925424	MF925454	MF925469	--
CB04370	China	VII	*Streptomyces* sp.	MF925443	MF925428	MF925458	MF925473	--
MJM7752[f]	Korea	VII	*Streptomyces* sp.	MF925450	MF925435	MF925465	MF925480	--
CB02209	China	VII	*Streptomyces* sp.	MF455343	MF455329	MF455358	MF455373	--
CB01373	TSRI	VII	*Streptomyces* sp.	MF455341	MF455327	MF455356	MF455371	NNBK00000000
CB01318	China	VII	*Streptomyces* sp.	MF455342	MF455328	MF455357	MF455372	--
CB02274	China	VII	*Streptomyces* sp.	MF455344	MF455330	MF455359	MF455374	--
CB01883	China	IX	*Streptomyces* sp.	MF455349	KT722854	MF455364	KT793843	LIWA00000000
CB04099	USA	IX	*Streptomyces* sp.	MF925440	MF925425	MF925455	MF925470	--
CB01201	China	X	*Streptomyces* sp.	MF455348	MF455334	MF455363	MF455378	NNBJ00000000
CB05481	China	XII	*Streptomyces* sp.	MF925444	MF925429	MF925459	MF925474	--
TSRI0429	TSRI	XII	*Streptomyces* sp.	MF455347	MF455333	MF455362	MF455377	--
CB02120-2	China	XII	*Streptomyces* sp.	MF455346	MF455332	MF455361	MF455376	NNBM00000000
CB04134	China	XII	*Streptomyces* sp.	MF925442	MF925427	MF925457	MF925472	--
TSRI0370	TSRI	XV	*Micromonospora* sp.	MF455345	MF455331	MF455360	MF455375	--
S-140[e]	Japan	I	*Streptomyces* sp.	--	MF511767	MF511768	MF511769	--

[a]Geographic distribution of the 30 hits, UAE abbreviated for United Arab Emirates, and TSRI denoted legacy strains in the TSRI strain collection whose geographic origins are not available (also see **Fig. S2**).

[b]See **Fig. 2** and **Fig. S3B** for detailed clade information.

[c]Taxonomy classification based on the three housekeeping genes including *16S rRNA, recA,* and *trpB* (also see **Fig. S3A**).

[d]The partial DUF-SH didomain sequences used to construct the phylogenetic tree (**Fig. 2** and **Fig. S3B**).

[e]Original leinamycin producing strain *Streptomyces atroolivaceus* S-140.

[f]Hit strains from Myongji University.

[g]Hit strain from the Naicons' collection.

Table S2. Strains and plasmids used in this study

Strains/Plasmids	Genotype description	Source
***E. coli* strains**		
DH5α	*E. coli* host for general cloning	Life Technologies
ET12567/pUZ8002	Methylation-deficient *E. coli* host for intergeneric conjugation	2
XL1-Blue MRF	*E. coli* host for cosmid library construction	Epicentre
BL21(DE3)	*E. coli* host for protein production	Life Technologies
***Streptomyces* strains**		
Streptomyces sp. CB01883	Producing strain of guangnanmycins (GNM), wild-type	This study
SB21001 (Δ*gnmB*)	*gnmB* in-frame deletion mutant strain of CB01883	This study
SB21002 (Δ*gnmB*/*gnmB*)	mutant strain SB21001 harboring plasmid pBS21005 (pSETgnmB)	This study
SB21003 (Δ*gnmO*)	*gnmO* in-frame deletion mutant strain of CB01883	This study
SB21004 (Δ*gnmO*/*gnmO*)	mutant strain SB21003 harboring plasmid pBS21007 (pSETgnmO)	This study
Streptomyces sp. NRRL F-5630	Alternative producer of GNM, wild-type	NRRL
Streptomyces aureofaciens NRRL B-2183	Alternative producer of GNM, wild-type	NRRL
Streptomyces atroolivaceus S-140	Producing strain of LNM, wild-type	Kyowa Hakko Kogyo Co. Ltd.
Streptomyces sp. CB01635	Alternative producer of LNM E1, wild-type	This study
Streptomyces sp. CB02120-2	Producing strain of weishanmycin (WSM) A1-A3, wild-type	This study
SB22001 (Δ*wsmW*)	*wsmW* in-frame deletion mutant strain of CB02120-2	This study
SB22002 (Δ*wsmZ3*)	*wsmZ3* in-frame deletion mutant strain of CB02120-2	This study
SB22003 (Δ*wsmZ3*/*wsmZ3*)	mutant strain SB22002 harboring plasmid pBS22006 (pSETwsmZ3)	This study
SB22004 (Δ*wsmZ4*)	*wsmZ4* in-frame deletion mutant strain of CB02120-2	This study
SB22005 (Δ*wsmZ4*/*wsmZ4*)	mutant strain SB22004 harboring plasmid pBS22008 (pSETwsmZ4)	This study
Plasmids		
SuperCos 1	Vector for the construction of cosmid library, $Kana^R$	Agilent Technologies
pOJ260	*E. coli-Streptomyces* shuttle vector, Am^R	10
pSETTurdR	pSET152 derived plasmid containing the promoter *kasO**, Am^R	11
pBS3080	pRSFDuet-1 derived plasmid containing *BsmF*I site for ligation-independent cloning	15
pBS3082	pRSFDuet-1 derived plasmid with insert encodes a TEV protease site after the His_6-tag	15
pBS21001	Cosmid 1C4 containing partial *gnm* gene cluster	This study
pBS21002	Cosmid 1B9 containing partial *gnm* gene cluster	This study
pBS21003 (pSET-kasO*)	pSETTurdR derived plasmid containing the promoter *kasO**, Am^R	This study
pBS21004 (pOJgnmB)	pOJ260 derived plasmid for inactivation of *gnmB* in CB01883	This study
pBS21005 (pSETgnmB)	pBS21003 derived plasmid for over-expression of *gnmB*	This study
pBS21006 (pOJgnmO)	pOJ260 derived plasmid for inactivation of *gnmO* in CB01883	This study
pBS21007 (pSETgnmO)	pBS21003 derived plasmid for over-expression of *gnmO*	This study
pBS21008 (pET28a-GnmS)	pET28a derived plasmid for production of GnmS	This study
pBS21009 (pBS3082-Maur_X-Ad)	pBS3082 derived plasmid for protein production of the A domain of Maur_X	This study
pBS21010 (pRSF-1373Q-Ad)	pBS3080 derived plasmid for protein production of the A domain of CB01373_Q	This study

pBS21011 (pRSF-2613Z17-Ad)	pBS3080 derived plasmid for protein production of the A domain of CB02613_Z17	This study
pBS21012 (pRSF-2613M-Ad)	pBS3080 derived plasmid for protein production of the A domain of CB02613_M	This study
pBS22001	Cosmid 8E4 containing partial *wsm* gene cluster	This study
pBS22002	Cosmid 7F6 containing partial *wsm* gene cluster	This study
pBS22003	Cosmid 3F12 containing partial *wsm* gene cluster	This study
pBS22004 (pOJwsmW)	pOJ260 derived plasmid for inactivation of *wsmW* in CB02120-2	This study
pBS22005 (pOJwsmZ3)	pOJ260 derived plasmid for inactivation of *wsmZ3* in CB02120-2	This study
pBS22006 (pSETwsmZ3)	pBS21003 derived plasmid for over-expression of *wsmZ3*	This study
pBS22007 (pOJwsmZ4)	pOJ260 derived plasmid for inactivation of *wsmZ4* in CB02120-2	This study
pBS22008 (pSETwsmZ4)	pBS21003 derived plasmid for over-expression of *wsmZ4*	This study
pBS22009 (pRSF-WsmQ)	pBS3080 derived plasmid for protein production of WsmQ	This study

Table S3. Primers used in this study

Name	Sequence (from 5' to 3') (restriction sites underlined)	Function
DUF-F	AYCTGCTSGACSTGGAGTTCTA	Real-time PCR screening
SH-sub-R	CSMGBTCSGMCCASGAGTCGGT	Real-time PCR screening
Seq-DUF-F	TGTAAAACGACGGCCAGTAYCTGCTSGACSTGGAGTTCTA	Amplifying partial sequence from DUF-SH didomain
Seq-SH-R	CAGGAAACAGCTATGACCCGSHGTCGASRCCGAAGTCCTT	Amplifying partial sequence from DUF-SH didomain
Seq-SH-2120-R	CAGGAAACAGCTATGACACGCCGAAGTCCTTSGASAGRCT	Amplifying partial sequence from DUF-SH didomain
16SrRNA_for	AGAGTTTGATCCTGGCTCAG	Amplifying partial 16S rRNA
16SrRNA_rev	ACGGCTACCTTGTTACGACTT	Amplifying partial 16S rRNA
recA-for	TAATACGACTCACTATAGGGCCGCRCTCGCACARATTGAACG	Amplifying partial *recA*
recA-rev	GCTAGTTATTGCTCAGCGGCGTCGGGGTTGTCCTTSAGGAAG	Amplifying partial *recA*
recA-micromono-F	TAATACGACTCACTATAGGGAGACGCCGGCGAAGGCAGG	Amplifying partial *recA* from *Micromonospora* sp. strains
recA-micromono-R	GCTAGTTATTGCTCAGCGGCCGACGCCGAGCTTCTCCA	Amplifying partial *recA* from *Micromonospora* sp. strains
recA-kitasatospora-F	TAATACGACTCACTATAGGGGYNCYDGRSRCCGSHCTYGC	Amplifying partial *recA* from *Kitasatospora* sp. strains
recA-kitasatospora-R	GCTAGTTATTGCTCAGCGGSRGCTGGWMSCCMTSGYAGG	Amplifying partial *recA* from *Kitasatospora* sp. strains
trpB-for	TAATACGACTCACTATAGGGGCGCGAGGACCTGAACCACAC	Amplifying partial *trpB*
trpB-rev	GCTAGTTATTGCTCAGCGGCATGGCCGGGATGATGCCC	Amplifying partial *trpB*
trpB-micromono-Fn	TAATACGACTCACTATAGGGACCGGCTCCCGGCCCAGCT	Amplifying partial *trpB* from *Micromonospora* sp. strains
trpB-micromono-Rn	GCTAGTTATTGCTCAGCGGCGCSCTGCTGCGCGACTAC	Amplifying partial *trpB* from *Micromonospora* sp. strains
1883-screen-U5	TACATGACCTGCCTGGAGCG	Screening for cosmids containing *gnm* gene cluster
1883-screen-U3	GGTCGAGGATCAGCAGTCCC	Screening for cosmids containing *gnm* gene cluster
1883-screen-M5	AGCACGCCAGAAAAGTCCCT	Screening for cosmids containing *gnm* gene cluster
1883-screen-M3	GTGTAGGTCCAGCTCCGGTC	Screening for cosmids containing *gnm* gene cluster
1883-screen-D5	CTACGTCGAGACACACGGCA	Screening for cosmids containing *gnm* gene cluster
1883-screen-D3	CAAGGGGTTCTCCGTCCTCG	Screening for cosmids containing *gnm* gene cluster
1883orf155-I5	CCCAAGCTTGGCACCGCCGGAGTGGCATC (*Hin*dIII)	Inactivation of *gnmB* in CB01883
1883orf155-I3	GCTCTAGACCTTGGGCGACCAGCGGCTC (*Xba*I)	Inactivation of *gnmB* in CB01883
1883orf155-II5	GCTCTAGACCCTGGTCAACGCGGTCTCG (*Xba*I)	Inactivation of *gnmB* in CB01883
1883orf155-II3	GGAATTCCGCTGCGTGCCCGATGTC (*Eco*RI)	Inactivation of *gnmB* in

		CB01883
orf155-SBlot-5n	CTGGAGCACCGGCAGGACCC	Southern blot analysis of SB21001 (Δ*gnmB*)
orf155-SBlot-3n	AAGTGCTCCGCGAACCGGC	Southern blot analysis of SB21001 (Δ*gnmB*)
orf155-kasO-5	CCCTTAAGCCTGCGGTGGTGCACAAAGTTGA (*Afl*II)	Complementation of SB21001 (Δ*gnmB*)
orf155-kasO-3	ACGTACTAGTCGGTGTCCCCTCACTCGGTTCC (*Spe*I)	Complementation of SB21001 (Δ*gnmB*)
PSET-kasO-F	TGCATGCATACGTACTAGTCTGACATATGACGTGAATTCGTAATC (*Nsi*I)	Construction of pBS21003 (pSET-kasO*)
PSET-kasO-R	TGCATGCATACGTCTTAAGAACTCCCCCAGTCCTGCA (*Nsi*I)	Construction of pBS21003 (pSET-kasO*)
1883orf132-I5	CCCAAGCTTTCTGTGTCGAGGGGGACTC (*Hin*dIII)	Inactivation of *gnmO* in CB01883
1883orf132-I3	GCTCTAGAATTCCATGTCGTCGGGATGC (*Xba*I)	Inactivation of *gnmO* in CB01883
1883orf132-II5	GCTCTAGACGATCGCCTTCGAGCTGATA (*Xba*I)	Inactivation of *gnmO* in CB01883
1883orf132-II3	GGAATTCACTTTTCTGGCGTGCTCAGT (*Eco*RI)	Inactivation of *gnmO* in CB01883
orf132-SBlot-5nn	AGAACTCCGCGAGAGCCTTC	Southern blot analysis of SB21003 (Δ*gnmO*)
orf132-SBlot-3nn	CCCGTTCGAGCAGACCGAC	Southern blot analysis of SB21003 (Δ*gnmO*)
orf132-kasO-5	CCCTTAAGCCTCTCGCTCACGGGAC (*Afl*II)	Complementation of SB21003 (Δ*gnmO*)
orf132-kasO-3	ACGTACTAGTGTCTACGCCAAGCGGTGA (*Spe*I)	Complementation of SB21003 (Δ*gnmO*)
2120-screen-U5	CCCTGACCCCGCTCACC	Screening for cosmids containing *wsm* gene cluster
2120-screen-U3	GTTGCTCGGTGAAGTCGTCG	Screening for cosmids containing *wsm* gene cluster
2120-screen-M5	CCGTCGAACGCGTACACATC	Screening for cosmids containing *wsm* gene cluster
2120-screen-M3	ATGGAGTCGAGCCCCATGTC	Screening for cosmids containing *wsm* gene cluster
2120-screen-D5	TTCGGCTTCGGAGGGTTCAA	Screening for cosmids containing *wsm* gene cluster
2120-screen-D3	ACACCAGGATGACGGACCTG	Screening for cosmids containing *wsm* gene cluster
2120orf25-I5	CCCAAGCTTGCGCAGGAACTCGGCCTTG (*Hin*dIII)	Inactivation of *wsmW* in CB02120-2
2120orf25-I3	GCTCTAGACGCACCACCGTCATCGAGTC (*Xba*I)	Inactivation of *wsmW* in CB02120-2
2120orf25-II5	GCTCTAGAGTGGACTGGGCGAACCGGAC (*Xba*I)	Inactivation of *wsmW* in

		CB02120-2
2120orf25-II3	GGAATTCTTCGAGACGATGGACACCGC (*Eco*RI)	Inactivation of *wsmW* in CB02120-2
orf25-SBlot-5	GTCAGCCACTGCGACCTCAT	Southern blot analysis of SB22001 (Δ*wsmW*)
orf25-SBlot-3	TGCGGAAGCGTTCGTCGTC	Southern blot analysis of SB22001 (Δ*wsmW*)
2120orfZ3-I5	GTAACAAGCTTTGCTCGTACATCGTCACCGC (*Hin*dIII)	Inactivation of *wsmZ3* in CB02120-2
2120orfZ3-I3	GCTCTAGATGCGCTCCTCGATCCATTCG (*Xba*I)	Inactivation of *wsmZ3* in CB02120-2
2120orfZ3-II5	GCTCTAGACGGAGACCTCGTCGTCATGG (*Xba*I)	Inactivation of *wsmZ3* in CB02120-2
2120orfZ3-II3	GGAATTCGACCAGAGCCGGACCTACGA (*Eco*RI)	Inactivation of *wsmZ3* in CB02120-2
orfZ3-SBlot-5	CTTCGACTACGTCGGCCGGG	Southern blot analysis of SB22002 (Δ*wsmZ3*)
orfZ3-SBlot-3	TCCATTCGGCGGTGATGCCG	Southern blot analysis of SB22002 (Δ*wsmZ3*)
orfZ3-kasO-5	CCCTTAAGCTCGTCCAGGAGAACCGGCAC (*Afl*II)	Complementation of SB22002 (Δ*wsmZ3*)
orfZ3-kasO-3	ACGTACTAGTTGGCACGGGAGAACGACCACG (*Spe*I)	Complementation of SB22002 (Δ*wsmZ3*)
2120orfZ4-I5	GTAACAAGCTTCTCCTGCGCTACCTGAGCAT (*Hin*dIII)	Inactivation of *wsmZ4* in CB02120-2
2120orfZ4-I3	GCTCTAGATGGCACGGGAGACGACCA (*Xba*I)	Inactivation of *wsmZ4* in CB02120-2
2120orfZ4-II5	GCTCTAGAGAGTTCGCCGAGCCGCAGTA (*Xba*I)	Inactivation of *wsmZ4* in CB02120-2
2120orfZ4-II3	GGAATTCAGCATGCTCTTCCCCGGTGC (*Eco*RI)	Inactivation of *wsmZ4* in CB02120-2
orfZ4-SBlot-5	ACCTCAAGCGGGTCATCGGC	Southern blot analysis of SB22004 (Δ*wsmZ4*)
orfZ4-SBlot-3	CGCAGGGCCGAGGTGAGG	Southern blot analysis of SB22004 (Δ*wsmZ4*)
orfZ4-kasO-5	CCCTTAAGAAGATCGGGATGTCCCACGC (*Afl*II)	Complementation of SB22004 (Δ*wsmZ4*)
orfZ4-kasO-3	ACGTACTAGTGTTCTCGGCCTGGCTGTACG (*Spe*I)	Complementation of SB22004 (Δ*wsmZ4*)
1883orf127-F	GGAATTCCATATGACGGCCGACGCGGCGGT (*Nde*I)	Production of GnmS
1883orf127-R	GTAACAAGCTTTCACCGGTCCGCCAGGCAGC (*Hin*dIII)	Production of GnmS
2120orf32-N-5	AAAACCTCTATTTCCAGTCGGTGCTGCTGGACCAGTGGTTC	Production of WsmQ
2120orf32-N-3	TACTTACTTAAATGTCACGAGGACGGAGTGGGG	Production of WsmQ
MaurX-Ad-N5	CGGGATCC ATGCCCGTGACTGACTGCCTGCAC (*Bam*HI)	Production of the A domain Maur_X
MaurX-Ad-N3	CCCAAGCTTACTCGGCCGGCTCGGCCAGA (*Hin*dIII)	Production of the A domain

		Maur_X
1373Q-Ad-N5	AAAACCTCTATTTCCAGTCGATGGTTTCCGGCAAGACGCTG	Production of the A domain CB01373_Q
1373Q-Ad-N3	TACTTACTTAAATGTTATCACGTCAGTTCCGCGCCGGC	Production of the A domain CB01373_Q
2613Z17-Ad-N5	AAAACCTCTATTTCCAGTCGGTGAGCGCACCACCCGACCC	Production of the A domain CB02613_Z17
2613Z17-Ad-N3	TACTTACTTAAATGTTATCAGCTCTCGGGATCCCCGCCG	Production of the A domain CB02613_Z17
2613M-Ad-N5	AAAACCTCTATTTCCAGTCGATGACCGACACGCAGCCGCTC	Production of the A domain CB02613_M
2613M-Ad-N3	TACTTACTTAAATGTTATCACAGGTACGCGGCGACGCCCT	Production of the A domain CB02613_M

Table S4. Predicted functions of ORFs in the *lnm*-type gene cluster from *Micromonospora tulbaghiae* DSM 45142 (Clade IV)

gene[a]	aa[b]	putative function[c]	protein homologue (accession #)	% identity/ % similarity
Mtul_orf(-1)	274	hypothetical protein	(WP_043960781)	86/90
Mtul_A	552	4-coumarate:CoA ligase	LnmW (AAN85536)	35/48
Mtul_B	193	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	46/67
Mtul_C	602	NRPS (A-PCP)	SyrB1 (AKF46132)	45/59
Mtul_D	315	halogenase	SyrB2 (AKF46131)	69/79
Mtul_E	379	alpha/beta hydrolase	ADZ36_03165 (KNE83935)	46/55
Mtul_F	244	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	50/60
Mtul_G	306	unknown	LnmE (AAN85518)	49/61
Mtul_H	257	enoyl-CoA hydratase	LnmF (AAN85519)	53/64
Mtul_I	778	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	65/76
Mtul_J	314	unknown	LnmH (AAN85521)	44/61
Mtul_K	4285	AT-less type I PKS	LnmI (AAN85522)	50/60
Mtul_L	7078	AT-less type I PKS	LnmJ (AAN85523)	50/60
Mtul_M	329	acyltransferase/decarboxylase	LnmK (AAN85524)	55/70
Mtul_N	86	acyl carrier protein	LnmL (AAN85525)	58/73
Mtul_O	413	HMG-CoA-synthase	LnmM (AAN85526)	65/79
Mtul_P	250	type II thioesterase	LnmN (AAN85527)	55/65
Mtul_Q	134	unknown	LnmZ' (AAN85540)	33/47
Mtul_R	476	dTDP-hexose 2,3-dehydratase	JadO (AAK01936)	45/56
Mtul_S	333	oxidoreductase	(WP_060727604)	49/60
Mtul_T	203	dTDP-4-dehydrorhamnose 3,5-epimerase	NivL (AGZ78381)	45/57
Mtul_U	374	dTDP-sugar aminotransferase	(SEP57347)	60/73
Mtul_V	214	TetR family transcriptional regulator	(WP_062008487)	46/57
Mtul_W	510	antibiotic efflux protein	LnmY (AAN85538)	44/61
Mtul_X	256	sugar O-methlytransferase	(WP_009950022)	64/77
Mtul_Y	411	glycosyltransferase, MGT family	(WP_063747047)	43/54
Mtul_Z1	378	hypothetical protein	(WP_054026420)	30/50
Mtul_Z2	403	cytochrome P450 oxygenase	LnmA (AAN85514)	40/55
Mtul_Z3	405	cytochrome P450 oxygenase	LnmA (AAN85514)	49/65
Mtul_Z4	126	glyoxalase/bleomycin resistance protein	BN971_01690 (CPR09988)	44/54
Mtul_Z5	527	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	34/50
Mtul_Z6	331	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	45/59
Mtul_Z7	803	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	43/55
Mtul_Z8	419	cytochrome P450 oxygenase	LnmA (AAN85514)	35/52
Mtul_Z9	261	methyltransferase	(WP_031183778)	32/44
Mtul_Z10	238	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	58/68
Mtul_Z11	469	sodium:proton exchanger	(WP_020553490)	34/52
Mtul_Z12	522	4-coumarate:CoA ligase	LnmW (AAN85536)	36/67
Mtul_Z13	138	unknown, NTF2 family protein	LnmV (AAN85535)	46/65
Mtul_Z14	432	hypothetical protein	(WP_067304298)	43/62
Mtul_Z15	229	4'-phosphopantetheinyl transferase	Sfp (P39135)	24/40

Mtul_orf(+1)	308	transcriptional regulator	(WP_009713296)	51/64

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S5. Predicted functions of ORFs in the *lnm*-type gene cluster from *Streptomyces* sp. NBRC 109436 (Clade V)

gene[a]	aa[b]	putative function[c]	protein homologue (accession #)	% identity/ % similarity
S109_orf(-1)	230	TetR family transcriptional regulator	(WP_062008487)	99/99
S109_A	572	4-coumarate:CoA ligase	LnmW (AAN85536)	40/56
S109_B	252	methyltransferase	(WP_031183778)	35/45
S109_C	293	unknown	LnmE (AAN85518)	49/63
S109_D	411	HMG-CoA-synthase	LnmM (AAN85526)	47/63
S109_E	261	enoyl-CoA hydratase	LnmF (AAN85519)	51/63
S109_F	845	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	67/78
S109_G	297	unknown	LnmH (AAN85521)	43/59
S109_H	626	adenylation protein-peptidyl carrier protein	SyrB1 (AKF46132)	45/59
S109_I	316	halogenase	SyrB2 (AKF46131)	65/75
S109_J	387	alpha/beta hydrolase	ADZ36_03165 (KNE83935)	44/55
S109_K	4266	AT-less type I PKS	LnmI (AAN85522)	47/57
S109_L	6593	AT-less type I PKS	LnmJ (AAN85523)	50/61
S109_M	327	acyltransferase/decarboxylase	LnmK (AAN85524)	52/64
S109_N	86	acyl carrier protein	LnmL (AAN85525)	63/77
S109_O	410	HMG-CoA synthase	LnmM (AAN85526)	64/76
S109_P	82	acyl carrier protein	(WP_062008561)	99/100
S109_Q	424	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	46/58
S109_R	260	type II thioesterase	LnmN (AAN85527)	54/65
S109_S	139	unknown, NTF2 family protein	LnmV (AAN85535)	43/61
S109_T	133	unknown	LnmZ' (AAN85540)	34/49
S109_U	396	cytochrome P450 oxygenase	LnmA (AAN85514)	44/58
S109_V	411	cytochrome P450 oxygenase	LnmA (AAN85514)	50/66
S109_W	144	glyoxalase/bleomycin resistance protein	BN971_01690 (CPR09988)	41/53
S109_X	238	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	53/64
S109_Y	530	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	32/46
S109_Z1	333	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	44/60
S109_Z2	817	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	42/54
S109_Z3	226	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	48/68
S109_orf(+1)	209	dTDP-4-dehydrorhamnose 3,5-epimerase	NivL (AGZ78381)	67/79

[a] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b] Number of amino acids. [c] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S6. Predicted functions of ORFs in the *lnm*-type gene cluster from *S. hygroscopicus* sp. NBRC 16556 (Clade V)

gene[a]	aa[b]	putative function[c]	protein homologue (accession #)	% identity/ % similarity
Shyg_orf(-1)	230	TetR family transcriptional regulator	(WP_062008487)	99/99
Shyg_A	572	4-coumarate:CoA ligase	LnmW (AAN85536)	40/56
Shyg_B	252	methyltransferase	(WP_031183778)	35/45
Shyg_C	293	unknown	LnmE (AAN85518)	49/63
Shyg_D	411	HMG-CoA-synthase	LnmM (AAN85526)	47/63
Shyg_E	261	enoyl-CoA hydratase	LnmF (AAN85519)	51/63
Shyg_F	845	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	67/78
Shyg_G	337	unknown	LnmH (AAN85521)	43/59
Shyg_H	626	adenylation protein-peptidyl carrier protein	SyrB1 (AKF46132)	45/59
Shyg_I	316	halogenase	SyrB2 (AKF46131)	65/75
Shyg_J	397	alpha/beta hydrolase	ADZ36_03165 (KNE83935)	44/55
Shyg_K	4258	AT-less type I PKS	LnmI (AAN85522)	47/57
Shyg_L	6571	AT-less type I PKS	LnmJ (AAN85523)	50/61
Shyg_M	327	acyltransferase/decarboxylase	LnmK (AAN85524)	52/64
Shyg_N	86	acyl carrier protein	LnmL (AAN85525)	63/77
Shyg_O	410	HMG-CoA synthase	LnmM (AAN85526)	64/76
Shyg_P	82	acyl carrier protein	(WP_062008561)	99/100
Shyg_Q	424	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	46/58
Shyg_R	260	type II thioesterase	LnmN (AAN85527)	54/65
Shyg_S	139	unknown, NTF2 family protein	LnmV (AAN85535)	43/61
Shyg_T	133	unknown	LnmZ' (AAN85540)	34/49
Shyg_U	396	cytochrome P450 oxygenase	LnmA (AAN85514)	44/58
Shyg_V	411	cytochrome P450 oxygenase	LnmA (AAN85514)	50/66
Shyg_W	144	glyoxalase/bleomycin resistance protein	BN971_01690 (CPR09988)	41/53
Shyg_X	238	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	53/64
Shyg_Y	530	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	32/46
Shyg_Z1	333	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	44/60
Shyg_Z2	817	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	42/54
Shyg_Z3	226	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	48/68
Shyg_orf(+1)	209	dTDP-4-dehydrorhamnose 3,5-epimerase	NivL (AGZ78381)	67/79

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S7. Predicted functions of ORFs in the *lnm*-type gene cluster from *S. leeuwenhoekii* DSM 42122 (Clade V)

gene[a]	aa[b]	putative function[c]	protein homologue (accession #)	% identity/ % similarity
Slee_orf(-1)	187	TetR family transcriptional regulator	(WP_062008487)	98/98
Slee_A	479	4-coumarate:CoA ligase	LnmW (AAN85536)	40/56
Slee_B	234	methyltransferase	(WP_031183778)	37/46
Slee_C	293	unknown	LnmE (AAN85518)	48/62
Slee_D	411	HMG-CoA-synthase	LnmM (AAN85526)	47/64
Slee_E	255	enoyl-CoA hydratase	LnmF (AAN85519)	50/63
Slee_F	808	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	67/78
Slee_G	303	unknown	LnmH (AAN85521)	44/60
Slee_H	630	adenylation protein-peptidyl carrier protein	SyrB1 (AKF46132)	45/59
Slee_I	316	halogenase	SyrB2 (AKF46131)	65/75
Slee_J	376	alpha/beta hydrolase	ADZ36_03165 (KNE83935)	44/56
Slee_K	4212	AT-less type I PKS	LnmI (AAN85522)	49/59
Slee_L	6495	AT-less type I PKS	LnmJ (AAN85523)	50/61
Slee_M	310	acyltransferase/decarboxylase	LnmK (AAN85524)	52/64
Slee_N	86	acyl carrier protein	LnmL (AAN85525)	50/67
Slee_O	410	HMG-CoA synthase	LnmM (AAN85526)	64/76
Slee_P	82	acyl carrier protein	(WP_062008561)	96/97
Slee_Q	424	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	46/58
Slee_R	260	type II thioesterase	LnmN (AAN85527)	52/64
Slee_S	139	unknown, NTF2 family protein	LnmV (AAN85535)	43/61
Slee_T	120	unknown	LnmZ' (AAN85540)	36/51
Slee_U	396	cytochrome P450 oxygenase	LnmA (AAN85514)	43/57
Slee_V	387	cytochrome P450 oxygenase	LnmA (AAN85514)	51/66
Slee_W	124	glyoxalase/bleomycin resistance protein	BN971_01690 (CPR09988)	43/53
Slee_X	238	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	53/65
Slee_Y	530	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	32/46
Slee_Z1	315	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	44/60
Slee_Z2	778	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	42/55
Slee_Z3	226	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	49/68
Slee_orf(+1)	203	dTDP-4-dehydrorhamnose 3,5-epimerase	NivL (AGZ78381)	68/80

[a] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b] Number of amino acids. [c] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S8. Predicted functions of ORFs in the *lnm*-type gene cluster from *Streptomyces canus* ATCC 12647 (Clade VII)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
Scan_orf(-1)	205	TetR family transcriptional regulator	(WP_058852673)	41/62
Scan_A	191	YdeI-like protein	SAVERM_307 (BAC68016)	71/80
Scan_B	454	cytochrome P450 oxygenase	VR46_04425 (KJY47279)	38/55
Scan_C	228	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	42/59
Scan_D	236	4'-phosphopantetheinyl transferase	Sfp (P39135)	22/41
Scan_E	433	cytochrome P450 oxygenase	LnmA (AAN85514)	36/57
Scan_F	238	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	43/62
Scan_G	437	sodium:proton exchanger	(WP_020553490)	58/74
Scan_H	298	unknown	LnmE (AAN85518)	46/63
Scan_I	410	HMG-CoA synthase	LnmM (AAN85526)	46/62
Scan_J	263	enoyl-CoA hydratase	LnmF (AAN85519)	51/65
Scan_K	867	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	57/69
Scan_L	316	unknown	LnmH (AAN85521)	43/59
Scan_M	599	adenylation protein-peptidyl carrier protein	SyrB1 (AKF46132)	46/62
Scan_N	311	halogenase	SyrB2 (AKF46131)	67/79
Scan_O	391	alpha/beta hydrolase	ADZ36_03165 (KNE83935)	43/54
Scan_P	4242	hybrid NRPS/type I PKS	LnmI (AAN85522)	51/61
Scan_Q	7157	AT-less type I PKS	LnmJ (AAN85523)	49/61
Scan_R	329	acyltransferase/decarboxylase	LnmK (AAN85524)	55/67
Scan_S	87	acyl carrier protein	LnmL (AAN85525)	60/69
Scan_T	412	HMG-CoA synthase	LnmM (AAN85526)	63/77
Scan_U	80	acyl carrier protein	(WP_062008561)	53/71
Scan_V	422	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	43/53
Scan_W	253	type II thioesterase	LnmN (AAN85527)	53/65
Scan_X	239	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	55/65
Scan_Y	132	unknown	LnmZ' (AAN85540)	38/57
Scan_Z1	488	4-coumarate:CoA ligase	LnmW (AAN85536)	44/57
Scan_Z2	395	FAD-dependent oxidoreductase	(WP_010982689)	61/70
Scan_Z3	121	glyoxalase/bleomycin resistance protein	(WP_062700437)	41/60
Scan_Z4	491	4-coumarate:CoA ligase	LnmW (AAN85536)	42/55
Scan_Z5	414	cytochrome P450 oxygenase	LnmA (AAN85514)	34/50
Scan_orf(+1)	1157	Repeat domain-containing protein	STTU_5173 (EGJ77962)	64/75

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S9. Predicted functions of ORFs in the *lnm*-type gene cluster from *Catenulispora acidiphila* DSM 44928 (Clade VIII)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
Caci_orf(-1)	311	Hypothetical protein	Not found	
Caci_A	397	FAD-dependent oxidoreductase	(WP_010982689)	44/58
Caci_B	483	amidase	(WP_011874565)	86/92
Caci_C	503	4-coumarate:CoA ligase	LnmW (AAN85536)	33/45
Caci_D	539	antibiotic efflux protein	LnmY (AAN85538)	29/48
Caci_E	360	oxidoreductase	(WP_026583996)	45/60
Caci_F	404	cytochrome P450 oxygenase	LnmA (AAN85514)	39/53
Caci_G	308	unknown	LnmH (AAN85521)	43/60
Caci_H	132	unknown	LnmZ' (AAN85540)	34/49
Caci_I	179	hypothetical protein	(WP_053707566)	50/65
Caci_J	80	ferredoxin	LnmB (AAN85515)	36/47
Caci_K	865	acyltransferase/oxidoreductase	LnmG	51/62
Caci_L	622	NRPS (A-PCP)	SyrB1 (AKF46132)	43/59
Caci_M	332	halogenase	SyrB2 (AKF46131)	67/79
Caci_N	397	alpha/beta hydrolase	ADZ36_03165 (KNE83935)	46/58
Caci_O	221	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	46/66
Caci_P	2454	AT-less type I PKS	LnmI (AAN85522)	46/58
Caci_Q	264	type II thioesterase	LnmN (AAN85527)	56/67
Caci_R	325	unknown	LnmE (AAN85518)	44/63
Caci_S	413	HMG-CoA synthase	LnmM (AAN85526)	45/61
Caci_T	268	enoyl-CoA hydratase	LnmF (AAN85519)	46/58
Caci_U	2284	AT-less type I PKS	LnmI (AAN85522)	49/58
Caci_V	7149	AT-less type I PKS	LnmJ (AAN85523)	44/56
Caci_W	324	acyltransferase/decarboxylase	LnmK (AAN85524)	50/68
Caci_X	89	acyl carrier protein	LnmL (AAN85525)	50/72
Caci_Y	424	HMG-CoA synthase	LnmM (AAN85526)	62/75
Caci_Z1	83	acyl carrier protein	(WP_062008561)	48/57
Caci_Z2	428	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	41/53
Caci_Z3	241	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	54/64
Caci_orf(+1)	98	aminotransferase	(WP_051733654)	50/59

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S10. Predicted functions of ORFs in the *gnm*-like gene cluster from *Streptomyces* sp. F5630 (Clade IX)

gene[a]	aa[b]	putative function[c]	protein homologue (accession #)	% identity/ % similarity
Sf56_orf(-1)	196	putative lanthionine synthetase	SAVERM_307 (BAC68016)	59/74
Sf56_A	145	truncated hemoglobin	SAVERM_3316 (BAC71027)	36/50
Sf56_B	2390	hybrid NRPS/type I PKS	LnmI (AAN85522)	47/56
Sf56_C	252	type II thioesterase	LnmN (AAN85527)	55/65
Sf56_D	246	4'-phosphopantetheinyl transferase	Sfp (P39135)	28/49
Sf56_E	410	HMG-CoA synthase	LnmM (AAN85526)	48/63
Sf56_F	262	enoyl-CoA hydratase	LnmF (AAN85519)	48/60
Sf56_G	759	acyltransferase/oxidoreductase	LnmG (AAN85520)	47/62
Sf56_H	2052	AT-less type I PKS	LnmI (AAN85522)	50/58
Sf56_I	518	antibiotic efflux protein	LnmY (AAN85538)	30/47
Sf56_J	279	unknown	LnmE (AAN85518)	48/62
Sf56_K	765	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	63/74
Sf56_L	310	unknown	LnmH (AAN85521)	40/56
Sf56_M	121	glyoxalase/bleomycin resistance protein	(WP_062700437)	47/67
Sf56_N	135	unknown	LnmZ' (AAN85540)	35/52
Sf56_O	237	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	54/67
Sf56_P	248	methyltransferase	(WP_031183778)	37/47
Sf56_Q	665	putative alpha/beta hydrolase	LnmD (AAN85517)	35/45
Sf56_R	84	peptidyl carrier protein	LnmP (AAN85529)	47/57
Sf56_S	542	discrete adenylation protein	LnmQ (AAN85530)	55/66
Sf56_T	7148	AT-less type I PKS	LnmJ (AAN85523)	47/57
Sf56_U	411	HMG-CoA-synthase	LnmM (AAN85526)	59/76
Sf56_V	74	acyl carrier protein	Fr9M (ADH01494)	40/65
Sf56_W	433	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	43/54
Sf56_X	413	cytochrome P450 oxygenase	LnmA (AAN85514)	42/53
Sf56_Y	434	aminotransferase	(KMS73037)	40/55
Sf56_Z	229	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	44/63
Sf56_orf(+1)	297	LysR family transcriptional regulator	(WP_031032372)	77/83

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *gnm*-like gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *gnm*-like gene cluster.

Table S11. Predicted functions of ORFs in the *gnm*-like gene cluster from *S. aureofaciens* NRRL B-2183 (Clade IX)

gene[a]	aa[b]	putative function[c]	protein homologue (accession #)	% identity/ % similarity
Saur_orf(-1)	195	putative lanthionine synthetase	SAVERM_307 (BAC68016)	60/77
Saur_A	145	truncated hemoglobin	SAVERM_3316 (BAC71027)	39/50
Saur_B	2271	hybrid NRPS/type I PKS	LnmI (AAN85522)	47/57
Saur_C	252	type II thioesterase	LnmN (AAN85527)	54/63
Saur_D	245	4'-phosphopantetheinyl transferase	Sfp (P39135)	28/47
Saur_E	410	HMG-CoA synthase	LnmM (AAN85526)	48/62
Saur_F	264	enoyl-CoA hydratase	LnmF (AAN85519)	49/61
Saur_G	738	acyltransferase/oxidoreductase	LnmG (AAN85520)	46/63
Saur_H	2042	AT-less type I PKS	LnmI (AAN85522)	49/58
Saur_I	489	antibiotic efflux protein	LnmY (AAN85538)	31/49
Saur_J	264	unknown	LnmE (AAN85518)	48/63
Saur_K	767	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	56/68
Saur_L	310	unknown	LnmH (AAN85521)	40/56
Saur_M	121	glyoxalase/bleomycin resistance protein	(WP_062700437)	45/66
Saur_N	133	unknown	LnmZ' (AAN85540)	31/47
Saur_O	238	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	54/67
Saur_P	248	methyltransferase	(WP_031183778)	54/67
Saur_Q	684	putative alpha/beta hydrolase	LnmD (AAN85517)	44/56
Saur_R	84	peptidyl carrier protein	LnmP (AAN85529)	46/58
Saur_S	542	discrete adenylation protein	LnmQ (AAN85530)	54/64
Saur_T	7020	AT-less type I PKS	LnmJ (AAN85523)	47/56
Saur_U	411	HMG-CoA-synthase	LnmM (AAN85526)	59/75
Saur_V	73	acyl carrier protein	Fr9M (ADH01494)	36/62
Saur_W	433	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	42/55
Saur_X	413	cytochrome P450 oxygenase	LnmA (AAN85514)	42/54
Saur_Y	433	aminotransferase	(KMS73037)	42/55
Saur_Z	193	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	44/61
Saur_orf(+1)	367	transposase	SAVERM_309 (BAC68018)	57/65

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *gnm*-like gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *gnm*-like gene cluster.

Table S12. Predicted functions of ORFs in the *lnm*-type gene cluster from *Saccharothrix* sp. ST-888 (Clade XI)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
Sast_orf(-1)	607	glucoamylase	(WP_010983371)	81/87
Sast_A	126	glyoxalase/bleomycin resistance protein	(WP_062700437)	34/48
Sast_B	411	cytochrome P450 oxygenase	LnmA (AAN85514)	36/50
Sast_C	464	ferredoxin reductase	SCO2106 (NP_626364)	35/49
Sast_D	78	ferredoxin	LnmB (AAN85515)	32/51
Sast_E	144	truncated hemoglobin	SAVERM_3316 (BAC71027)	37/50
Sast_F	2290	hybrid NRPS/type I PKS	LnmI (AAN85522)	50/60
Sast_G	251	thioesterase type II	LnmN (AAN85527)	56/67
Sast_H	410	HMG-CoA synthase	LnmM (AAN85526)	48/63
Sast_I	260	enoyl-CoA hydratase	LnmF (AAN85519)	48/58
Sast_J	769	acyltransferase/oxidoreductase	LnmG (AAN85520)	42/55
Sast_K	2063	AT-less type I PKS	LnmI (AAN85522)	50/59
Sast_L	503	antibiotic efflux protein	LnmY (AAN85538)	33/49
Sast_M	301	unknown	LnmE (AAN85518)	45/61
Sast_N	764	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	66/78
Sast_O	306	unknown	LnmH (AAN85521)	43/59
Sast_P	375	hypothetical protein	(WP_054026420)	45/67
Sast_Q	119	glyoxalase/bleomycin resistance protein	(WP_062700437)	50/65
Sast_R	135	unknown	LnmZ' (AAN85540)	35/45
Sast_S	238	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	55/68
Sast_T	248	methyltransferase	(WP_031183778)	37/50
Sast_U	424	putative alpha/beta hydrolase	LnmD (AAN85517)	46/55
Sast_V	193	hypothetical protein	NA	NA
Sast_W	82	peptidyl carrier protein	LnmP (AAN85529)	43/59
Sast_X	531	discrete adenylation protein	LnmQ (AAN85530)	56/66
Sast_Y	7401	AT-less type I PKS	LnmJ (AAN85523)	48/58
Sast_Z1	411	HMG-CoA-synthase	LnmM (AAN85526)	63/76
Sast_Z2	82	acyl carrier protein	Fr9M (ADH01494)	31/60
Sast_Z3	412	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	45/58
Sast_Z4	419	cytochrome P450 oxygenase	LnmA (AAN85514)	43/56
Sast_Z5	64	ferredoxin	LnmB (AAN85515)	32/47
Sast_Z6	212	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	49/67
Sast_orf(+1)	179	SMI1/KNR4 family protein	(WP_030749561)	44/52

[a] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b] Number of amino acids. [c] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S13. Predicted functions of ORFs in the *lnm*-type gene cluster from *Streptomyces* sp. NBRC 110035 (Clade XII)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
S110_orf(-1)	132	transposase	(ANZ13531)	89/90
S110_A	237	Crp/Fnr/Fnr family transcriptional regulator	LnmO (AAN85528)	37/58
S110_B	189	YdeI-like protein	SAVERM_307 (BAC68016)	55/71
S110_C	244	phosphopantetheinyl transferase	(WP_059300333)	36/47
S110_D	409	HMG-CoA-synthase	LnmM (AAN85526)	44/59
S110_E	255	enoyl-CoA hydratase	LnmF (AAN85519)	47/59
S110_F	290	acyltransferase	LnmG (AAN85520)	42/54
S110_G	1993	AT-less type I PKS	LnmI (AAN85522)	47/56
S110_H	274	transposase	(WP_011029009)	79/85
S110_I	131	glyoxalase/bleomycin resistance protein	(WP_062700437)	28/46
S110_J	496	antibiotic efflux protein	LnmY (AAN85538)	32/49
S110_K	233	unknown	LnmE (AAN85518)	38/54
S110_L	776	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	60/72
S110_M	307	unknown	LnmH (AAN85521)	35/52
S110_N	120	glyoxalase/bleomycin resistance protein	(WP_062700437)	42/62
S110_O	236	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	50/62
S110_P	249	methyltransferase	(WP_031183778)	41/49
S110_Q	418	putative alpha/beta hydrolase	LnmD (AAN85517)	36/46
S110_R	81	peptidyl carrier protein	LnmP (AAN85529)	45/60
S110_S	525	discrete adenylation protein	LnmQ (AAN85530)	52/62
S110_T	7438	AT-less type I PKS	LnmJ (AAN85523)	43/54
S110_U	416	HMG-CoA-synthase	LnmM (AAN85526)	56/69
S110_V	81	acyl carrier protein	Fr9M (ADH01494)	31/53
S110_W	408	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	41/54
S110_X	248	type II thioesterase	LnmN (AAN85527)	55/64
S110_Y	2337	hybrid NRPS/type I PKS	LnmI (AAN85522)	45/54
S110_Z1	143	truncated hemoglobin	SAVERM_3316 (BAC71027)	35/48
S110_Z2	408	cytochrome P450 oxygenase	LnmA (AAN85514)	39/55
S110_Z3	74	ferredoxin	LnmB (AAN85515)	31/55
S110_Z4	453	crotonyl-CoA carboxylase/reductase	DivR (CCP20056)	63/77
S110_Z5	330	ketoacyl-acyl carrier protein synthase III (KASIII)	DivS (CCP20057)	56/77
S110_Z6	289	3-hydroxyacyl-CoA dehydrogenase	DivT (CCP20058)	48/66
S110_orf(+1)	127	transposase	(WP_069173729)	75/82

[a] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b] Number of amino acids. [c] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S14. Predicted functions of ORFs in the *lnm*-type gene cluster from *S. novaecaesareae* NRRL B-1267 (Clade XIII)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
Snov_orf(-1)	126	glyoxalase/bleomycin resistance protein	(WP_062700437)	37/54
Snov_A	408	HMG-CoA-synthase	LnmM (AAN85526)	44/62
Snov_B	278	enoyl-CoA hydratase	LnmF (AAN85519)	45/59
Snov_C	764	acyltransferase/oxidoreductase	LnmG (AAN85520)	63/79
Snov_D	281	unknown	LnmH (AAN85521)	41/58
Snov_E	193	YdeI-like protein	SAVERM_307 (BAC68016)	61/74
Snov_F	150	truncated hemoglobin	SAVERM_3316 (BAC71027)	43/59
Snov_G	404	cytochrome P450 oxygenase	LnmA (AAN85514)	46/61
Snov_H	402	cytochrome P450 oxygenase	LnmA (AAN85514)	39/55
Snov_I	70	ferredoxin	LnmB (AAN85515)	39/50
Snov_J	223	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	41/59
Snov_K	542	putative alpha/beta hydrolase	LnmD (AAN85517)	40/52
Snov_L	83	peptidyl carrier protein	LnmP (AAN85529)	37/50
Snov_M	406	diaminopimelate decarboxylase	(WP_064072918)	54/62
Snov_N	495	discrete adenylation protein	LnmQ (AAN85530)	28/47
Snov_O	341	hypothetical protein	(WP_012392161)	46/61
Snov_P	376	hypothetical protein	(WP_040258597)	58/71
Snov_Q	340	hypothetical protein	(WP_052482831)	49/64
Snov_R	564	acyl-CoA synthetase	(WP_052482829)	45/59
Snov_S	82	ACP-like PKS domain	LnmJ (AAN85523)	38/54
Snov_T	973	KS-AT-ACP	(WP_052482825)	52/62
Snov_U	798	KS-ACP	(WP_052482823)	48/58
Snov_V	803	putative class-III aminotransferase	(AJE87366)	54/68
Snov_W	594	oxidoreductase	(AJE87364)	55/74
Snov_X	2844	AT-less type I PKS	LnmI (AAN85522)	48/57
Snov_Y	7724	AT-less type I PKS	LnmJ (AAN85523)	45/56
Snov_Z1	408	HMG-CoA synthase	LnmM (AAN85526)	55/70
Snov_Z2	82	acyl carrier protein	Fr9M (ADH01494)	29/56
Snov_Z3	405	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	41/54
Snov_Z4	258	type II thioesterase	LnmN (AAN85527)	48/60
Snov_Z5	266	ABC transporter	(WP_069169226)	49/64
Snov_Z6	314	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	33/48
Snov_Z7	152	RDD family protein, transporter	WP_011730428	35/49
Snov_Z8	451	sodium:proton exchanger	(WP_020553490)	28/47
Snov_Z9	263	methyltransferase	(WP_051803680)	46/63
Snov_orf(+1)	280	transglutaminase	(WP_052145777)	64/74

[a] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b] Number of amino acids. [c] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S15. Predicted functions of ORFs in the *lnm*-type gene cluster from *Saccharothrix espanaensis* DSM 44229 (Clade XIV)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
Saes_orf(-1)	314	RNA polymerase sigma24 factor	(WP_009948905)	78/86
Saes_A	136	unknown	LnmZ' (AAN85540)	35/56
Saes_B	121	glyoxalase/bleomycin resistance protein	(WP_062700437)	37/53
Saes_C	270	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	42/61
Saes_D	500	4-coumarate:CoA ligase	LnmW (AAN85536)	40/56
Saes_E	183	hypothetical protein	(WP_053707566)	48/60
Saes_F	741	acyltransferase/oxidoreductase	LnmG (AAN85520)	49/60
Saes_G	254	enoyl-CoA hydratase	LnmF (AAN85519)	51/62
Saes_H	412	HMG-CoA synthase	LnmM (AAN85526)	46/62
Saes_I	264	unknown	LnmE (AAN85518)	43/59
Saes_J	1999	hybrid NRPS/type I PKS	LnmI (AAN85522)	49/57
Saes_K	6706	AT-less type I PKS	LnmJ (AAN85523)	46/57
Saes_L	312	acyltransferase/decarboxylase	LnmK (AAN85524)	51/63
Saes_M	92	acyl carrier protein	LnmL (AAN85525)	50/66
Saes_N	412	HMG-CoA synthase	LnmM (AAN85526)	63/77
Saes_O	406	alpha/beta hydrolase	ADZ36_03165 (KNE83935)	45/57
Saes_P	416	hypothetical protein	(WP_067304298)	43/59
Saes_Q	267	ABC transporter	(WP_069169226)	37/59
Saes_R	333	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	33/43
Saes_S	412	cytochrome P450 oxygenase	LnmA (AAN85514)	40/55
Saes_T	499	putative class-III aminotransferase	(AJE87366)	33/47
Saes_U	202	thioesterase type II	LnmN (AAN85527)	53/61
Saes_V	2223	AT-less type I PKS	LnmI (AAN85522)	50/61
Saes_W	784	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	60/72
Saes_X	313	unknown	LnmH (AAN85521)	43/59
Saes_Y	600	NRPS (A-PCP)	SyrB1 (AKF46132)	45/59
Saes_Z1	318	halogenase	SyrB2 (AKF46131)	67/78
Saes_Z2	76	acyl carrier protein	(WP_062008561)	53/70
Saes_Z3	409	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	41/54
Saes_Z4	404	FAD-dependent oxidoreductase	(WP_010982689)	58/69
Saes_Z5	520	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	51/65
Saes_Z6	315	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	58/70
Saes_Z7	263	ABC transporter component, membrane spanning protein	LnmS (AAN85532)	60/70
Saes_Z8	516	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	57/67
Saes_Z9	246	methyltransferase	(WP_026218883)	46/62
Saes_Z10	640	Radical SAM superfamily protein	(SED67026)	54/68
Saes_orf(+1)	271	lipase	(WP_040253406)	68/79

[a] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b] Number of amino acids. [c] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S16. Predicted functions of ORFs in the *lnm*-type gene cluster from *Micromonospora* sp. CNB394 (Clade XV)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
Mcnb_orf(-1)	210	peptidase 23	(SCL52691)	82/88
Mcnb_A	494	4-coumarate:CoA ligase	LnmW (AAN85536)	41/57
Mcnb_B	413	HMG-CoA-synthase	LnmM (AAN85526)	49/63
Mcnb_C	179	hypothetical protein	(WP_053707566)	42/57
Mcnb_D	426	sodium:proton exchanger	(WP_020553490)	33/51
Mcnb_E	253	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	46/59
Mcnb_F	744	acyltransferase/oxidoreductase	LnmG (AAN85520)	45/57
Mcnb_G	252	enoyl-CoA hydratase	LnmF (AAN85519)	52/63
Mcnb_H	292	unknown	LnmE (AAN85518)	43/60
Mcnb_I	1960	hybrid NRPS/type I PKS	LnmI (AAN85522)	50/58
Mcnb_J	6709	AT-less type I PKS	LnmJ (AAN85523)	46/57
Mcnb_K	310	acyltransferase/decarboxylase	LnmK (AAN85524)	55/67
Mcnb_L	80	acyl carrier protein	LnmL (AAN85525)	62/75
Mcnb_M	411	HMG-CoA synthase	LnmM (AAN85526)	61/75
Mcnb_N	397	putative alpha/beta hydrolase	LnmD (AAN85517)	27/36
Mcnb_O	418	hypothetical protein	(WP_067304298)	41/59
Mcnb_P	266	ABC transporter	(WP_069169226)	38/59
Mcnb_Q	328	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	33/45
Mcnb_R	412	cytochrome P450 oxygenase	LnmA (AAN85514)	42/58
Mcnb_S	453	putative class-III aminotransferase	(AJE87366)	33/49
Mcnb_T	245	thioesterase type II	LnmN (AAN85527)	56/63
Mcnb_U	2161	AT-less type I PKS	LnmI (AAN85522)	48/58
Mcnb_V	752	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	60/74
Mcnb_W	313	unknown	LnmH (AAN85521)	44/59
Mcnb_X	629	NRPS (A-PCP)	SyrB1 (AKF46132)	43/58
Mcnb_Y	321	halogenase	SyrB2 (AKF46131)	66/79
Mcnb_Z1	76	acyl carrier protein	LnmL (AAN85525)	37/63
Mcnb_Z2	412	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	42/56
Mcnb_Z3	394	FAD-dependent oxidoreductase	(WP_010982689)	59/70
Mcnb_Z4	518	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	53/65
Mcnb_Z5	318	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	60/73
Mcnb_Z6	263	ABC transporter component, membrane spanning protein	LnmS (AAN85532)	60/71
Mcnb_Z7	518	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	57/69
Mcnb_Z8	136	unknown	LnmZ' (AAN85540)	34/55
Mcnb_Z9	121	glyoxalase/bleomycin resistance protein	(WP_062700437)	37/53
Mcnb_Z10	245	methyltransferase	(WP_026218883)	44/61
Mcnb_Z11	644	Radical SAM superfamily protein	(SED67026)	52/66
Mcnb_orf(+1)	1028	transcriptional regulator	(WP_073892172)	60/68

[a] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b] Number of amino acids. [c] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S17. Predicted functions of ORFs in the *lnm*-type gene cluster from *Micromonospora marina* DSM 45555 (Clade XV)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
Mmar_orf(-1)	214	peptidase 23	(SCL52691)	82/88
Mmar_A	494	4-coumarate:CoA ligase	LnmW (AAN85536)	41/57
Mmar_B	413	HMG-CoA-synthase	LnmM (AAN85526)	49/63
Mmar_C	161	hypothetical protein	(WP_053707566)	41/55
Mmar_D	426	sodium:proton exchanger	(WP_020553490)	33/51
Mmar_E	253	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	46/59
Mmar_F	744	acyltransferase/oxidoreductase	LnmG (AAN85520)	38/51
Mmar_G	252	enoyl-CoA hydratase	LnmF (AAN85519)	52/63
Mmar_H	292	unknown	LnmE (AAN85518)	43/60
Mmar_I	1954	hybrid NRPS/type I PKS	LnmI (AAN85522)	49/57
Mmar_J	6700	AT-less type I PKS	LnmJ (AAN85523)	46/57
Mmar_K	313	acyltransferase/decarboxylase	LnmK (AAN85524)	54/67
Mmar_L	82	acyl carrier protein	LnmL (AAN85525)	62/75
Mmar_M	411	HMG-CoA synthase	LnmM (AAN85526)	61/75
Mmar_N	397	putative alpha/beta hydrolase	LnmD (AAN85517)	27/36
Mmar_O	418	hypothetical protein	(WP_067304298)	41/59
Mmar_P	266	ABC transporter	(WP_069169226)	38/59
Mmar_Q	328	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	33/46
Mmar_R	412	cytochrome P450 oxygenase	LnmA (AAN85514)	42/58
Mmar_S	453	putative class-III aminotransferase	(AJE87366)	33/49
Mmar_T	245	thioesterase type II	LnmN (AAN85527)	55/63
Mmar_U	2165	AT-less type I PKS	LnmI (AAN85522)	48/58
Mmar_V	749	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	59/74
Mmar_W	313	unknown	LnmH (AAN85521)	44/59
Mmar_X	635	NRPS (A-PCP)	SyrB1 (AKF46132)	42/58
Mmar_Y	321	halogenase	SyrB2 (AKF46131)	66/79
Mmar_Z1	76	acyl carrier protein	LnmL (AAN85525)	37/63
Mmar_Z2	412	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	41/56
Mmar_Z3	394	FAD-dependent oxidoreductase	(WP_010982689)	59/70
Mmar_Z4	519	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	53/66
Mmar_Z5	318	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	60/73
Mmar_Z6	263	ABC transporter component, membrane spanning protein	LnmS (AAN85532)	60/71
Mmar_Z7	518	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	57/69
Mmar_Z8	136	unknown	LnmZ' (AAN85540)	34/55
Mmar_Z9	121	glyoxalase/bleomycin resistance protein	(WP_062700437)	37/53
Mmar_Z10	253	methyltransferase	(WP_026218883)	44/61
Mmar_Z11	644	Radical SAM superfamily protein	(SED67026)	52/66
Mmar_orf(+1)	1028	transcriptional regulator	(WP_073892172)	60/68

[a] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b] Number of amino acids. [c] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S18. Predicted functions of ORFs in the *lnm*-type gene cluster from *Micromonospora* sp. L5 (Clade XV)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
Ml5_orf(-1)	213	peptidase 23	(SCL52691)	83/90
Ml5_A	494	4-coumarate:CoA ligase	LnmW (AAN85536)	42/57
Ml5_B	413	HMG-CoA-synthase	LnmM (AAN85526)	49/63
Ml5_C	179	hypothetical protein	(WP_053707566)	41/58
Ml5_D	426	sodium:proton exchanger	(WP_020553490)	33/50
Ml5_E	253	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	46/59
Ml5_F	717	acyltransferase/oxidoreductase	LnmG (AAN85520)	38/50
Ml5_G	252	enoyl-CoA hydratase	LnmF (AAN85519)	52/63
Ml5_H	292	unknown	LnmE (AAN85518)	43/60
Ml5_I	2006	hybrid NRPS/type I PKS	LnmI (AAN85522)	65/74
Ml5_J	6765	AT-less type I PKS	LnmJ (AAN85523)	45/57
Ml5_K	311	acyltransferase/decarboxylase	LnmK (AAN85524)	54/66
Ml5_L	82	acyl carrier protein	LnmL (AAN85525)	59/71
Ml5_M	411	HMG-CoA synthase	LnmM (AAN85526)	61/75
Ml5_N	397	putative alpha/beta hydrolase	LnmD (AAN85517)	27/37
Ml5_O	418	hypothetical protein	(WP_067304298)	43/61
Ml5_P	266	ABC transporter	(WP_069169226)	38/59
Ml5_Q	327	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	33/45
Ml5_R	412	cytochrome P450 oxygenase	LnmA (AAN85514)	41/58
Ml5_S	453	putative class-III aminotransferase	(AJE87366)	32/48
Ml5_T	246	thioesterase type II	LnmN (AAN85527)	55/63
Ml5_U	2174	AT-less type I PKS	LnmI (AAN85522)	48/58
Ml5_V	767	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	59/73
Ml5_W	313	unknown	LnmH (AAN85521)	44/58
Ml5_X	621	NRPS (A-PCP)	SyrB1 (AKF46132)	43/59
Ml5_Y	321	halogenase	SyrB2 (AKF46131)	66/79
Ml5_Z1	76	acyl carrier protein	LnmL (AAN85525)	37/63
Ml5_Z2	412	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	42/56
Ml5_Z3	394	FAD-dependent oxidoreductase	(WP_010982689)	58/70
Ml5_Z4	519	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	53/65
Ml5_Z5	318	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	59/72
Ml5_Z6	263	ABC transporter component, membrane spanning protein	LnmS (AAN85532)	60/71
Ml5_Z7	518	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	57/69
Ml5_Z8	136	unknown	LnmZ' (AAN85540)	34/55
Ml5_Z9	121	glyoxalase/bleomycin resistance protein	(WP_062700437)	37/53
Ml5_Z10	245	methyltransferase	(WP_026218883)	44/61
Ml5_Z11	644	Radical SAM superfamily protein	(SED67026)	52/66
Ml5_orf(+1)	1030	transcriptional regulator	(WP_073892172)	60/68

[a] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b] Number of amino acids. [c] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S19. Predicted functions of ORFs in the *lnm*-type gene cluster from *Micromonospora globosa* NRRL B-2673 (Clade XV)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
Mglo_orf(-1)	213	peptidase 23	(SCL52691)	83/90
Mglo_A	494	4-coumarate:CoA ligase	LnmW (AAN85536)	42/57
Mglo_B	413	HMG-CoA-synthase	LnmM (AAN85526)	49/63
Mglo_C	179	hypothetical protein	(WP_053707566)	41/58
Mglo_D	426	sodium:proton exchanger	(WP_020553490)	33/50
Mglo_E	253	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	46/59
Mglo_F	735	acyltransferase/oxidoreductase	LnmG (AAN85520)	38/50
Mglo_G	252	enoyl-CoA hydratase	LnmF (AAN85519)	52/63
Mglo_H	292	unknown	LnmE (AAN85518)	43/60
Mglo_I	1986	hybrid NRPS/type I PKS	LnmI (AAN85522)	65/73
Mglo_J	6771	AT-less type I PKS	LnmJ (AAN85523)	46/55
Mglo_K	308	acyltransferase/decarboxylase	LnmK (AAN85524)	55/67
Mglo_L	82	acyl carrier protein	LnmL (AAN85525)	59/71
Mglo_M	411	HMG-CoA synthase	LnmM (AAN85526)	61/75
Mglo_N	397	putative alpha/beta hydrolase	LnmD (AAN85517)	27/36
Mglo_O	418	hypothetical protein	(WP_067304298)	43/61
Mglo_P	266	ABC transporter	(WP_069169226)	38/59
Mglo_Q	328	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	33/45
Mglo_R	412	cytochrome P450 oxygenase	LnmA (AAN85514)	41/59
Mglo_S	453	putative class-III aminotransferase	(AJE87366)	32/48
Mglo_T	246	thioesterase type II	LnmN (AAN85527)	55/63
Mglo_U	2186	AT-less type I PKS	LnmI (AAN85522)	48/59
Mglo_V	767	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	59/73
Mglo_W	313	unknown	LnmH (AAN85521)	44/58
Mglo_X	614	NRPS (A-PCP)	SyrB1 (AKF46132)	43/60
Mglo_Y	321	halogenase	SyrB2 (AKF46131)	66/79
Mglo_Z1	76	acyl carrier protein	LnmL (AAN85525)	37/63
Mglo_Z2	412	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	42/56
Mglo_Z3	394	FAD-dependent oxidoreductase	(WP_010982689)	58/70
Mglo_Z4	519	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	53/65
Mglo_Z5	318	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	59/72
Mglo_Z6	263	ABC transporter component, membrane spanning protein	LnmS (AAN85532)	60/71
Mglo_Z7	518	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	57/69
Mglo_Z8	136	unknown	LnmZ' (AAN85540)	34/55
Mglo_Z9	121	glyoxalase/bleomycin resistance protein	(WP_062700437)	37/53
Mglo_Z10	245	methyltransferase	(WP_026218883)	44/61
Mglo_Z11	644	Radical SAM superfamily protein	(SED67026)	52/66
Mglo_orf(+1)	1030	transcriptional regulator	(WP_073892172)	60/68

[a] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b] Number of amino acids. [c] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S20. Predicted functions of ORFs in the *lnm*-type gene cluster from *Micromonospora aurantiaca* ATCC 27029 (Clade XV)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
Maur_orf(-1)	213	peptidase 23	(SCL52691)	83/89
Maur_A	494	4-coumarate:CoA ligase	LnmW (AAN85536)	42/57
Maur_B	413	HMG-CoA-synthase	LnmM (AAN85526)	49/63
Maur_C	179	hypothetical protein	(WP_053707566)	41/58
Maur_D	426	sodium:proton exchanger	(WP_020553490)	33/50
Maur_E	253	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	46/59
Maur_F	735	acyltransferase/oxidoreductase	LnmG (AAN85520)	38/50
Maur_G	252	enoyl-CoA hydratase	LnmF (AAN85519)	52/63
Maur_H	292	unknown	LnmE (AAN85518)	43/60
Maur_I	2005	hybrid NRPS/type I PKS	LnmI (AAN85522)	65/73
Maur_J	6727	AT-less type I PKS	LnmJ (AAN85523)	46/56
Maur_K	311	acyltransferase/decarboxylase	LnmK (AAN85524)	55/67
Maur_L	82	acyl carrier protein	LnmL (AAN85525)	59/71
Maur_M	411	HMG-CoA synthase	LnmM (AAN85526)	61/75
Maur_N	397	putative alpha/beta hydrolase	LnmD (AAN85517)	27/36
Maur_O	418	hypothetical protein	(WP_067304298)	43/60
Maur_P	266	ABC transporter	(WP_069169226)	38/59
Maur_Q	328	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	33/45
Maur_R	412	cytochrome P450 oxygenase	LnmA (AAN85514)	41/59
Maur_S	453	putative class-III aminotransferase	(AJE87366)	32/48
Maur_T	246	thioesterase type II	LnmN (AAN85527)	55/63
Maur_U	2174	AT-less type I PKS	LnmI (AAN85522)	48/58
Maur_V	767	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	59/73
Maur_W	313	unknown	LnmH (AAN85521)	44/58
Maur_X	621	NRPS (A-PCP)	SyrB1 (AKF46132)	43/59
Maur_Y	321	halogenase	SyrB2 (AKF46131)	67/80
Maur_Z1	76	acyl carrier protein	LnmL (AAN85525)	37/63
Maur_Z2	412	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	42/56
Maur_Z3	394	FAD-dependent oxidoreductase	(WP_010982689)	58/70
Maur_Z4	519	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	53/65
Maur_Z5	318	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	59/72
Maur_Z6	263	ABC transporter component, membrane spanning protein	LnmS (AAN85532)	60/71
Maur_Z7	518	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	57/69
Maur_Z8	136	unknown	LnmZ' (AAN85540)	34/55
Maur_Z9	121	glyoxalase/bleomycin resistance protein	(WP_062700437)	37/53
Maur_Z10	245	methyltransferase	(WP_026218883)	44/61
Maur_Z11	644	Radical SAM superfamily protein	(SED67026)	52/66
Maur_orf(+1)	1030	transcriptional regulator	(WP_073892172)	60/68

[a] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b] Number of amino acids. [c] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S21. Predicted functions of ORFs in the *lnm*-type gene cluster from *Salinispora arenicola* CNH964 (Clade XVII)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
Sal964_orf(-1)	101	hypothetical protein	(WP_040795638)	66/79
Sal964_A	513	4-coumarate:CoA ligase	LnmW (AAN85536)	43/59
Sal964_B	278	hypothetical protein	(AGP55856)	38/52
Sal964_C	468	serine hydroxymethyltransferase	AmiS (AEF16058)	61/73
Sal964_D	802	NRPS (A-PCP)	SyrB1 (AKF46132)	35/52
Sal964_E	236	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	52/67
Sal964_F	504	4-coumarate:CoA ligase	LnmW (AAN85536)	38/52
Sal964_G	431	hypothetical protein	(WP_067304298)	46/64
Sal964_H	308	unknown	LnmH (AAN85521)	44/59
Sal964_I	405	FAD-dependent oxidoreductase	(WP_010982689)	61/70
Sal964_J	120	glyoxalase/bleomycin resistance protein	(WP_062700437)	44/60
Sal964_K	133	unknown	LnmZ' (AAN85540)	39/56
Sal964_L	291	phosphopantetheinyl transferase	(WP_059300333)	41/50
Sal964_M	287	unknown	LnmE (AAN85518)	37/52
Sal964_N	410	HMG-CoA-synthase	LnmM (AAN85526)	48/64
Sal964_O	262	enoyl-CoA hydratase	LnmF (AAN85519)	51/64
Sal964_P	783	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	63/75
Sal964_Q	4311	hybrid NRPS/AT-less PKS	LnmI (AAN85522)	49/59
Sal964_R	336	taurine dioxygenase-like protein	(WP_060732555)	55/65
Sal964_S	311	kinase	(WP_040836916)	66/77
Sal964_T	1086	NRPS (A-PCP)	SyrB1 (AKF46132)	31/45
Sal964_U	81	peptidyl carrier protein	SyrB1 (AKF46132)	30/56
Sal964_V	317	halogenase	SyrB2 (AKF46131)	60/74
Sal964_W	381	putative alpha/beta hydrolase	LnmD (AAN85517)	26/37
Sal964_X	7105	AT-less type I PKS	LnmJ (AAN85523)	47/59
Sal964_Y	320	acyltransferase/decarboxylase	LnmK (AAN85524)	56/67
Sal964_Z1	90	acyl carrier protein	LnmL (AAN85525)	59/69
Sal964_Z2	407	HMG-CoA-synthase	LnmM (AAN85526)	64/76
Sal964_Z3	86	acyl carrier protein	Fr9M (ADH01494)	32/61
Sal964_Z4	433	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	41/55
Sal964_Z5	270	thioesterase type II	LnmN (AAN85527)	56/69
Sal964_Z6	511	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	58/74
Sal964_Z7	318	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	55/72
Sal964_Z8	263	ABC transporter component, membrane spanning protein	LnmS (AAN85532)	62/72
Sal964_Z9	547	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	57/69
Sal964_Z10	297	unknown	LnmE (AAN85518)	44/59
Sal964_Z11	499	4-coumarate:CoA ligase	LnmW (AAN85536)	41/54
Sal964_Z12	407	diaminobutyric acid synthase C	DabC (AEG64688)	53/64
Sal964_Z13	485	diaminobutyric acid synthase B	DabB (AEG64687)	56/67
Sal964_Z14	354	diaminobutyric acid synthase A	DabA (AEG64686)	62/75

Sal964_orf(+1)	349	LuxR family transcriptional regulator	(WP_063766468)	37/56

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S22. Predicted functions of ORFs in the *lnm*-type gene cluster from *Burkholderia ubonensis* RF25-BP1 (Clade XVIII)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
Bubo_orf(-1)	503	outer membrane efflux protein	OprM (WP_057405643)	49/64
Bubo_A	330	unknown	LnmH (AAN85521)	43/60
Bubo_B	396	FAD-dependent oxidoreductase	(WP_010982689)	54/69
Bubo_C	510	4-coumarate:CoA ligase	LnmW (AAN85536)	34/51
Bubo_D	130	unknown	LnmZ' (AAN85540)	36/53
Bubo_E	298	unknown	LnmE (AAN85518)	36/55
Bubo_F	176	hypothetical protein	(WP_053707566)	39/52
Bubo_G	129	transposase	WQE_51432 (EIM93082)	56/77
Bubo_H	134	hypothetical protein	(WP_060317165)	95/96
Bubo_I	267	4'-phosphopantetheinyl transferase	Sfp (P39135)	25/40
Bubo_J	410	HMG-CoA synthase	LnmM (AAN85526)	45/65
Bubo_K	268	enoyl-CoA hydratase	LnmF (AAN85519)	44/57
Bubo_L	295	acyltransferase/oxidoreductase	LnmG (AAN85520)	50/63
Bubo_M	769	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	53/66
Bubo_N	609	adenylation protein-peptidyl carrier protein	SyrB1 (AKF46132)	41/57
Bubo_O	313	halogenase	SyrB2 (AKF46131)	66/79
Bubo_P	365	alpha/beta hydrolase	ADZ36_03165 (KNE83935)	36/48
Bubo_Q	4169	AT-less type I PKS	LnmI (AAN85522)	39/49
Bubo_R	6941	AT-less type I PKS	LnmJ (AAN85523)	42/54
Bubo_S	318	acyltransferase/decarboxylase	LnmK (AAN85524)	44/58
Bubo_T	89	acyl carrier protein	LnmL (AAN85525)	49/67
Bubo_U	409	HMG-CoA synthase	LnmM (AAN85526)	58/72
Bubo_V	88	acyl carrier protein	LnmL (AAN85525)	29/57
Bubo_W	419	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	38/52
Bubo_X	419	type II thioesterase	LnmN (AAN85527)	50/65
Bubo_orf(+1)	712	methylmalonyl-CoA mutase	(WP_060136708)	97/98

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S23. Predicted functions of ORFs in the *lnm*-type gene cluster from *Streptomyces* sp. CB01635 (Clade I) in comparison with the knonw LNM producer *S. atroolivaceus* S-140

gene[a]	aa[b]	putative function[c]	protein homologue (accession #)	% identity/ % similarity
CB01635_orf(-1)	462	NRPS	Orf(-1) (AAN85513)	96/99
CB01635_A	399	cytochrome P450 oxygenase	LnmA (AAN85514)	99/99
CB01635_B	78	ferredoxin	LnmB (AAN85515)	99/100
CB01635_C	115	unknown	LnmC (AAN85516)	100/100
CB01635_D	438	putative alpha/beta hydrolase	LnmD (AAN85517)	97/98
CB01635_E	305	unknown	LnmE (AAN85518)	99/99
CB01635_F	265	enoyl-CoA hydratase	LnmF (AAN85519)	99/99
CB01635_G	795	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	99/99
CB01635_H	274	unknown	LnmH (AAN85521)	100/100
CB01635_I	4447	hybrid NRPS/AT-less PKS	LnmI (AAN85522)	98/98
CB01635_J	7351	AT-less PKS	LnmJ (AAN85523)	97/97
CB01635_K	319	acyltransferase/decarboxylase	LnmK (AAN85524)	99/99
CB01635_L	86	acyl carrier protein	LnmL (AAN85525)	99/100
CB01635_M	416	HMG-CoA-synthase	LnmM (AAN85526)	99/99
CB01635_N	267	type II thioesterase	LnmN (AAN85527)	99/99
CB01635_O	227	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	100/100
CB01635_P	82	peptidyl carrier protein	LnmP (AAN85529)	99/99
CB01635_Q	516	discrete adenylation protein	LnmQ (AAN85530)	99/99
CB01635_R	575	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	98/98
CB01635_S	291	ABC transporter component, membrane spanning protein	LnmS (AAN85532)	98/98
CB01635_T	321	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	99/99
CB01635_U	513	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	99/99
CB01635_V	120	unknown, NTF2 family protein	LnmV (AAN85535)	98/100
CB01635_W	516	4-coumarate:CoA ligase	LnmW (AAN85536)	99/99
CB01635_X	243	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	99/100
CB01635_Y	474	antibiotic efflux protein	LnmY (AAN85538)	99/99
CB01635_Z1	400	cytochrome P450 oxygenase	LnmZ (AAN85539)	99/100
CB01635_Z2	134	unknown	LnmZ' (AAN85540)	99/99
CB01635_orf(+1)	216	unknown	Orf(+1) (AAN85541)	97/97

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S24. Predicted functions of ORFs in the *lnm*-type gene cluster from *Streptomyces* sp. CB02959 (Clade II)

gene[a]	aa[b]	putative function[c]	protein homologue (accession #)	% identity/ % similarity
CB02959_orf(-1)	174	hypothetical protein	(WP_018488465)	84/89
CB02959_A	402	cytochrome P450 oxygenase	LnmA (AAN85514)	49/64
CB02959_B	247	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	54/66
CB02959_C	446	putative alpha/beta hydrolase	LnmD (AAN85517)	62/68
CB02959_D	321	unknown	LnmE (AAN85518)	74/81
CB02959_E	271	enoyl-CoA hydratase	LnmF (AAN85519)	82/88
CB02959_F	762	acyltransferase/oxidoreductase	LnmG (AAN85520)	69/76
CB02959_G	280	unknown	LnmH (AAN85521)	83/89
CB02959_H	4391	AT-less type I PKS	LnmI (AAN85522)	68/74
CB02959_I	7490	AT-less type I PKS	LnmJ (AAN85523)	68/75
CB02959_J	328	acyltransferase/decarboxylase	LnmK (AAN85524)	74/83
CB02959_K	86	acyl carrier protein	LnmL (AAN85525)	76/81
CB02959_L	419	HMG-CoA-synthase	LnmM (AAN85526)	83/88
CB02959_M	264	thioesterase type II	LnmN (AAN85527)	70/79
CB02959_N	228	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	76/85
CB02959_O	377	glucosyltransferase	Cpz31 (ACQ63639)	45/59
CB02959_P	87	peptidyl carrier protein	LnmP (AAN85529)	61/68
CB02959_Q	517	discrete adenylation protein	LnmQ (AAN85530)	76/81
CB02959_R	586	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	69/76
CB02959_S	289	ABC transporter component, membrane spanning protein	LnmS (AAN85532)	85/90
CB02959_T	331	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	81/87
CB02959_U	514	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	75/84
CB02959_V	305	glucose-1-phosphate thymidylyltransferase	MydA (BAC57039)	75/82
CB02959_W	323	mycothiol conjugate amidase Mca	(WP_003974009)	56/70
CB02959_X	634	methoxymalonyl-ACP biosynthesis protein FkbH	ChlD1 (AAZ77703)	56/63
CB02959_Y	524	oxygenase	OxyS (AAZ78342)	43/52
CB02959_orf(+1)	209	type III effector protein	SVEN_3170 (CCA56456)	85/89

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S25. Predicted functions of ORFs in the *lnm*-type gene cluster from *Streptomyces* sp. CB02891 (Clade III)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
CB02891_orf(-1)	220	peptidase S51	ABB07_38760 (AKJ15883)	70/81
CB02891_A	304	unknown	LnmE (AAN85518)	65/74
CB02891_B	514	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	66/75
CB02891_C	324	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	67/77
CB02891_D	272	ABC transporter component, membrane spanning protein	LnmS (AAN85532)	72/85
CB02891_E	543	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	65/73
CB02891_F	528	discrete adenylation protein	LnmQ (AAN85530)	63/71
CB02891_G	91	peptidyl carrier protein	LnmP (AAN85529)	55/70
CB02891_H	228	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	60/73
CB02891_I	76	ferredoxin	LnmB (AAN85515)	71/80
CB02891_J	399	cytochrome P450 oxygenase	LnmA (AAN85514)	82/90
CB02891_K	449	putative alpha/beta hydrolase	LnmD (AAN85517)	53/63
CB02891_L	277	enoyl-CoA hydratase	LnmF (AAN85519)	61/73
CB02891_M	757	acyltransferase/oxidoreductase	LnmG (AAN85520)	67/77
CB02891_N	290	unknown	LnmH (AAN85521)	73/83
CB02891_O	4535	AT-less type I PKS	LnmI (AAN85522)	62/70
CB02891_P	7244	AT-less type I PKS	LnmJ (AAN85523)	61/69
CB02891_Q	328	acyltransferase/decarboxylase	LnmK (AAN85524)	67/77
CB02891_R	87	acyl carrier protein	LnmL (AAN85525)	65/78
CB02891_S	418	HMG-CoA synthase	LnmM (AAN85526)	77/85
CB02891_T	251	thioesterase type II	LnmN (AAN85527)	62/71
CB02891_U	266	4'-phosphopantetheinyl transferase	Sfp (P39135)	25/43
CB02891_V	136	unknown	LnmZ' (AAN85540)	82/85
CB02891_W	400	cytochrome P450 oxygenase	LnmZ (AAN85539)	75/86
CB02891_X	130	hypothetical protein	(WP_052398215)	36/53
CB02891_Y	475	antibiotic efflux protein	LnmY (AAN85538)	77/87
CB02891_Z1	242	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	79/88
CB02891_Z2	512	4-coumarate:CoA ligase	LnmW (AAN85536)	74/80
CB02891_Z3	120	unknown, NTF2 family protein	LnmV (AAN85535)	83/93
CB02891_orf(+1)	191	hypothetical protein	SAZU_6586 (GAP51713)	65/73

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S26. Predicted functions of ORFs in the *lnm*-type gene cluster from *Streptomyces* sp. TSRI0384-2 (Clade V)

gene[a]	aa[b]	putative function[c]	protein homologue (accession #)	% identity/ % similarity
TSRI0384-2_orf(-1)	229	TetR family transcriptional regulator	(WP_062008487)	77/83
TSRI0384-2_A	572	4-coumarate:CoA ligase	LnmW (AAN85536)	41/56
TSRI0384-2_B	252	methyltransferase	(WP_031183778)	34/45
TSRI0384-2_C	292	unknown	LnmE (AAN85518)	47/61
TSRI0384-2_D	411	HMG-CoA-synthase	LnmM (AAN85526)	47/63
TSRI0384-2_E	268	enoyl-CoA hydratase	LnmF (AAN85519)	49/62
TSRI0384-2_F	764	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	65/78
TSRI0384-2_G	303	unknown	LnmH (AAN85521)	43/58
TSRI0384-2_H	623	adenylation protein-peptidyl carrier protein	SyrB1 (AKF46132)	45/58
TSRI0384-2_I	316	halogenase	SyrB2 (AKF46131)	65/76
TSRI0384-2_J	397	alpha/beta hydrolase	ADZ36_03165 (KNE83935)	42/53
TSRI0384-2_K	4202	AT-less type I PKS	LnmI (AAN85522)	48/58
TSRI0384-2_L	6510	AT-less type I PKS	LnmJ (AAN85523)	51/61
TSRI0384-2_M	327	acyltransferase/decarboxylase	LnmK (AAN85524)	50/64
TSRI0384-2_N	86	acyl carrier protein	LnmL (AAN85525)	51/67
TSRI0384-2_O	410	HMG-CoA synthase	LnmM (AAN85526)	64/76
TSRI0384-2_P	82	acyl carrier protein	(WP_062008561)	84/93
TSRI0384-2_Q	425	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	44/57
TSRI0384-2_R	249	type II thioesterase	LnmN (AAN85527)	52/63
TSRI0384-2_S	139	unknown, NTF2 family protein	LnmV (AAN85535)	44/61
TSRI0384-2_T	133	unknown	LnmZ' (AAN85540)	37/54
TSRI0384-2_U	397	cytochrome P450 oxygenase	LnmA (AAN85514)	42/57
TSRI0384-2_V	410	cytochrome P450 oxygenase	LnmA (AAN85514)	49/66
TSRI0384-2_W	125	glyoxalase/bleomycin resistance protein	BN971_01690 (CPR09988)	43/55
TSRI0384-2_X	238	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	55/65
TSRI0384-2_Y	528	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	33/47
TSRI0384-2_Z1	328	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	44/60
TSRI0384-2_Z2	811	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	43/55
TSRI0384-2_Z3	201	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	47/68
TSRI0384-2_orf(+1)	209	dTDP-4-dehydrorhamnose 3,5-epimerase	NivL (AGZ78381)	69/80

[a] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b] Number of amino acids. [c] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S27. Predicted functions of ORFs in the *lnm*-type gene cluster from *Streptomyces* sp. CB01373 (Clade VII)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
CB01373_orf(-1)	140	hypothetical protein	(WP_059300313)	63/72
CB01373_A	486	antibiotic efflux protein	LnmY (AAN85538)	41/58
CB01373_B	205	TetR family transcriptional regulator	(WP_058852673)	42/63
CB01373_C	75	hypothetical protein	(WP_062204104)	45/52
CB01373_D	191	YdeI-like protein	SAVERM_307 (BAC68016)	71/78
CB01373_E	454	cytochrome P450 oxygenase	VR46_04425 (KJY47279)	38/54
CB01373_F	97	hypothetical protein	(WP_028566161)	38/50
CB01373_G	222	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	43/59
CB01373_H	250	phosphopantetheinyl transferase	(WP_059300333)	87/88
CB01373_I	433	cytochrome P450 oxygenase	LnmA (AAN85514)	35/56
CB01373_J	252	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	43/62
CB01373_K	437	sodium:proton exchanger	(WP_020553490)	57/73
CB01373_L	299	unknown	LnmE (AAN85518)	45/60
CB01373_M	410	HMG-CoA-synthase	LnmM (AAN85526)	45/61
CB01373_N	264	enoyl-CoA hydratase	LnmF (AAN85519)	53/67
CB01373_O	808	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	63/77
CB01373_P	316	unknown	LnmH (AAN85521)	43/59
CB01373_Q	602	NRPS (A-PCP)	SyrB1 (AKF46132)	46/61
CB01373_R	314	halogenase	SyrB2 (AKF46131)	67/80
CB01373_S	391	alpha/beta hydrolase	ADZ36_03165 (KNE83935)	42/54
CB01373_T	4209	AT-less type I PKS	LnmI (AAN85522)	51/61
CB01373_U	7120	AT-less type I PKS	LnmJ (AAN85523)	50/61
CB01373_V	329	acyltransferase/decarboxylase	LnmK (AAN85524)	55/66
CB01373_W	87	acyl carrier protein	LnmL (AAN85525)	63/71
CB01373_X	413	HMG-CoA synthase	LnmM (AAN85526)	62/76
CB01373_Y	80	acyl carrier protein	(WP_062008561)	55/72
CB01373_Z1	423	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	43/56
CB01373_Z2	253	thioesterase type II	LnmN (AAN85527)	54/64
CB01373_Z3	239	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	54/65
CB01373_Z4	132	unknown	LnmZ' (AAN85540)	38/58
CB01373_Z5	488	4-coumarate:CoA ligase	LnmW (AAN85536)	45/58
CB01373_Z6	226	antibiotic efflux protein	LnmY (AAN85538)	30/43
CB01373_Z7	201	TetR family transcriptional regulator	(WP_058852673)	42/65
CB01373_Z8	395	FAD-dependent oxidoreductase	(WP_010982689)	61/70
CB01373_Z9	121	glyoxalase/bleomycin resistance protein	(WP_062700437)	42/59
CB01373_Z10	491	4-coumarate:CoA ligase	LnmW (AAN85536)	42/56
CB01373_Z11	414	cytochrome P450 oxygenase	LnmA (AAN85514)	38/54
CB01373_orf(+1)	169	hypothetical protein	(WP_053707566)	60/69

[a] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b] Number of amino acids. [c] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S28. Predicted functions of ORFs in the *gnm* gene cluster (accession number: MF925481) from *Streptomyces* sp. CB01883 (Clade IX)

gene[a]	aa[b]	putative function[c]	protein homologue (accession #)	% identity/ % similarity
gnm_orf(-1)	441	putative lanthionine synthetase	AQJ30_24115 (KUN35765)	77/85
gnmA	145	truncated hemoglobin	SAVERM_3316 (BAC71027)	37/51
gnmB	2339	hybrid NRPS/type I PKS	LnmI (AAN85522)	46/56
gnmC	252	type II thioesterase	LnmN (AAN85527)	53/63
gnmD	255	4'-phosphopantetheinyl transferase	Sfp (P39135)	26/45
gnmE	410	HMG-CoA synthase	LnmM (AAN85526)	49/63
gnmF	268	enoyl-CoA hydratase	LnmF (AAN85519)	47/58
gnmG	829	acyltransferase/oxidoreductase	LnmG (AAN85520)	45/62
gnmH	2175	AT-less type I PKS	LnmI (AAN85522)	49/58
gnmI	518	antibiotic efflux protein	LnmY (AAN85538)	31/49
gnmJ	299	unknown	LnmE (AAN85518)	45/60
gnmK	803	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	61/73
gnmL	310	unknown	LnmH (AAN85521)	39/56
gnmM	121	glyoxalase/bleomycin resistance protein	(WP_062700437)	51/66
gnmN	135	unknown	LnmZ' (AAN85540)	34/48
gnmO	238	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	53/67
gnmP	248	methyltransferase	(WP_031183778)	39/51
gnmQ	702	putative alpha/beta hydrolase	LnmD (AAN85517)	47/57
gnmR	84	peptidyl carrier protein	LnmP (AAN85529)	45/59
gnmS	542	discrete adenylation protein	LnmQ (AAN85530)	55/66
gnmT	7422	AT-less type I PKS	LnmJ (AAN85523)	46/56
gnmU	415	HMG-CoA-synthase	LnmM (AAN85526)	60/75
gnmV	91	acyl carrier protein	Fr9M (ADH01494)	36/68
gnmW	423	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	44/57
gnmX	417	cytochrome P450 oxygenase	LnmA (AAN85514)	41/53
gnmY	433	aminotransferase	(KMS73037)	41/56
gnmZ	229	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	42/62
gnm_orf(+1)	761	AfsR-like regulatory protein	(AKN70481)	78/84

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *gnm* gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *gnm* gene cluster.

Table S29. Predicted functions of ORFs in the *lnm*-type gene cluster from *Streptomyces* sp. CB01201 (Clade X)

gene[a]	aa[b]	putative function[c]	protein homologue (accession #)	% identity/ % similarity
CB01201_orf(-1)	367	hypothetical protein	(WP_045300259)	63/74
CB01201_A	145	truncated hemoglobin	SAVERM_3316 (BAC71027)	36/52
CB01201_B	2338	hybrid NRPS/type I PKS	LnmI (AAN85522)	46/57
CB01201_C	252	thioesterase type II	LnmN (AAN85527)	53/62
CB01201_D	290	4'-phosphopantetheinyl transferase	Sfp (P39135)	27/45
CB01201_E	410	HMG-CoA synthase	LnmM (AAN85526)	47/62
CB01201_F	257	enoyl-CoA hydratase	LnmF (AAN85519)	50/62
CB01201_G	727	acyltransferase/oxidoreductase	LnmG (AAN85520)	48/62
CB01201_H	2063	AT-less type I PKS	LnmI (AAN85522)	51/60
CB01201_I	515	antibiotic efflux protein	LnmY (AAN85538)	32/49
CB01201_J	299	unknown	LnmE (AAN85518)	43/59
CB01201_K	756	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	57/67
CB01201_L	310	unknown	LnmH (AAN85521)	42/58
CB01201_M	121	glyoxalase/bleomycin resistance protein	(WP_062700437)	45/66
CB01201_N	135	unknown	LnmZ' (AAN85540)	35/49
CB01201_O	231	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	53/66
CB01201_P	248	methyltransferase	(WP_031183778)	38/48
CB01201_Q	664	putative alpha/beta hydrolase	LnmD (AAN85517)	36/47
CB01201_R	85	peptidyl carrier protein	LnmP (AAN85529)	43/60
CB01201_S	516	discrete adenylation protein	LnmQ (AAN85530)	55/66
CB01201_T	7016	AT-less type I PKS	LnmJ (AAN85523)	47/58
CB01201_U	418	HMG-CoA-synthase	LnmM (AAN85526)	62/78
CB01201_V	97	acyl carrier protein	Fr9M (ADH01494)	38/63
CB01201_W	433	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	44/57
CB01201_X	412	cytochrome P450 oxygenase	LnmA (AAN85514)	41/54
CB01201_Y	446	aminotransferase	(KMS73037)	42/57
CB01201_Z	229	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	42/60
CB01201_orf(+1)	581	methylthioribulose 1-phosphate dehydratase	(WP_014143993)	63/72

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S30. Predicted functions of ORFs in the *wsm* gene cluster (accession number: MF925482) from *Streptomyces* sp. CB02120-2 (Clade XII)

gene[a]	aa[b]	putative function[c]	protein homologue (accession #)	% identity/ % similarity
wsm_orf(-1)	294	SAM-dependent methyltransferase	ADK53_27075 (KOU30977)	
wsmA	219	Crp/Fnr/Fnr family transcriptional regulator	LnmO (AAN85528)	38/57
wsmB	190	YdeI-like protein	SAVERM_307 (BAC68016)	54/72
wsmC	251	phosphopantetheinyl transferase	(WP_059300333)	37/48
wsmD	398	HMG-CoA-synthase	LnmM (AAN85526)	46/59
wsmE	268	enoyl-CoA hydratase	LnmF (AAN85519)	52/61
wsmF	290	acyltransferase	LnmG (AAN85520)	42/55
wsmG	2147	AT-less type I PKS	LnmI (AAN85522)	45/54
wsmH	521	antibiotic efflux protein	LnmY (AAN85538)	35/52
wsmI	268	unknown	LnmE (AAN85518)	37/53
wsmJ	773	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	61/72
wsmK	302	unknown	LnmH (AAN85521)	38/54
wsmL	121	glyoxalase/bleomycin resistance protein	(WP_062700437)	44/63
wsmM	236	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	49/60
wsmN	249	methyltransferase	(WP_031183778)	36/44
wsmO	455	putative alpha/beta hydrolase	LnmD (AAN85517)	35/45
wsmP	83	peptidyl carrier protein	LnmP (AAN85529)	46/68
wsmQ	526	discrete adenylation protein	LnmQ (AAN85530)	52/61
wsmR	7443	AT-less type I PKS	LnmJ (AAN85523)	43/53
wsmS	414	HMG-CoA-synthase	LnmM (AAN85526)	57/71
wsmT	81	acyl carrier protein	Fr9M (ADH01494)	31/50
wsmU	408	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	43/56
wsmV	251	type II thioesterase	LnmN (AAN85527)	55/62
wsmW	2244	hybrid NRPS/type I PKS	LnmI (AAN85522)	45/54
wsmX	141	truncated hemoglobin	SAVERM_3316 (BAC71027)	32/45
wsmY	408	cytochrome P450 oxygenase	LnmA (AAN85514)	39/55
wsmZ1	74	ferredoxin	LnmB (AAN85515)	36/54
wsmZ2	452	crotonyl-CoA carboxylase/reductase	DivR (CCP20056)	64/77
wsmZ3	330	ketoacyl-acyl carrier protein synthase III (KASIII)	DivS (CCP20057)	56/73
wsmZ4	294	3-hydroxyacyl-CoA dehydrogenase	DivT (CCP20058)	52/67
wsm_orf(+1)	475	peptidase M1	VR46_04345 (KJY47303)	86/90

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *wsm* gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *wsm* gene cluster.

Table S31. Predicted functions of ORFs in the *lnm*-type gene cluster from *Streptomyces* sp. CB02613 (Clade XVI)

gene[a]	aa[b]	putative function[c]	protein homologue	% identity/ % similarity
CB02613_orf(-1)	188	phosphotransferase	SPW_1180 (EHM30421)	91/93
CB02613_A	514	ABC transporter component, periplasmic oligopeptide binding protein	LnmU (AAN85534)	59/74
CB02613_B	321	ABC transporter component, membrane spanning protein	LnmT (AAN85533)	53/67
CB02613_C	269	ABC transporter component, membrane spanning protein	LnmS (AAN85532)	62/74
CB02613_D	551	ABC transporter component, ATP hydrolase	LnmR (AAN85531)	54/66
CB02613_E	426	cytochrome P450 oxygenase	LnmA (AAN85514)	40/57
CB02613_F	391	methyltransferase domain of nonribosomal peptide synthetase	(WP_052853671)	41/50
CB02613_G	289	hypothetical protein	TrdC (ADY38535)	33/49
CB02613_H	271	unknown	LnmE (AAN85518)	33/50
CB02613_I	783	AT-less acyltransferase/oxidoreductase	LnmG (AAN85520)	64/77
CB02613_J	4128	AT-less type I PKS	LnmI (AAN85522)	50/59
CB02613_K	315	SyrP-like protein	(BAB69364)	54/64
CB02613_L	306	kinase	PduX_2 (GAQ50744)	64/73
CB02613_M	1059	NRPS (A-PCP-TE)	(ADK54899)	55/65
CB02613_N	402	diaminobutyric acid synthase C	DabC (AEG64688)	58/66
CB02613_O	487	diaminobutyric acid synthase B	DabB (AEG64687)	62/72
CB02613_P	360	diaminobutyric acid synthase A	DabA (AEG64686)	62/71
CB02613_Q	111	peptidyl carrier protein	(WP_059300325)	39/61
CB02613_R	317	halogenase	SyrB2 (AKF46131)	63/77
CB02613_S	283	alpha/beta hydrolase	ADZ36_03165 (KNE83935)	52/64
CB02613_T	6970	AT-less type I PKS	LnmJ (AAN85523)	50/60
CB02613_U	299	acyltransferase/decarboxylase	LnmK (AAN85524)	54/67
CB02613_V	86	acyl carrier protein	LnmL (AAN85525)	58/68
CB02613_W	406	HMG-CoA synthase	LnmM (AAN85526)	66/77
CB02613_X	79	acyl carrier protein	LnmL (AAN85525)	39/61
CB02613_Y	418	beta-ketoacyl-ACP synthase	SiaA (AFS33443)	43/55
CB02613_Z1	266	thioesterase type II	LnmN (AAN85527)	53/64
CB02613_Z2	212	phosphatidylethanolamine-binding protein	Mesil_1512 (ADH63402)	50/64
CB02613_Z3	127	unknown	LnmZ' (AAN85540)	37/55
CB02613_Z4	261	enoyl-CoA hydratase	LnmF (AAN85519)	50/66
CB02613_Z5	415	HMG-CoA synthase	LnmM (AAN85526)	47/63
CB02613_Z6	457	sodium:proton exchanger	(WP_020553490)	56/70
CB02613_Z7	242	Crp/Fnr family transcriptional regulator	LnmO (AAN85528)	43/61
CB02613_Z8	134	unknown	LnmZ' (AAN85540)	39/55
CB02613_Z9	314	unknown	LnmH (AAN85521)	44/59
CB02613_Z10	514	4-coumarate:CoA ligase	LnmW (AAN85536)	42/55
CB02613_Z11	236	N-acetylglucosaminyl deacetylase	LnmX (AAN85537)	52/65
CB02613_Z12	294	unknown	LnmE (AAN85518)	47/65
CB02613_Z13	94	peptidyl carrier protein	PyrB (AFV71306)	53/62
CB02613_Z14	293	hypothetical protein	TrdC (ADY38535)	37/49

CB02613_Z15	500	4-coumarate:CoA ligase	LnmW (AAN85536)	43/57
CB02613_Z16	470	serine hydroxymethyltransferase	AmiS (AEF16058)	62/75
CB02613_Z17	824	nonribosomal peptide synthetase (A-PCP)	LnmQ (AAN85530)	36/46
CB02613_Z18	118	hypothetical protein	SSHG_04587 (EFE84146)	60/66
CB02613_Z19	362	hypothetical protein	(WP_028812750)	78/87
CB02613_Z20	975	4-coumarate:CoA ligase	LnmW (AAN85536)	41/55
CB02613_Z21	400	FAD-dependent oxidoreductase	(WP_010982689)	59/69
CB02613_orf(+1)	290	UDP pyrophosphate phosphatase	(WP_030588347)	98/99

[a]*orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene cluster.
[b]Number of amino acids. [c]Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S32. Comparison of the *lnm*-type gene clusters from *Streptomyces* sp. NBRC 109436, *S. hygroscopicus* sp. NBRC 16556, and *S. leeuwenhoekii* DSM 42122 with the cluster from *Streptomyces* sp. TSRI0384-2 (Clade V)

Streptomyces sp. TSRI0384-2[a]	*Streptomyces* sp. NBRC 109436[b]	*Streptomyces hygroscopicus* sp. NBRC 16556[c]	*S. leeuwenhoekii* DSM 42122[d]
TSRI0384-2_orf(-1)	*S109_orf(-1)* (78/83)	*Shyg_orf(-1)* (78/83)	*Slee_orf(-1)* (82/86)
TSRI0384-2_A	*S109_A* (82/88)	*Shyg_A* (82/88)	*Slee_A* (86/90)
TSRI0384-2_B	*S109_B* (89/93)	*Shyg_B* (89/93)	*Slee_B* (88/93)
TSRI0384-2_C	*S109_C* (90/95)	*Shyg_C* (90/95)	*Slee_C* (90/95)
TSRI0384-2_D	*S109_D* (92/95)	*Shyg_D* (92/95)	*Slee_D* (92/95)
TSRI0384-2_E	*S109_E* (86/88)	*Shyg_E* (86/88)	*Slee_E* (85/88)
TSRI0384-2_F	*S109_F* (86/90)	*Shyg_F* (86/90)	*Slee_F* (78/83)
TSRI0384-2_G	*S109_G* (87/92)	*Shyg_G* (87/92)	*Slee_G* (90/94)
TSRI0384-2_H	*S109_H* (79/85)	*Shyg_H* (79/85)	*Slee_H* (80/85)
TSRI0384-2_I	*S109_I* (92/95)	*Shyg_I* (91/95)	*Slee_I* (91/95)
TSRI0384-2_J	*S109_J* (77/83)	*Shyg_J* (76/82)	*Slee_J* (76/82)
TSRI0384-2_K	*S109_K* (78/82)	*Shyg_K* (78/82)	*Slee_K* (78/82)
TSRI0384-2_L	*S109_L* (76/80)	*Shyg_L* (76/80)	*Slee_L* (76/81)
TSRI0384-2_M	*S109_M* (85/91)	*Shyg_M* (85/91)	*Slee_M* (87/92)
TSRI0384-2_N	*S109_N* (81/89)	*Shyg_N* (81/89)	*Slee_N* (80/90)
TSRI0384-2_O	*S109_O* (91/94)	*Shyg_O* (91/95)	*Slee_O* (91/94)
TSRI0384-2_P	*S109_P* (79/90)	*Shyg_P* (80/90)	*Slee_P* (79/91)
TSRI0384-2_Q	*S109_Q* (82/86)	*Shyg_Q* (82/86)	*Slee_Q* (81/86)
TSRI0384-2_R	*S109_R* (84/88)	*Shyg_R* (84/88)	*Slee_R* (84/88)
TSRI0384-2_S	*S109_S* (88/93)	*Shyg_S* (88/93)	*Slee_S* (88/93)
TSRI0384-2_T	*S109_T* (88/93)	*Shyg_T* (87/93)	*Slee_T* (89/94)
TSRI0384-2_U	*S109_U* (86/92)	*Shyg_U* (86/92)	*Slee_U* (85/91)
TSRI0384-2_V	*S109_V* (90/94)	*Shyg_V* (90/93)	*Slee_V* (93/96)
TSRI0384-2_W	*S109_W* (89/95)	*Shyg_W* (89/95)	*Slee_W* (86/94)
TSRI0384-2_X	*S109_X* (92/94)	*Shyg_X* (92/94)	*Slee_X* (92/94)
TSRI0384-2_Y	*S109_Y* (85/91)	*Shyg_Y* (85/91)	*Slee_Y* (86/92)
TSRI0384-2_Z1	*S109_Z1* (84/90)	*Shyg_Z1* (84/90)	*Slee_Z1* (86/93)
TSRI0384-2_Z2	*S109_Z2* (82/87)	*Shyg_Z2* (82/87)	*Slee_Z2* (84/88)
TSRI0384-2_Z3	*S109_Z3* (88/94)	*Shyg_Z3* (88/94)	*Slee_Z3* (88/94)
TSRI0384-2_orf(+1)	*S109_orf(+1)* (86/91)	*Shyg_orf(+1)* (86/91)	*Slee_orf(+1)* (87/93)

[a,b,c,d]Genes within the *lnm*-type gene clusters, *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene clusters. Numbers insides brackets correspond to the protein sequence similarity/identity to those of *Streptomyces* sp. TSRI0384-2. Also see **Fig. 2** for the genetic organization of the *lnm*-type gene clusters.

Table S33. Comparison of the *lnm*-type gene clusters from *Streptomyces* sp. CB01373 and *S. canus* ATCC 12647 (Clade VII)

CB01373[a]	putative function[b]	*S. canus* ATCC 12647[c]	% identity/ % similarity
CB01373_orf(-1)	hypothetical protein		
CB01373_A	antibiotic efflux protein		
CB01373_B	TetR family transcriptional regulator	*Scan_orf(-1)*	91/94
CB01373_C	hypothetical protein	*Scan_Z2*	44/47
CB01373_D	YdeI-like protein	*Scan_A*	91/95
CB01373_E	cytochrome P450 oxygenase	*Scan_B*	91/94
CB01373_F	hypothetical protein		
CB01373_G	Crp/Fnr family transcriptional regulator	*Scan_C*	91/93
CB01373_H	phosphopantetheinyl transferase	*Scan_D*	87/88
CB01373_I	cytochrome P450 oxygenase	*Scan_E*	97/98
CB01373_J	Crp/Fnr family transcriptional regulator	*Scan_F*	92/93
CB01373_K	sodium:proton exchanger	*Scan_G*	92/95
CB01373_L	unknown	*Scan_H*	93/96
CB01373_M	HMG-CoA-synthase	*Scan_I*	96/97
CB01373_N	enoyl-CoA hydratase	*Scan_J*	94/96
CB01373_O	AT-less acyltransferase/oxidoreductase	*Scan_K*	84/86
CB01373_P	unknown	*Scan_L*	96/98
CB01373_Q	NRPS (A-PCP)	*Scan_M*	94/95
CB01373_R	halogenase	*Scan_N*	96/97
CB01373_S	alpha/beta hydrolase	*Scan_O*	91/93
CB01373_T	AT-less type I PKS	*Scan_P*	89/92
CB01373_U	AT-less type I PKS	*Scan_Q*	87/90
CB01373_V	acyltransferase/decarboxylase	*Scan_R*	96/97
CB01373_W	acyl carrier protein	*Scan_S*	94/96
CB01373_X	HMG-CoA synthase	*Scan_T*	93/95
CB01373_Y	acyl carrier protein	*Scan_U*	98/98
CB01373_Z1	beta-ketoacyl-ACP synthase	*Scan_V*	91/94
CB01373_Z2	thioesterase type II	*Scan_W*	92/96
CB01373_Z3	N-acetylglucosaminyl deacetylase	*Scan_X*	95/98
CB01373_Z4	unknown	*Scan_Y*	95/96
CB01373_Z5	4-coumarate:CoA ligase	*Scan_Z1*	93/95
CB01373_Z6	antibiotic efflux protein		
CB01373_Z7	TetR family transcriptional regulator	*Scan_orf(-1)*	66/81
CB01373_Z8	FAD-dependent oxidoreductase	*Scan_Z2*	93/95
CB01373_Z9	glyoxalase/bleomycin resistance protein	*Scan_Z3*	97/99
CB01373_Z10	4-coumarate:CoA ligase	*Scan_Z4*	95/97
CB01373_Z11	cytochrome P450 oxygenase	*Scan_Z5*	94/97
CB01373_orf(+1)	hypothetical protein		60/69
	Repeat domain-containing protein	*Scan_orf(+1)*	

[a,c]Genes within the *lnm*-type gene clusters, *orf(-1)* and *orf(+1)* are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene clusters. [b]Also see **Fig. 2** for the genetic organization of the *lnm*-type gene clusters.

Table S34. Comparison of the *lnm*-type gene clusters from *Streptomyces* sp. F5630 and *S. aureofaciens* NRRL B-2183 with the *gnm* gene cluster from *Streptomyces* sp. CB01883 (Clade IX)

gene[a]	putative function[c]	Streptomyces sp. F5630	S. aureofaciens NRRL B-2183
	YdeI-like protein	*Sf56_orf(-1)*	*Saur_orf(-1)*
gnm_orf(-1)	putative lanthionine synthetase		
gnmA	truncated hemoglobin	*Sf56_A* (79/84)	*Saur_A* (82/87)
gnmB	hybrid NRPS/type I PKS	*Sf56_B* (76/80)	*Saur_B* (78/82)
gnmC	type II thioesterase	*Sf56_C* (87/93)	*Saur_C* (86/92)
gnmD	4'-phosphopantetheinyl transferase	*Sf56_D* (72/78)	*Saur_D* (79/80)
gnmE	HMG-CoA synthase	*Sf56_E* (87/90)	*Saur_E* (90/94)
gnmF	enoyl-CoA hydratase	*Sf56_F* (77/85)	*Saur_F* (81/86)
gnmG	acyltransferase/oxidoreductase	*Sf56_G* (67/71)	*Saur_G* (70/75)
gnmH	AT-less type I PKS	*Sf56_H* (72/77)	*Saur_H* (71/76)
gnmI	antibiotic efflux protein	*Sf56_I* (91/94)	*Saur_I* (91/94)
gnmJ	unknown	*Sf56_J* (91/94)	*Saur_J* (90/96)
gnmK	AT-less acyltransferase/oxidoreductase	*Sf56_K* (81/86)	*Saur_K* (90/93)
gnmL	unknown	*Sf56_L* (88/93)	*Saur_L* (90/94)
gnmM	glyoxalase/bleomycin resistance protein	*Sf56_M* (88/95)	*Saur_M* (91/98)
gnmN	unknown	*Sf56_N* (86/92)	*Saur_N* (88/95)
gnmO	N-acetylglucosaminyl deacetylase	*Sf56_O* (89/93)	*Saur_O* (95/98)
gnmP	methyltransferase	*Sf56_P* (87/91)	*Saur_P* (86/90)
gnmQ	putative alpha/beta hydrolase	*Sf56_Q* (69/75)	*Saur_Q* (71/76)
gnmR	peptidyl carrier protein	*Sf56_R* (86/92)	*Saur_R* (90/95)
gnmS	discrete adenylation protein	*Sf56_S* (83/88)	*Saur_S* (83/86)
gnmT	AT-less type I PKS	*Sf56_T* (70/75)	*Saur_T* (71/76)
gnmU	HMG-CoA-synthase	*Sf56_U* (90/94)	*Saur_U* (91/94)
gnmV	acyl carrier protein	*Sf56_V* (85/94)	*Saur_V* (79/91)
gnmW	beta-ketoacyl-ACP synthase	*Sf56_W* (78/85)	*Saur_W* (78/86)
gnmX	cytochrome P450 oxygenase	*Sf56_X* (89/92)	*Saur_X* (90/93)
gnmY	aminotransferase	*Sf56_Y* (85/90)	*Saur_Y* (89/94)
gnmZ	Crp/Fnr family transcriptional regulator	*Sf56_Z* (84/92)	*Saur_Z* (92/96)
gnm_orf(+1)	AfsR-like regulatory protein		
	LysR family transcriptional regulator	*Sf56_orf(+1)*	
	transposase		*Saur_orf(+1)*

[a,b,c,d]Genes within the *lnm*-type gene clusters, *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene clusters. Numbers insides brackets correspond to the protein sequence similarity/identity to that of *gnm* gene cluster. Also see **Fig. 2** for the genetic organization of the *lnm*-type gene clusters.

Table S35. Comparison of the *lnm*-type gene clusters from *Streptomyces* sp. NBRC 110035 with the *wsm* gene cluster from *Streptomyces* sp. CB02120-2 (Clade XII)

CB02120-2[a]	putative function[b]	*Streptomyces* sp. NBRC 110035[c]	% identity/ % similarity
	transposase	*S110_orf(-1)*	
wsm_orf(-1)	SAM-dependent methyltransferase		
wsmA	Crp/Fnr/Fnr family transcriptional regulator	*S110_A*	81/88
wsmB	YdeI-like protein	*S110_B*	81/89
wsmC	phosphopantetheinyl transferase	*S110_C*	73/78
wsmD	HMG-CoA-synthase	*S110_D*	89/94
wsmE	enoyl-CoA hydratase	*S110_E*	81/88
wsmF	acyltransferase	*S110_F*	83/86
wsmG	AT-less type I PKS	*S110_G*	77/81
	transposase	*S110_H*	
	glyoxalase/bleomycin resistance protein	*S110_I*	
wsmH	antibiotic efflux protein	*S110_J*	84/89
wsmI	unknown	*S110_K*	91/95
wsmJ	AT-less acyltransferase/oxidoreductase	*S110_L*	81/87
wsmK	unknown	*S110_M*	86/91
wsmL	glyoxalase/bleomycin resistance protein	*S110_N*	81/91
wsmM	N-acetylglucosaminyl deacetylase	*S110_O*	83/90
wsmN	methyltransferase	*S110_P*	88/93
wsmO	putative alpha/beta hydrolase	*S110_Q*	69/76
wsmP	peptidyl carrier protein	*S110_R*	78/87
wsmQ	discrete adenylation protein	*S110_S*	80/85
wsmR	AT-less type I PKS	*S110_T*	74/79
wsmS	HMG-CoA-synthase	*S110_U*	84/88
wsmT	acyl carrier protein	*S110_V*	91/93
wsmU	beta-ketoacyl-ACP synthase	*S110_W*	78/84
wsmV	type II thioesterase	*S110_X*	81/88
wsmW	hybrid NRPS/type I PKS	*S110_Y*	70/76
wsmX	truncated hemoglobin	*S110_Z1*	82/89
wsmY	cytochrome P450 oxygenase	*S110_Z2*	89/94
wsmZ1	ferredoxin	*S110_Z3*	78/83
wsmZ2	crotonyl-CoA carboxylase/reductase	*S110_Z4*	86/93
wsmZ3	ketoacyl-acyl carrier protein synthase III (KASIII)	*S110_Z5*	82/91
wsmZ4	3-hydroxyacyl-CoA dehydrogenase	*S110_Z6*	77/86
wsm_orf(+1)	peptidase M1		
	transposase	*S110_orf(+1)*	

[a,c] *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene clusters.
[b] Also see **Fig. 2** for the genetic organization of the *lnm*-type gene cluster.

Table S36. Comparison of the *lnm*-type gene clusters from *Micromonospora globosa* NRRL B-2673*, Micromonospora marina* DSM 45555, *Micromonospora* sp. CNB394 and *Micromonospora* sp. L5 with the cluster from *Micromonospora aurantiaca* ATCC 27029 (Clade XV)

M. aurantiaca ATCC 27029[a]	*M. globosa* NRRL B-2673[b]	*M. marina* DSM 45555[c]	*Micromonospora*. sp. CNB394[d]	*Micromonospora* sp. L5[e]
Maur_orf(-1)	*Mglo_orf(-1)* (99/99)	*Mmar_orf(-1)* (99/99)	*Mcnb_orf(-1)* (94/95)	*Ml5_orf(-1)* (94/95)
Maur_A	*Mglo_A* (100/100)	*Mmar_A* (98/99)	*Mcnb_A* (98/98)	*Ml5_A* (99/99)
Maur_B	*Mglo_B* (100/100)	*Mmar_B* (98/98)	*Mcnb_B* (98/98)	*Ml5_B* (99/100)
Maur_C	*Mglo_C* (100/100)	*Mmar_C* (96/98)	*Mcnb_C* (97/98)	*Ml5_C* (100/100)
Maur_D	*Mglo_D* (100/100)	*Mmar_D* (95/96)	*Mcnb_D* (96/96)	*Ml5_D* (99/99)
Maur_E	*Mglo_E* (100/100)	*Mmar_E* (99/99)	*Mcnb_E* (98/99)	*Ml5_E* (100/100)
Maur_F	*Mglo_F* (100/100)	*Mmar_F* (95/97)	*Mcnb_F* (95/96)	*Ml5_F* (96/96)
Maur_G	*Mglo_G* (100/100)	*Mmar_G* (98/100)	*Mcnb_G* (98/100)	*Ml5_G* (100/100)
Maur_H	*Mglo_H* (100/100)	*Mmar_H* (98/97)	*Mcnb_H* (98/97)	*Ml5_H* (100/100)
Maur_I	*Mglo_I* (98/97)	*Mmar_I* (89/90)	*Mcnb_I* (93/94)	*Ml5_I* (98/98)
Maur_J	*Mglo_J* (99/99)	*Mmar_J* (91/93)	*Mcnb_J* (91/92)	*Ml5_J* (98/98)
Maur_K	*Mglo_K* (99/100)	*Mmar_K* (97/98)	*Mcnb_K* (97/98)	*Ml5_K* (98/98)
Maur_L	*Mglo_L* (100/100)	*Mmar_L* (92/93)	*Mcnb_L* (91/92)	*Ml5_L* (95/96)
Maur_M	*Mglo_M* (100/100)	*Mmar_M* (98/98)	*Mcnb_M* (98/98)	*Ml5_M* (99/99)
Maur_N	*Mglo_N* (100/100)	*Mmar_N* (95/96)	*Mcnb_N* (95/96)	*Ml5_N* (99/98)
Maur_O	*Mglo_O* (99/99)	*Mmar_O* (96/97)	*Mcnb_O* (96/97)	*Ml5_O* (99/99)
Maur_P	*Mglo_P* (100/100)	*Mmar_P* (99/99)	*Mcnb_P* (99/99)	*Ml5_P* (99/99)
Maur_Q	*Mglo_Q* (100/100)	*Mmar_Q* (97/98)	*Mcnb_Q* (97/97)	*Ml5_Q* (98/98)
Maur_R	*Mglo_R* (100/100)	*Mmar_R* (98/99)	*Mcnb_R* (98/99)	*Ml5_R* (98/99)
Maur_S	*Mglo_S* (100/100)	*Mmar_S* (99/99)	*Mcnb_S* (99/99)	*Ml5_S* (100/100)
Maur_T	*Mglo_T* (100/100)	*Mmar_T* (97/96)	*Mcnb_T* (96/96)	*Ml5_T* (100/100)
Maur_U	*Mglo_U* (99/98)	*Mmar_U* (93/94)	*Mcnb_U* (93/94)	*Ml5_U* (99/99)
Maur_V	*Mglo_V* (99/99)	*Mmar_V* (96/97)	*Mcnb_V* (96/96)	*Ml5_V* (99/99)
Maur_W	*Mglo_W* (99/100)	*Mmar_W* (98/99)	*Mcnb_W* (98/99)	*Ml5_W* (99/100)
Maur_X	*Mglo_X* (98/98)	*Mmar_X* (92/93)	*Mcnb_X* (93/94)	*Ml5_X* (99/99)
Maur_Y	*Mglo_Y* (99/100)	*Mmar_Y* (98/98)	*Mcnb_Y* (98/98)	*Ml5_Y* (99/100)
Maur_Z1	*Mglo_Z1* (100/100)	*Mmar_Z1* (100/100)	*Mcnb_Z1* (100/100)	*Ml5_Z1* (100/100)
Maur_Z2	*Mglo_Z2* (99/99)	*Mmar_Z2* (97/98)	*Mcnb_Z2* (98/99)	*Ml5_Z2* (99/99)
Maur_Z3	*Mglo_Z3* (99/99)	*Mmar_Z3* (96/98)	*Mcnb_Z3* (97/98)	*Ml5_Z3* (100/100)
Maur_Z4	*Mglo_Z4* (99/100)	*Mmar_Z4* (99/99)	*Mcnb_Z4* (99/99)	*Ml5_Z4* (99/100)
Maur_Z5	*Mglo_Z5* (100/100)	*Mmar_Z5* (99/99)	*Mcnb_Z5* (99/99)	*Ml5_Z5* (100/100)
Maur_Z6	*Mglo_Z6* (100/100)	*Mmar_Z6* (98/98)	*Mcnb_Z6* (98/98)	*Ml5_Z6* (100/100)
Maur_Z7	*Mglo_Z7* (99/99)	*Mmar_Z7* (99/99)	*Mcnb_Z7* (98/98)	*Ml5_Z7* (99/99)
Maur_Z8	*Mglo_Z8* (99/100)	*Mmar_Z8* (98/100)	*Mcnb_Z8* (99/100)	*Ml5_Z8* (100/100)

Maur_Z9	*Mglo_Z9* (100/100)	*Mmar_Z9* (99/100)	*Mcnb_Z9* (99/100)	*Ml5_Z9* (100/100)
Maur_Z10	*Mglo_Z10* (100/100)	*Mmar_Z10* (99/100)	*Mcnb_Z10* (100/100)	*Ml5_Z10* (100/100)
Maur_Z11	*Mglo_Z11* (99/100)	*Mmar_Z11* (97/98)	*Mcnb_Z11* (97/98)	*Ml5_Z11* (99/100)
Maur_orf(+1)	*Mglo_orf(+1)* (97/98)	*Mmar_orf(+1)* (91/92)	*Mcnb_orf(+1)* (91/92)	*Ml5_orf(+1)* (97/97)

[a,b,c,d,e] Genes within the *lnm*-type gene clusters, *orf(-1)* and *orf(+1*) are predicted to represent the upstream and downstream boundaries of the *lnm*-type gene clusters. Numbers insides brackets correspond to the protein sequence similarity/identity. Also see **Fig. 2** for the genetic organization of the *lnm*-type gene clusters.

Table S37. Substrate specificity-conferring codes and predicted substrates of the adenylation proteins/domains from LNM-tpe biosynthetic pathways.

Adenylation protein/domain[a]	8 angstrom signature[b]	Stachelhaus code[c]	Predicted substrates
Group I			
WsmQ	FTAYFDLSMSELHAVFCGEDNLYGPTEASVACTG	DLSHVCLVAK	D-Ala[d]
S110_S	FTAYFDLSMSELHAIFCGEDNLYGPTEASVACTG	DLSHICLVAK	Unknown
LnmQ	FASYFDLSMFECTSVFCGEDNLYGPTELTIACTA	DLFTVCLIAK	D-Ala[d]
CB01635_Q	FASYFDLSMFECTSVFCGEDNLYGPTELTIACTA	DLFTVCLIAK	D-Ala[d]
CB02959_Q	FATYFDLSMFEATSVFCGEDNLYGPTELTVACTG	DLFTVCLVAK	Unknown
CB02891_F	FAVYFDLSMFEATSMFCGEDNLYGPTELAVACTG	DLFTMCLVAK	Unknown
Sast_X	FAAYFDLAMFDMHTVFCGEDNIYGPTELTIACTA	DLFHVCIIAK	Unknown
CB01201_S	FSSHFDLALADPHSVFAGEDNIYGPTELTIACSA	DLAHVAIIAK	Unknown
GnmS	FTSHFDLAMCDPHSVFAGEDNIYGPTELTIACTA	DLCHVAIIAK	ACC[d]
Sf56_S	FASHFDLALCDPHSVFAGEDNIYGPTELTIACTA	DLCHVAIIAK	ACC[d]
Saur_S	FASHFDLAMCDPHSVFAGEDNIYGPTELTIACTA	DLCHVAIIAK	ACC[d]
Group II			
S109_H	LSSHFDFSVWEGSQVFGGEVNMYGITETCVHVTH	DFWSVGMVHK	L-Thr
Shyg_H	LSSHFDFSVWEGSQVFGGEVNMYGITETCVHVTH	DFWSVGMVHK	L-Thr
Slee_H	LSSHFDFSVWEGSQVFGGEVNMYGITETCVHVTH	DFWSVGMVHK	L-Thr
TSRI0384-2_H	LSSHFDFSVWEGSQVLGGEVNMYGITETCVHVTH	DFWSVGMVHK	L-Thr
Mtul_C	LQLHFDFSVWEGSQVLGGEINMYGITETTVHVTH	DFWSVGMVHK	L-Thr
Caci_L	LQTHFDFSVWEGSQVFGGEVNMYGITETTVHVTY	DFWSVGMVHK	L-Thr
Saes_Y	LRAHFDFSVWEGSQVFGGEVNMYGITEITVHATH	DFWSVGMVHK	L-Thr
Mcnb_X	LHGHFDFSVWEGNQVFGGEVNMYGITEVTVHATA	DFWNVGMVHK	L-Thr
Mmar_X	LHGHFDFSVWEGNQVFGGEVNMYGITEVTVHATA	DFWNVGMVHK	L-Thr
Maur_X	LHGHFDFSVWEGNQVFGGEVNMYGITEVTVHATE	DFWNVGMVHK	L-Thr
Mglo_X	LHGHFDFSVWEGNQVFGGEVNMYGITEVTVHATA	DFWNVGMVHK	L-Thr
Ml5_X	LHGHFDFSVWEGNQVFGGEVNMYGITEVTVHATA	DFWNVGMVHK	L-Thr
CB01373_Q	LQTHFDFSVWEGNQVFGGEVNMYGITETTVHVTA	DFWNVGMVHK	L-Thr
Scan_M	LQTHFDFSVWEGNQVFGGEVNMYGITETTVHVTA	DFWNVGMVHK	L-Thr
Bubo_N	LASHFDFSVWEGSLVMGGEVNMYGITETTVHVTY	DFWSVGMVHK	L-Thr
others			
CB02613_Z17	NDISFDMSVYEGNSLCGGDIALGGATEASVHSTV	DMYNLGLVHK	Unknown
Sal964_D	NDLSFDMSVYEGNTLTGGDISLGGVTEASIHSTV	DMYNLGLIHK	Unknown
CB02613_M	HAKAFDISLWQAQISVTGEVNAYGATEVSDDTMH	DIWQSTADDK	Unknown
Sal964_T	HIKAFDISLWQAQISVTGEVNAYGATEVSDDTMH	DIWQSTADDK	Unknown
Snov_N	AAILFDYGLYQLPLTNTGAVSMFGLTECKRVTIA	DYYPTTMRVK	Unknown

[a]Adenylation proteins/domains from the LNM and 28 LNM-type pathways. The analysis was performed with NRPSpredictor2 (18).
[b]The extracted 8 angstrom signature (19).
[c]The extracted Stachelhaus code (20).
[d]The adenylation proteins/domains whose substrates have been identified experimentally.

Table S38. ^{1}H NMR (700 MHz) and ^{13}C NMR (175 MHz) data of GNM A in DMSO-d_6

No.	Major rotamer		Minor rotamer	
	δ_C	δ_H	δ_C	δ_H
1	169.9		174.4	
2	40.4	2.89 (1H, d, 14.6) 2.39 (1H, d. 14.6)	39.0	3.13 (1H, d, 15.2) 3.00 (1H, d, 15.2)
3	53.6		53.9	
4	34.3	1.62 (1H, m) 1.46 (1H, m)	36.2	1.53 (1H, m) 0.48 (1H, m)
5	33.2	1.84 (2H, m)	33.5	1.87 (1H, m) 1.74 (1H, m)
6	134.9		134.8	
7	122.8	4.96 (1H, t, 6.4)	124.1	4.81 (1H, d, 6.5)
8	32.68	2.97 (1H, m) 2.90 (1H, m)	32.73	3.22 (1H, dd, 8.3, 15.2) 2.81 (1H, d, 15.6)
9	145.9		146.0	
10	132.1	6.38 (1H, d, 16.2)	132.0	6.40 (1H, d, 16.0)
11	128.1	7.00 (1H, d, 16.1)	128.5	7.10 (1H, d, 15.1)
12	136.2		136.1	
13[a]	124.2	6.36 (1H, s)	124.2	6.36 (1H, s)
14	150.4		150.0	
15	117.9	7.40 (1H, s)	118.1	7.51 (1H, s)
16	170.8		173.4	
17	33.1		35.9	
18[b]	15.0	1.39 (1H, m) 1.19 (1H, m)	19.2	1.36 (1H, m) 1.14 (1H, m)
19[b]	14.6	1.19 (1H, m) 1.13 (1H, m)	17.4	2.06 (1H, m) 1.36 (1H, m)
20	19.8	1.96 (3H, s)	19.6	1.99 (3H, s)
21	118.5	5.10 (2H, m)	118.5	5.11 (2H, m)
22	16.25	1.57 (3H, s)	16.31	1.58 (3H, s)
23	41.0	2.96 (1H, d, 15.5) 2.58 (1H, d, 15.5)	38.6	3.20 (1H, d, 15.5) 2.33 (1H, d, 15.5)
24	171.5		171.4	
25	24.2	2.37 (3H, s)	24.0	2.16 (3H, s)
CO$_2$*H* [a]		12.12 (1H, br s)		12.12 (1H, br s)
N*H*		8.92 (1H, br s)		8.78 (1H, br s)

[a]overlapped. [b]interchangeable.

Table S39. ^{1}H NMR (700 MHz) and ^{13}C NMR (175 MHz) data of GNM B and GNM B3 in DMSO-d_6

No.	GNM B[a]		GNM B3	
	δ_C	δ_H	δ_C	δ_H
1	170.8		170.4	
2	50.1	2.55 (1H, d, 13.1) 2.21 (1H, d, 13.0)	43.6	2.76 (1H, d, 11.8) 2.18 (1H, d, 12.0)
3	46.6		53.4	
4	41.6	1.09 (1H, m) 1.00 (1H, m)	39.0	1.41 (1H, m) 0.62 (1H, t, 12.9)
5	33.5	1.94 (1H, m) 1.69 (1H, dt, 4.1, 12.4)	33.4	1.81 (1H, m) 1.62 (1H, m)
6	135.9		135.8	
7	122.0	4.73 (1H, t, 5.3)	122.3	4.71 (1H, t, 5.0)
8	32.62	2.96 (1H, dd, 7.2, 15.9) 2.84 (1H, dd, 5.8, 15.8)	32.5[c]	3.00 (1H, dd, 6.5, 15.8) 2.79 (1H, d, 14.6)
9	145.9		146.0	
10	131.8	6.38 (1H, d, 16.0)	131.7	6.37 (1H, d, 16.0)
11	128.1	6.73 (1H, d, 16.0)	128.3	6.72 (1H, d, 16.0)
12	136.5		136.4	
13	124.9	6.37 (1H, s)	124.9	6.36 (1H, s)
14	149.7		149.7	
15	117.6	7.29 (1H, s)	118.0	7.29 (1H, s)
16	170.1		169.8	
17	32.56		32.5[c]	
18[b]	14.4	1.33 (1H, m) 1.04 (1H, m)	15.4	1.26 (1H, m) 1.08 (1H, m)
19[b]	13.9	1.08 (1H, m) 1.06 (1H, m)	12.7	1.12 (1H, m) 1.07 (1H, m)
20	19.6	1.96 (3H, s)	19.7	1.95 (3H, d, 1.1)
21	118.4	5.09 (1H, d, 2.1) 5.08 (1H, d, 2.1)	118.4	5.08 (1H, s) 5.07 (1H, s)
22	16.4	1.49 (3H, s)	16.0	1.45 (3H, s)
23	50.8	2.37 (1H, d, 13.1) 2.10 (1H, d, 13.0)	46.4	2.25 (1H, d, 12.7) 2.09 (1H, d, 13.5)
24	174.4		174.2	
CO_2*H*				11.32 (1H, br s)
N*H*		10.70 (1H, br s)		

[a]Only the major rotamer is shown. [b]interchangeable. [c]overlapped

Table S40. ^{1}H NMR (700 MHz) and ^{13}C NMR (175 MHz) data of (*S*)- and (*R*)-PGME-GNM B2 in DMSO-d_6

No.	(*S*)-PGME-GNM B2		(*R*)-PGME-GNM B2	
	δ_C	δ_H	δ_C	δ_H
1	170.4		171.3[e]	
2	38.4	2.96 (1H, d, 15.3) 2.26 (1H, d, 15.3)	38.1	3.14 (1H, d, 14.8) 2.72 (1H, d, 14.8)
3	44.6		44.7	
4	37.0	1.93 (1H, td, 3.5, 14.4) 1.79 (1H, dt, 4.3, 13.9)	37.4	1.81 (1H, td, 4.3, 15.0) 1.78 (1H, dt, 3.5, 14.4)
5	32.8	1.64-1.61 (1H, m) 1.62-1.58 (1H, m)	33.0	1.68 (1H, td, 3.7, 13.6) 1.61 (1H, dt, 3.9, 13.0)
6	76.4		76.6	
7	42.6	3.64 (1H, dd, 2.8, 9.0)	43.2	3.74 (1H, d, 6.9)
8	31.3	2.59 (1H, br s) 2.50[b] (1H, m)	30.4	2.73 (1H, br s) 2.54 (1H, dd, 8.6, 13.9)
9	144.4		144.5	
10	133.6	6.37 (1H, d, 16.2)	133.7	6.36 (1H, d, 16.2)
11	128.3[a]	8.05 (1H, d, 16.1)	128.9[f]	8.16 (1H, d, 13.6)
12	133.5		133.5	
13	121.8	6.27 (1H, s)	121.8	6.26 (1H, s)
14	153.1		153.2	
15	117.1	7.33 (1H, s)	117.0	7.32 (1H, s)
16	173.3		173.0	
17	34.8		35.1	
18[d]	19.5[c]	1.43-1.41 (2H, m)	19.7	1.58-1.55 (1H, m) 1.21-1.18 (1H, m)
19[d]	19.5[c]	1.29-1.27 (1H, m) 1.21-1.18 (1H, m)	19.5	1.51 (1H, br s) 1.38[g] (1H, m)
20	21.1	1.96 (3H, s)	21.1	1.95 (3H, s)
21	120.9	5.20 (1H, d, 1.1) 5.11 (1H, d, 1.1)	121.3	5.24 (1H, s) 5.12 (1H, s)
22	16.7	1.38 (3H, s)	16.8	1.38 (3H, s)
23	44.0	2.85 (1H, d, 13.4) 2.20 (1H, d, 13.4)	45.1	2.83 (1H, d, 12.5) 2.04 (1H, d, 12.4)
24	168.6		168.7	
25	48.3	3.12 (3H, s)	48.4	3.19 (3H, s)
1′	170.9		171.3[e]	
2′	56.1	5.36 (1H, d, 7.1)	56.7	5.27 (1H, d, 6.7)
3′	136.2		135.6	
4′	127.6	7.35-7.33 (2H, m)	127.6	7.35-7.33 (2H, m)
5′	128.7	7.39-7.37 (2H, m)	128.9[f]	7.41-7.38 (2H, m)
6′	128.3[a]	7.36-7.34 (1H, m)	128.5	7.38-7.36 (1H, m)
7′	52.2	3.61 (3H, s)	52.5	3.58 (3H, s)
1-N*H*		8.98 (1H, s)		9.14 (1H, s)
24-N*H*		8.81 (1H, d, 7.1)		9.04 (1H, d, 6.4)

[a,c,e,f]overlapped. [b]overlapped with solvent. [d]interchangeable. [g]overlapped by 22-CH_3.

Table S41. ^{1}H NMR (700 MHz) and ^{13}C NMR (175 MHz) data of (*S*)- and (*R*)-PGME-LNM E2 in DMSO-d_6

No.	(*S*)-PGME-LNM E2		(*R*)-PGME-LNM E2	
	δ_C	δ_H	δ_C	δ_H
1	169.0		169.0	
2	38.2	3.16 (1H, d, 14.8) 2.32 (1H, d, 14.9)	38.2	3.14 (1H, d, 14.8) 2.37 (1H, d, 14.9)
3	48.2		48.2	
4	30.5	2.58 (1H, dt, 3.4, 14.4) 1.39 (1H, td, 3.5, 14.4)	31.0	2.61 (1H, dt, 3.2, 14.3) 1.45 (1H, td, 3.4, 14.3)
5	37.9	1.78 (1H, dt, 2.5, 14.0) 1.36 (1H, td, 3.7, 13.7)	38.0	1.82 (1H, dt, 2.5, 14.2) 1.40 (1H, td, 3.5, 13.8)
6	71.0		71.1	
7	45.2	3.75 (1H, dd, 3.7, 12.5)	45.4	3.79 (1H, dd, 3.9, 12.5)
8	40.9	2.69 (1H, dd, 3.6, 12.7) 2.03 (1H, t, 12.7)	41.2	2.78 (1H, dd, 3.9, 12.7) 2.12 (1H, t, 12.6)
9	198.9		198.9	
10	133.7	6.02 (1H, d, 16.1)	133.8	6.04 (1H, d, 16.1)
11	140.4	8.55 (1H, dd, 11.3, 16.3)	140.4	8.61 (1H, dd, 11.3, 16.7)
12	128.24	6.35 (1H, t, 11.5)	128.3[a]	6.36 (1H, t, 11.3)
13	127.6	6.71 (1H, d, 11.5)	127.6	6.72 (1H, d, 11.6)
14	152.5		152.6	
15	122.5	7.85 (1H, s)	122.5	7.86 (1H, s)
16	172.9		173.0	
17	47.4	5.16 (1H, quintet, 6.4)	47.4	5.20 (1H, quintet, 6.5)
18	21.4	1.57 (3H, d, 6.7)	21.4	1.59 (3H, d, 6.7)
19	18.8	1.14 (3H, s)	18.8	1.26 (3H, s)
20	44.1	3.45 (1H, q, 7.0)	44.5	3.41 (1H, q, 7.0)
21	173.5		173.5	
22	12.7	1.11 (3H, d, 7.0)	13.0	1.06 (3H, d, 7.0)
1′	171.1		170.8	
2′	56.3	5.40 (1H, d, 7.1)	56.1	5.40 (1H, d, 7.4)
3′	136.2		137.0	
4′	128.21	7.41-7.39 (2H, m)	127.8[b]	7.43-7.42 (2H, m)
5′	128.1	7.34-7.32 (2H, m)	128.3[a]	7.35-7.33 (2H, m)
6′	127.9	7.32-7.30 (1H, m)	127.8[b]	7.31-7.30 (1H, m)
7′	52.1	3.62 (3H, s)	52.0	3.60 (3H, s)
1-N*H*		8.51 (1H, d, 6.3)		8.54 (1H, d, 6.4)
21-N*H*		8.57 (1H, d, 6.9)		8.67 (1H, d, 7.3)
6-O*H*		4.18 (1H, br s)		4.24 (1H, br s)

[a,b]overlapped

Table S42. ^{1}H NMR (700 MHz) and ^{13}C NMR (175 MHz) data of WSM A1, WSM A2 and WSM A3 in DMSO-d_6

No.	WSM A1		WSM A2		WSM A3	
	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H
1	169.1		169.2		168.8	
2	40.5	3.04 (1H, d, 13.9) 2.40 (1H, d, 13.9)	41.2	3.25 (1H, d, 14.4) 2.47 (1H, d, 14.5)	40.7	3.08 (1H, d, 14.4) 2.64 (1H, d, 13.9)
3	44.9		46.4		46.2	
4	36.6[b]	2.02 (1H, td, 3.5, 14.1) 1.94 (1H, dt, 2.5, 14.1)	42.0	2.27 (1H, m) 1.92 (1H, dt, 5.1, 14.1)	35.2	2.09 (1H, d, 14.4) 2.00 (1H, t, 12.7)
5	38.0	1.77 (1H, dt, 2.3, 13.8) 1.42 (1H, td, 3.4, 14.1)	32.6	2.26-2.20 (2H, m)	35.3	1.72 (1H, dt, 2.9, 12.7) 1.68 (1H, td, 3.5, 12.8)
6	71.2		147.5		81.2	
7	44.7	3.47 (1H, dd, 5.4, 9.2)	38.9	4.05 (1H, dd, 2.2, 11.4)	45.0	3.13 (1H, dd, 6.2, 13.5)
8	33.9	2.54-2.50 (1H, m) 2.25 (1H, dd, 4.7, 13.7)	34.4	2.75 (1H, dd, 12.0, 13.7) 2.53 (1H, dd, 2.1, 12.4)	32.6	2.15 (1H, dd, 6.1, 12.1) 1.86 (1H, t, 12.8)
9	144.9		142.6		83.6	
10	132.9	6.38 (1H, d, 16.5)	132.7	6.46 (1H, d, 16.6)	133.6	5.93 (1H, d, 16.5)
11	126.7	7.78 (1H, d, 13.2)	125.9	7.96 (1H, d, 16.6)	125.9	7.50 (1H, d, 16.4)
12	138.8		138.0		139.1	
13	120.2[a]	6.35 (1H, s)	120.9	6.36 (1H, s)	119.9	6.34 (1H, s)
14	152.7		152.7		151.9	
15	118.4	7.48 (1H, s)	118.9	7.50 (1H, s)	118.1	7.51 (1H, s)
16	170.5		171.2		170.1	
17	47.5	5.16-5.13 (1H, m)	47.4	5.21 (1H, quintet, 6.9)	48.1	5.06 (1H, quintet, 5.9)
18	23.0	1.50 (3H, d, 6.5)	22.9	1.54 (3H, d, 6.7)	23.2	1.52 (3H, d, 6.5)
19	36.6[b]	2.37-2.33 (1H, m) 2.31-2.28 (1H, m)	36.2	2.36-2.31 (2H, m)	36.3	2.37 (1H, ddd, 6.2, 8.3, 14.3) 2.25 (1H, ddd, 6.8, 8.3, 14.7)
20	22.3	1.53-1.50 (2H, m)	22.4	1.54-1.50 (2H, m)	21.6	1.55-1.49 (2H, m)
21	14.0	0.94 (3H, t, 7.3)	14.0	0.94 (3H, t, 7.4)	14.0	0.92 (3H, t, 7.3)
22	120.2[a]	5.16 (1H, s) 5.11 (1H, s)	123.0	5.35 (1H, s) 5.31 (1H, s)	69.0	3.42 (2H, s)
23	19.2	1.32 (3H, s)	107.8	5.09 (1H, s) 4.88 (1H, s)	19.9	1.23 (3H, s)
24	43.1	2.64 (1H, d, 15.4) 2.39 (1H, d, 15.4)	43.2	2.63 (1H, d, 15.3) 2.40 (1H, d, 15.3)	44.9	2.73 (1H, d, 14.7) 2.54 (1H, d, 13.9)
25	170.9		170.6		171.6	
CO_2*H*				12.27 (1H, br s)		12.46 (1H, br s)
N*H*		7.89 (1H, d, 6.6)		8.21 (1H, d, 7.2)		7.80 (1H, d, 5.0)
O*H*		3.81 (1H, br s)				4.87 (1H, br s)

[a,b]overlapped

Table S43. ^{1}H NMR (700 MHz) and ^{13}C NMR (175 MHz) data of (*S*)- and (*R*)-PGME-WSM A2 in DMSO-d_6

No.	(*S*)-PGME-WSM A2		(*R*)-PGME-WSM A2	
	δ_C	δ_H	δ_C	δ_H
1	169.8		170.7	
2	40.7	3.26 (1H, d, 14.3) 2.31 (1H, d, 14.4)	40.3	3.49 (1H, d, 13.9) 2.80 (1H, d, 14.0)
3	47.1		47.4	
4	42.4	2.02 (1H, td, 3.6, 13.6) 1.88 (1H, ddd, 6.0, 11.3, 13.5)	43.0	1.96 (1H, td, 3.8, 13.5) 1.88 (1H, dt, 4.1, 13.1)
5	32.7	2.21-2.19 (2H, m)	32.7	2.27 (1H, dt, 2.7, 13.2) 2.23 (1H, td, 4.1, 13.7)
6	147.8		148.0	
7	38.8	4.07 (1H, dd, 2.3, 11.5)	38.8	4.15 (1H, dd, 2.3, 11.4)
8	34.4	2.74 (1H, dd, 12.0, 13.7) 2.52 (1H, dd, 2.2, 13.9)	34.3	2.74 (1H, dd, 12.0, 13.7) 2.51-2.49[b] (1H, m)
9	142.6		142.6	
10	132.7	6.47 (1H, d, 16.7)	133.0	6.49 (1H, d, 16.6)
11	125.9	7.99 (1H, d, 16.6)	125.9	8.00 (1H, d, 16.6)
12	138.0		138.2	
13	120.8	6.36 (1H, s)	120.7	6.38 (1H, s)
14	152.8		152.8	
15	118.8	7.50 (1H, s)	118.8	7.52 (1H, s)
16	170.89[a]		170.2	
17	47.6	5.19 (1H, quintet, 9.6)	48.0	5.21 (1H, quintet, 7.0)
18	22.8	1.55 (3H, d, 6.6)	22.9	1.54 (3H, d, 6.6)
19	36.2	2.36 (1H, td, 6.8, 14.1) 2.32 (1H, td, 6.8, 14.1)	36.3	2.37 (1H, td, 7.3, 14.2) 2.34 (1H, td, 7.4, 14.2)
20	22.4	1.54 (2H, m)	22.5	1.55-1.51 (2H, m)
21	14.0	0.94 (3H, t, 7.4)	14.1	0.95 (3H, t, 7.4)
22	123.1	5.35 (1H, s) 5.32 (1H, s)	123.4	5.37 (1H, s) 5.33 (1H, s)
23	107.7	5.10 (1H, s) 4.88 (1H, s)	107.8	5.12 (1H, s) 4.90 (1H, s)
24	44.5	2.71 (1H, d, 13.3) 2.27 (1H, d, 13.2)	45.5	2.68 (1H, d, 12.5) 2.12 (1H, d, 12.5)
25	168.0		168.2	
1′	170.91[a]		171.4	
2′	56.4	5.39 (1H, d, 6.8)	56.8	5.35 (1H, d, 7.0)
3′	136.1		135.5	
4′	127.7	7.39-7.38 (2H, m)	127.6	7.42-7.39 (2H, m)
5′	128.8	7.39-7.38 (2H, m)	129.0	7.43-7.40 (2H, m)
6′	128.3	7.36-7.33 (1H, m)	128.5	7.40-7.37 (1H, m)
7′	52.2	3.62 (3H, s)	52.3	3.60 (3H, s)
1-N*H*		8.21 (1H, d, 6.9)		8.11 (1H, d, 6.4)
25-N*H*		9.01 (1H, d, 6.8)		9.29 (1H, d, 7.1)

[a]interchangeable. [b]overlapped with solvent.

Table S44. ^{1}H NMR (700 MHz) and ^{13}C NMR (175 MHz) data of GNM B1 and GNM B2 in DMSO-d_6

No.	GNM B1		GNM B2	
	δ_C	δ_H	δ_C	δ_H
1	170.2		170.9	
2	38.6	3.03 (1H, d, 15.4) 2.43 (1H, d, 15.5)	39.3[b]	2.88 (1H, d, 16.4) 2.76 (1H, d, 15.8)
3	43.9		44.9	
4	37.5	1.98 (1H, m) 1.79 (1H, dt, 2.9, 13.9)	37.4	2.23 (1H, d, 14.1) 1.73 (1H, m)
5	38.8	1.69 (1H, t, 13.9) 1.42 (1H, m)	32.7	1.66 (1H, m) 1.56 (1H, m)
6	71.4		76.4	
7	45.0	3.51 (1H, dd, 3.3, 8.8)	41.6	3.53[c] (1H, m)
8	31.7	2.64 (1H, d, 13.7) 2.57 (1H, dd, 9.0, 13.7)	33.5	2.45 (1H, dd, 10.6, 13.8) 2.32 (1H, d, 13.6)
9	145.0		144.8	
10	133.4	6.36 (1H, d, 16.2)	133.5[d]	6.38 (1H, d, 16.2)
11	128.6	8.09 (1H, d, 16.2)	127.2	7.96 (1H, d, 16.5)
12	133.6		133.5[d]	
13	121.7	6.27 (1H, s)	121.7	6.28 (1H, s)
14	153.1		153.1	
15	117.0	7.32 (1H, s)	117.2	7.33 (1H, s)
16	173.4		173.9	
17	34.7		34.4	
18[a]	19.5	1.42 (1H, m) 1.24 (1H, m)	19.4	1.45 (1H, m) 1.09 (1H, m)
19[a]	19.2	1.47 (1H, m) 1.42 (1H, m)	19.0	1.41 (2H, m)
20	21.1	1.97 (3H, s)	21.3	1.98 (3H, s)
21	120.9	5.20 (2H, s)	119.6	5.12 (1H, s) 5.07 (1H, s)
22	20.1	1.33 (3H, s)	16.8	1.37 (3H, s)
23	43.3	2.77 (1H, d, 15.1) 2.27 (1H, d, 15.1)	47.0	2.50[b] (1H, m) 1.98 (1H, m)
24	171.1		172.3	
25			48.1	3.01 (3H, s)
CO_2*H*		12.10 (1H, br s)		
N*H*		8.97 (1H, br s)		9.31 (1H, br s)
O*H*		4.33 (1H, br s)		

[a]interchangeable. [b]overlapped with the solvent. [c]overlapped with water peak. [d]overlapped.

Fig. S1. Proposed biosynthetic pathway of leinamycin (LNM) in *S. atroolivaceus* S-140 and the modes of action for LNM and LNM E1 as DNA damaging antitumor antibiotics. (**A**) Genetic organization of *lnm* gene cluster. (**B**) The biosynthetic machinery of LNM featuring a hybrid NRPS-AT-less type I PKS (LnmQPIJ) with a discrete acyl transferase (AT, LnmG) loading malonyl-CoA as the extender unit to all six PKS modules, and a set of enzymes consisting of LnmKLMF responsible for the installation of the β-alkyl branch at C3 position. The DUF-SH didomain in PKS module-8 was proposed to be involved in the incorporation of 3-thiol group during LNM biosynthesis. After off-loaded from the PKS by TE domain, the nascent product LNM E1 was modified by tailoring enzymes to afford the natural product LNM. Colored structural moieties indicate their biosynthetic origin installed by NRPS (blue), PKS (red), β-alkyl branching cassette (green), and other tailoring enzymes (black). A, adenylation protein; Cy, condensation/cyclization domain; DH, dehydratase domain; DUF, domain of unknown function; KR, ketoreductase domain; KS, ketosynthase domain; KS^0, nonelongating KS domain; MT, methyltransferase domain; Ox, oxidation domain; SH, cysteine lyase domain; TE, thioesterase domain; solid blue circle, peptidyl carrier protein (PCP); solid red circle, acyl carrier protein (ACP). Gaps between domains denote protein boundaries, and solid line-connect domains are in the same protein. (**C**) LNM can be activated by cellular thiols to form a hydrodisulfide derivative, causing oxidative DNA damage, and an episulfonium ion intermediate that can alkylate DNA, thus leading to cell death. LNM E1 can be oxidatively activated by reactive oxygen species (ROS) to give a similar episulfonium ion intermediate, thereby alkylating DNA and leading to eventual cell death.

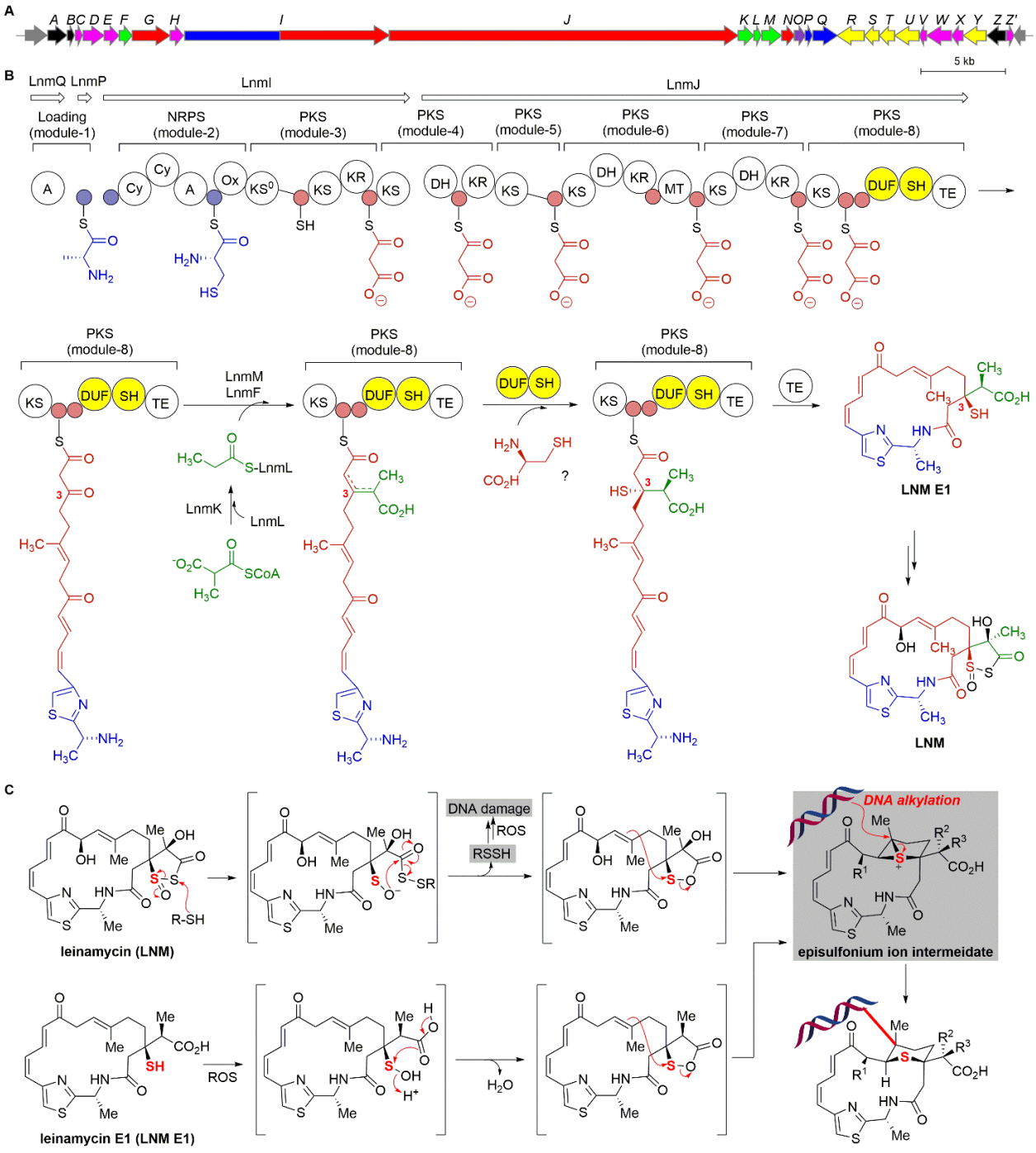

A
A BC D E F G H I J K L M NOP Q R S T U V W X Y Z Z'
5 kb
B
LnmQ LnmP
LnmI
LnmJ
Loading (module-1)
NRPS (module-2)
PKS (module-3)
PKS (module-4)
PKS (module-5)
PKS (module-6)
PKS (module-7)
PKS (module-8)
LnmM LnmF
LnmK
LnmL
S-LnmL
SCoA
LNM E1
LNM
C
leinamycin (LNM)
R-SH
DNA damage
ROS
RSSH
DNA alkylation
episulfonium ion intermeidate
leinamycin E1 (LNM E1)
ROS
H2O

Fig. S2. Discovery-based approach to combinatorial biosynthesis for the LNM family of natural products by genome mining for the *lnm*-type gene clusters. (**A**) Identification of 19 *lnm*-type gene clusters from public databases including ~12,000 actinobacterial genomes using DUF-SH didomain as a probe. (**B**) A high-throughput method for the identification of *lnm*-type gene clusters from 5,000 strains of the actinomycetes collection at TSRI (including ~100 from the Naicons' collection and ~500 from Myongji University). Real-time PCR with degenerate primers targeting the 0.6-kb internal fragment of DUF-SH didomain-coding sequence afforded 51 hits based on the analysis of melting curves and gel electrophoresis of the corresponding PCR products. All the 72 hits were further confirmed to be true by sequencing a 1.2-kb PCR-amplified internal fragment of DUF-SH didomain-coding sequence. The 72 hits were subsequently dereplicated to 30 based on the analysis of the 1.2-kb DNA sequences, taxonomy and geographic locations of the hit strains. (**C**) Representative melting curve analysis in real-time PCR. Each peak represents a specific PCR product. Solid line with open circles represents the positive control with gDNA of *S. atroolivaceus* S-140 (LNM producer) as the template, and dashed line represents the negative control without a template. An inset shows PCR products of the representative hits that were analyzed by agarose gel electrophoresis. (**D**) Pie chart representing the taxonomic distribution of the 49 newly identified potential producers, and the original LNM producer *S. atroolivaceus* S-140. Approximately 74% of the hit strains belong to the *Streptomyces* genus. (**E**) Geographic distribution of the hit strains showing *lnm*-type gene clusters are found worldwide.

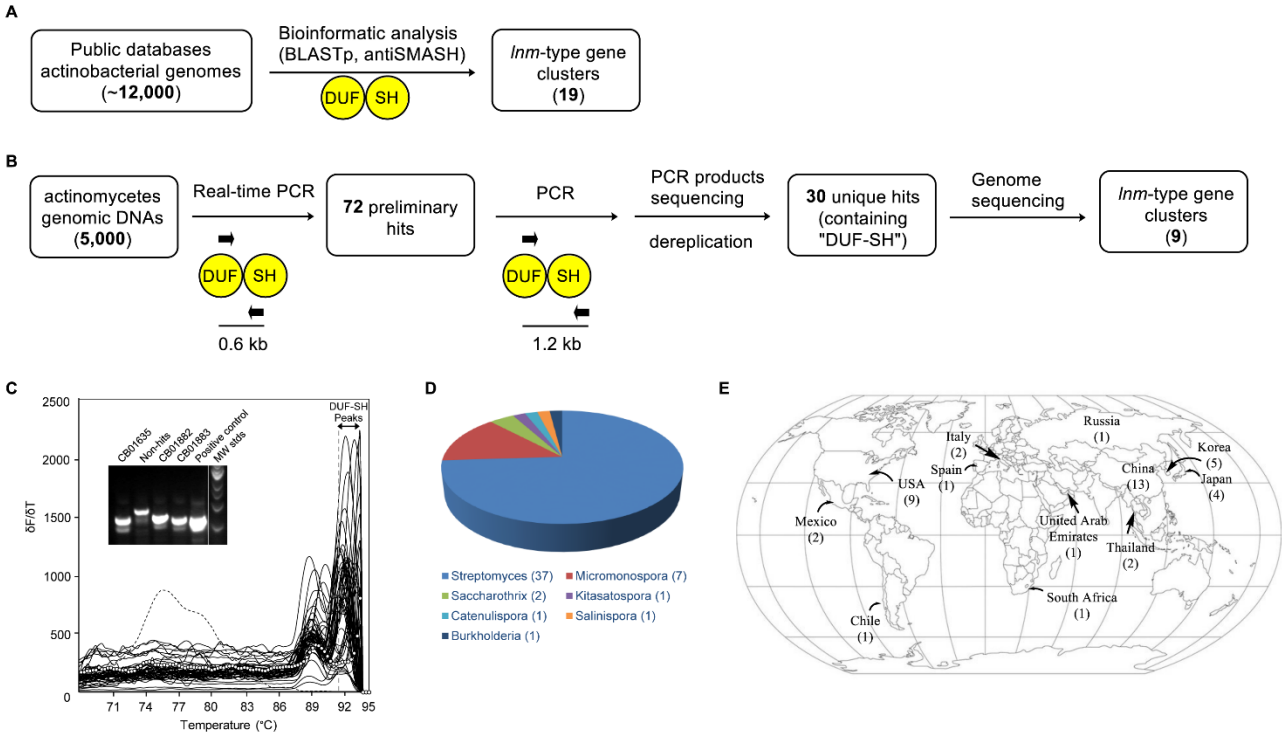
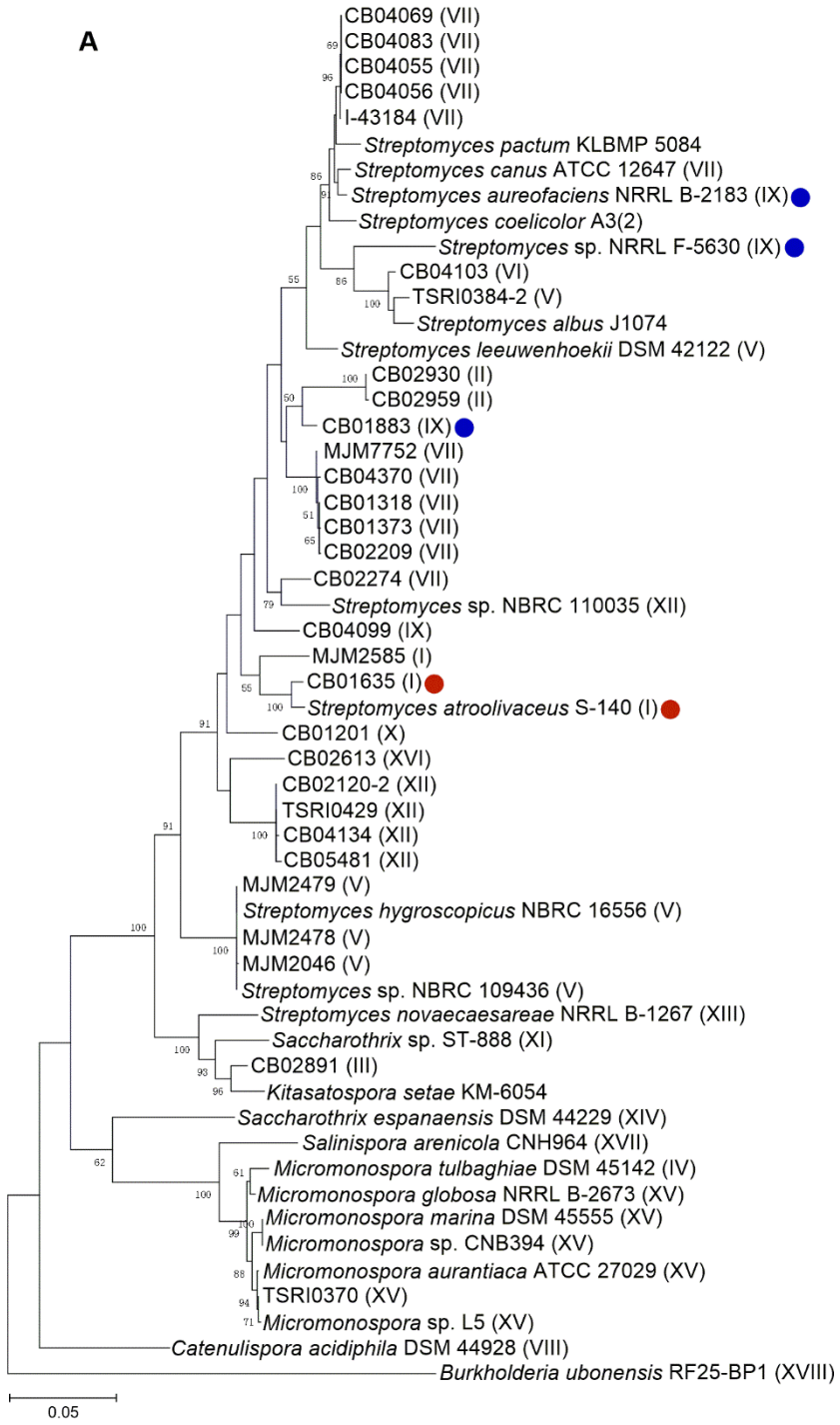
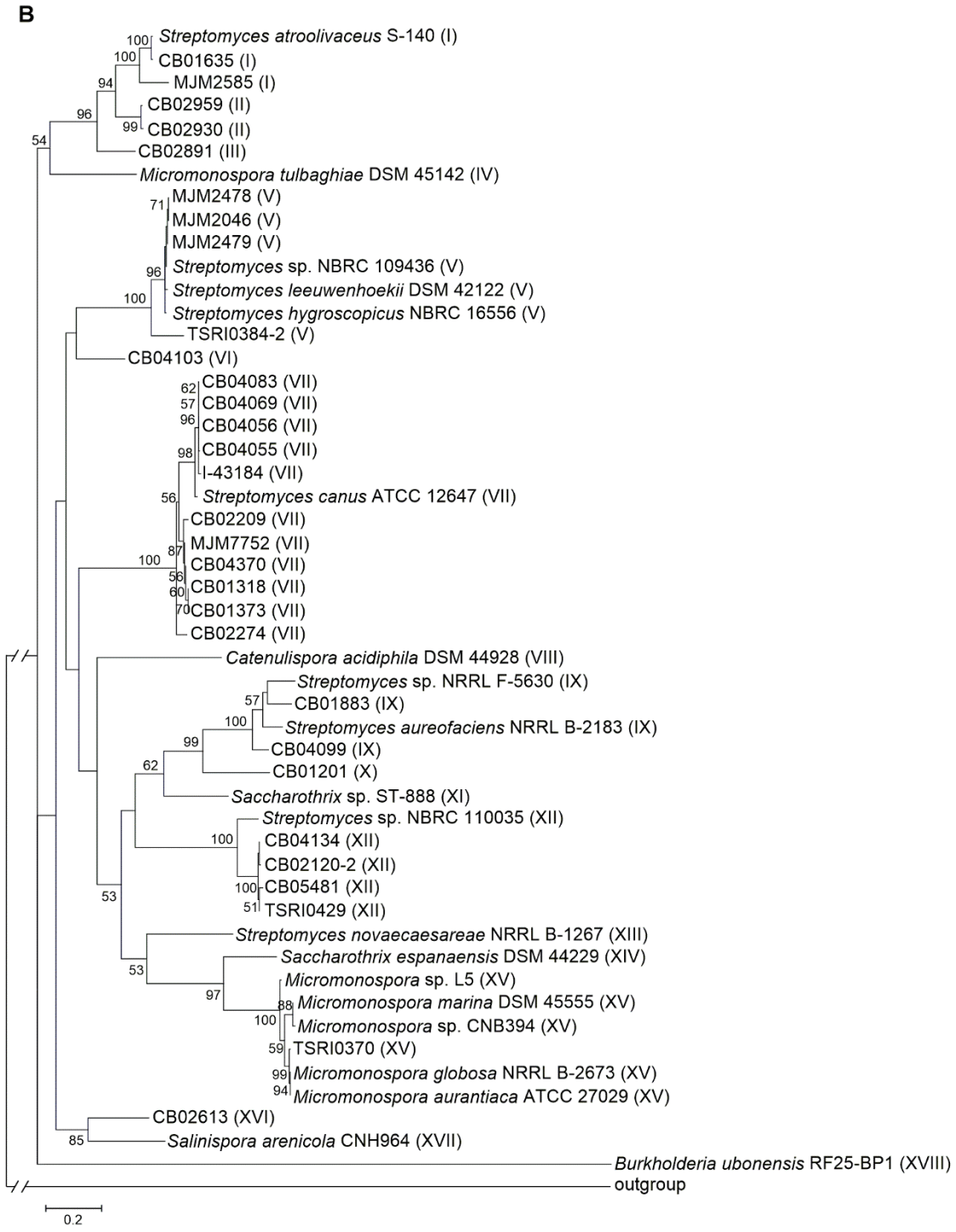
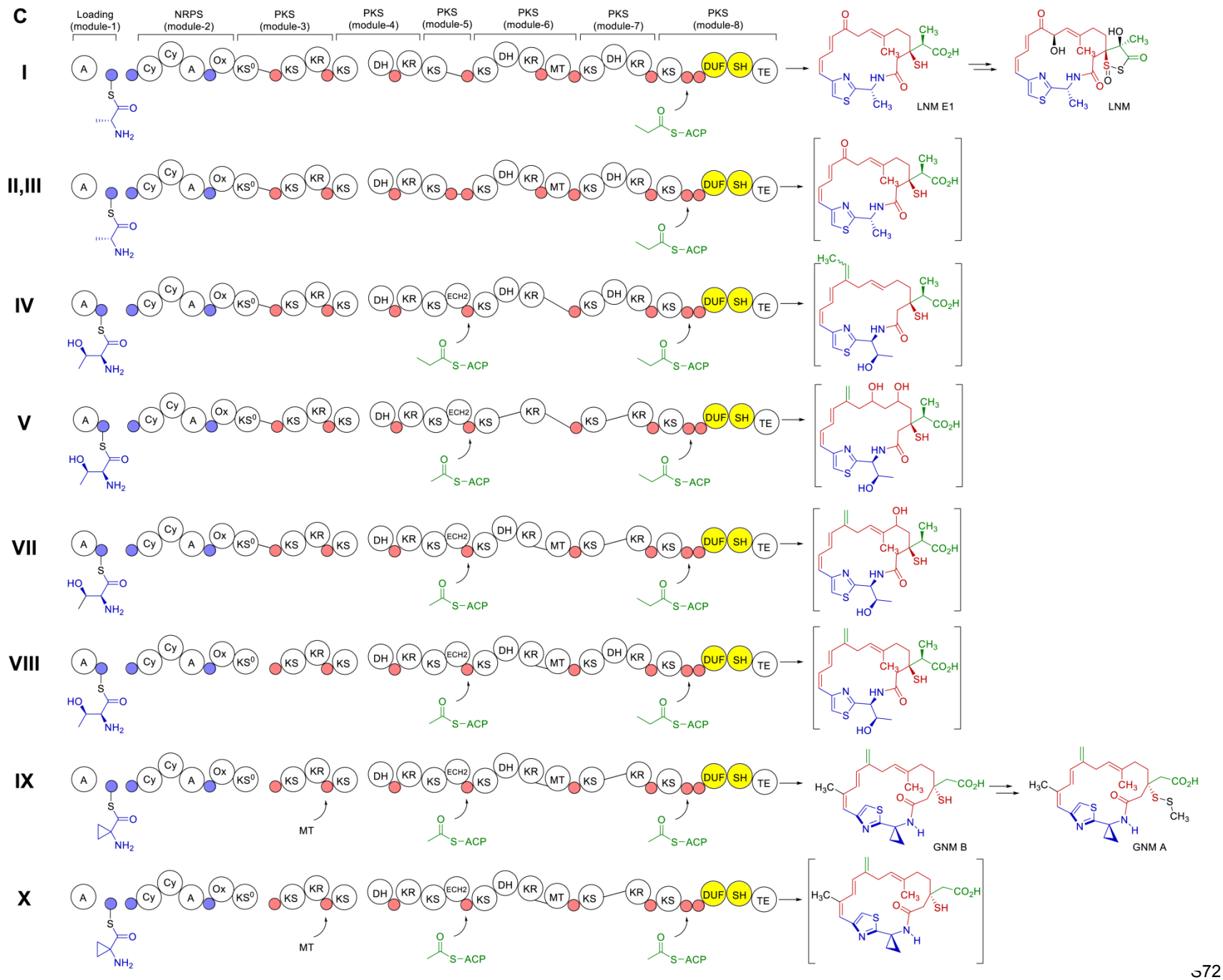


Fig. S3. Phylogenetic analysis of the hit strains, and domain organization and the proposed/identified nascent products of the hybrid NRPS-PKS biosynthetic machineries of LNM-type natural products from different clades. (**A**) Phylogenetic analysis of the taxonomy of the 49 hit strains, as well as the known LNM producer S. atroolivacues S-140, based on the concatenated partial sequences of the three housekeeping genes including 16S rRNA, *recA* and *trpB* (**Table S1**), in comparison with several known strains from databases. Selected strains included *Streptomyces pactum* KLBMP 5084 (NZ_CP016795); *Streptomyces coelicolor* A3(2) (NC_003888); *Streptomyces albus* J1074 (NC_020990); *Kitasatospora setae* KM-6054 (NC_016109). LNM E1 producers are highlighted with red dots. GNMs producers are highlighted with blue dots. The Roman numerals in the parentheses denote the different clades of the *lnm*-type gene clusters (see **Figs. 2** and **S3B** for detailed information). (**B**) Phylogenetic analysis of the 49 *lnm*-type gene cluster harboring strains based on the translated 1.2-kb internal fragment of DUF-SH didomain with S. atroolivaceus S-140 as a reference. CalE6 (AAM94792), methionine γ-lyase from *Micromonospora echiospora*, was used as an outgroup reference. Also see **Fig. 2** in main text for detailed information of *lnm*-type gene clusters. The accession number of LnmJ homologues of hit strains from public databases are as follows: clade IV, *Micromonospora tulbaghiae* DSM 45142 (SCE73393); clade V, *Streptomyces* sp. NBRC 109436 (WP_064455881); *S. hygroscopicus* NBRC 16556 (WP_063816964); *S. leeuwenhoekii* DSM 42122 (CQR60407); clade VII, *S. canus* ATCC 12647 (WP_059300321); clade VIII, *Catenulispora acidiphila* DSM 44928 (ACU71516.1); clade IX, *Streptomyces* sp. NRRL F-5630 (WP_037826198); *S. aureofaciens* NRRL B-2183 (WP_052841990); clade XI, *Saccharothrix* sp. ST-888 (KJK59202.1); clade XII, *Streptomyces* sp. NBRC 110035 (WP_042163782); clade XIII, *S. novaecaesareae* NRRL B-1267 (WP_033331212); clade XIV, *Saccharothrix espanaensis* DSM 44229 (WP_015102008); clade XV, *Micromonospora* sp. L5 (WP_013476297); *Micromonospora* sp. CNB394 (WP_026267551); *M. marina* DSM 45555 (SCE83186); *M. aurantiaca* ATCC 27029 (WP_013287043); *M. globosa* NRRL B-2673 (WP_030268223); clade XVII, *Salinispora arenicola* CNH964 (not available); clade XVIII, *Burkholderia ubonensis* RF25-BP1 (WP_059615564). (**C**) Compared to LNM pathway, variations at module-1 (different priming amino acids), module-3 (incorporation of unusual extender unit and/or trans-acting modification), module-5 (β-alkylation), module-6 and -7 (α-/β-modifications) and module-8 (β-alkylation and C3-stereochemistry) are identified in the newly discovered pathways. Gaps between domains denote protein boundaries; solid lines connect domains in the same protein; “X” denotes the amino acids whose structures are not bioinformatically predicted.


A
CB04069 (VII)
CB04083 (VII)
CB04055 (VII)
CB04056 (VII)
I-43184 (VII)
Streptomyces pactum KLBMP 5084
Streptomyces canus ATCC 12647 (VII)
Streptomyces aureofaciens NRRL B-2183 (IX)
Streptomyces coelicolor A3(2)
Streptomyces sp. NRRL F-5630 (IX)
CB04103 (VI)
TSRI0384-2 (V)
Streptomyces albus J1074
Streptomyces leeuwenhoekii DSM 42122 (V)
CB02930 (II)
CB02959 (II)
CB01883 (IX)
MJM7752 (VII)
CB04370 (VII)
CB01318 (VII)
CB01373 (VII)
CB02209 (VII)
CB02274 (VII)
Streptomyces sp. NBRC 110035 (XII)
CB04099 (IX)
MJM2585 (I)
CB01635 (I)
Streptomyces atroolivaceus S-140 (I)
CB01201 (X)
CB02613 (XVI)
CB02120-2 (XII)
TSRI0429 (XII)
CB04134 (XII)
CB05481 (XII)
MJM2479 (V)
Streptomyces hygroscopicus NBRC 16556 (V)
MJM2478 (V)
MJM2046 (V)
Streptomyces sp. NBRC 109436 (V)
Streptomyces novaecaesareae NRRL B-1267 (XIII)
Saccharothrix sp. ST-888 (XI)
CB02891 (III)
Kitasatospora setae KM-6054
Saccharothrix espanaensis DSM 44229 (XIV)
Salinispora arenicola CNH964 (XVII)
Micromonospora tulbaghiae DSM 45142 (IV)
Micromonospora globosa NRRL B-2673 (XV)
Micromonospora marina DSM 45555 (XV)
Micromonospora sp. CNB394 (XV)
Micromonospora aurantiaca ATCC 27029 (XV)
TSRI0370 (XV)
Micromonospora sp. L5 (XV)
Catenulispora acidiphila DSM 44928 (VIII)
Burkholderia ubonensis RF25-BP1 (XVIII)
0.05


B
Streptomyces atroolivaceus S-140 (I)
CB01635 (I)
MJM2585 (I)
CB02959 (II)
CB02930 (II)
CB02891 (III)
Micromonospora tulbaghiae DSM 45142 (IV)
MJM2478 (V)
MJM2046 (V)
MJM2479 (V)
Streptomyces sp. NBRC 109436 (V)
Streptomyces leeuwenhoekii DSM 42122 (V)
Streptomyces hygroscopicus NBRC 16556 (V)
TSRI0384-2 (V)
CB04103 (VI)
CB04083 (VII)
CB04069 (VII)
CB04056 (VII)
CB04055 (VII)
I-43184 (VII)
Streptomyces canus ATCC 12647 (VII)
CB02209 (VII)
MJM7752 (VII)
CB04370 (VII)
CB01318 (VII)
CB01373 (VII)
CB02274 (VII)
Catenulispora acidiphila DSM 44928 (VIII)
Streptomyces sp. NRRL F-5630 (IX)
CB01883 (IX)
Streptomyces aureofaciens NRRL B-2183 (IX)
CB04099 (IX)
CB01201 (X)
Saccharothrix sp. ST-888 (XI)
Streptomyces sp. NBRC 110035 (XII)
CB04134 (XII)
CB02120-2 (XII)
CB05481 (XII)
TSRI0429 (XII)
Streptomyces novaecaesareae NRRL B-1267 (XIII)
Saccharothrix espanaensis DSM 44229 (XIV)
Micromonospora sp. L5 (XV)
Micromonospora marina DSM 45555 (XV)
Micromonospora sp. CNB394 (XV)
TSRI0370 (XV)
Micromonospora globosa NRRL B-2673 (XV)
Micromonospora aurantiaca ATCC 27029 (XV)
CB02613 (XVI)
Salinispora arenicola CNH964 (XVII)
Burkholderia ubonensis RF25-BP1 (XVIII)
outgroup
0.2


C
Loading (module-1)
NRPS (module-2)
PKS (module-3)
PKS (module-4)
PKS (module-5)
PKS (module-6)
PKS (module-7)
PKS (module-8)
I
II,III
IV
V
VII
VIII
IX
X
LNM E1
LNM
GNM B
GNM A
MT
S–ACP

C (continued)

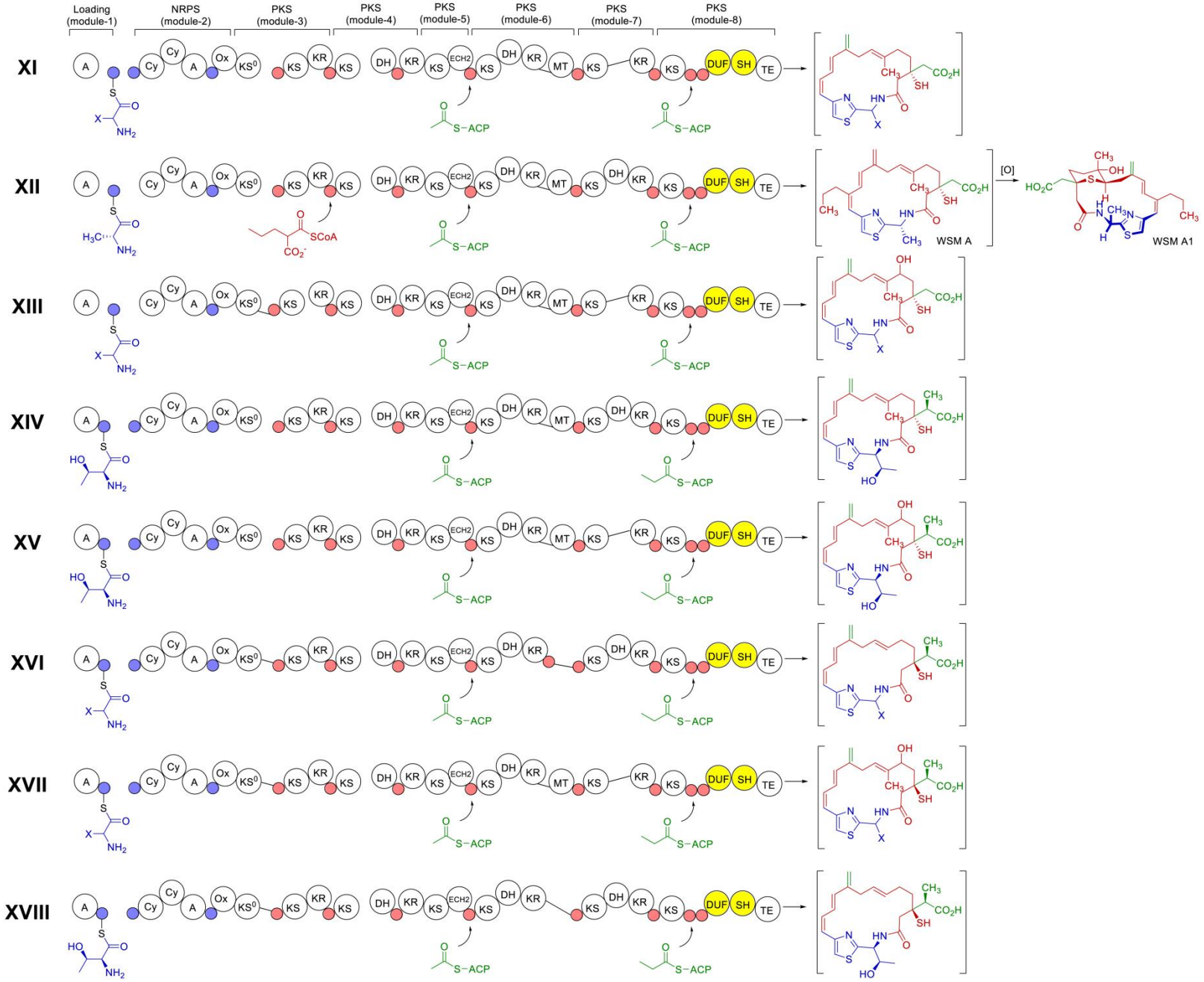
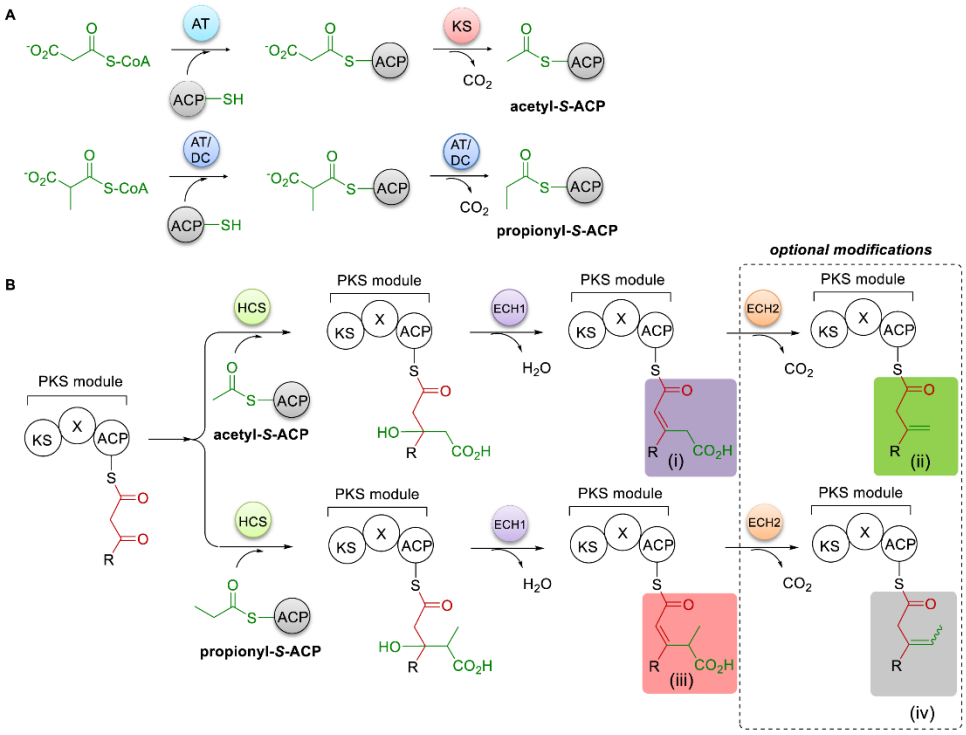

Loading (module-1)
NRPS (module-2)
PKS (module-3)
PKS (module-4)
PKS (module-5)
PKS (module-6)
PKS (module-7)
PKS (module-8)
XI
XII
XIII
XIV
XV
XVI
XVII
XVIII
WSM A
WSM A1
[O]

Fig. S4. The variations of β-branch pathway in the biosynthesis of LNM-type natural products. (**A**) Biosynthetic pathways of acetyl-*S*-ACP and propionyl-*S*-ACP. (**B**) Installation of different structural moieties by β-alkyl branching processes during polyketide biosynthesis. HCS, hydroxymethylglutaryl-CoA synthase; ECH1: enoyl-CoA hydratase 1; ECH2: enoyl-CoA hydratase 2. "X" denotes the potential accessory domains in PKS module. (**C**) β-alkylation related enzymes from the 17 representative LNM-type pathways (one representative pathway from each clade, except clade VI, was shown) and the corresponding structural outcomes at C-3 and/or C-9 of the final products. "**#**" denotes the trans-acting AT that was proposed to be responsible for loading malonyl-CoA to each ACP of the six PKS modules. The "star" mark indicates that ECH2 was integrated into the PKS as a domain. The Roman numerals in the first column denote the different clades of the *lnm*-type gene clusters (see **Fig. S3B** and **Fig. 2** in main text for detailed information).



C

	AT	KS	AT/DC	ACP	HCS	ECH1	ECH2*	Structural outcome
I	-	-	LnmK	LnmL	LnmM	LnmF	-	C-3 (iii)
II	-	-	CB02959_J	CB02959_K	CB02959_L	CB02959_E	-	C-3 (iii)
III	-	-	CB02891_Q	CB02891_R	CB02891_S	CB02891_L	-	C-3 (iii)
IV	-	-	Mtul_M	Mtul_N	Mtul_O	Mtul_H	-	C-3 (iii)
	-	-	Mtul_M	Mtul_N	Mtul_O	Mtul_H	Mtul_L	C-9 (iv)
V	-	-	TSRI0384-2_M	TSRI0384-2_N	TSRI0384-2_O	TSRI0384-2_E	-	C-3 (iii)
	TSRI0384-2_F#	TSRI0384-2_Q	-	TSRI0384-2_P	TSRI0384-2_D	TSRI0384-2_E	TSRI0384-2_L	C-9 (ii)
VII	-	-	CB01373_V	CB01373_W	CB01373_X	CB01373_N	-	C-3 (iii)
	CB01373_O#	CB01373_Z	-	CB01373_Y	CB01373_M	CB01373_N	CB01373_U	C-9 (ii)
VIII	-	-	Caci_W	Caci_X	Caci_Y	Caci_T	-	C-3 (iii)
	Caci_K#	Caci_Z2	-	Caci_Z1	Caci_S	Caci_T	Caci_V	C-9 (ii)
IX	GnmG	GnmW	-	GnmV	GnmU	GnmF	-	C-3 (i)
	GnmG	GnmW	-	GnmV	GnmE	GnmF	GnmT	C-9 (ii)
X	CB01201_G	CB01201_W	-	CB01201_V	CB01201_U	CB01201_F	-	C-3 (i)
	CB01201_G	CB01201_W	-	CB01201_V	CB01201_E	CB01201_F	CB01201_T	C-9 (ii)
XI	Sast_J	Sast_Z3	-	Sast_Z2	Sast_Z1	Sast_I	-	C-3 (i)
	Sast_J	Sast_Z3	-	Sast_Z2	Sast_H	Sast_I	Sast_Y	C-9 (ii)
XII	WsmF	WsmU	-	WsmT	WsmS	WsmE	-	C-3 (i)
	WsmF	WsmU	-	WsmT	WsmD	WsmE	WsmR	C-9 (ii)
XIII	Snov_C	Snov_Z3	-	Snov_Z2	Snov_Z1	Snov_B	-	C-3 (i)
	Snov_C	Snov_Z3	-	Snov_Z2	Snov_A	Snov_B	Snov_Y	C-9 (ii)
XIV	-	-	Saes_L	Saes_M	Saes_N	Saes_G	-	C-3 (iii)
	Saes_F	Saes_Z3	-	Saes_Z2	Saes_H	Saes_G	Saes_K	C-9 (ii)
XV	-	-	Maur_K	Maur_L	Maur_M	Maur_G	-	C-3 (iii)
	Maur_F	Maur_Z2	-	Maur_Z1	Maur_B	Maur_G	Maur_J	C-9 (ii)
XVI	-	-	CB02613_U	CB02613_V	CB02613_W	CB02613_Z4	-	C-3 (iii)
	CB02613_I#	CB02613_Y	-	CB02613_X	CB02613_Z5	CB02613_Z4	CB02613_T	C-9 (ii)
XVII	-	-	Sal964_Y	Sal964_Z1	Sal964_Z2	Sal964_O	-	C-3 (iii)
	Sal964_P#	Sal964_Z4	-	Sal964_Z3	Sal964_N	Sal964_O	Sal964_X	C-9 (ii)
XVIII	-	-	Bubo_S	Bubo_T	Bubo_U	Bubo_K	-	C-3 (iii)
	Bubo_L	Bubo_W		Bubo_V	Bubo_J	Bubo_K	Bubo_R	C-9 (ii)

Fig. S5. Phylogenetic analysis of the hydroxymethylglutaryl-CoA synthases (HCSs) from LNM and 28 newly discovered LNM-type biosynthetic pathways. The analysis revealed two major groups of HCSs, which are proposed for the installation of the β-alkyl branches at C3 and C9, respectively. The group of HCSs responsible for the installation of β-alkyl branch at C3 tends to be clustered into two subgroups that using propionyl-*S*-ACP and acetyl-*S*-ACP as substrates, respectively. The second group of HCSs responsible for the installation of β-alkyl branch at C9 is predicted to use acetyl-*S*-ACP as the substrate. The HCS from *Staphylococcus aureus* (AAG02422) was used as an outgroup reference. HCSs from the pathways with their corresponding natural products being discovered are highlighted in red. The Roman numerals in parentheses denote the different clades of the *lnm*-type gene clusters (see **Fig. S3B and Fig.2** in main text for detailed information).

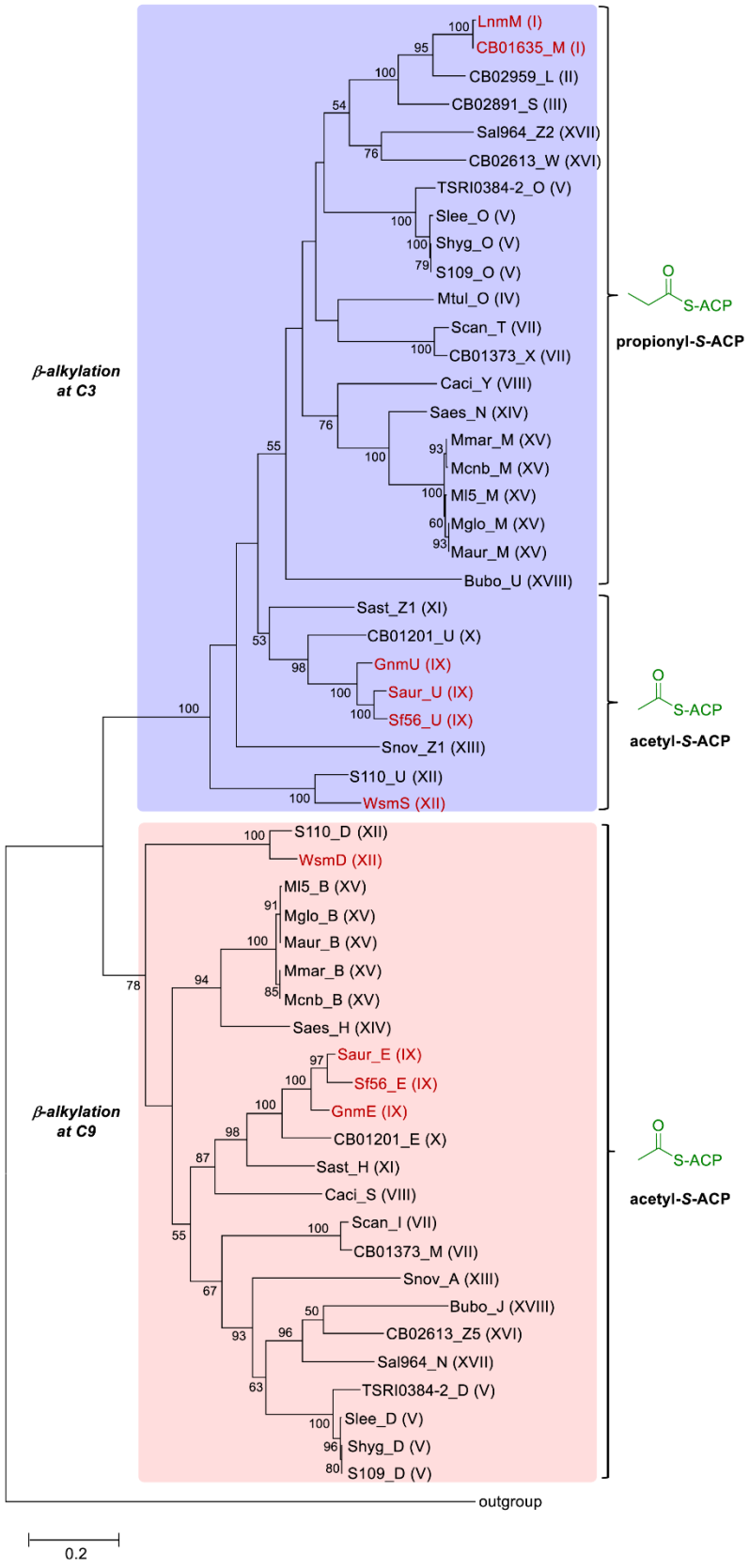

β-alkylation at C3
β-alkylation at C9
LnmM (I)
CB01635_M (I)
CB02959_L (II)
CB02891_S (III)
Sal964_Z2 (XVII)
CB02613_W (XVI)
TSRI0384-2_O (V)
Slee_O (V)
Shyg_O (V)
S109_O (V)
Mtul_O (IV)
Scan_T (VII)
CB01373_X (VII)
Caci_Y (VIII)
Saes_N (XIV)
Mmar_M (XV)
Mcnb_M (XV)
Ml5_M (XV)
Mglo_M (XV)
Maur_M (XV)
Bubo_U (XVIII)
Sast_Z1 (XI)
CB01201_U (X)
GnmU (IX)
Saur_U (IX)
Sf56_U (IX)
Snov_Z1 (XIII)
S110_U (XII)
WsmS (XII)
S110_D (XII)
WsmD (XII)
Ml5_B (XV)
Mglo_B (XV)
Maur_B (XV)
Mmar_B (XV)
Mcnb_B (XV)
Saes_H (XIV)
Saur_E (IX)
Sf56_E (IX)
GnmE (IX)
CB01201_E (X)
Sast_H (XI)
Caci_S (VIII)
Scan_I (VII)
CB01373_M (VII)
Snov_A (XIII)
Bubo_J (XVIII)
CB02613_Z5 (XVI)
Sal964_N (XVII)
TSRI0384-2_D (V)
Slee_D (V)
Shyg_D (V)
S109_D (V)
outgroup
0.2
O
S-ACP
propionyl-S-ACP
acetyl-S-ACP
acetyl-S-ACP

Fig. S6. Phylogenetic analysis of domain of unknown function (DUF domain), and cysteine lyase (SH) domains from LNM and 28 LNM-type biosynthetic pathways. (**A**) The analysis revealed that DUF domains tend to cluster into two groups (highlighted with light blue and light red, respectively). One group of DUF domains are from pathways with C3 β-alkylation using propionyl-S-ACP as the predicted substrate, and the other group are from pathways with C3 β-alkylation using acetyl-S-ACP as the predicted substrate. MprF (EWS52466), the phosphatidylglycerol lysyltransferase from *Methylibium* sp. T29, was used as an outgroup reference. (**B**) The analysis revealed that the SH domains tend to cluster into two groups (highlighted with light red and light blue, respectively). One group of SH domains are from pathways with C3 β-alkylation using acetyl-S-ACP as the predicted substrate, and the other group are from pathways with C3 β-alkylation using propionyl-S-ACP as the predicted substrate. TPL-1 (ABI75028), the tyrosine phenol-lyase from *Citrobacter freundii*, was used as an outgroup reference. The DUF domains/SH domains from the pathways with their corresponding natural products being discovered are highlighted in red. The Roman numerals in parentheses denote the different clades of the *lnm*-type gene clusters (see **Fig. S3B and Fig. 2** in main text for detailed information). It should be noted that the configurations at C3 of LNM-type natural products discovered from clade I (LNM) are opposite to those discovered from clades IX (GNM) and XII (WSM).

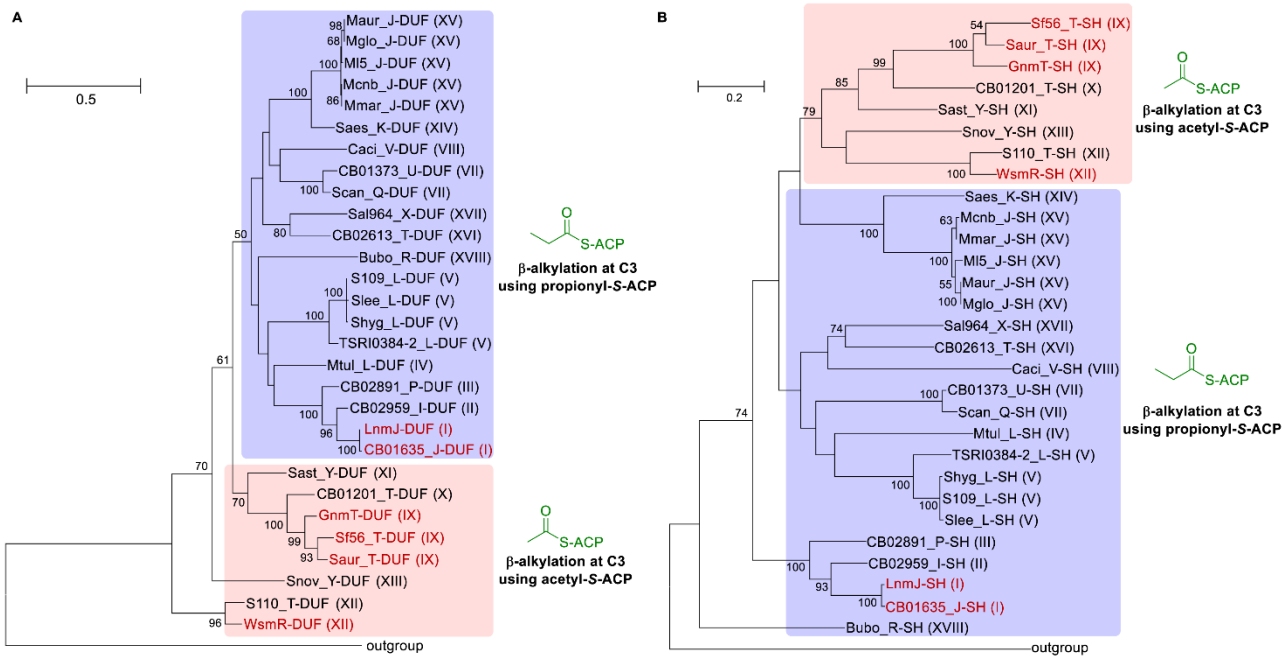


Fig. S7. Genome neighborhood network (GNN) analysis (*E* value of 10^{-6}) of *lnm* and other 28 new discovered *lnm*-type gene clusters (see **Fig. 2** in main text for detailed information) shows that both conserved and diverse proteins are involved in the biosynthesis of LNM-type natural products. The enzymes from the LNM, GNM and WSM biosynthetic pathways are highlighted with red, yellow, and green dots, respectively. The relatively conserved enzymes with unknown functions marked with "**1**" are proposed to be involved in the post-assembly-line modification steps. In comparison with the LNM pathway, the newly discovered pathways feature various new enzymes, e.g. **2**: discrete methyltransferases; **5**: truncated hemoglobin; **8**: halogenases; **9**: FAD-dependant oxidoreductases; **14**: class-III aminotransferases; **15**: radical SAM superfamily proteins; **19**: aminotransferases; **11** and **12**: unknown functions. Other distinct families of enzymes in the newly discovered pathways include **3** and **18**: glyoxalase/bleomycin resistance protein; **4**: phosphopantetheinyl transferases; **6**: YdeI-like protein; **7**: ferredoxin; **10**: sodium:proton exchanger; **13**: ABC transporter; **16**: TetR family transcriptional regulator; **17**: dTDP-4-dehydrorhamnose 3,5-epimerase.

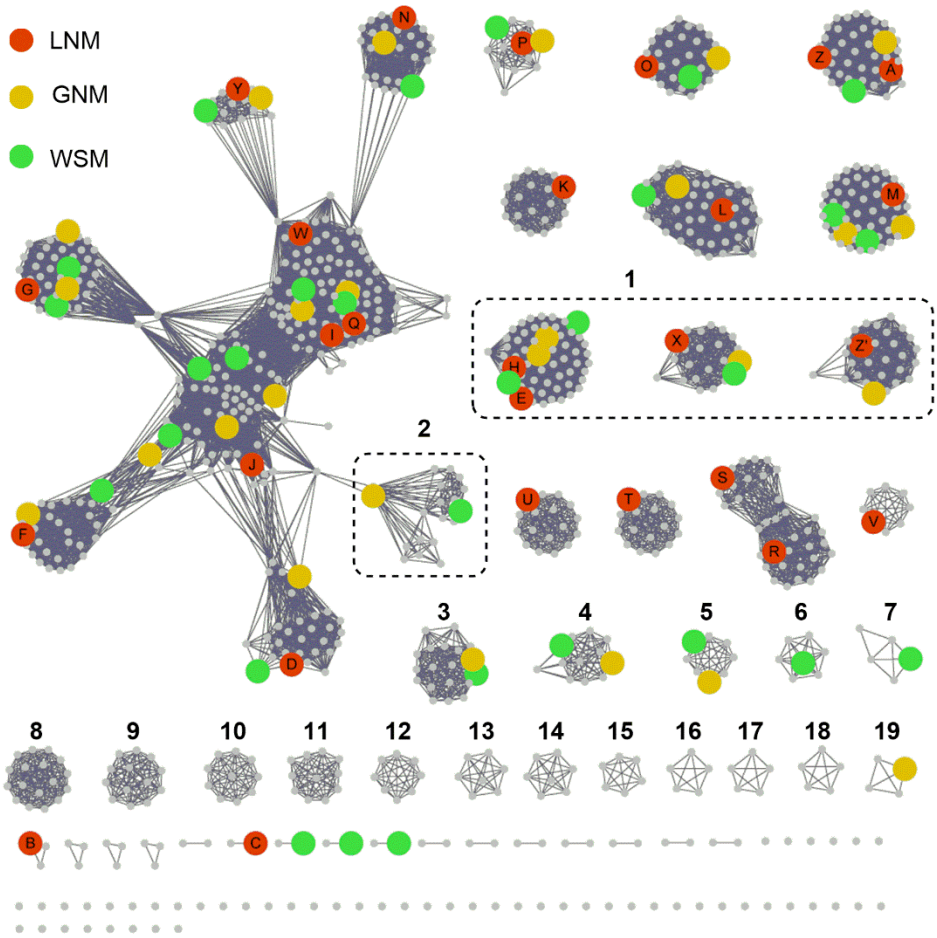


Fig. S8. Construction of the in-frame deletion mutant strain SB21001 (i.e. Δ*gnmB*) and confirmation of its genotype by Southern analysis. (**A**) Schematic representation for the deletion of *gnmB* in strain CB01883 by homologous recombination. The probe (517 bp) for Southern blot was amplified with primers orf155-SBlot-5n and orf155-SBlot-3n using genomic DNA of strain CB01883 as the template. Apr^S, apramycin sensitive; Apr^R, apramycin resistance; *acc(3)IV*, apramycin resistance gene. (**B**) Southern blot verification of strain CB01883 wild-type (4.8 kb) and SB21001 (2.7 kb). Genomic DNAs were digested with *Eco*NI and then used for Southern analysis. Lane 1, DNA marker VII, DIG-labeled (Roche); lane 2, strain CB01883 wild type; lane 3, SB21001.

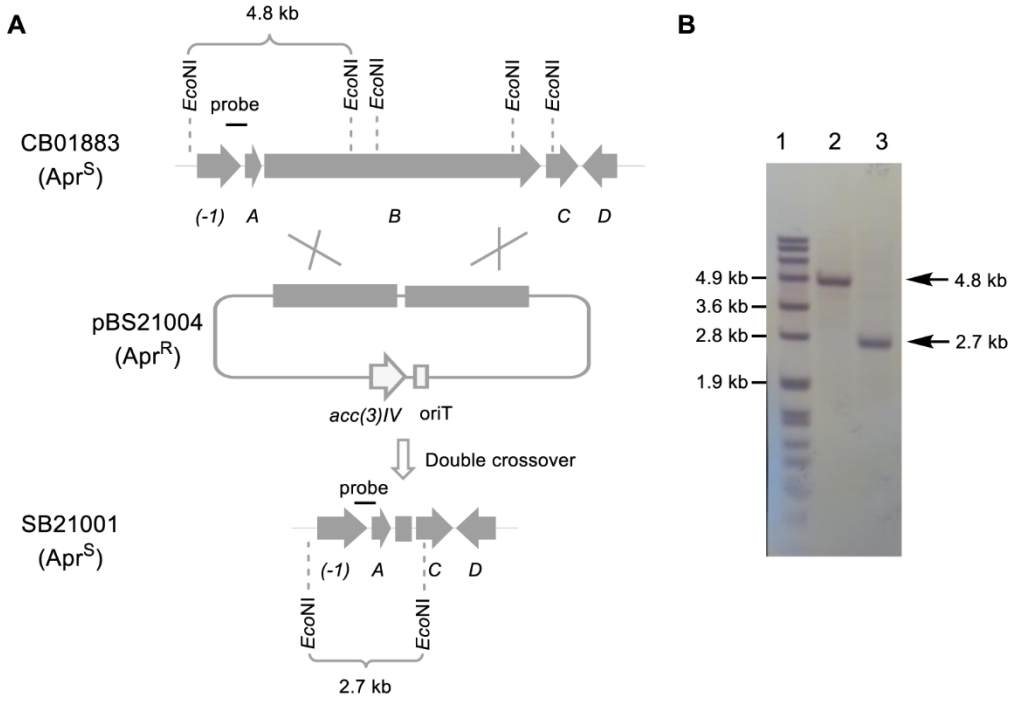


Fig. S9. Proposed biosynthetic pathway of guangnanmycin (GNM) in *S*. sp. CB01883. (**A**) Genetic organization of *gnm* gene cluster, together with two cosmids pBS21001 (1C4) and pBS21002 (1B9) covering the whole gene cluster. (**B**) The biosynthetic machinery of GNM featuring a similar hybrid NRPS-AT-less type I PKS to that of LNM consisting of GnmSRBHT and a discrete AT GnmK, producing GNM B as the nascent product, which was further converted to GNM by tailoring enzymes. (**C**) A set of enzymes including GnmWGVEF and ECH2 domain in PKS module-5 are proposed to be responsible for the installation of the β-alkyl branch at module-5, leading to the formation of a terminal double bond at C9. (**D**) A set of enzymes including GnmWGVUF are proposed to be responsible for the installation of the β-alkyl branch at module-8, leading to the formation of carboxymethyl moiety at C3. Also, the DUF-SH didomain was proposed to be responsible for the installation of 3-thiol group.

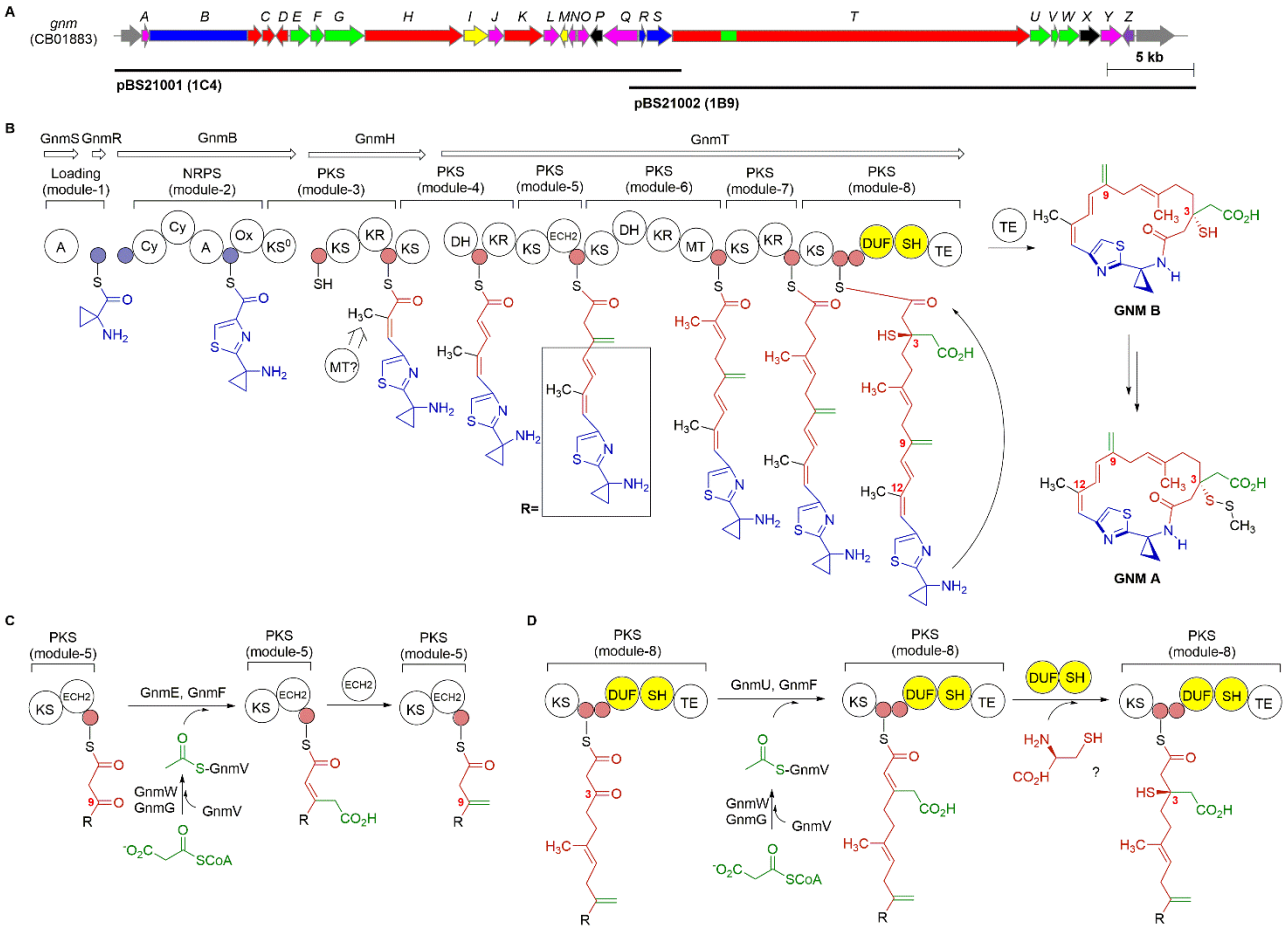


Fig. S10. Elucidation of new metabolites isolated from strain CB01883 wild type and mutant strain SB21003 (i.e. Δ*gnmO*). Key ^{1}H-^{1}H COSY, HMBC and ROESY correlations are shown.

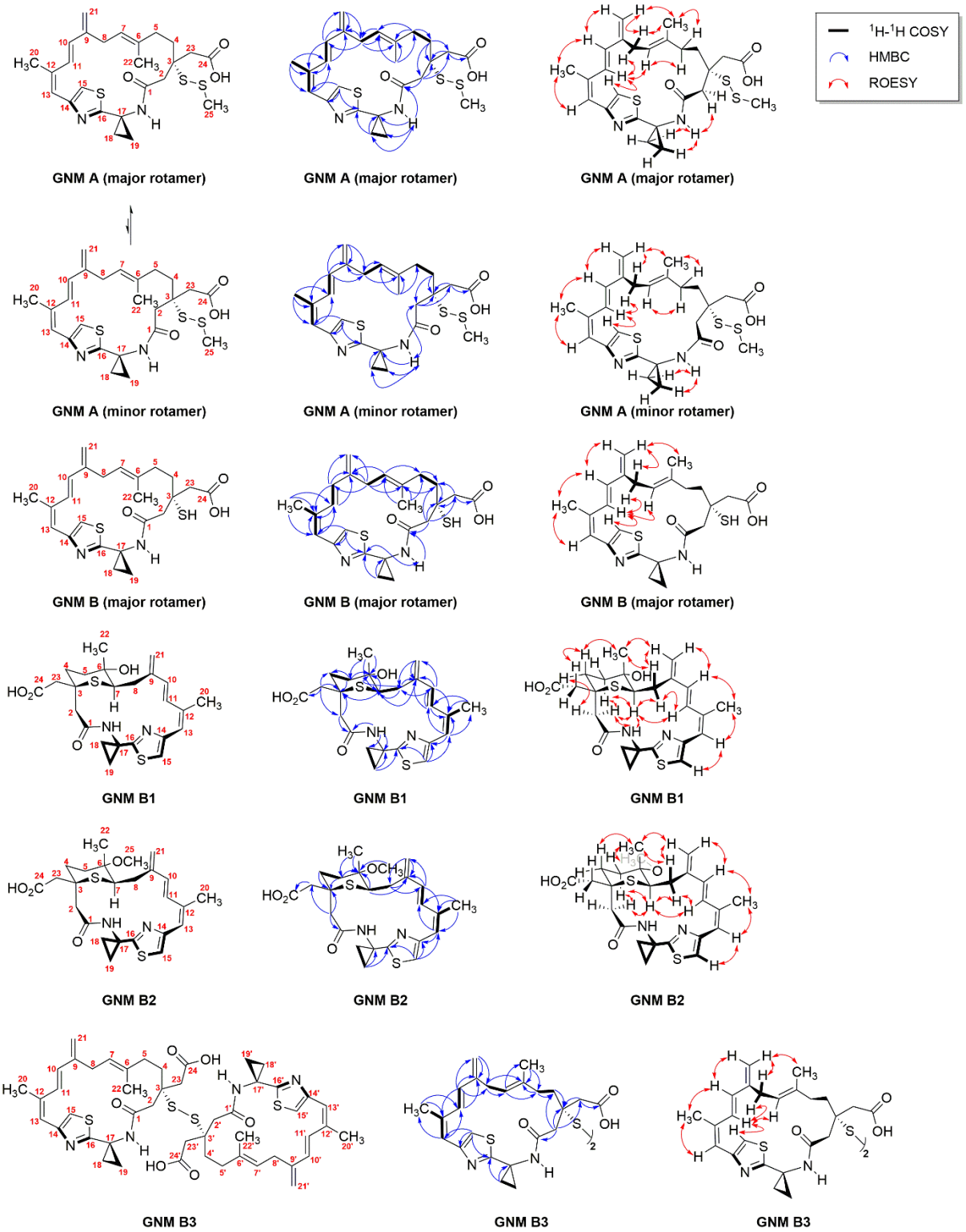


Fig. S11A. ^{1}H NMR spectra of GNM A (700 MHz, DMSO-d_6)

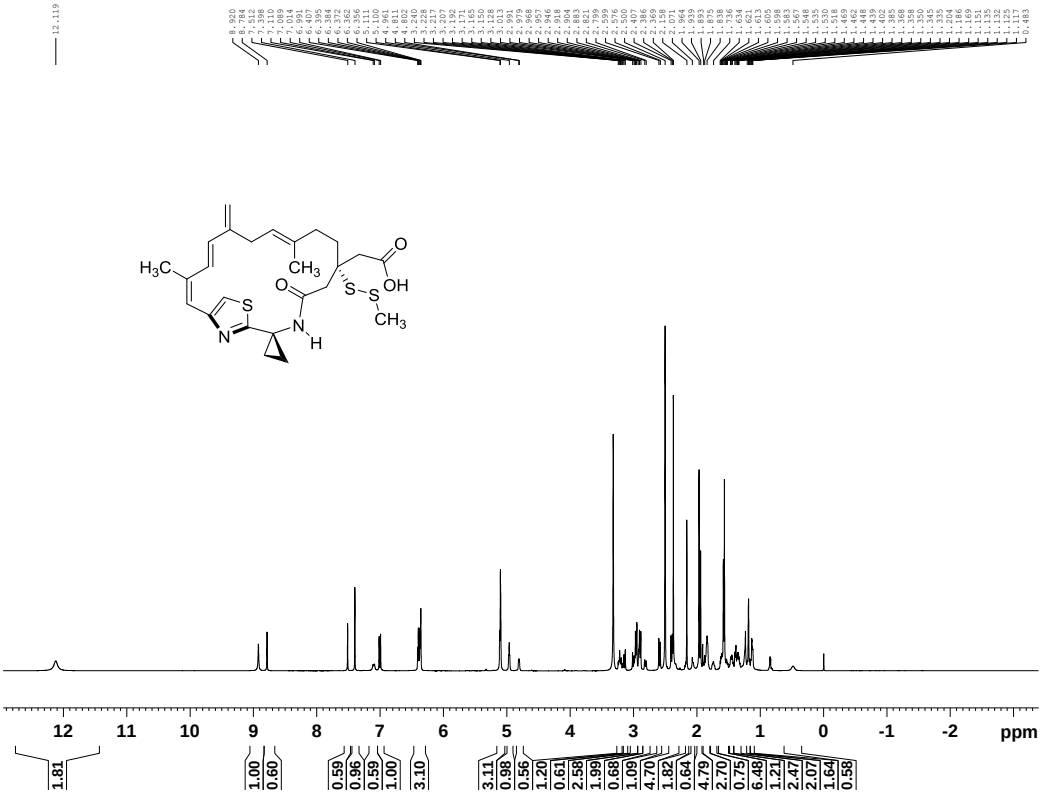


Fig. S11B. ^{13}C NMR spectra of GNM A (175 MHz, DMSO-d_6)

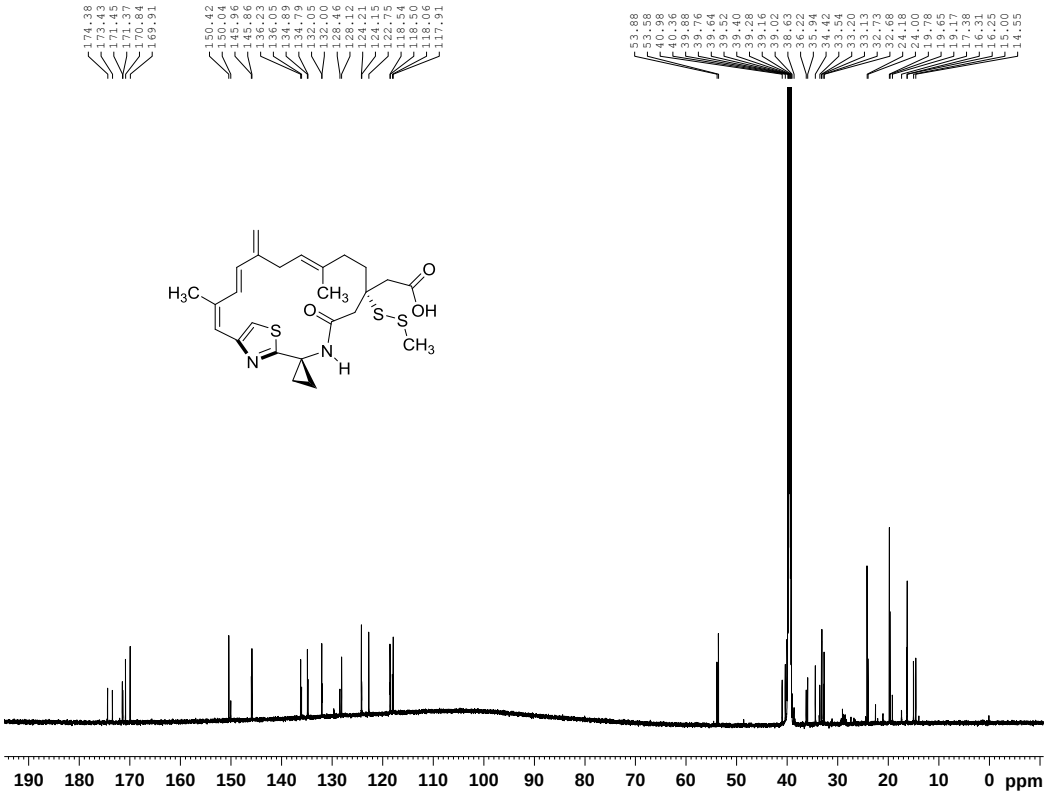


Fig. S11C. DEPT-135 spectra of GNM A (DMSO-d_6)

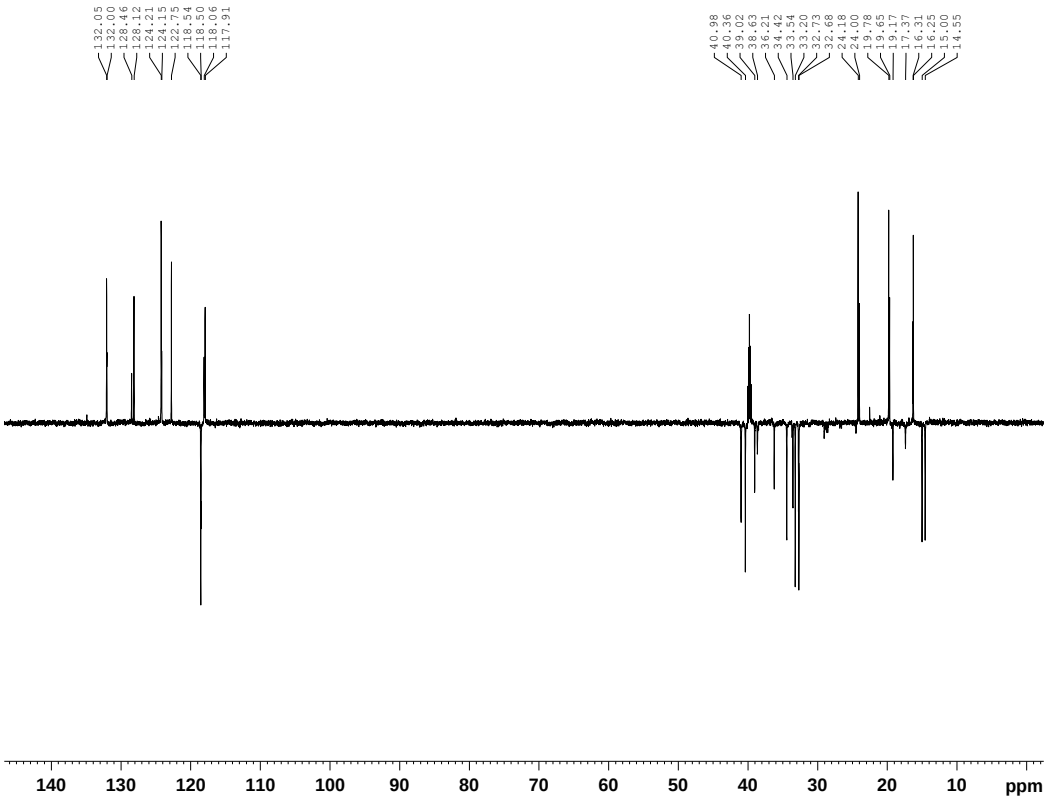


Fig. S11D. HSQC spectra of GNM A (DMSO-d_6)

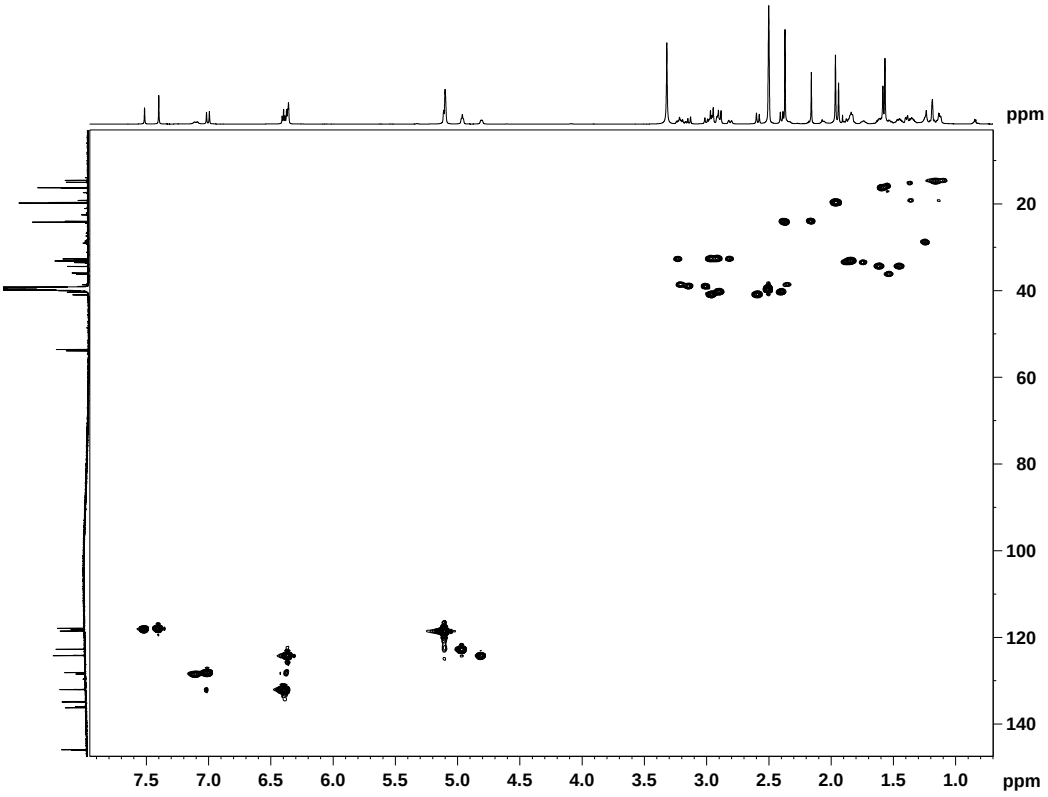


Fig. S11E. ^{1}H-^{1}H COSY spectra of GNM A (DMSO-d_6)

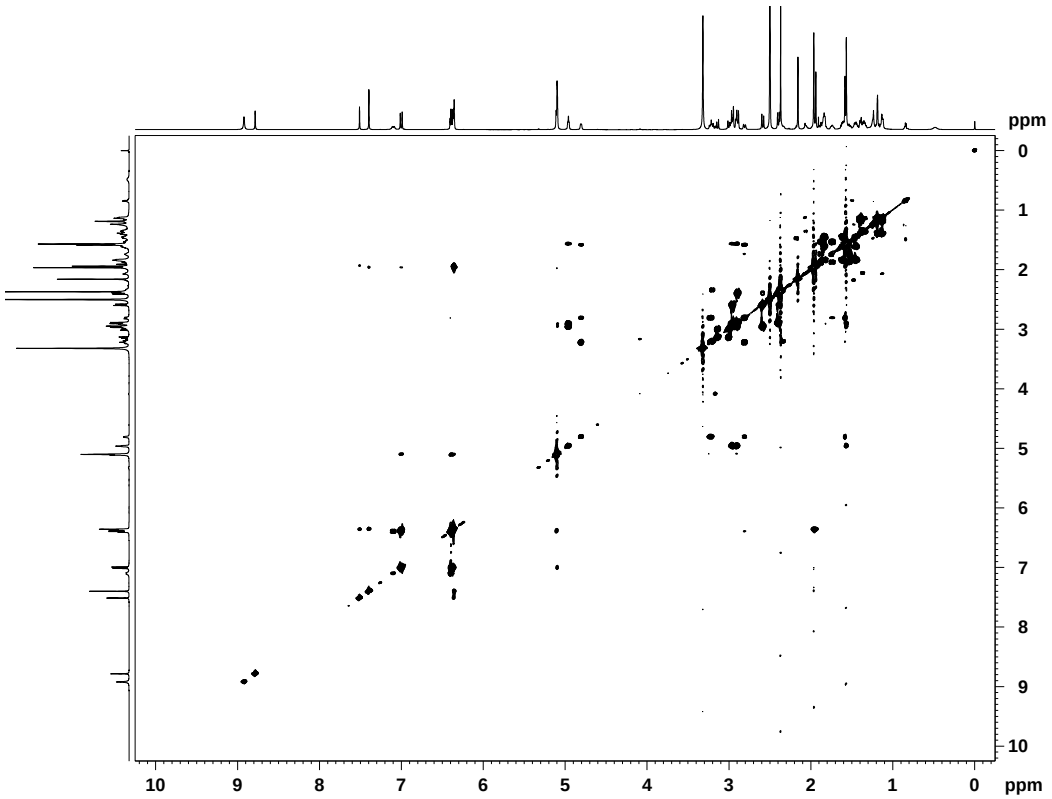


Fig. S11F. HMBC spectra of GNM A (DMSO-d_6)

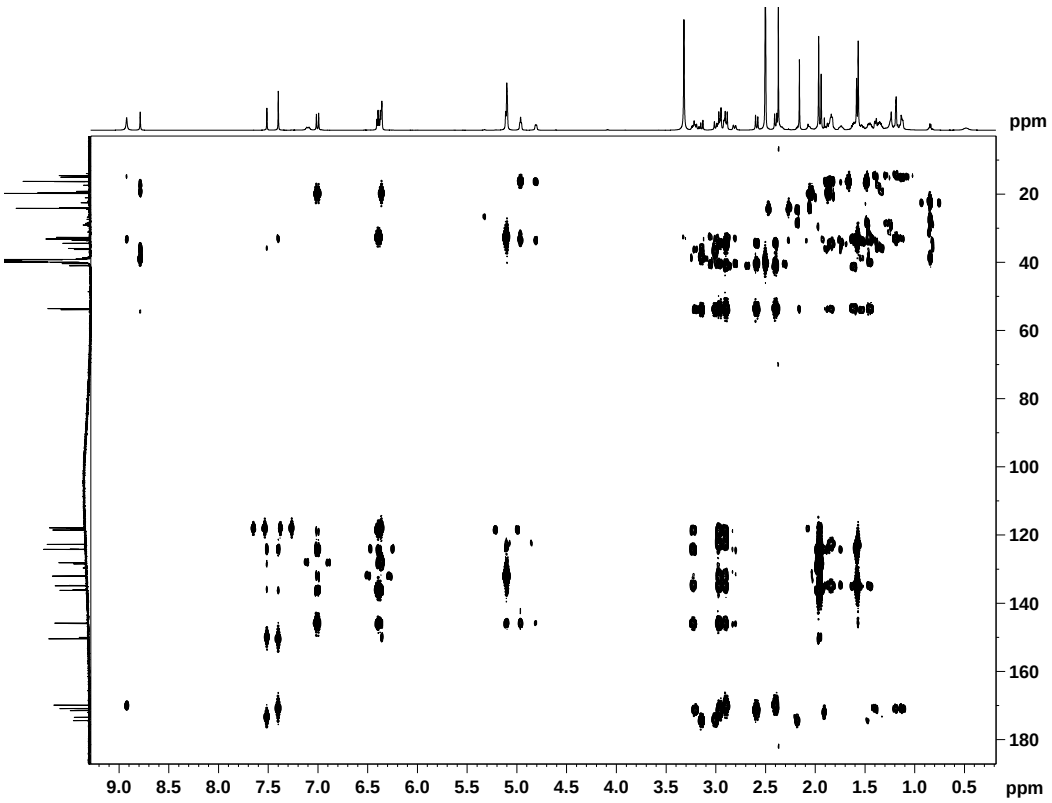


Fig. S11G. ROESY spectra of GNM A (DMSO-d_6)

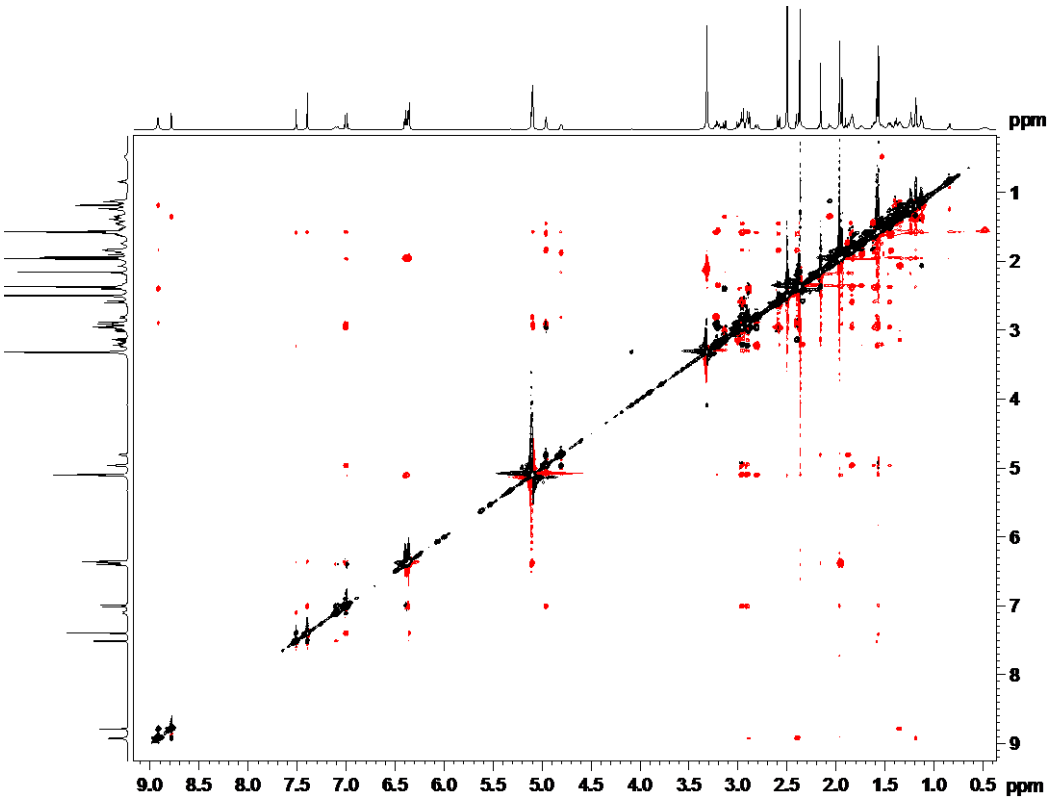


Fig. S11H. 2D-EXSY (21) signals (black) on ROESY spectra (red) of GNM A (DMSO-d_6)

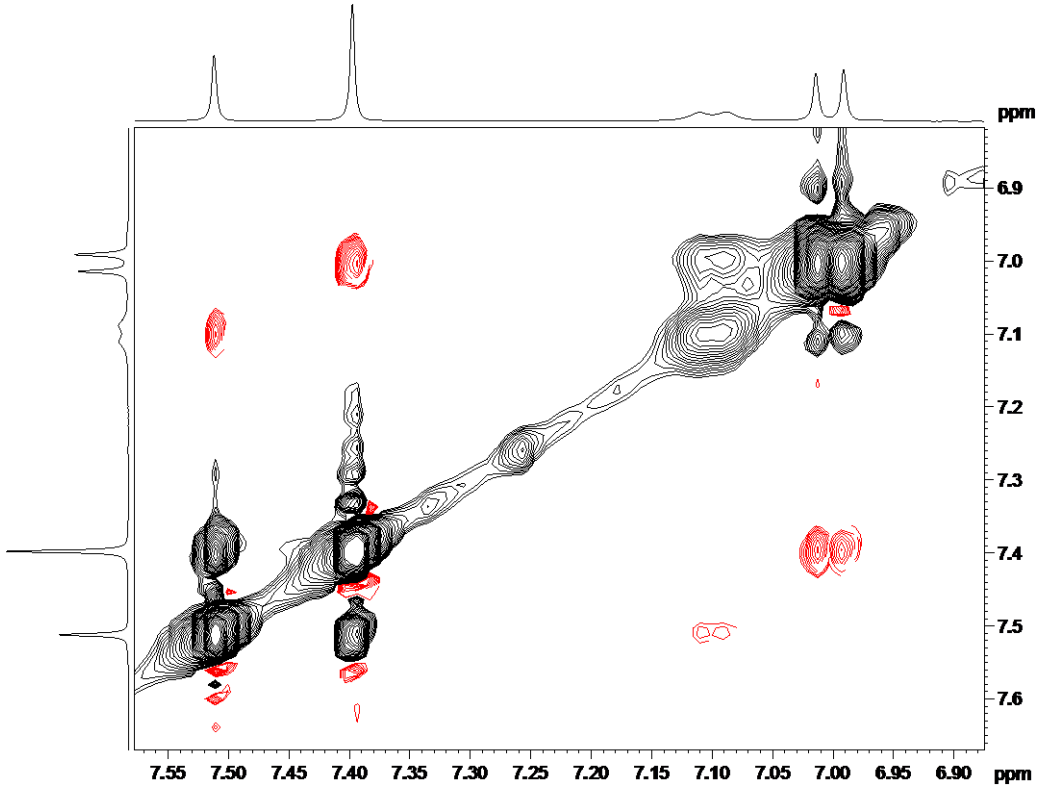


Fig. S11I. 2D-EXSY signals (black) on ROESY spectra (red) of GNM A (DMSO-d_6)

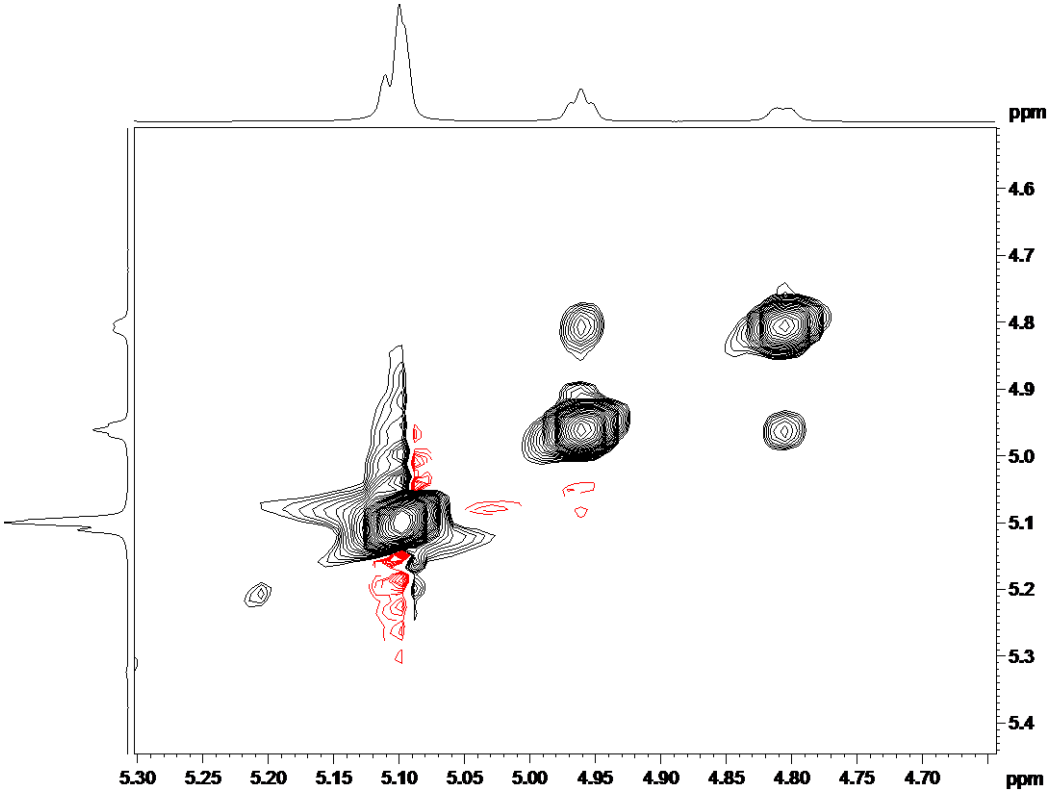


Fig. S11J. 2D-EXSY signals (black) on ROESY spectra (red) of GNM A (DMSO-d_6)

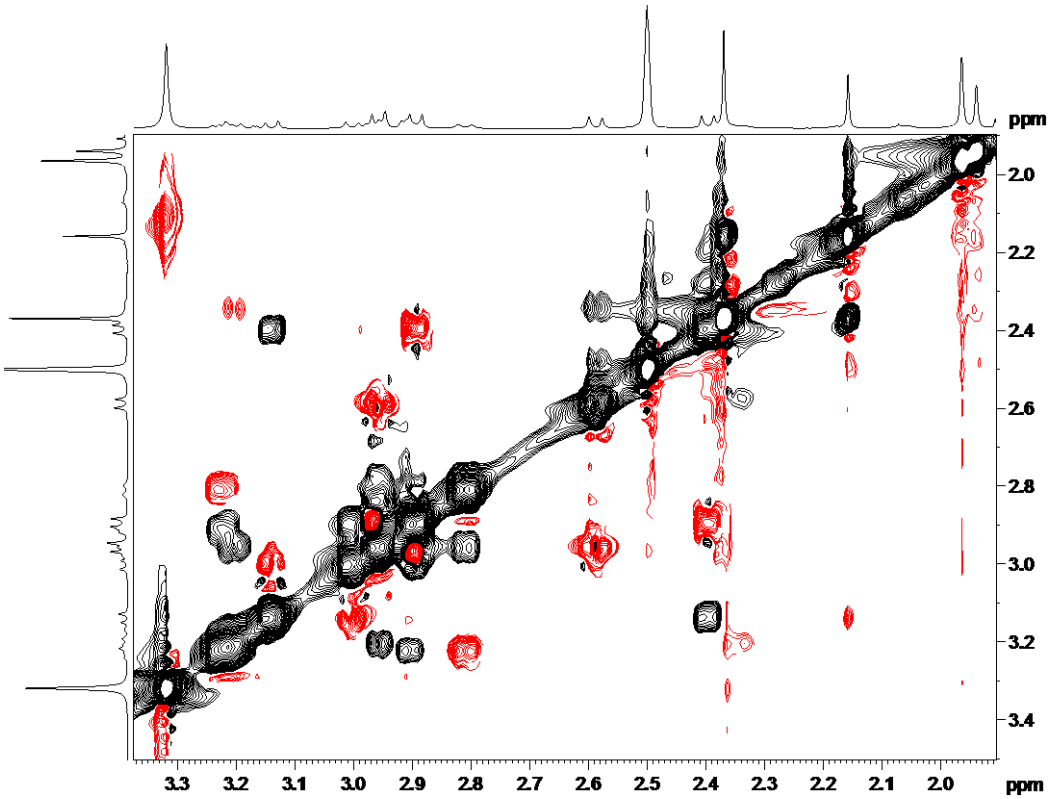


Fig. S11K. 1D-EXSY signals on 1D-NOE spectra (22) of GNM A (DMSO-d_6, 400 MHz) (red arrow showing the irradiated proton)

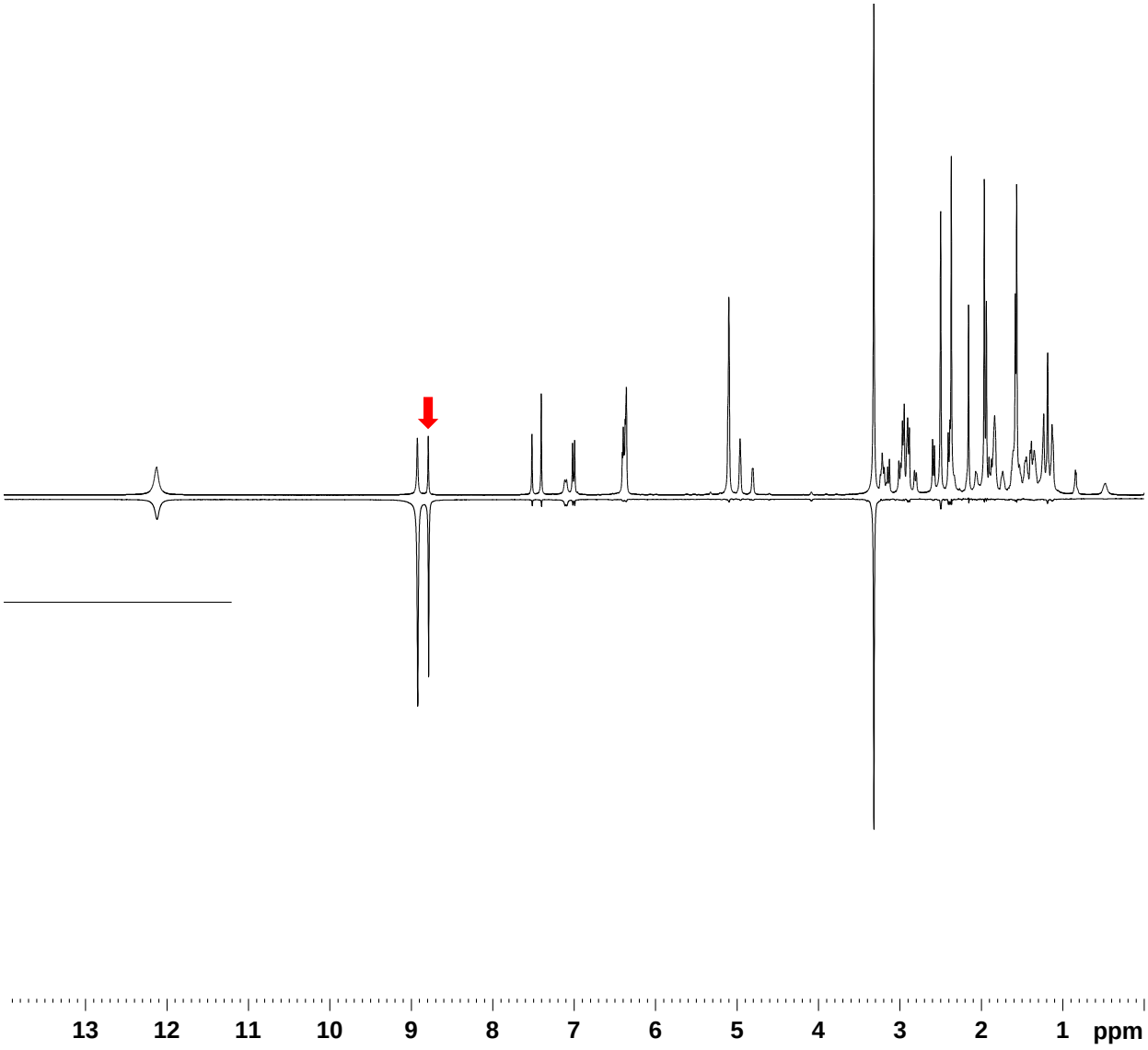


Fig. S11L. 1D-EXSY signals on 1D-NOE spectra of GNM A (DMSO-d_6, 400 MHz) (red arrow showing the irradiated proton)

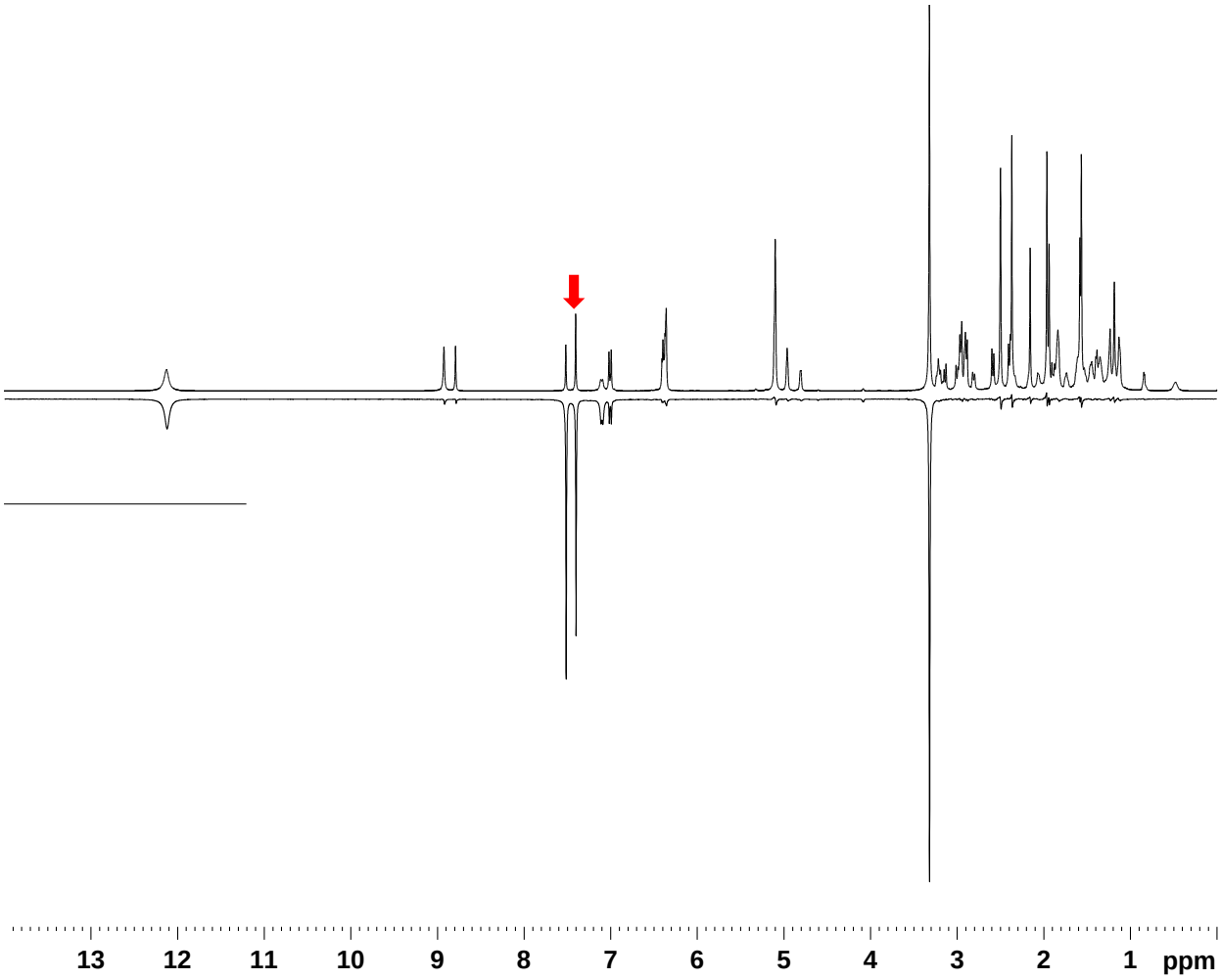


Fig. S11M. ^{1}H NMR spectra of GNM A at varied temperatures (400 MHz, DMSO-d_6)

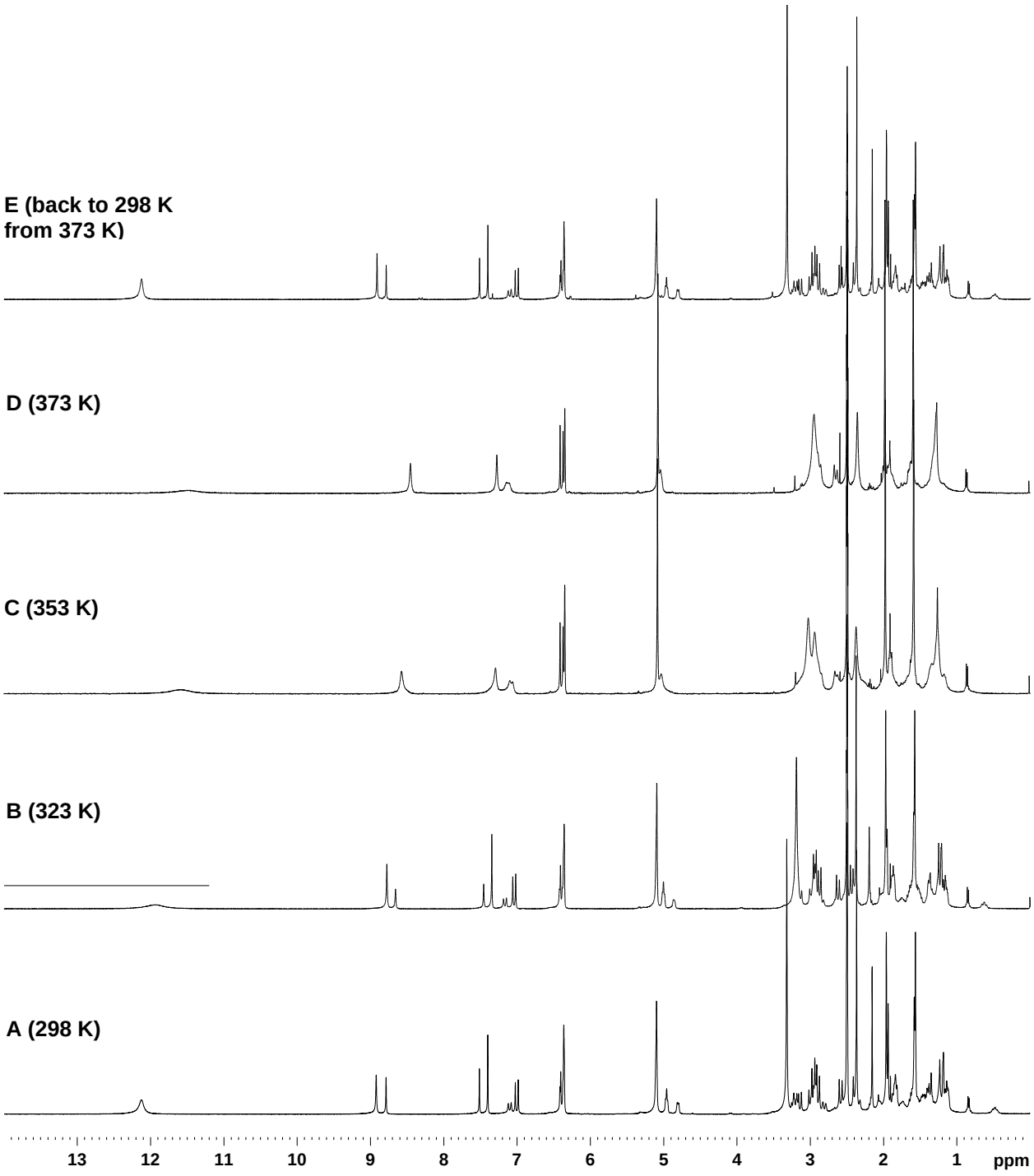


Fig. S11N. HR-MS (ESI) spectra of GNM A

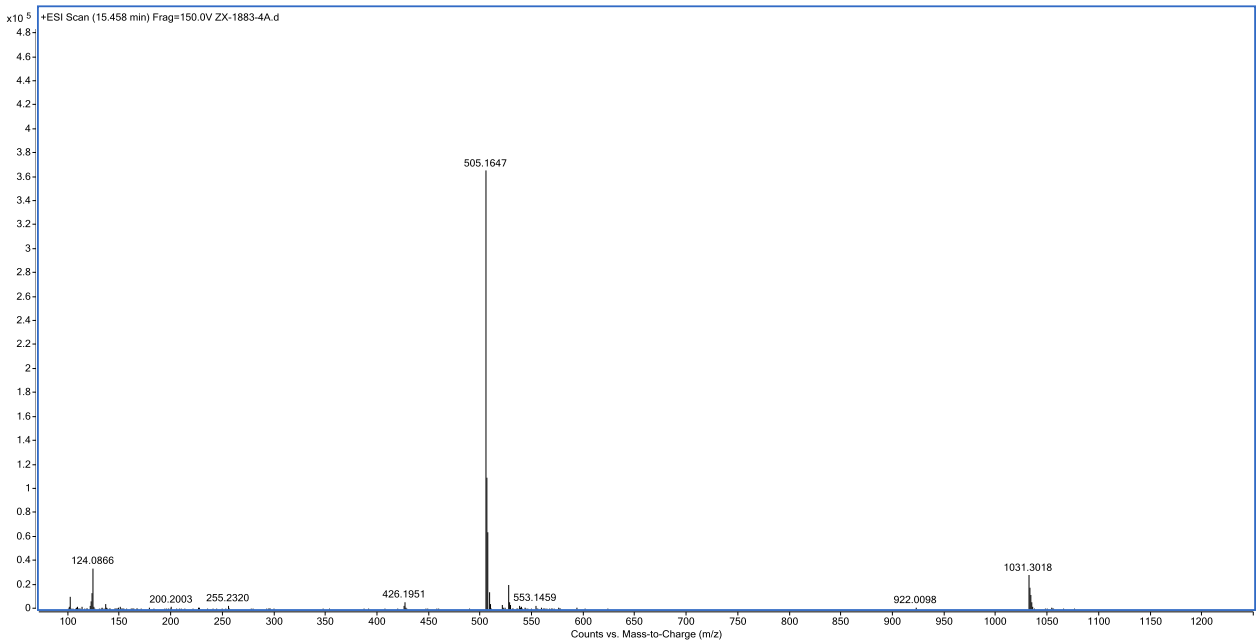


Fig. S12A. ^{1}H NMR spectra of GNM B (700 MHz, DMSO-d_6)

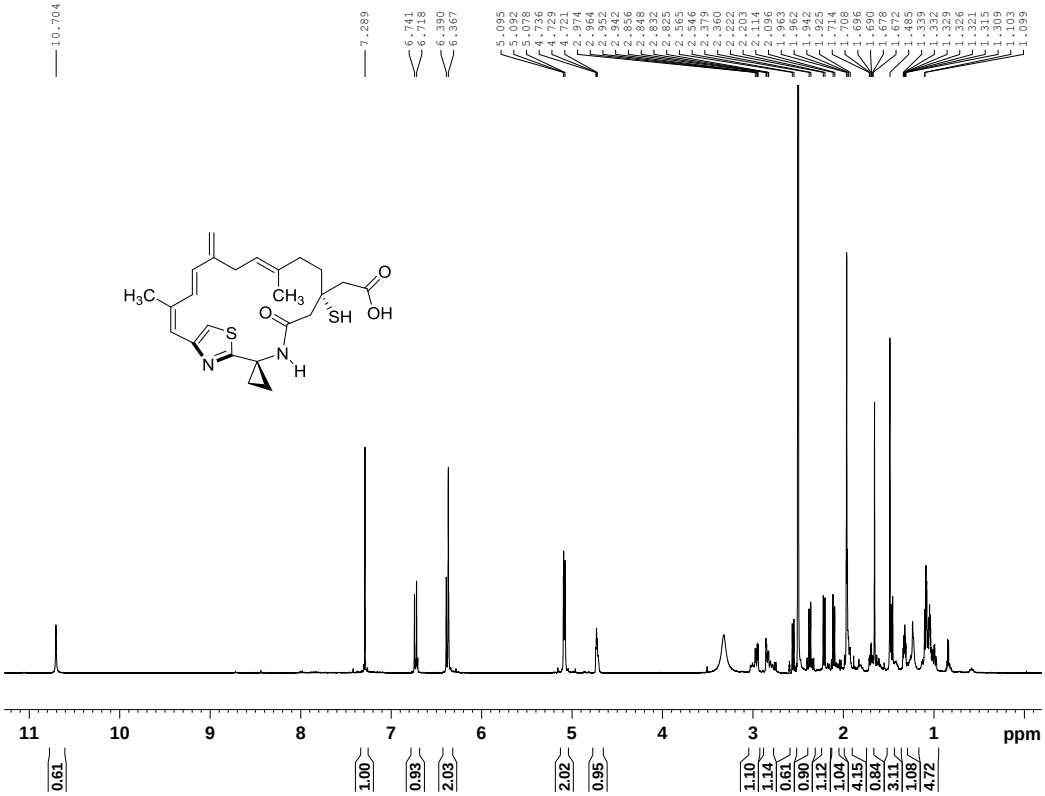


Fig. S12B. ^{13}C NMR spectra of GNM B (175 MHz, DMSO-d_6)

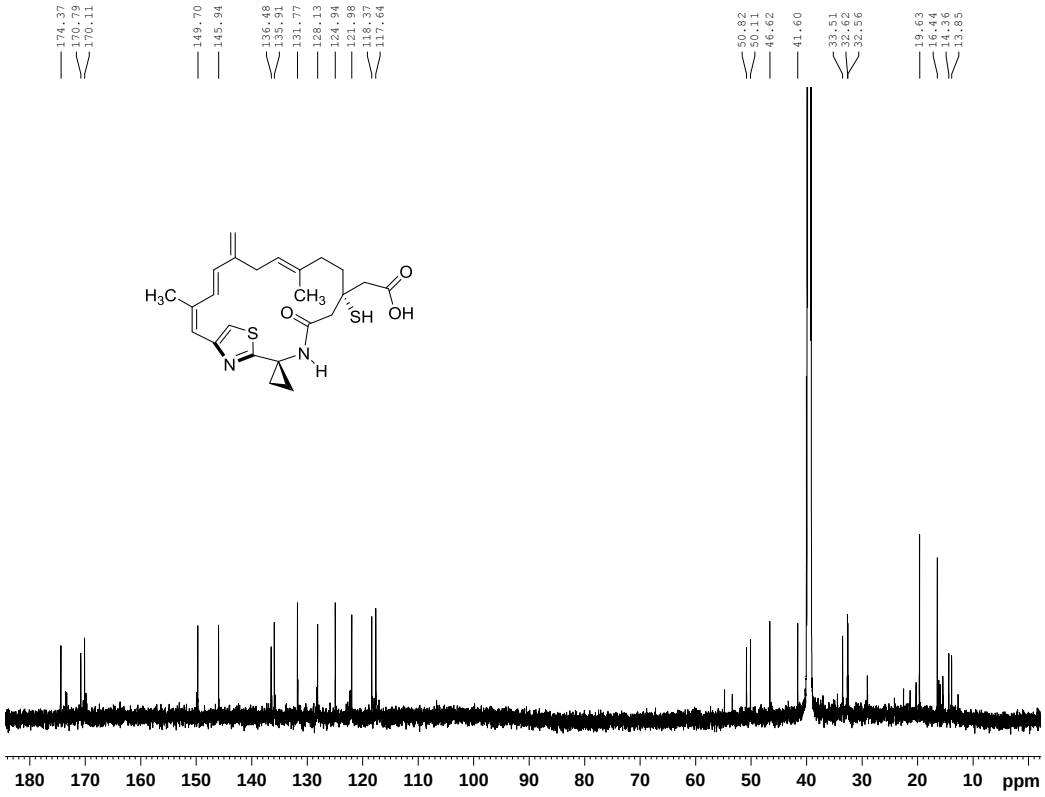


Fig. S12C. DEPT-135 spectra of GNM B (DMSO-d_6)

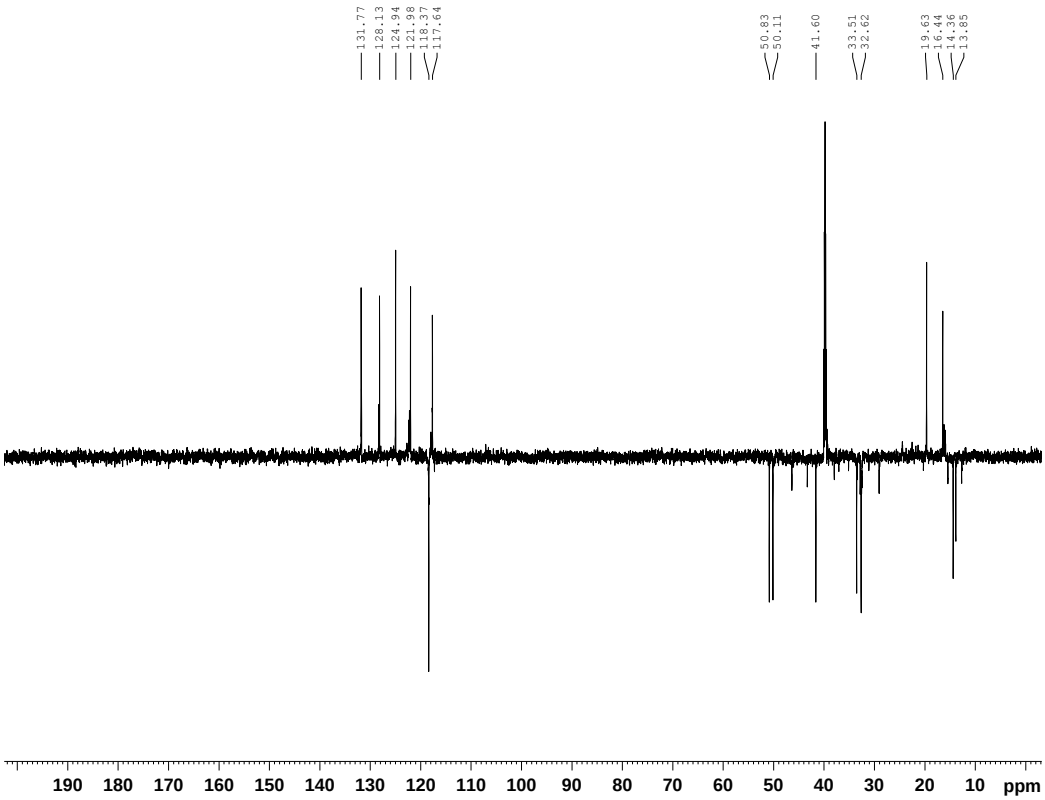


Fig. S12D. HSQC spectra of GNM B (DMSO-d_6)

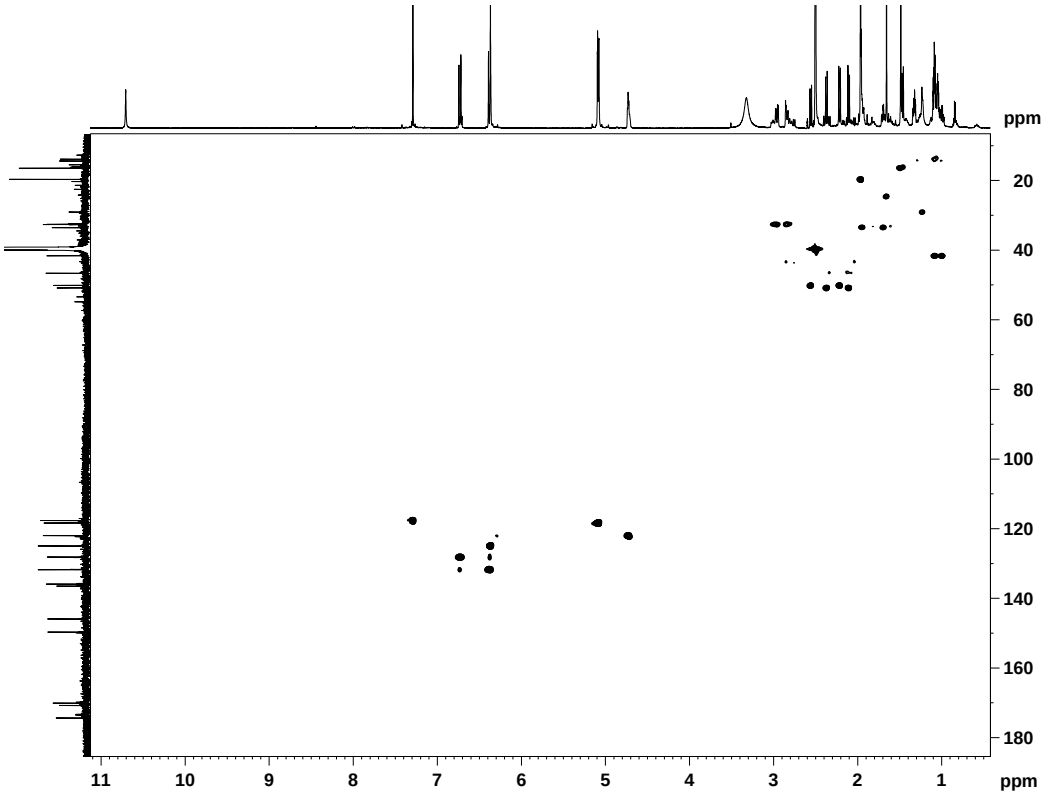


Fig. S12E. ^{1}H-^{1}H COSY spectra of GNM B (DMSO-d_6)

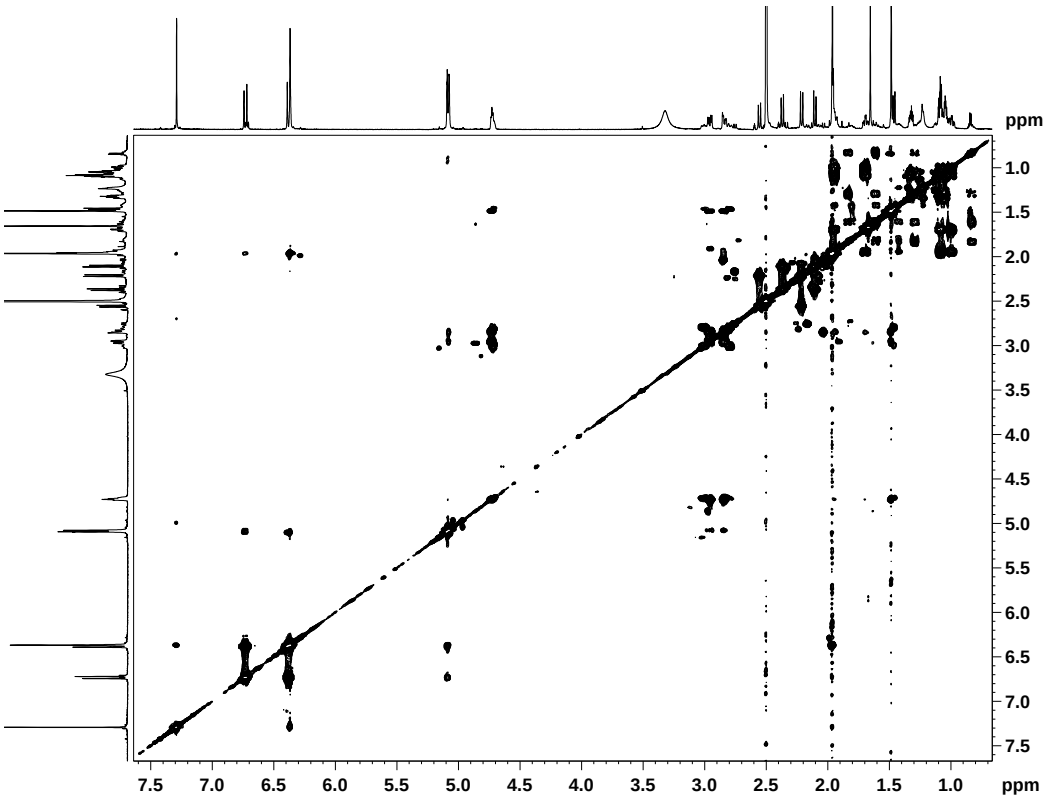


Fig. S12F. HMBC spectra of GNM B (DMSO-d_6)

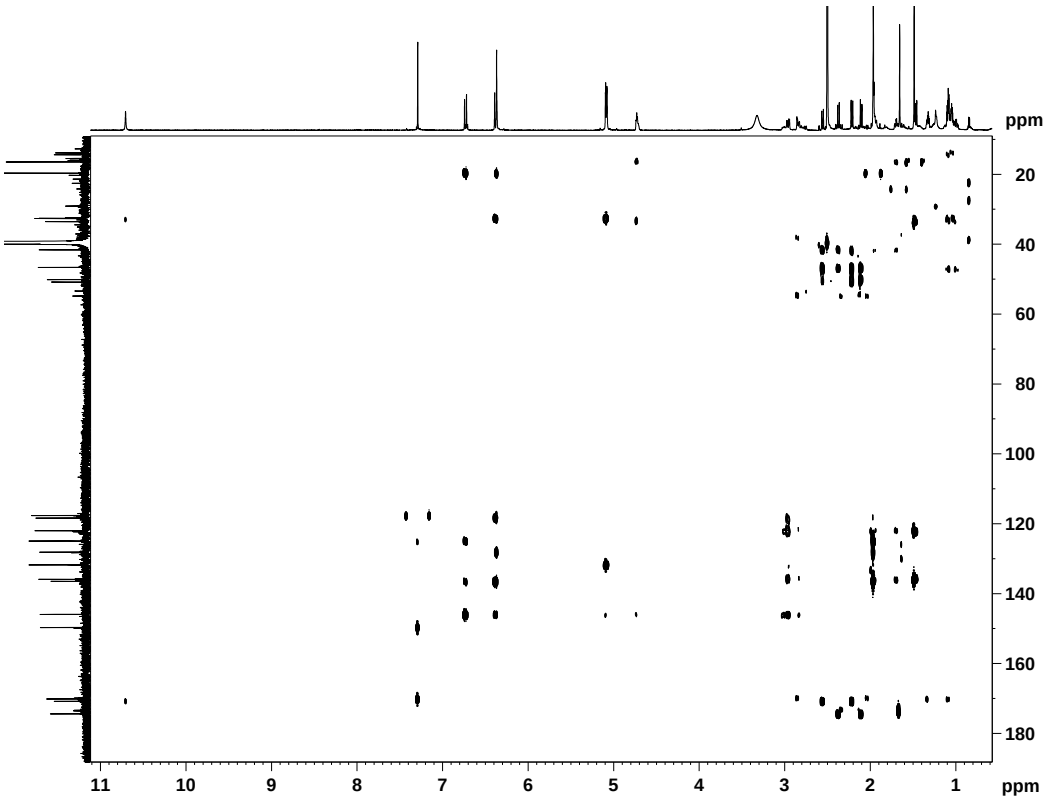


Fig. S12G. ROESY spectra of GNM B (DMSO-d_6)

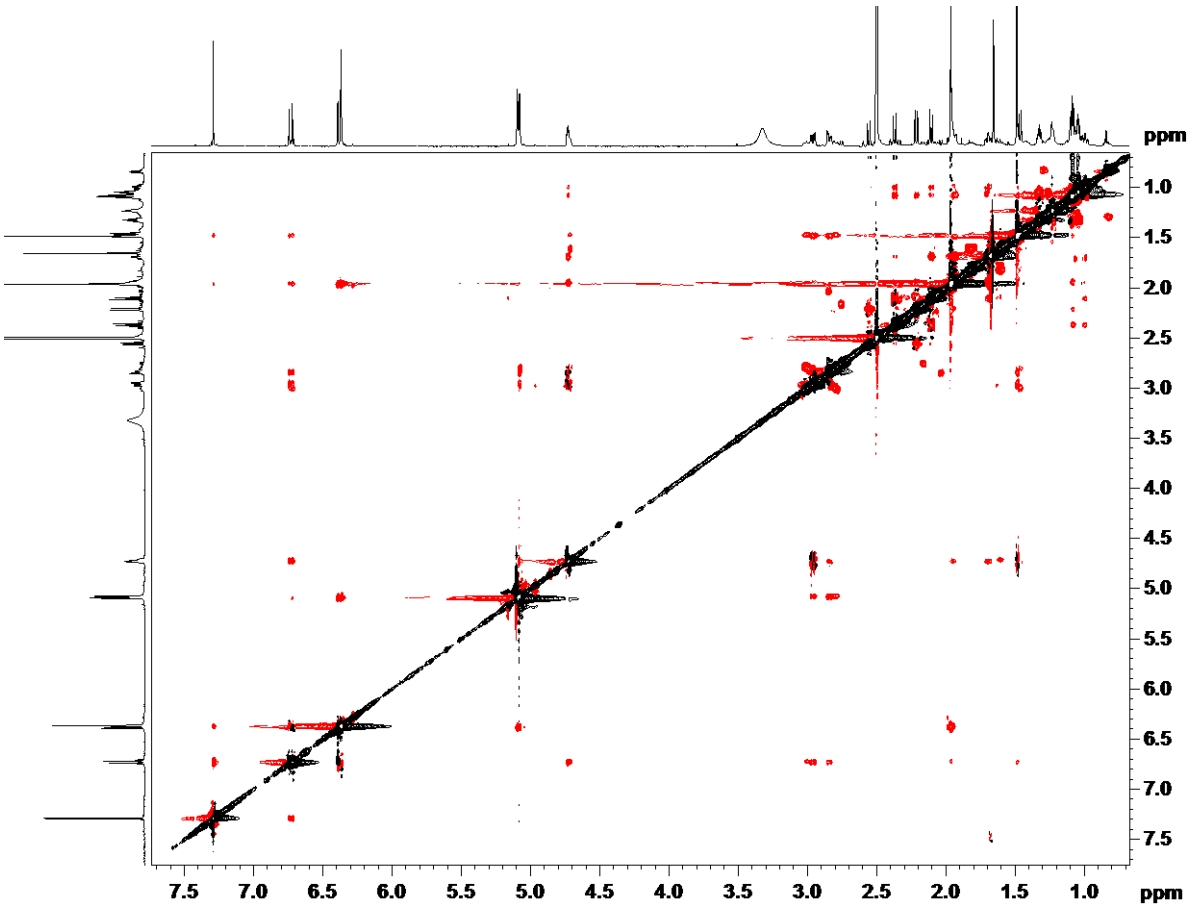


Fig. S12H. HR-MS (ESI) spectra of GNM B

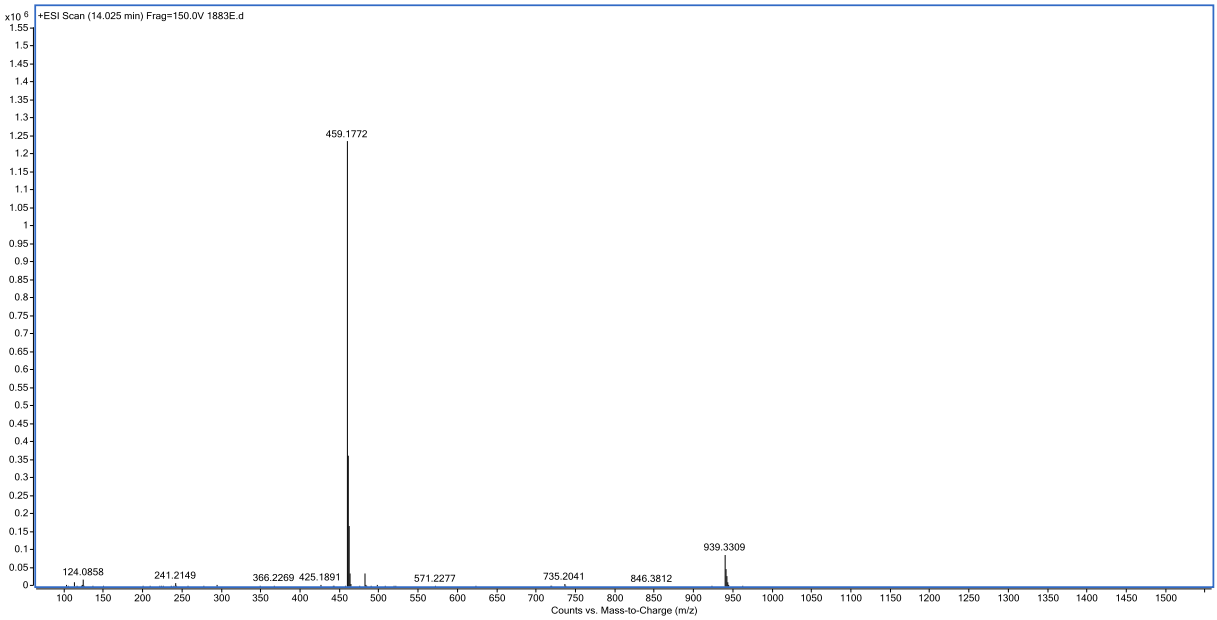


Fig. S13. Elucidation of (*S*)- and (*R*)-PGME derivatives of LNM E2, GNM B2 and WSM A2. Key ^{1}H-^{1}H COSY, HMBC and ROESY correlations are shown.

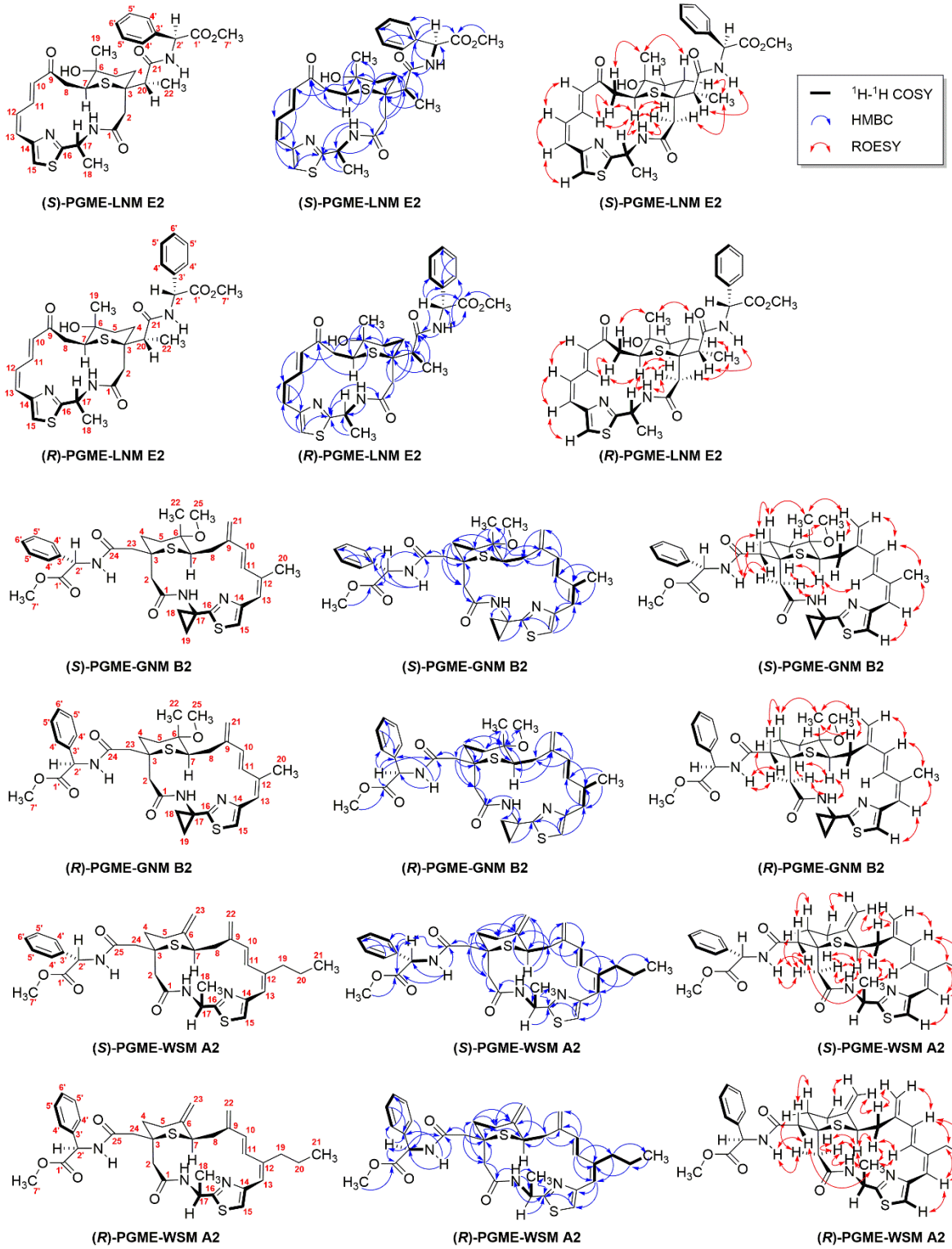


Fig. S14. Determination of the absolute configuration of GNM B2 using its (*S*)- and (*R*)-PGME derivatives. Two major conformations of (*S*)- and (*R*)-PGME-GNM B2 are shown in **A**, which are different from that of the corresponding derivatives of LNM E2 probably due to the missing methyl group at C23 position. In the ROESY spectra (**B**), the presence of correlations between 23-H_b and 4-H_b, between 2-H_b and 4-H_a, as well as the absent correlations between 23-H_a and 2-H/4-H, support that C23-C24 and C3-S are *anti*-oriented, and 23-H_b takes a pseudo-axial position in the preferred conformation at C23. Two major conformations (**A**, conformations I and II), due to the rotation of C23-C24 single bond, could be observed based on the ROESY correlations between 24-N*H* and 23-H_a/23-H_b/2-H_b (**C** and **D**). The relatively weak correlation between 24-N*H* and 23-H_b (**C**) supports that conformation I is the favored conformation of (*R*)-PGME-GNM B2, while the relatively weak correlation between 24-N*H* and 2-H_b (**D**) supports that conformation II is the favored conformation of (*S*)-PGME-GNM B2. Although several hydrogens in the macrolactam backbone are positioned in different sides of PGME plane in two different conformations, 23-H_b and 2-H_a/H_b are always positioned at the two different sides of the PGME plane in both conformations I and II, hence making them diagnostic useful for determination of the absolute configuration. Based on the differences of chemical shift in ^{1}H NMR between (*R*)- and (*S*)-PGME-GNM B2 ($\Delta\delta = \delta_{(R)} - \delta_{(S)}$) as shown in **A**, the absolute configuration of C3 is elucidated to be 3*S* for GNM B2 [3*R* for (*R*)/(*S*)-PGME-GNM B2]. H_a denotes one of the geminal hydrogens appearing at lower field in ^{1}H NMR, and H_b denotes one of the geminal hydrogens appearing at higher field in ^{1}H NMR.

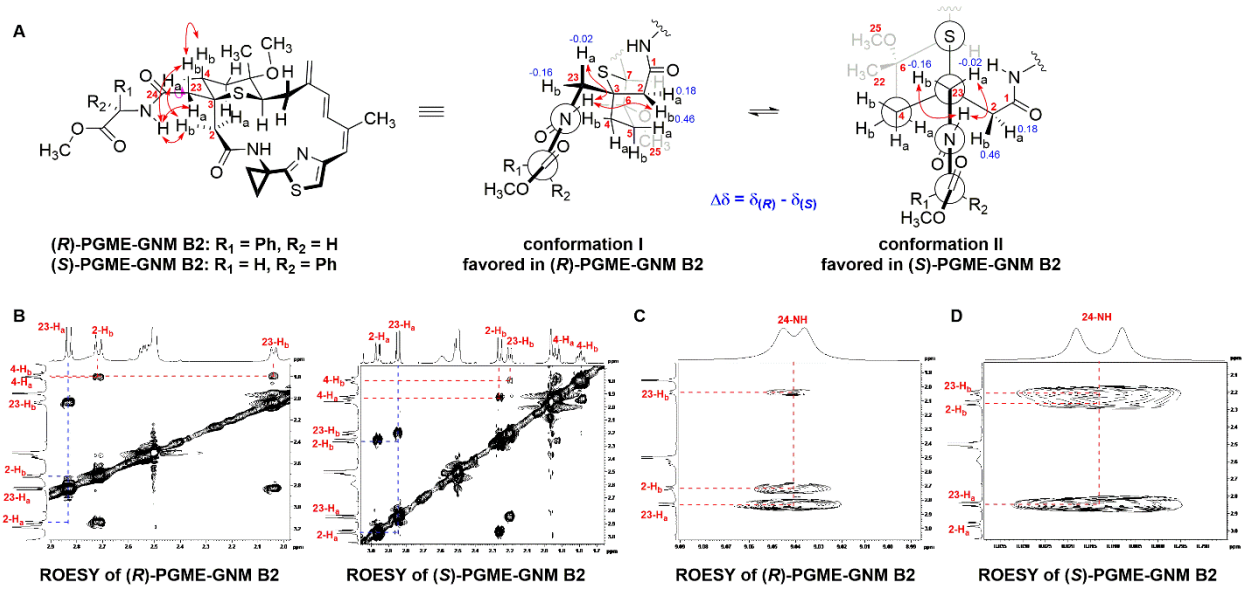


Fig. S15-SA. ^{1}H NMR spectra of (*S*)-PGME-GNM B2 (700 MHz, DMSO-d_6)

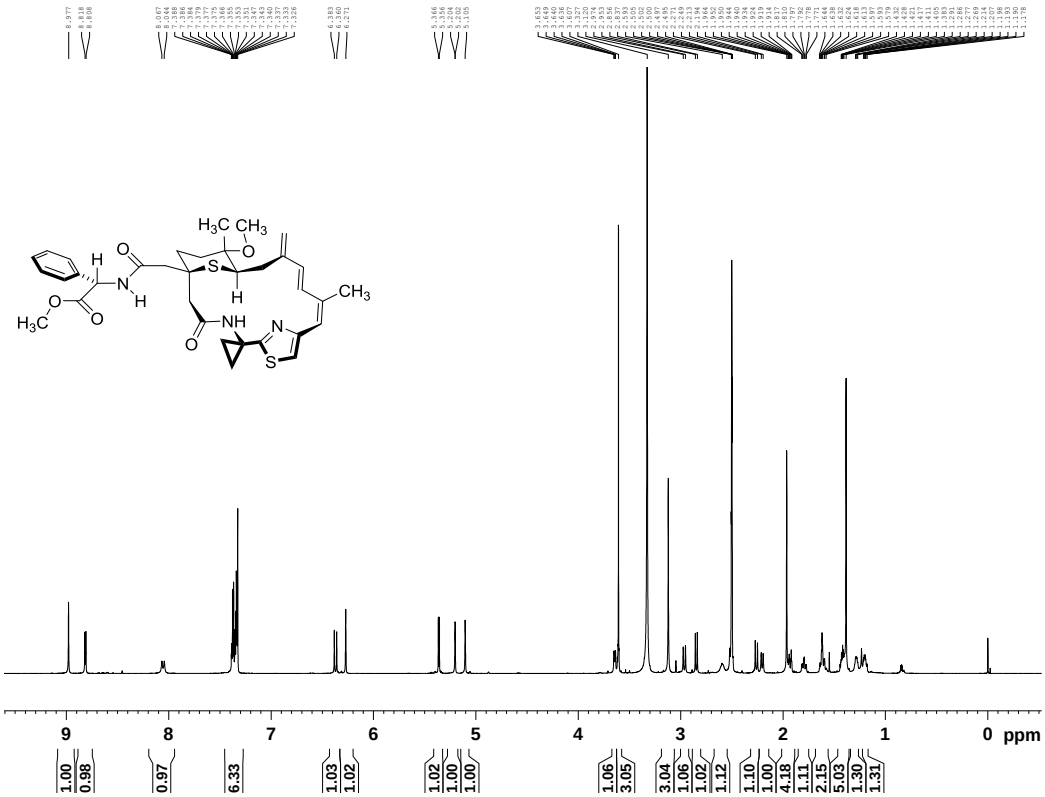


Fig. S15-SB. ^{13}C NMR spectra of (*S*)-PGME-GNM B2 (175 MHz, DMSO-d_6)

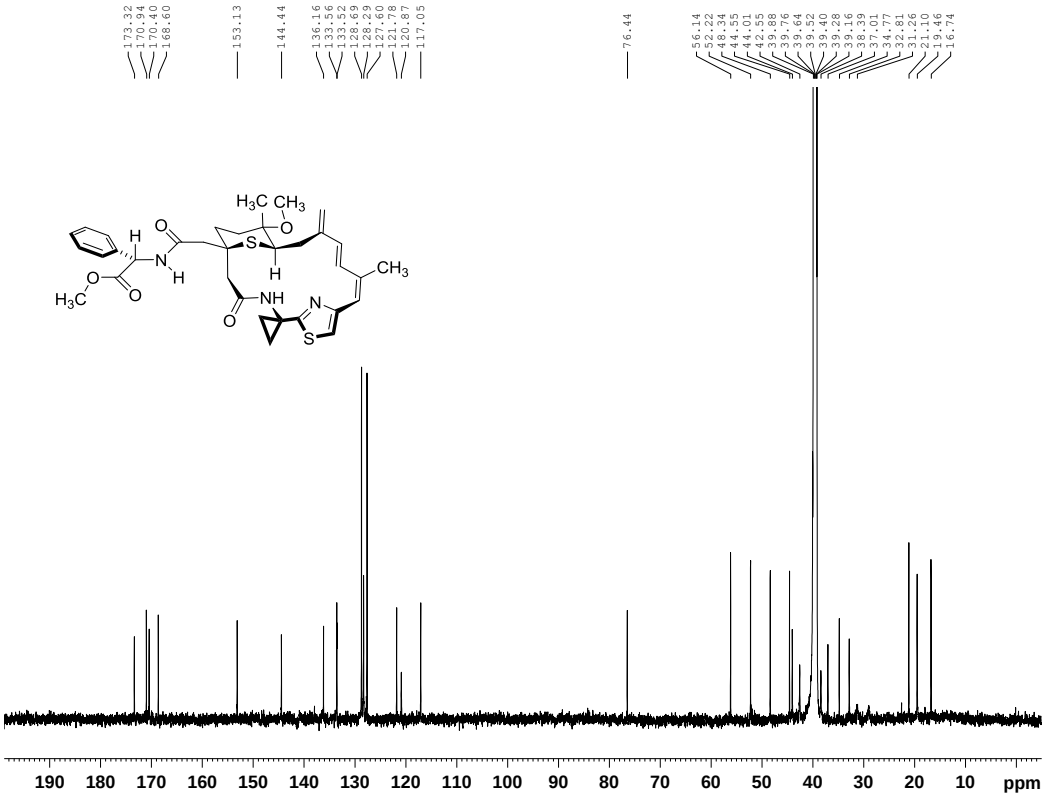


Fig. S15-SC. DEPT-135 spectra of (*S*)-PGME-GNM B2 (DMSO-d_6)

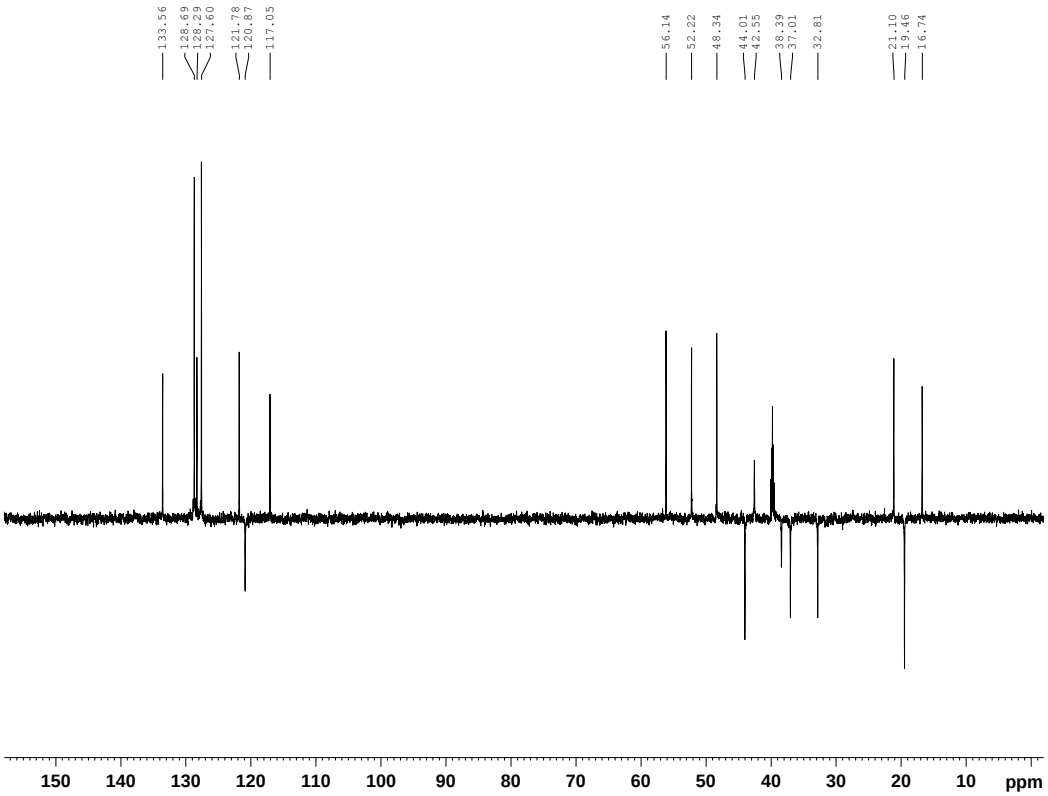


Fig. S15-SD. HSQC spectra of (*S*)-PGME-GNM B2 (DMSO-d_6)

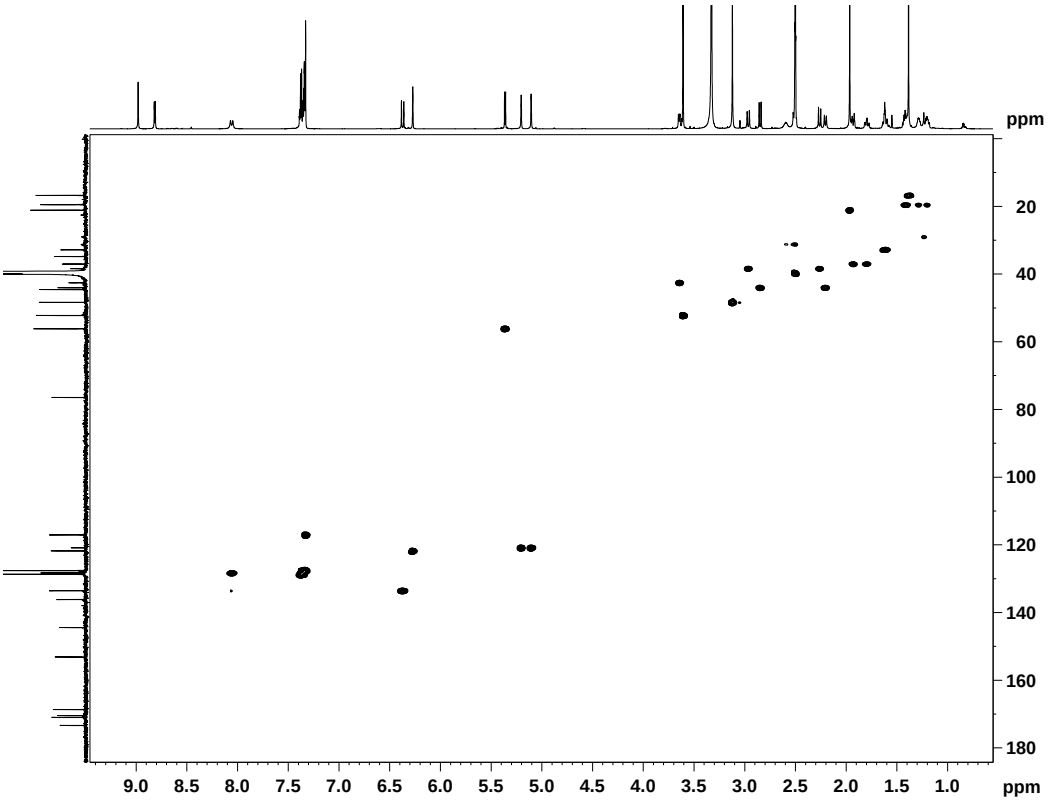


Fig. S15-SE. ^{1}H-^{1}H COSY spectra of (*S*)-PGME-GNM B2 (DMSO-d_6)

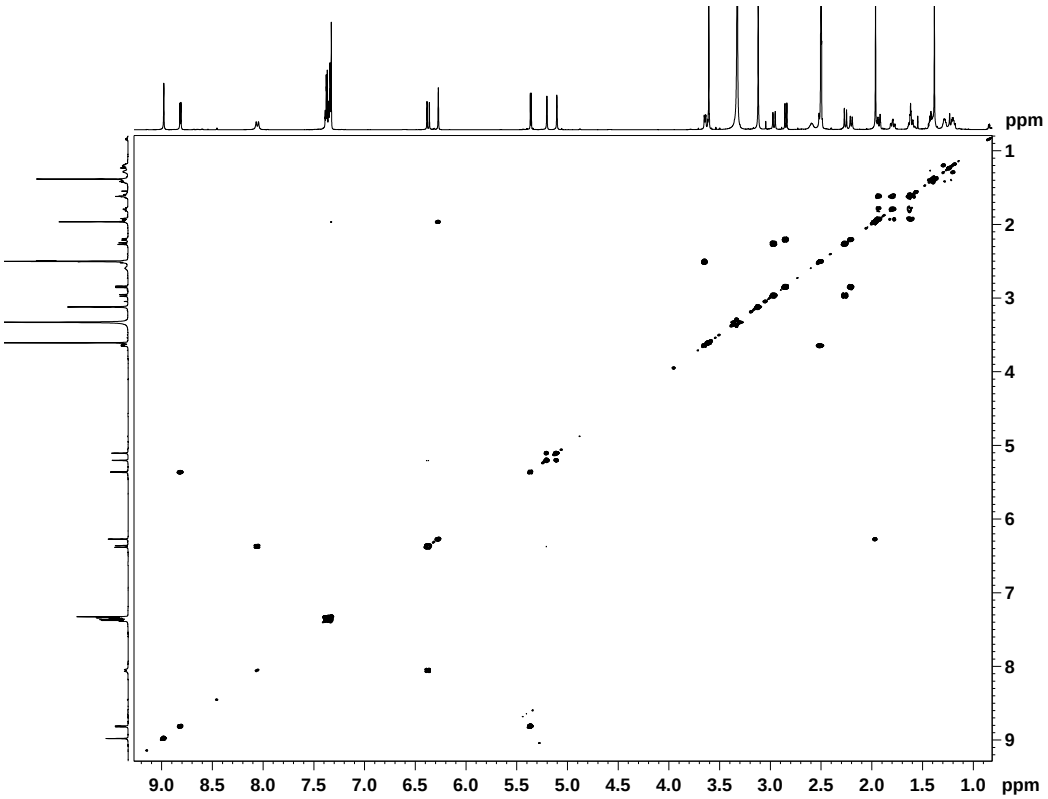


Fig. S15-SF. HMBC spectra of (*S*)-PGME-GNM B2 (DMSO-d_6)

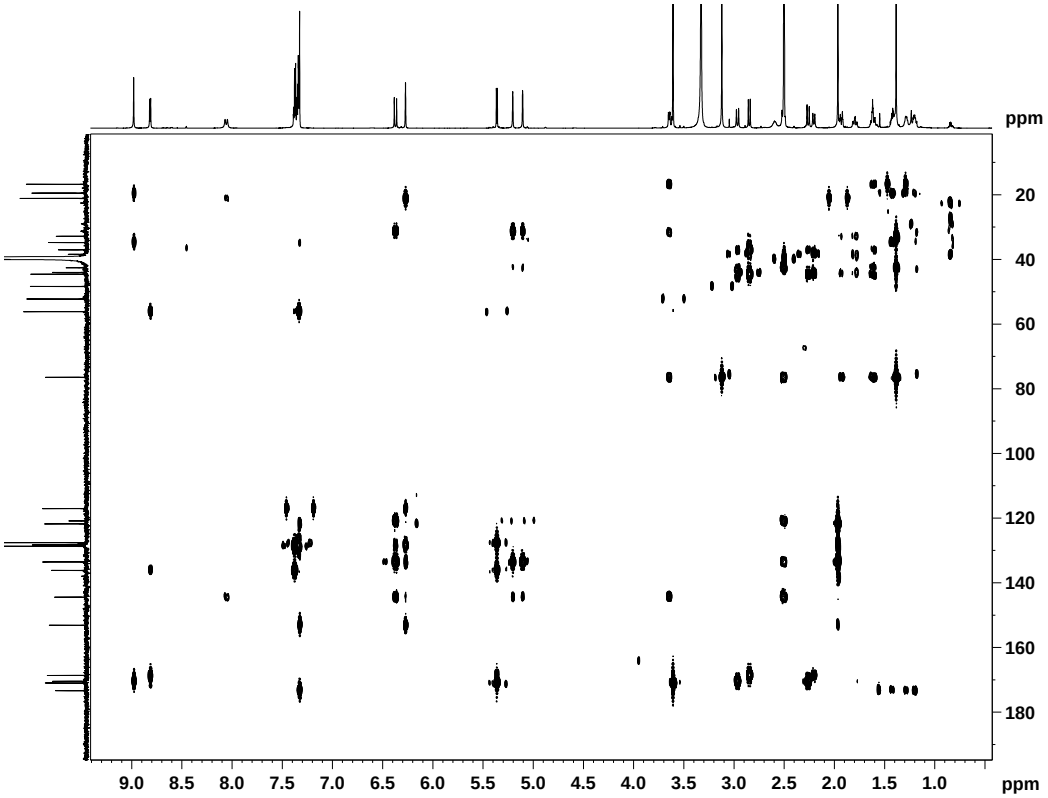


Fig. S15-SG. ROESY spectra of (*S*)-PGME-GNM B2 (DMSO-d_6)

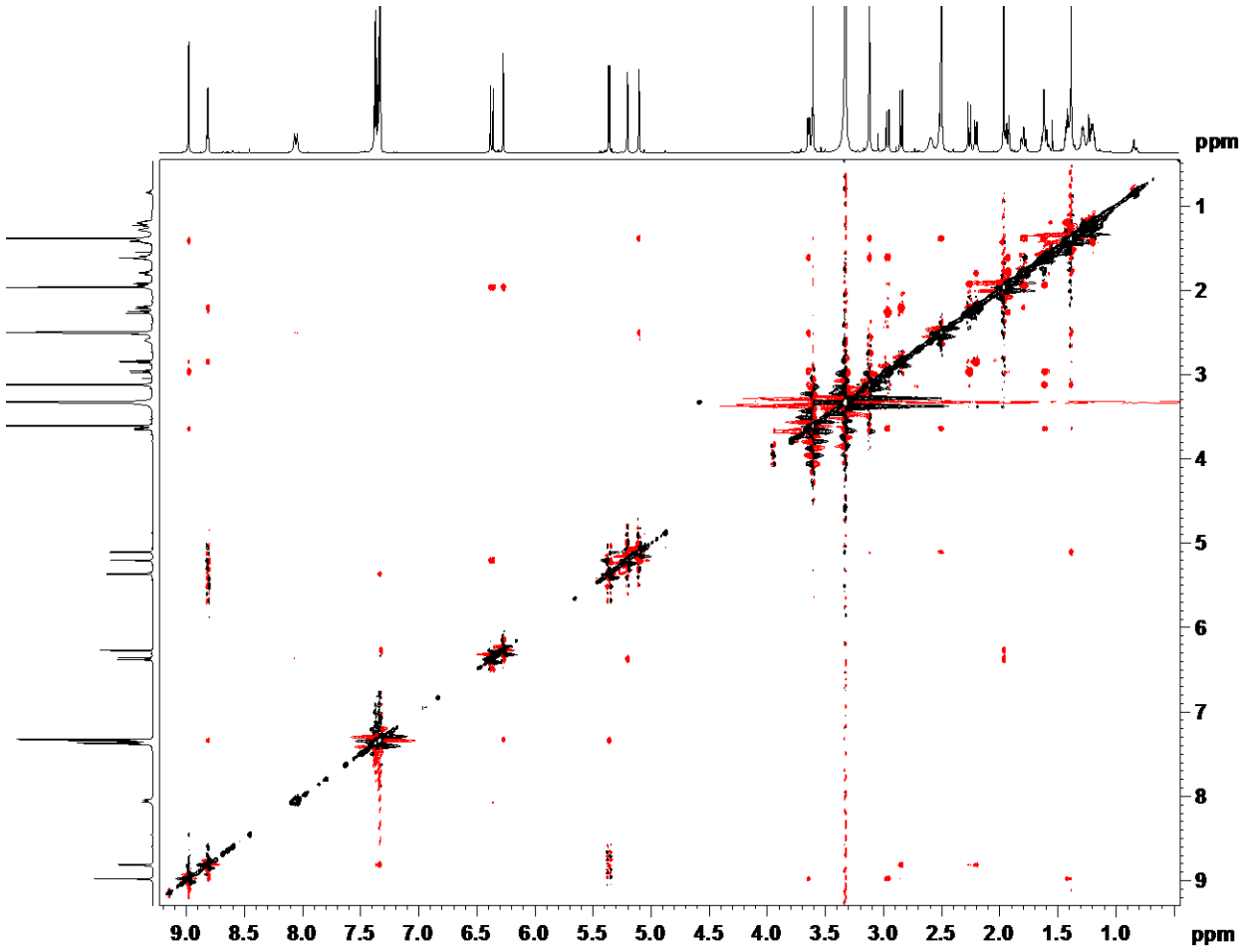


Figure S15-SH. HR-MS (ESI) spectra of (*S*)-PGME-GNM B2

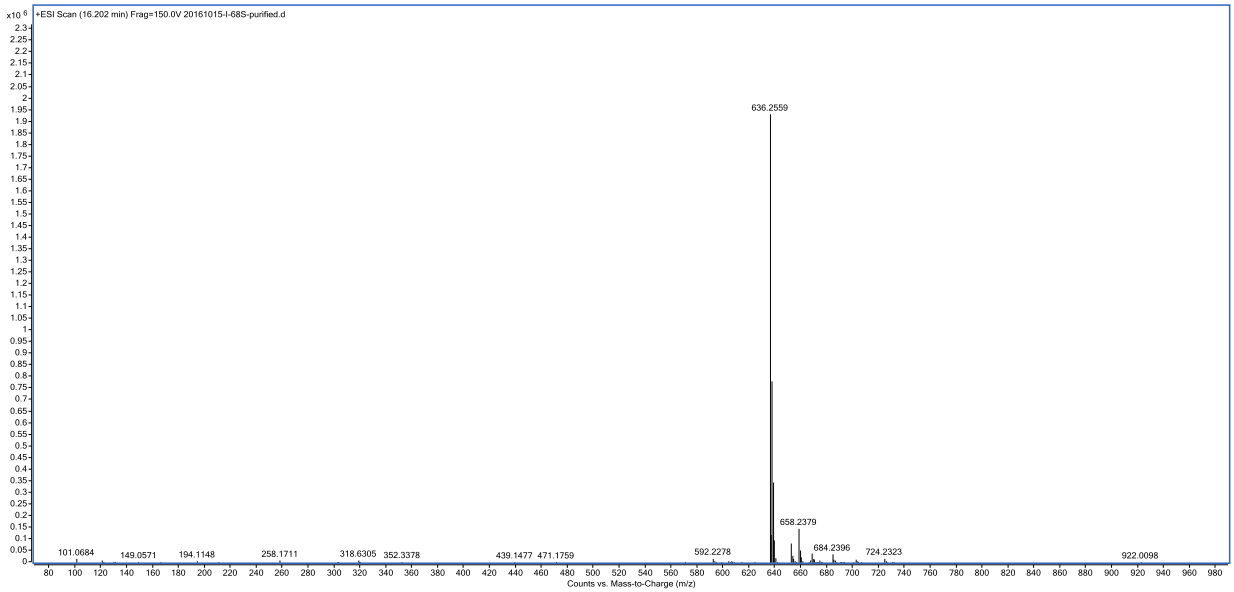


Fig. S15-RA. ^{1}H NMR spectra of (*R*)-PGME-GNM B2 (700 MHz, DMSO-d_6)

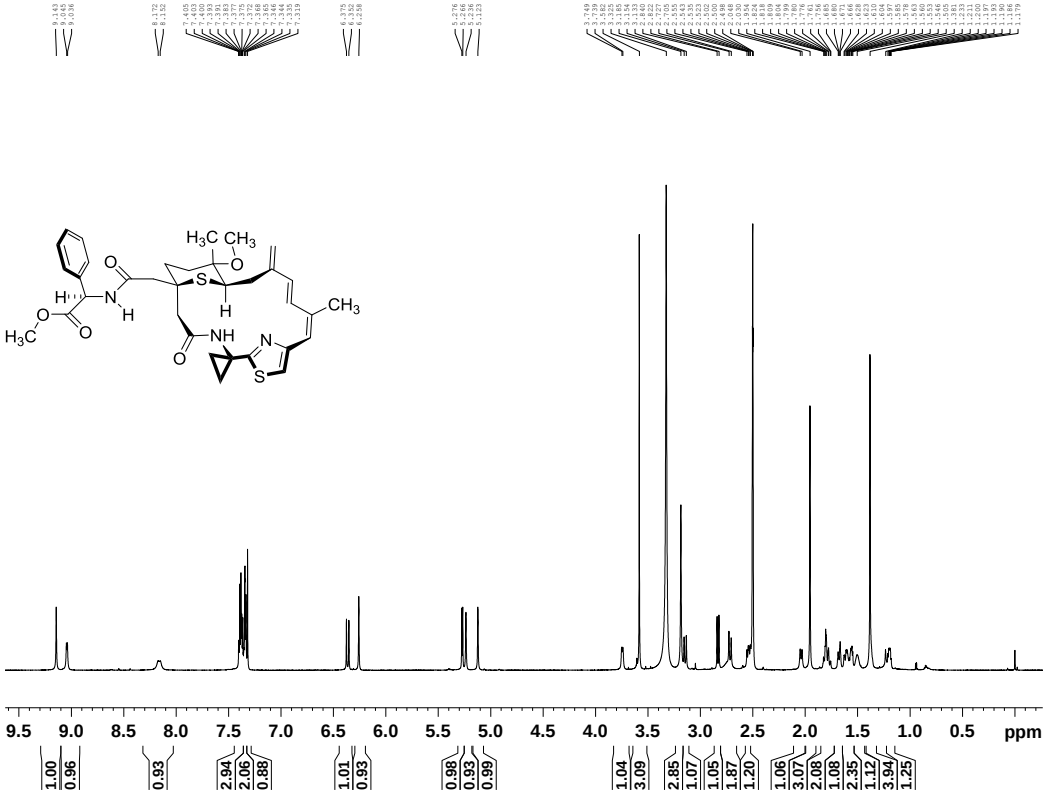


Fig. S15-RB. ^{13}C NMR spectra of (*R*)-PGME-GNM B2 (175 MHz, DMSO-d_6)

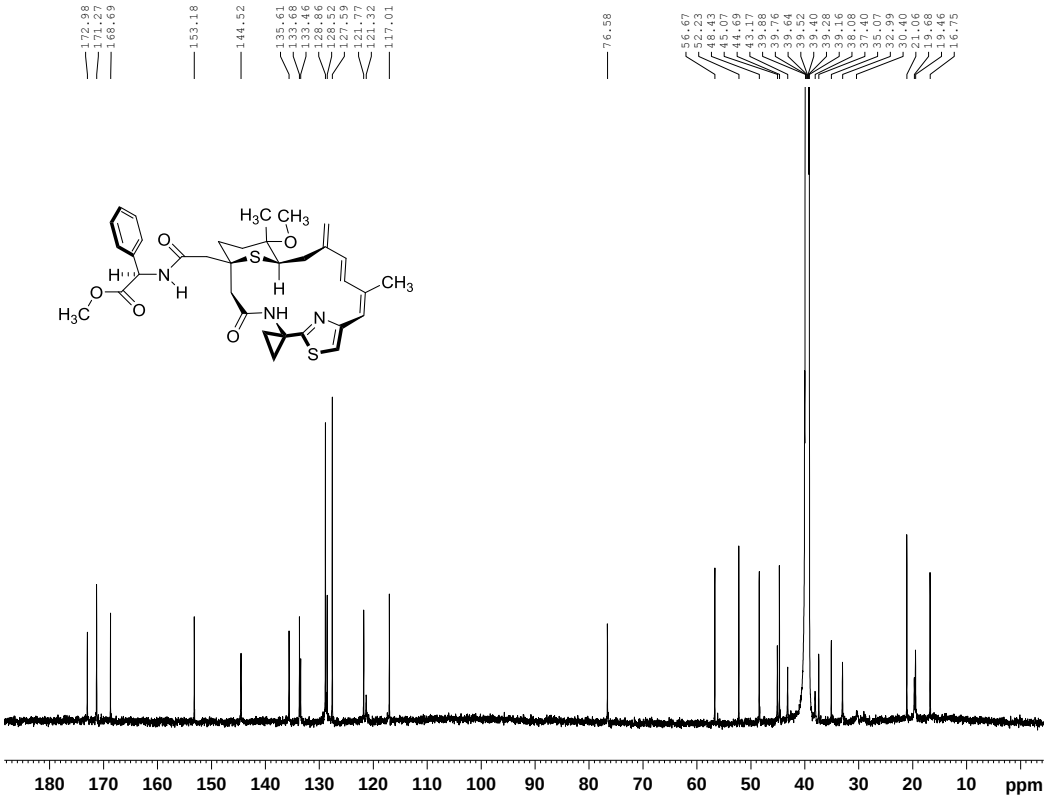


Fig. S15-RC. DEPT-135 spectra of (*R*)-PGME-GNM B2 (DMSO-d_6)

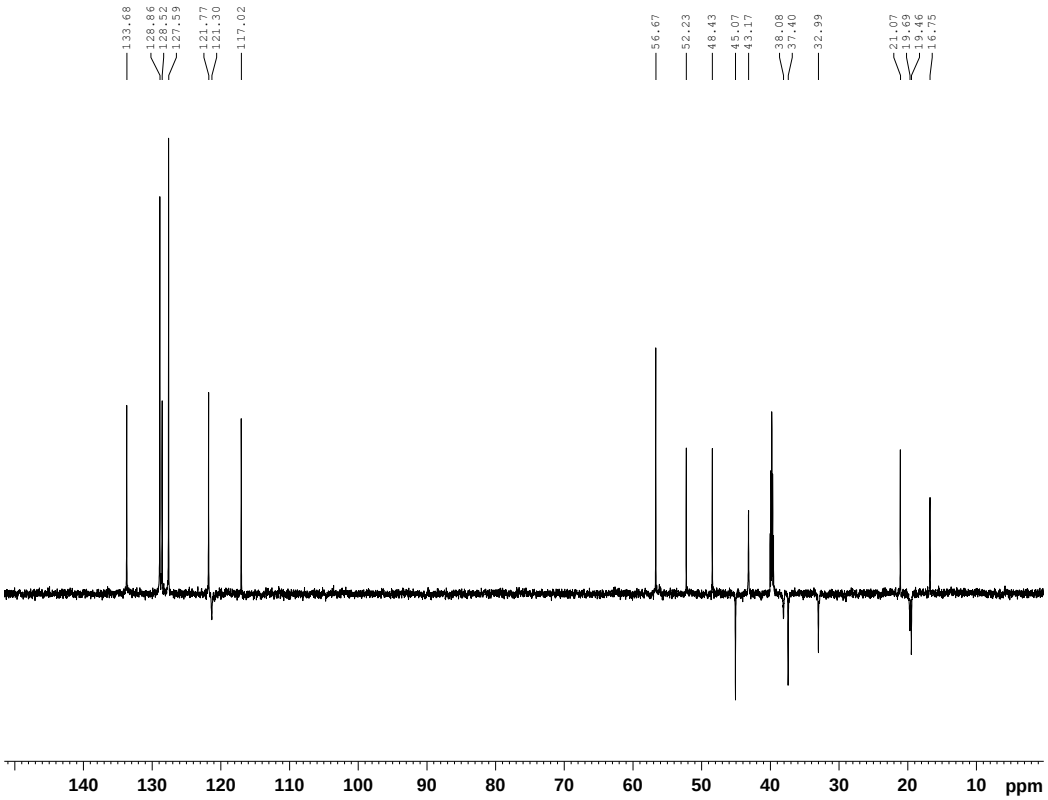


Fig. S15-RD. HSQC spectra of (*R*)-PGME-GNM B2 (DMSO-d_6)

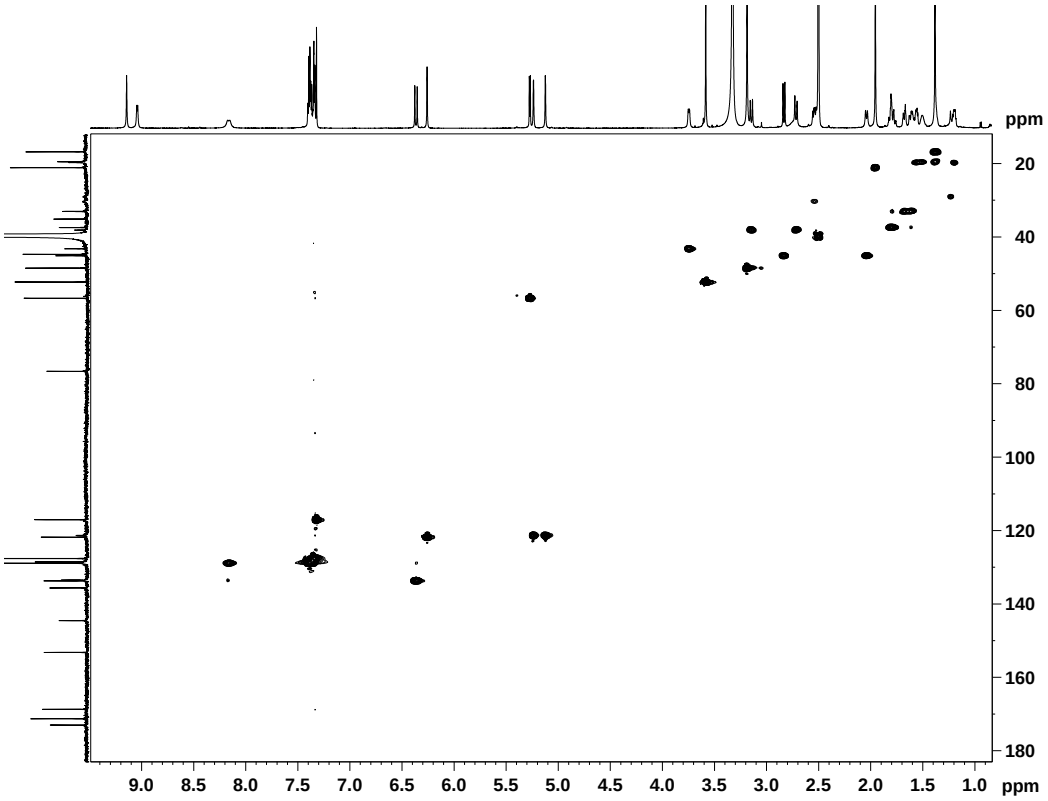


Fig. S15-RE. ^{1}H-^{1}H COSY spectra of (*R*)-PGME-GNM B2 (DMSO-d_6)

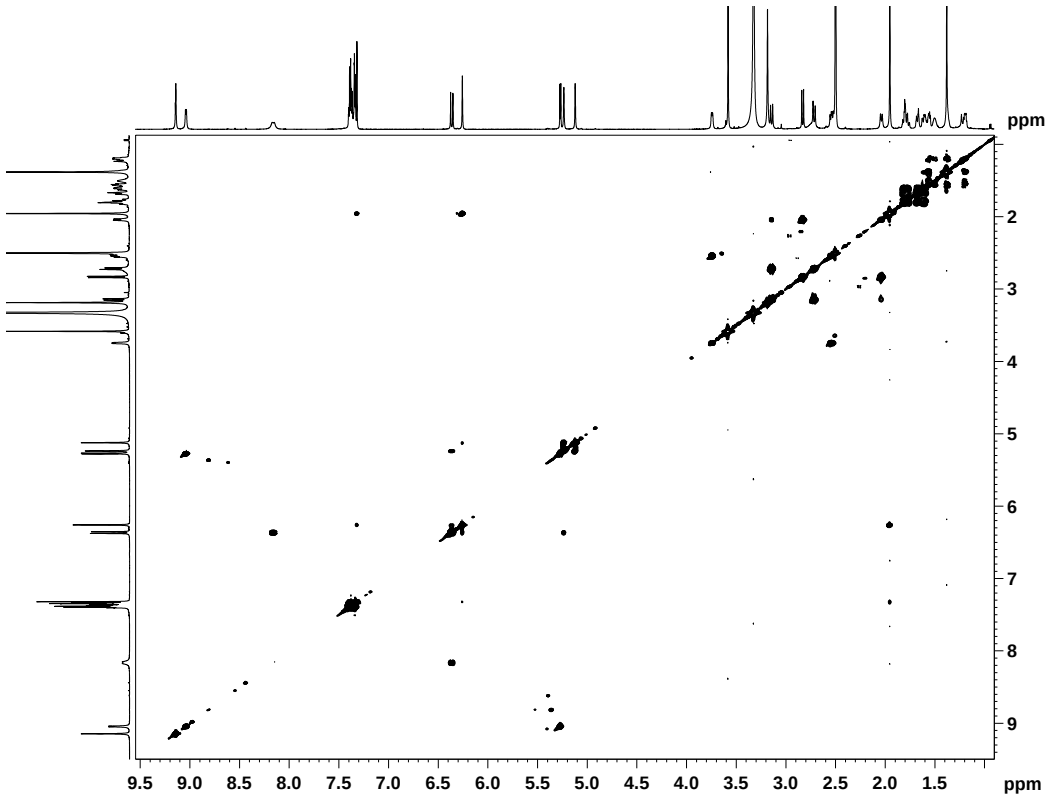


Fig. S15-RF. HMBC spectra of (*R*)-PGME-GNM B2 (DMSO-d_6)

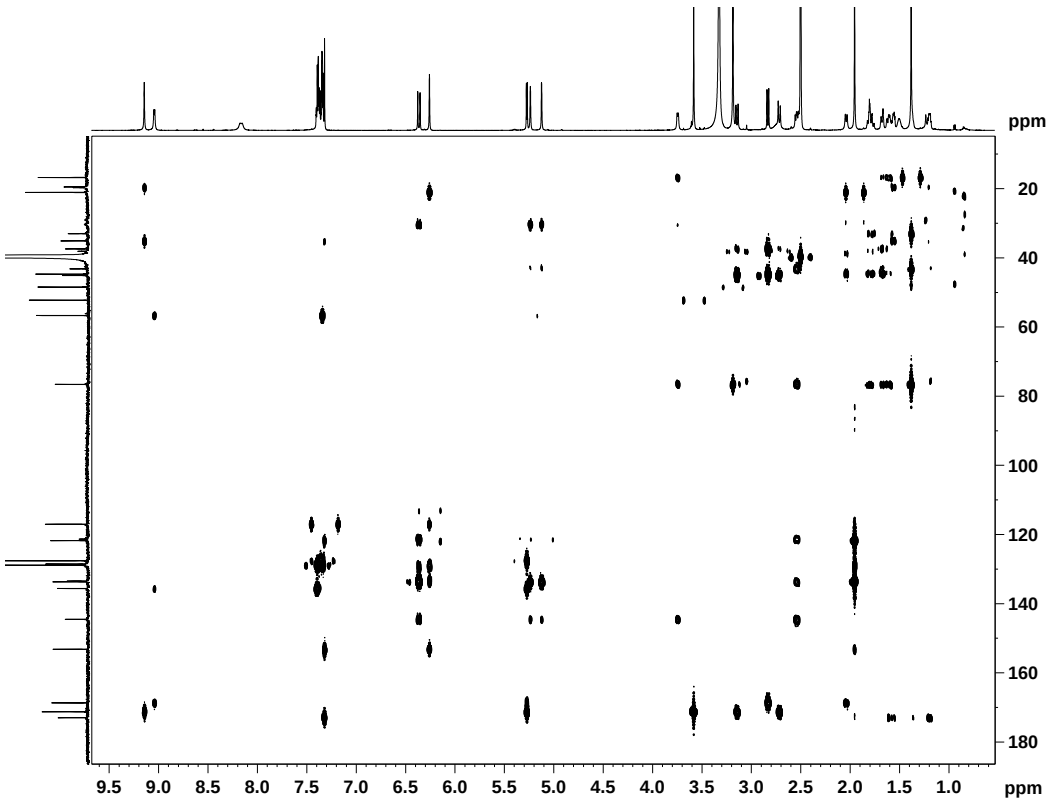


Fig. S15-RG. ROESY spectra of (*R*)-PGME-GNM B2 (DMSO-d_6)

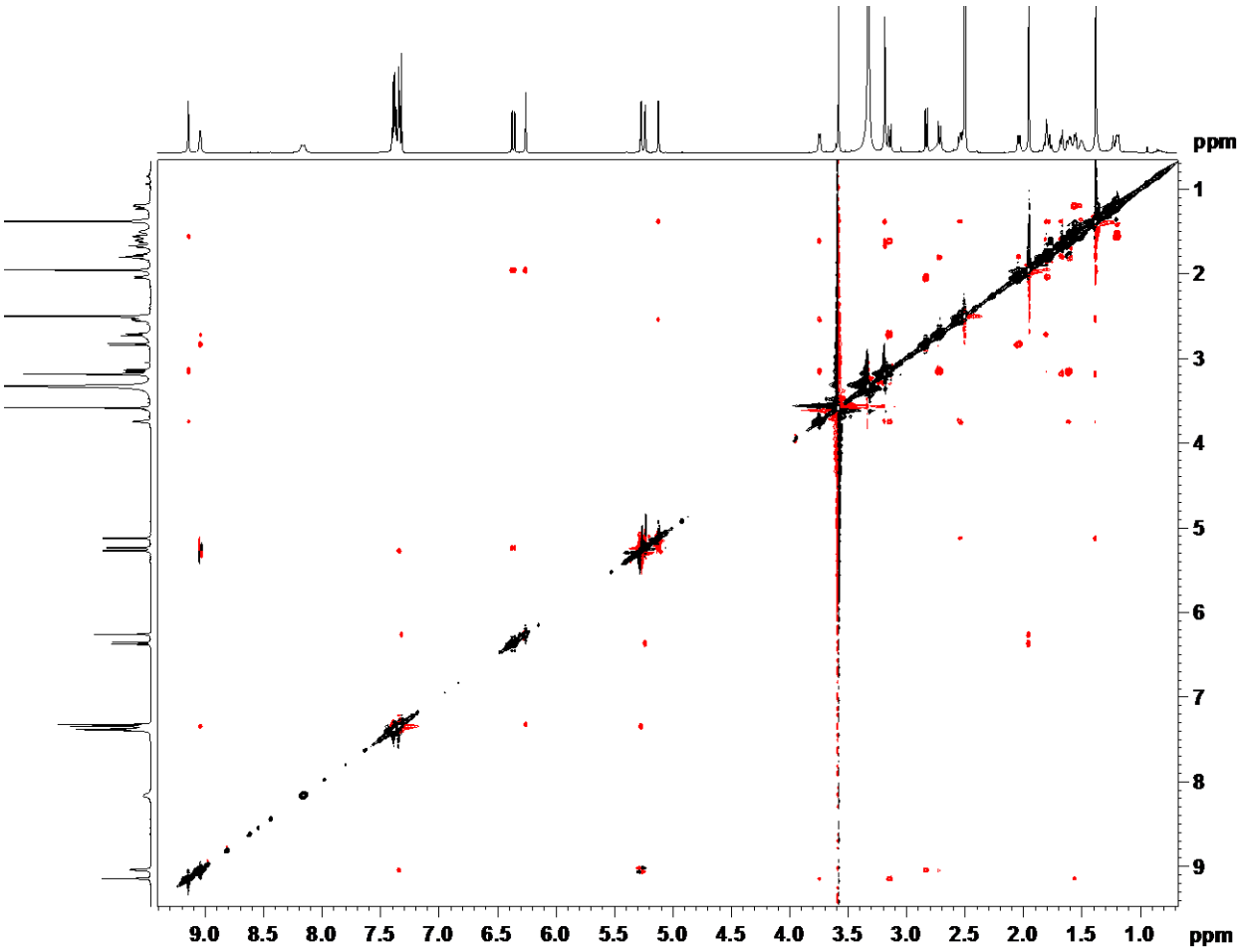


Fig. S15-RH. HR-MS (ESI) spectra of compound (*R*)-PGME-GNM B2

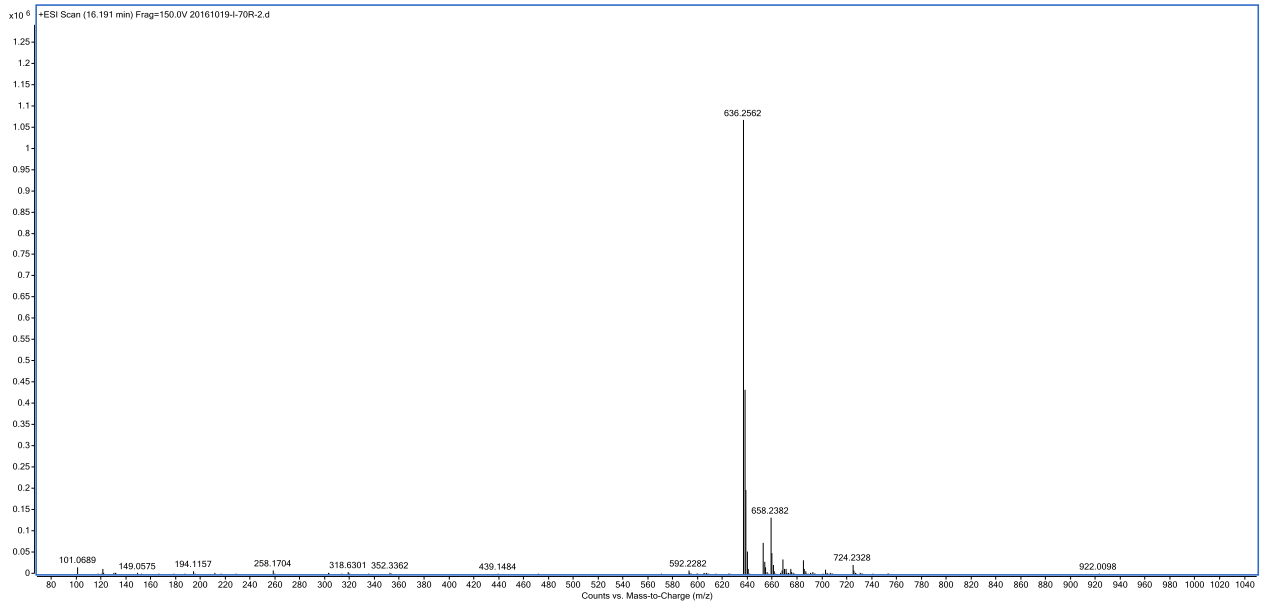


Fig. S16. Using (*S*)- and (*R*)-PGME derivatives of LNM E2 as models for the determination of the absolute configuration of LNM E2 (14, 23). The preferred conformations, especially the conformation of the C3 side chain, of both (*S*)- and (*R*)-PGME-LNM E2 (**A**) resembled that of LNM E2 (**C**) based on detailed analysis of their ROESY correlations. The presence of strong ROESY correlations between 20-H and 21-N*H* (**D**), as well as the residual correlations between 22-CH_3 and 21-N*H* (**E**), indicate in the preferred conformation, 20-H and 21-N*H* are *syn*-positioned, and 22-CH_3 and the macrolactam skeleton are positioned at the two sides of the plane defined by PGME. Based on the differences of chemical shift in ^{1}H NMR between (*R*)- and (*S*)-PGME-LNM E2 ($\Delta\delta = \delta_{(R)} - \delta_{(S)}$) as shown in **B**, the absolute configuration at C20 of LNM E2 is elucidated as (20*S*). In this case, $\Delta\delta$ could not be directly applied to determine the absolute configuration of C3, since the macrolactam backbone will be at the same side when C3 takes an opposite configuration as shown in **H**. This problem could be circumvented by correlating 20-H and 22-CH_3 with the hydrogens in the macrolactam backbone. The presence of ROESY correlations between 22-CH_3 and 2-H_b/4-H_b (**F**), support that C20-C22 and C3-S are *anti*-oriented in the preferred conformation. On the other hand, the absent correlations between 20-H and 2-H/4-H (**G**) is consistent with the 3*R* configurational assignment (**A**). Otherwise, strong correlation between 20-H and 4-H_a should be found in ROESY spectra due to their 1,3-diaxial relation (**H**), which is observed in the cases of GNM B2 (**Fig. S14**) and WSM A2 (**Fig. S24**). Furthermore, since the priming amino acid has been determined to be D-alanine in previous studies, the absent correlations between 18-CH_3 and 8-H_a/11-H (**I**), which is observed in the case of WSM A2 (**Fig. S20**), is supportive for 3*R* configurational assignment rather than 3*S*. Therefore, the absolute configuration of LNM E2 determined by PGME-derivatives is the same as that determined previously by total synthesis and X-ray diffractions. We thus conclude that the process by combination of conformational analysis based on ROESY correlations, and determination of the chirality of the α-carbon with PGME-derivatives, would provide reliable approach for the elucidation of the absolute configurations of other related natural products discovered from strains CB01883 and CB02120-2. H_a denotes one of the geminal hydrogens appearing at lower field in ^{1}H NMR, and H_b denotes one of the geminal hydrogens appearing at higher field in ^{1}H NMR.

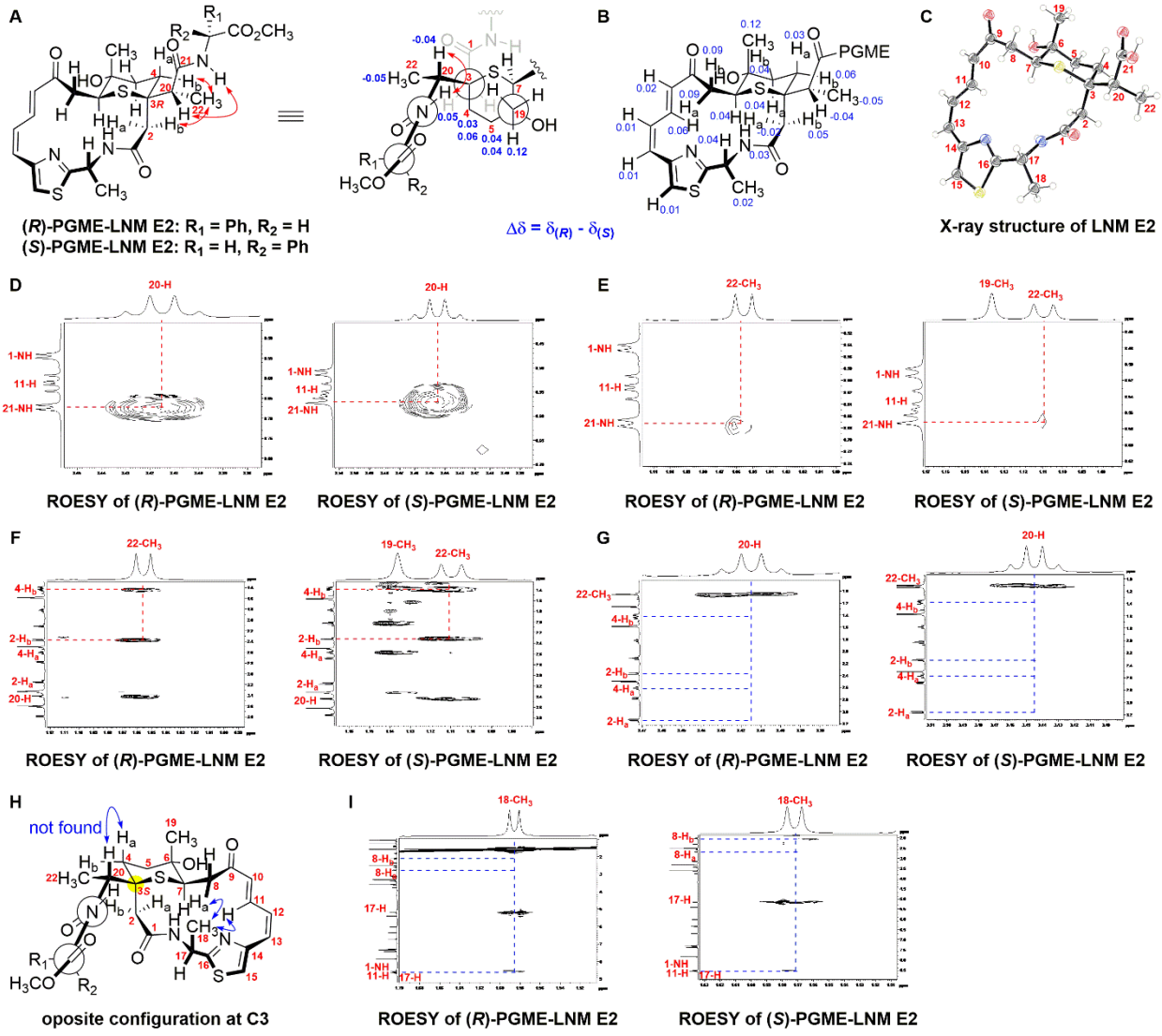

A
B
C
(R)-PGME-LNM E2: R1 = Ph, R2 = H
(S)-PGME-LNM E2: R1 = H, R2 = Ph
Δδ = δ(R) - δ(S)
X-ray structure of LNM E2
D
E
F
G
H
I
ROESY of (R)-PGME-LNM E2
ROESY of (S)-PGME-LNM E2
oposite configuration at C3
not found

Fig. S17-SA. ^{1}H NMR spectra of (*S*)-PGME-LNM E2 (700 MHz, DMSO-d_6)

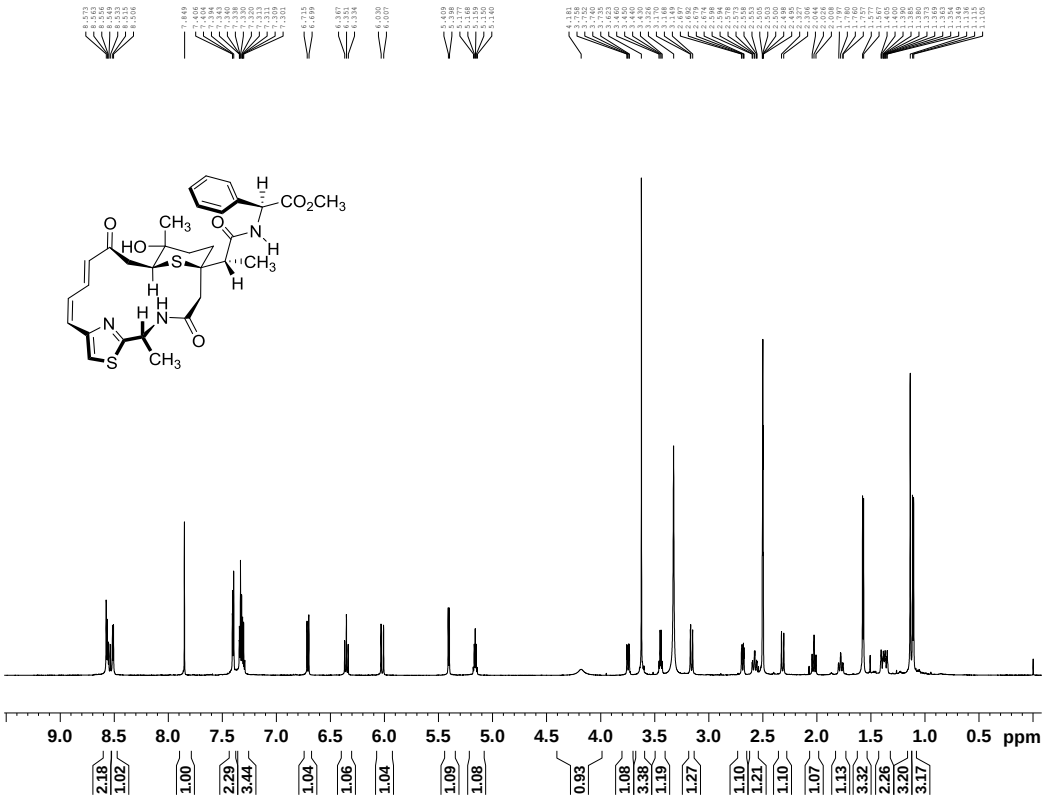


Fig. S17-SB. ^{13}C NMR spectra of (*S*)-PGME-LNM E2 (175 MHz, DMSO-d_6)

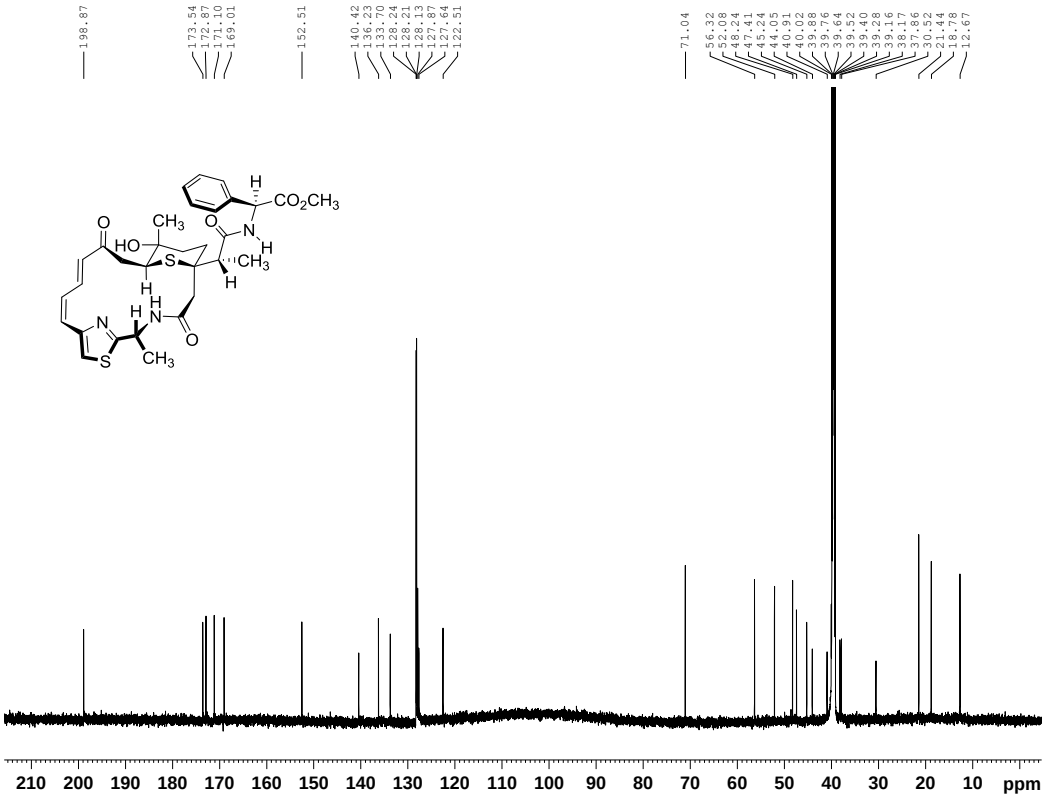


Fig. S17-SC. DEPT-135 spectra of (*S*)-PGME-LNM E2 (DMSO-d_6)

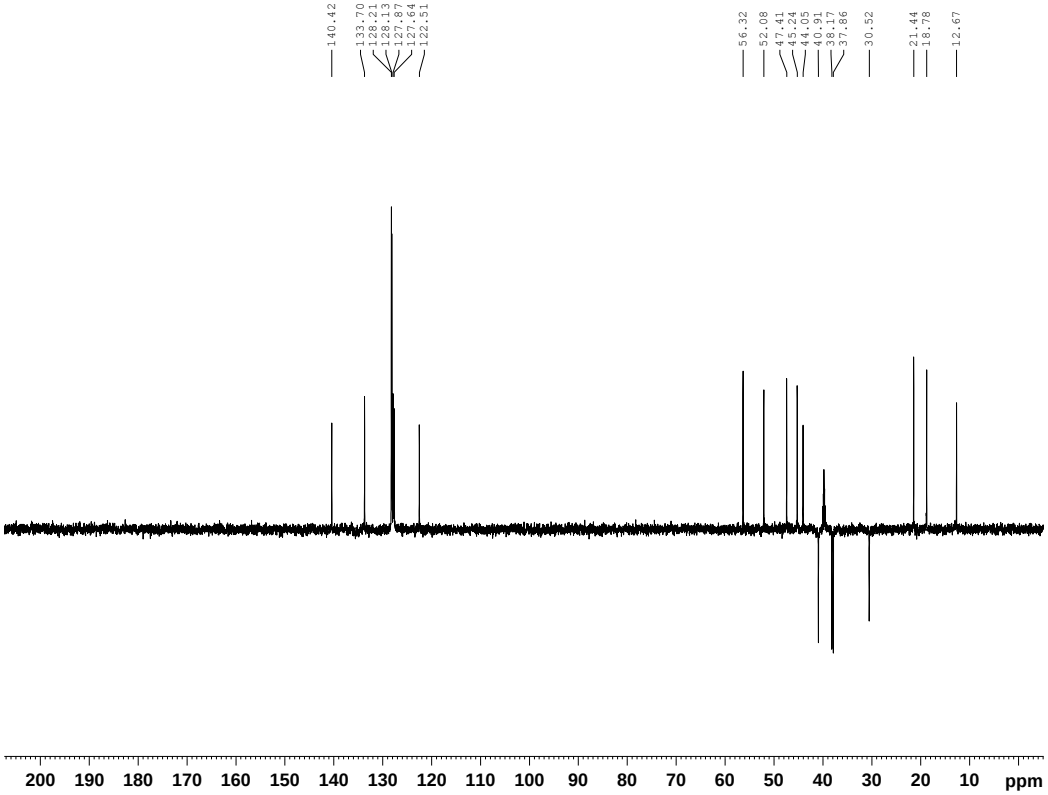


Fig. S17-SD. HSQC spectra of (*S*)-PGME-LNM E2 (DMSO-d_6)

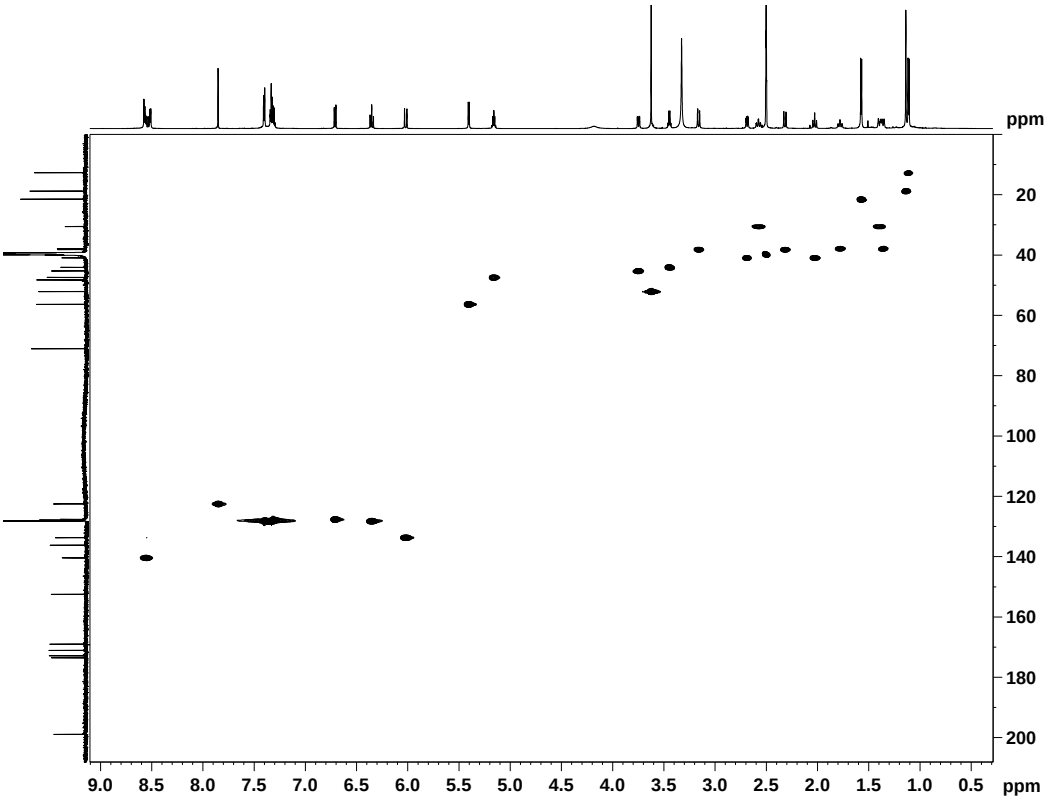


Fig. S17-SE. ^{1}H-^{1}H COSY spectra of (*S*)-PGME-LNM E2 (DMSO-d_6)

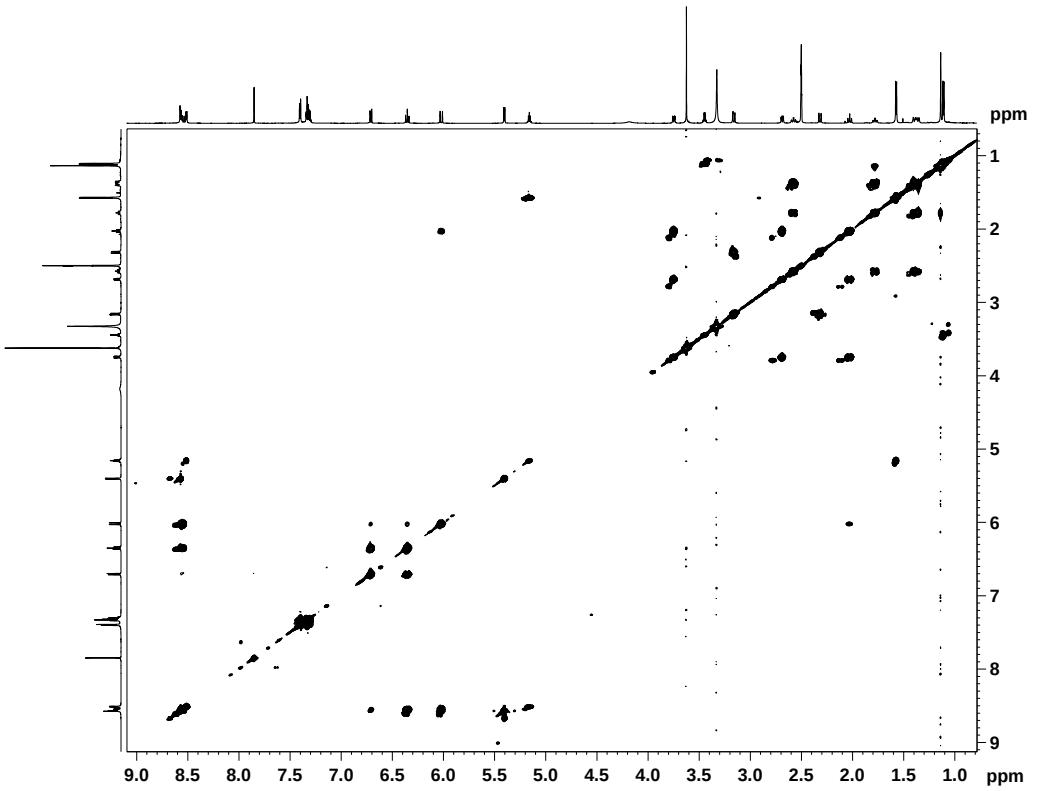


Fig. S11-SF. HMBC spectra of (*S*)-PGME-LNM E2 (DMSO-d_6)

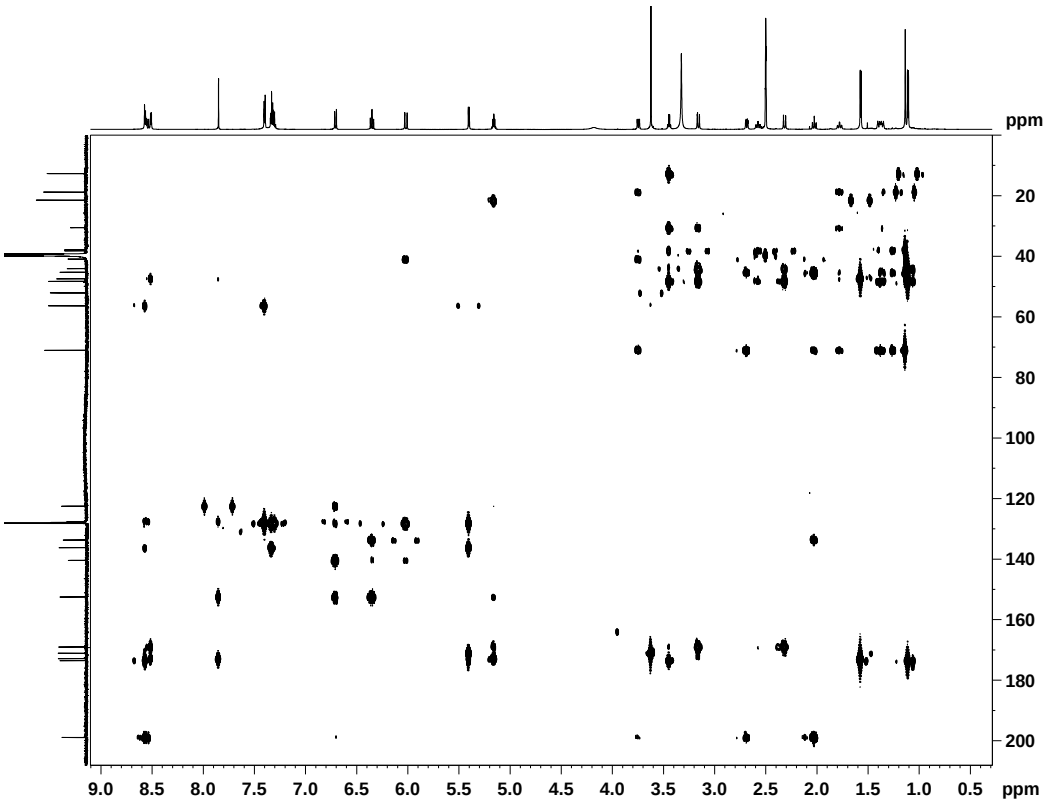


Fig. S17-SG. ROESY spectra of (*S*)-PGME-LNM E2 (DMSO-d_6)

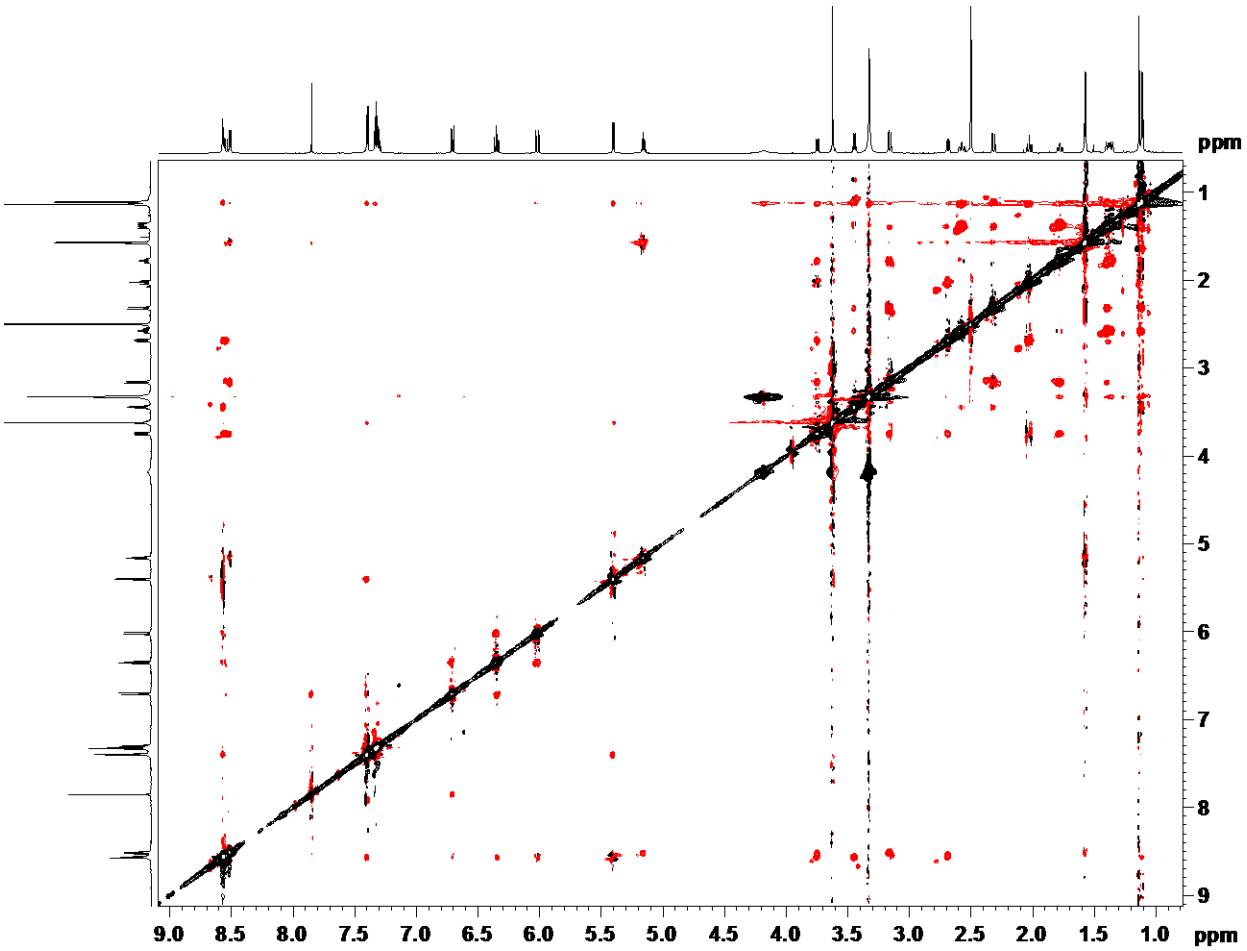


Fig. S17-SH. HR-MS (ESI) spectra of (*S*)-PGME-LNM E2

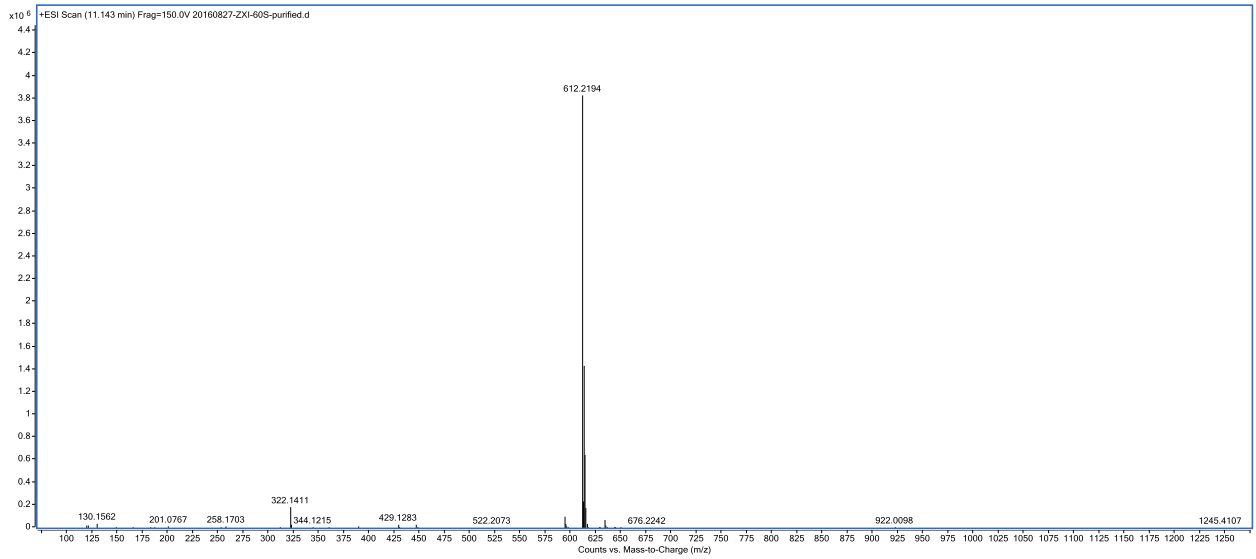


Fig. S17-RA. ^{1}H NMR spectra of (*R*)-PGME-LNM E2 (700 MHz, DMSO-d_6)

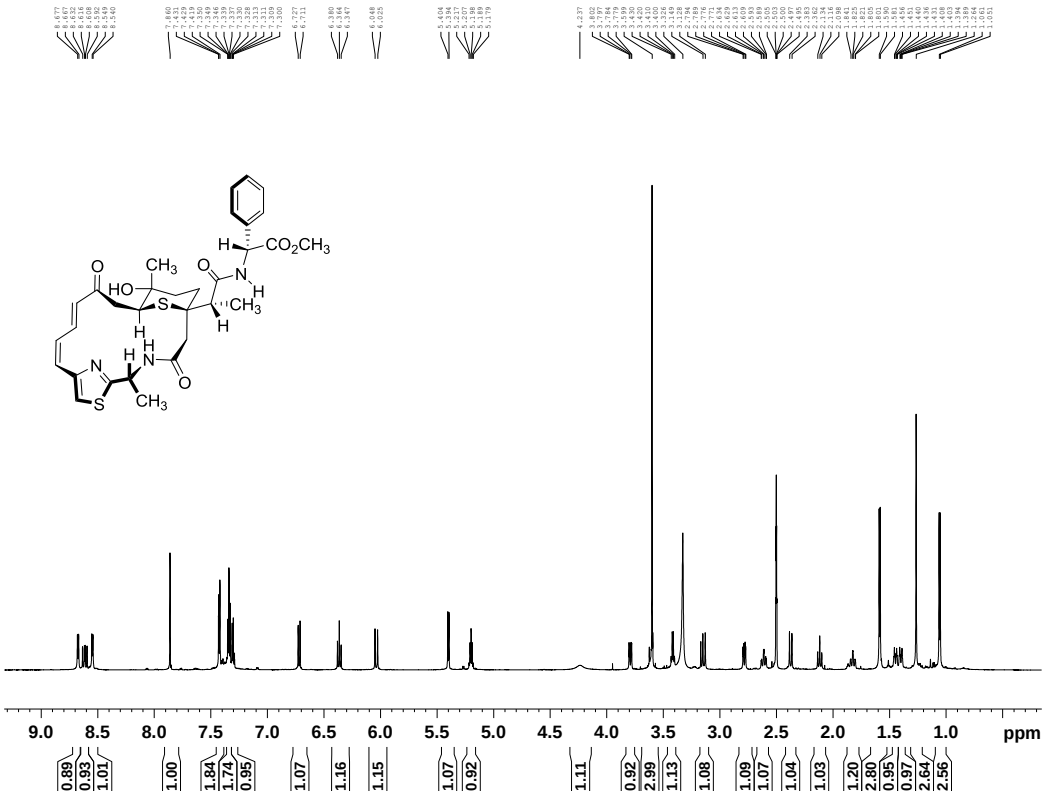


Fig. S17-RB. ^{13}C NMR spectra of (*R*)-PGME-LNM E2 (175 MHz, DMSO-d_6)

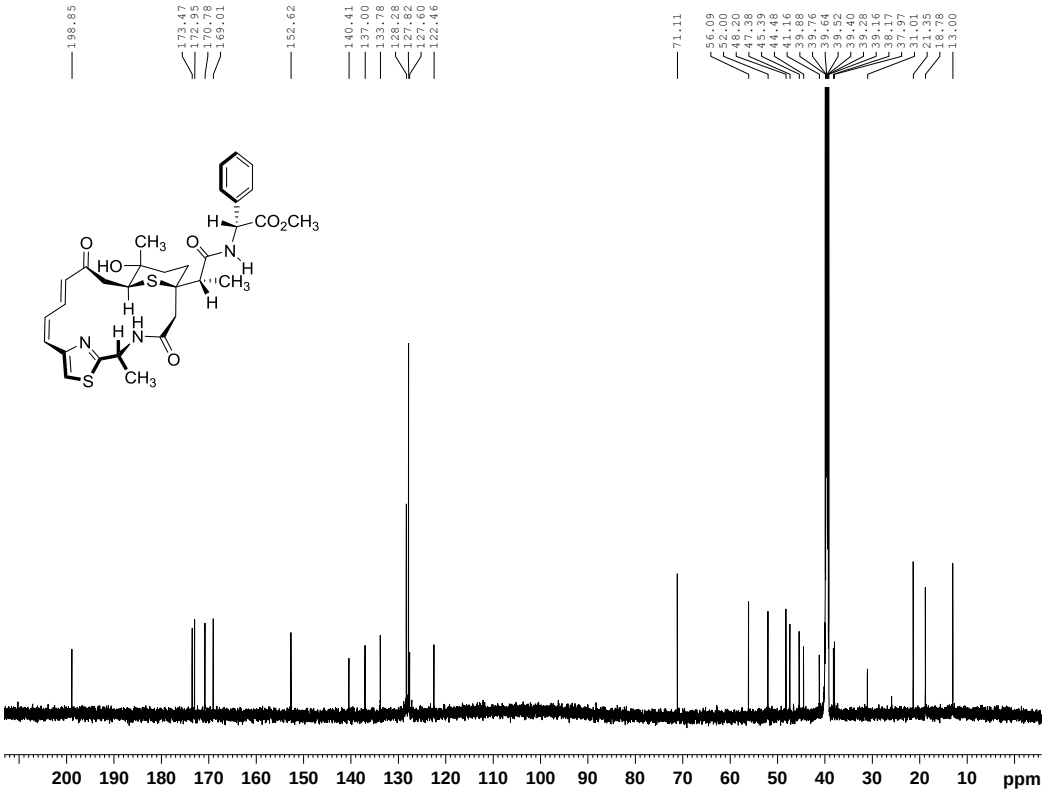


Fig. S17-RC. DEPT-135 spectra of (*R*)-PGME-LNM E2 (DMSO-d_6)

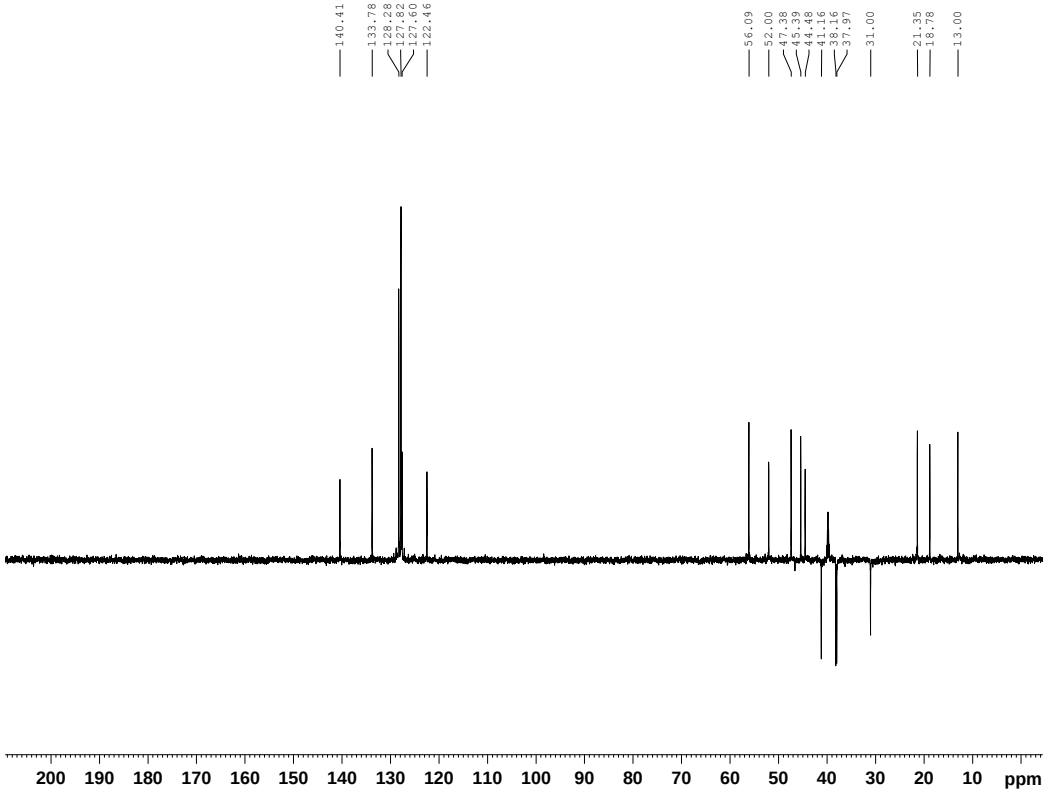


Fig. S17-RD. HSQC spectra of (*R*)-PGME-LNM E2 (DMSO-d_6)

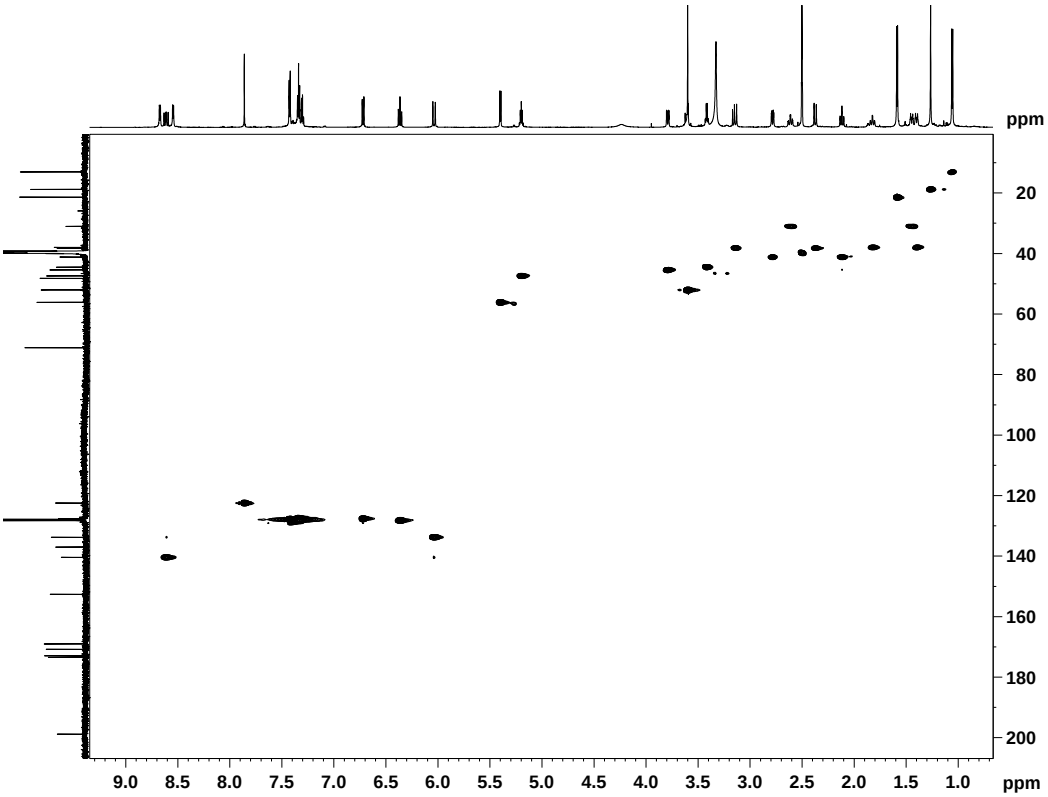


Fig. S17-RE. ^{1}H-^{1}H COSY spectra of (*R*)-PGME-LNM E2 (DMSO-d_6)

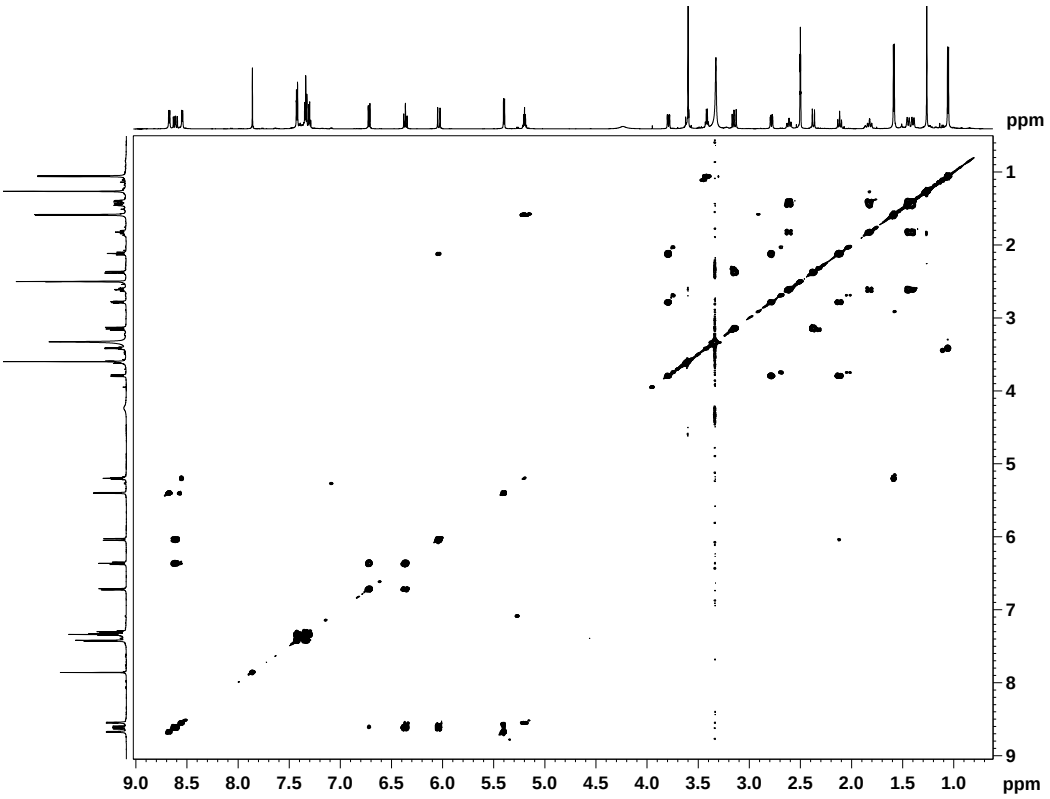


Fig. S17-RF. HMBC spectra of (*R*)-PGME-LNM E2 (DMSO-d_6)

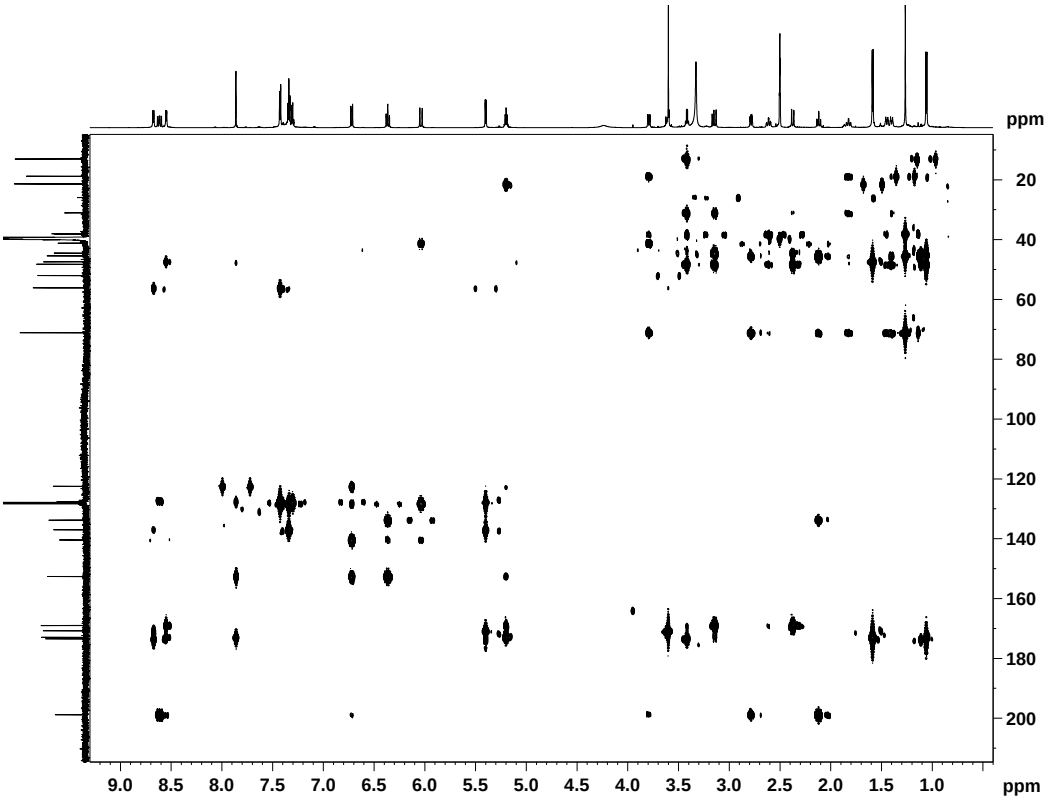


Fig. S17-RG. ROESY spectra of (*R*)-PGME-LNM E2 (DMSO-d_6)

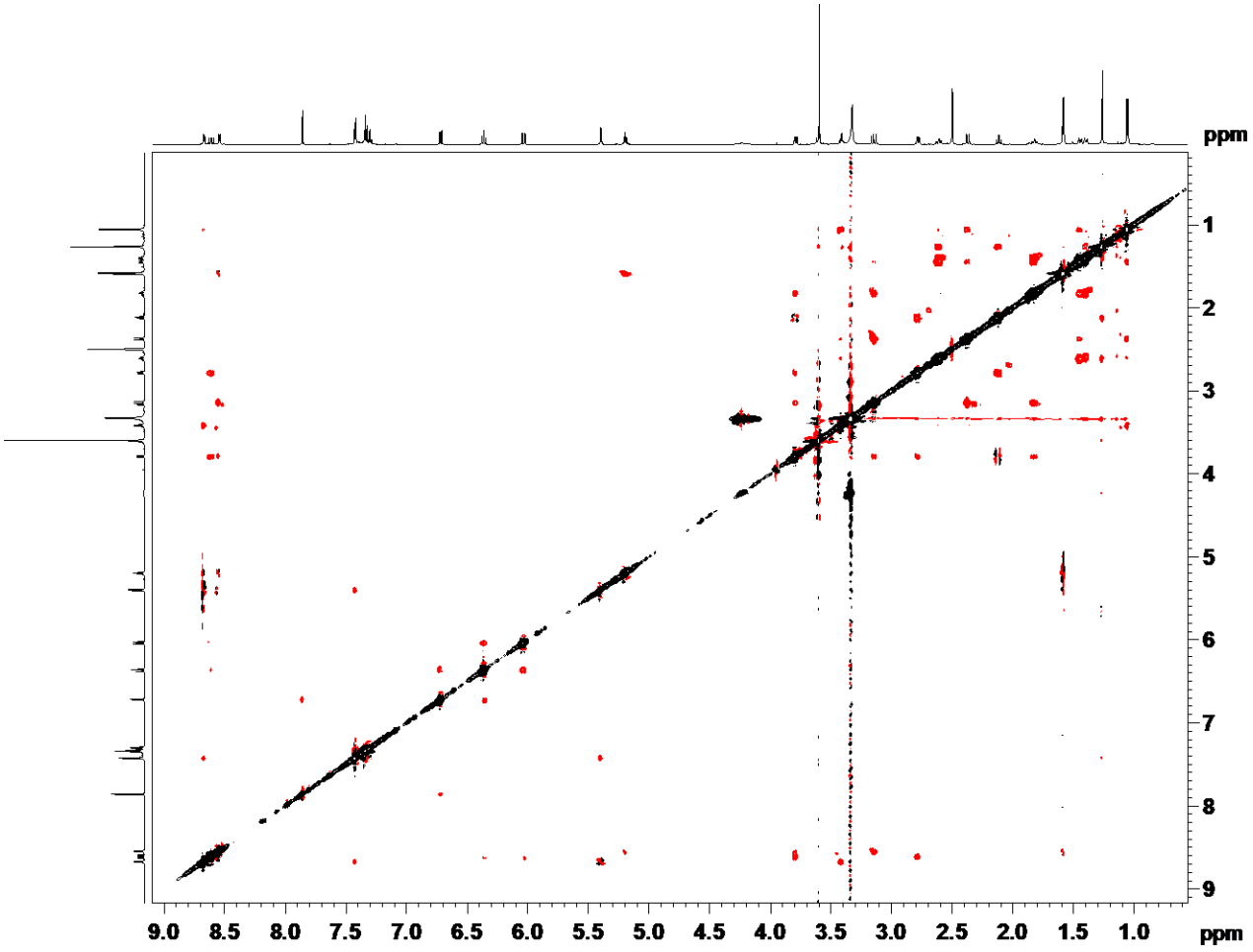


Fig. S17-RH. HR-MS (ESI) spectra of (*R*)-PGME-LNM E2

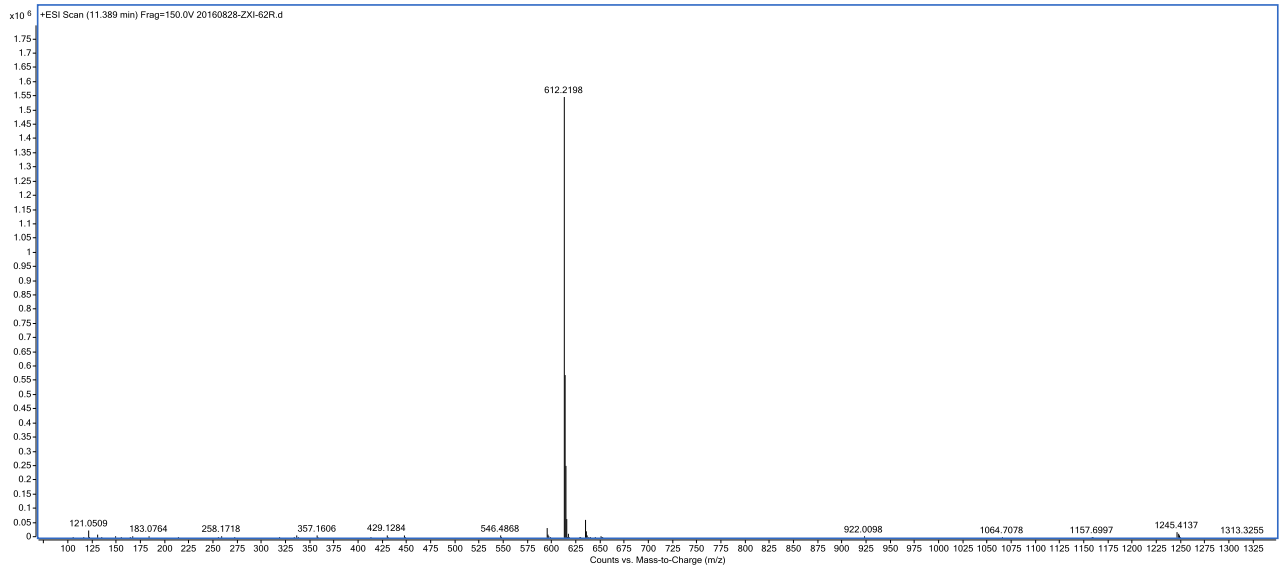


Fig. S18. Construction of the in-frame deletion mutant strain SB22001 (i.e. Δ*wsmW*) and confirmation of its genotype by Southern analysis. (**A**) Schematic representation for the deletion of *wsmW* in strain CB02120-2 by homologous recombination. The probe (503 bp) for Southern blot was amplified with primers orf25-SBlot-5 and orf25-SBlot-3 using genomic DNA of strain CB02120-2 as the template. Apr^S, apramycin sensitive; Apr^R, apramycin resistance; *acc(3)IV*, apramycin resistance gene. (**B**) Southern blot verification of strain CB02120-2 wild type (3.3 kb) and mutant strain SB22001 (1.3 kb). Genomic DNAs were digested with *Kpn*I and then used for Southern analysis. Lane 1, DNA marker VII, DIG-labeled (Roche); lane 2, strain CB02120-2 wild type; lane 3, SB22001.

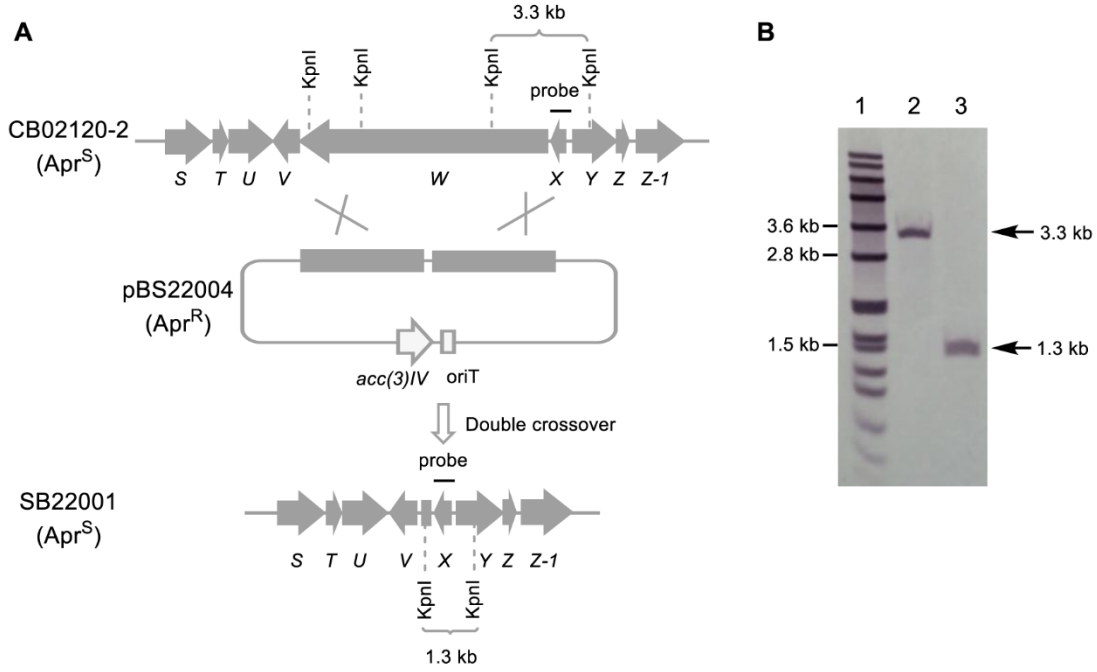


Fig. S19. Proposed biosynthetic pathway of weishanmycins (WSMs) in *S*. sp. CB002120-2. (**A**) Genetic organization of *wsm* gene cluster, together with three cosmids pBS22001 (8E4), pBS22002 (7F6) and pBS22003 (3F12) covering the whole gene cluster. (**B**) The biosynthetic machinery of WSMs featuring a similar hybrid NRPS-AT-less type I PKS to that of LNM consisting of WsmQPWGR and a discrete AT WsmJ, producing the nascent product with 3-thiol group, which was further oxidized to give WSM A1, A2, and A3. (**C**) Two possible pathways for the formation of propylmalonyl-*S*-ACP as the extender unit during the biosynthesis of WSMs. The propylmalonyl-*S*-ACP can be formed either by loading propylmalonyl-CoA, which is biosynthesized from malonyl-CoA and propionyl-CoA by WsmZ2Z3Z4, directly to the ACP of PKS module-3 by an unidentified trans-acting AT, or by loading malonyl-CoA to the ACP of PKS module-3 by trans-acting AT WsmJ, followed by modification of the resultant malonyl-*S*-ACP by WsmZ2Z3Z4 within the PKS before being incorporated into the growing intermediate.

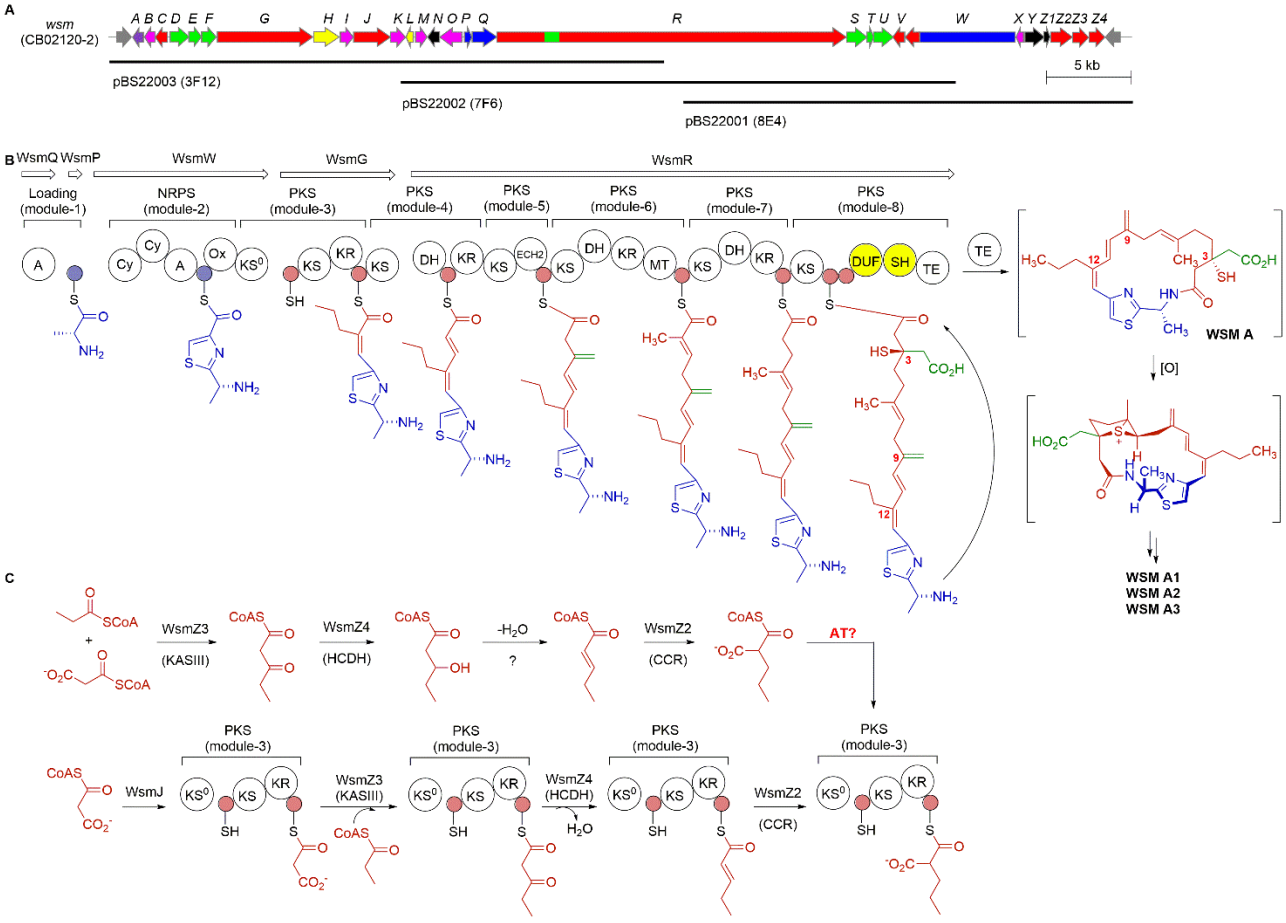


Fig. S20. Elucidation of new metabolites WSM A1, A2, and A3 isolated from strain CB02120-2 wild type. Key ^{1}H-^{1}H COSY, HMBC and ROESY correlations are shown.

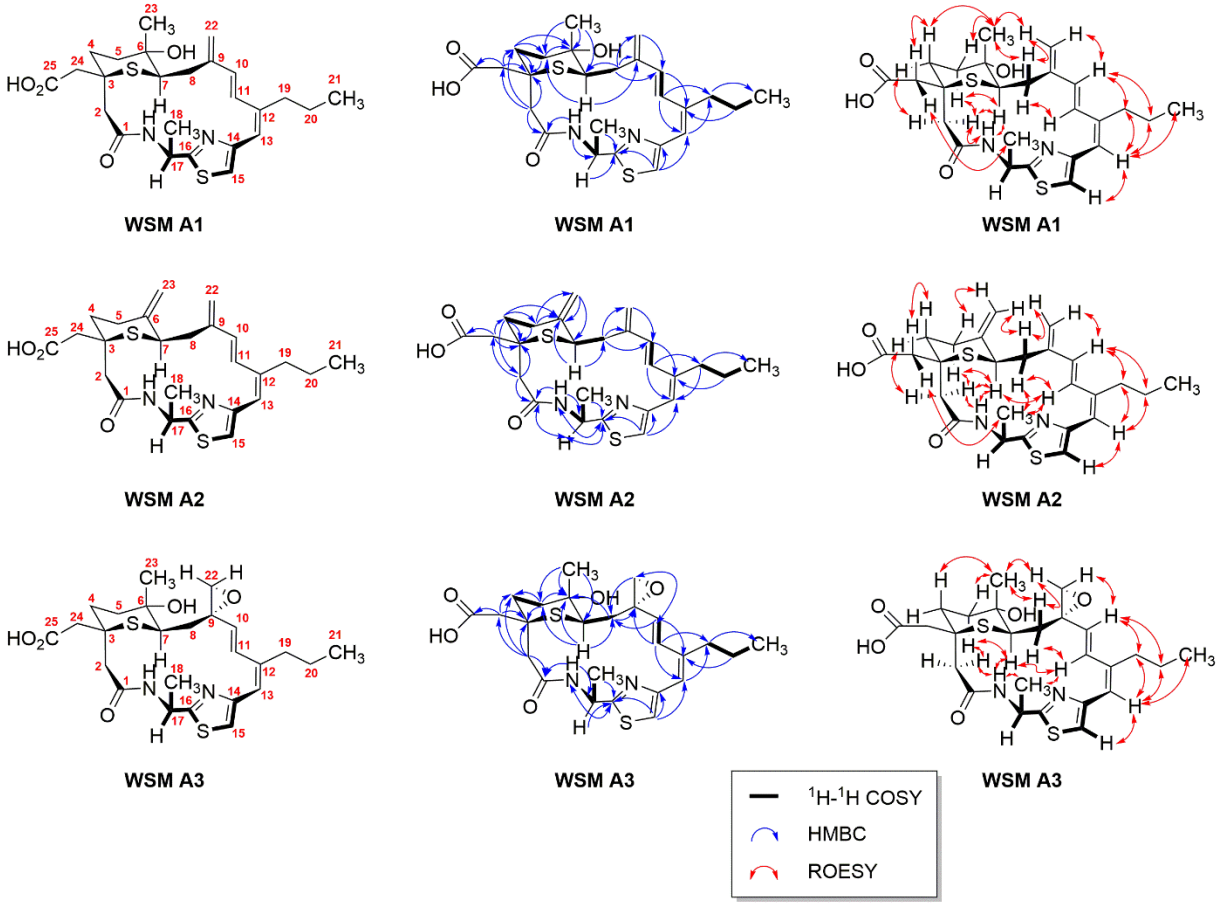


Fig. S21A. ^{1}H NMR spectra of WSM A1 (700 MHz, DMSO-d_6)

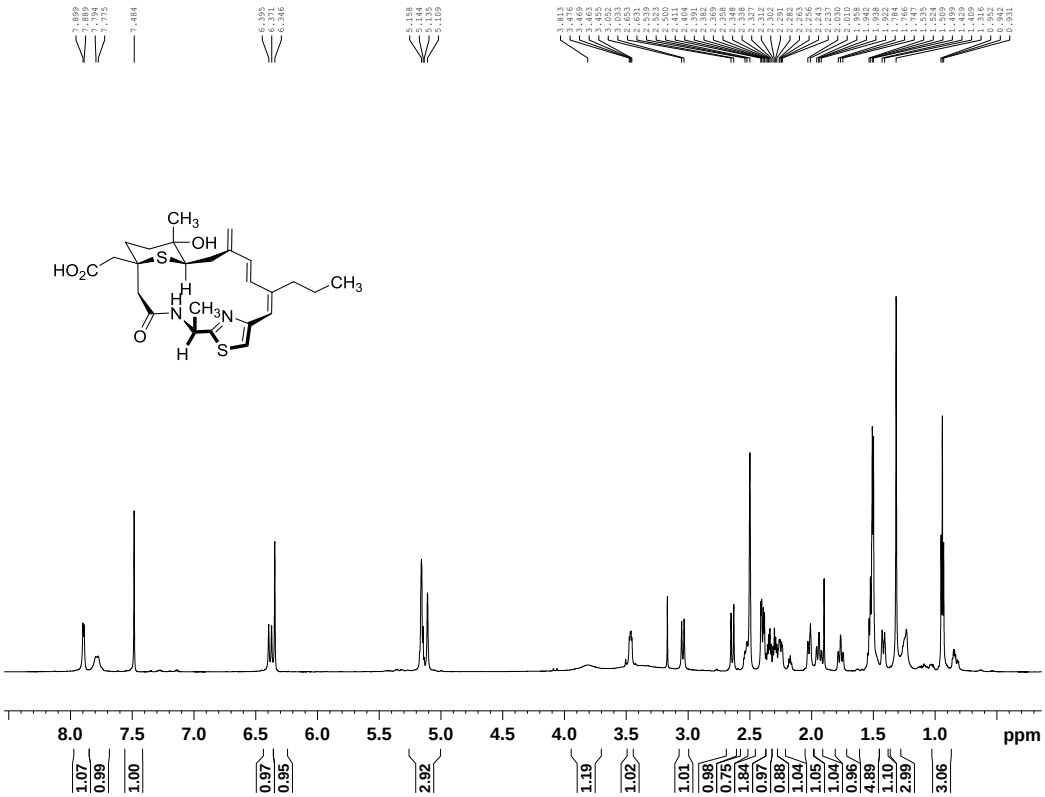


Fig. S21B. ^{13}C NMR spectra of WSM A1 (175 MHz, DMSO-d_6)

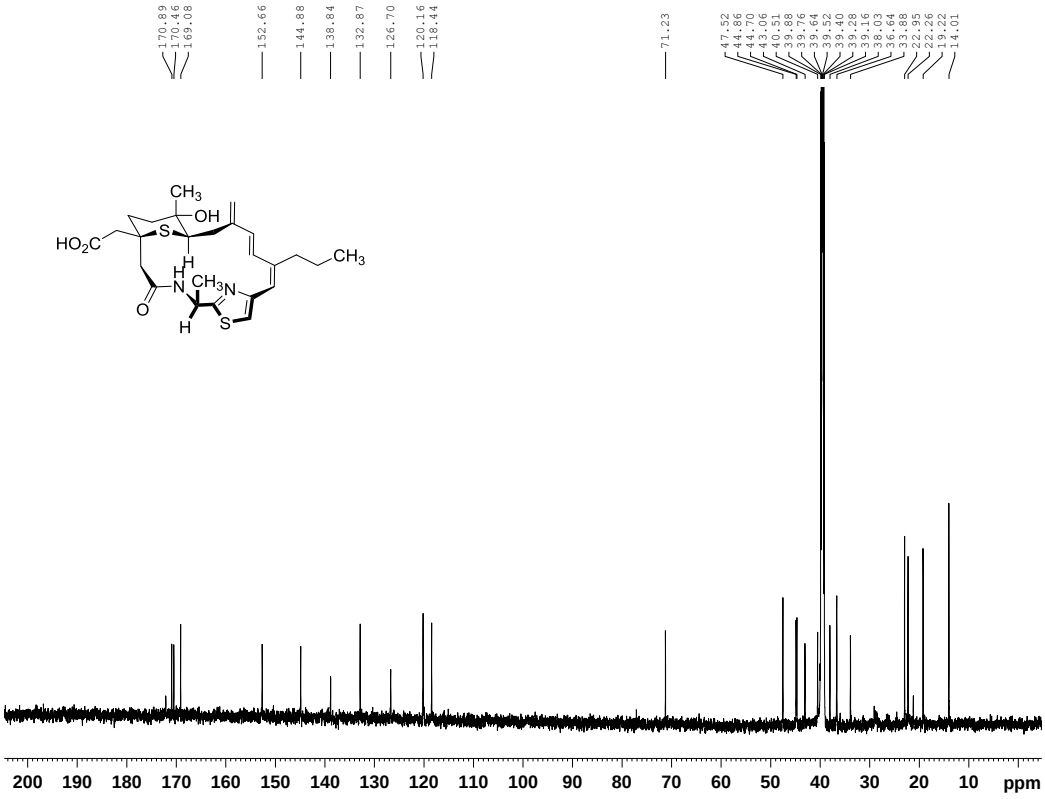


Fig. S21C. DEPT-135 spectra of WSM A1 (DMSO-d_6)

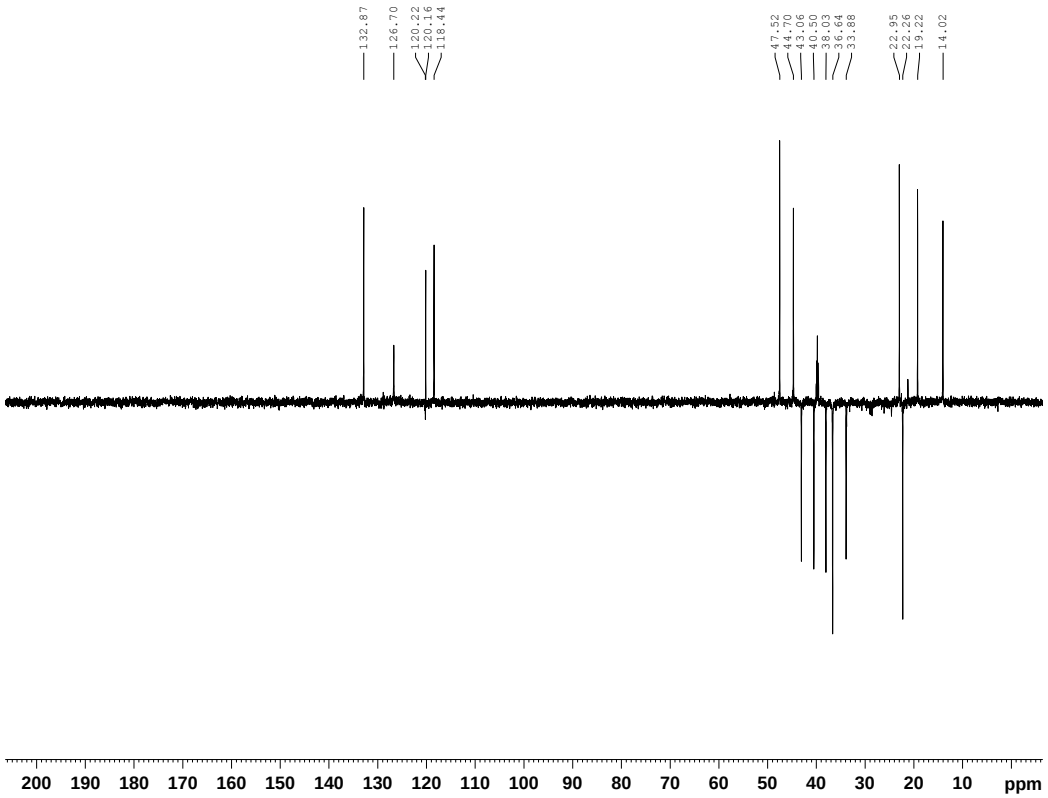


Fig. S21D. HSQC spectra of WSM A1 (DMSO-d_6)

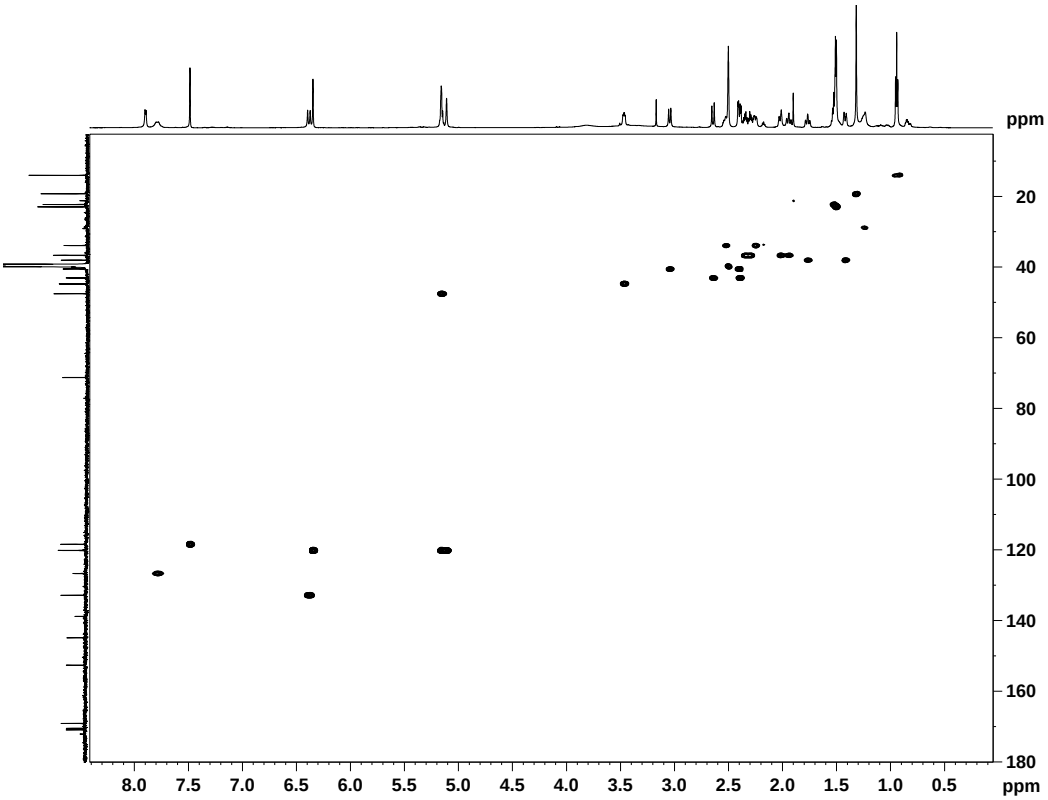


Fig. S21E. ^{1}H-^{1}H COSY spectra of WSM A1 (DMSO-d_6)

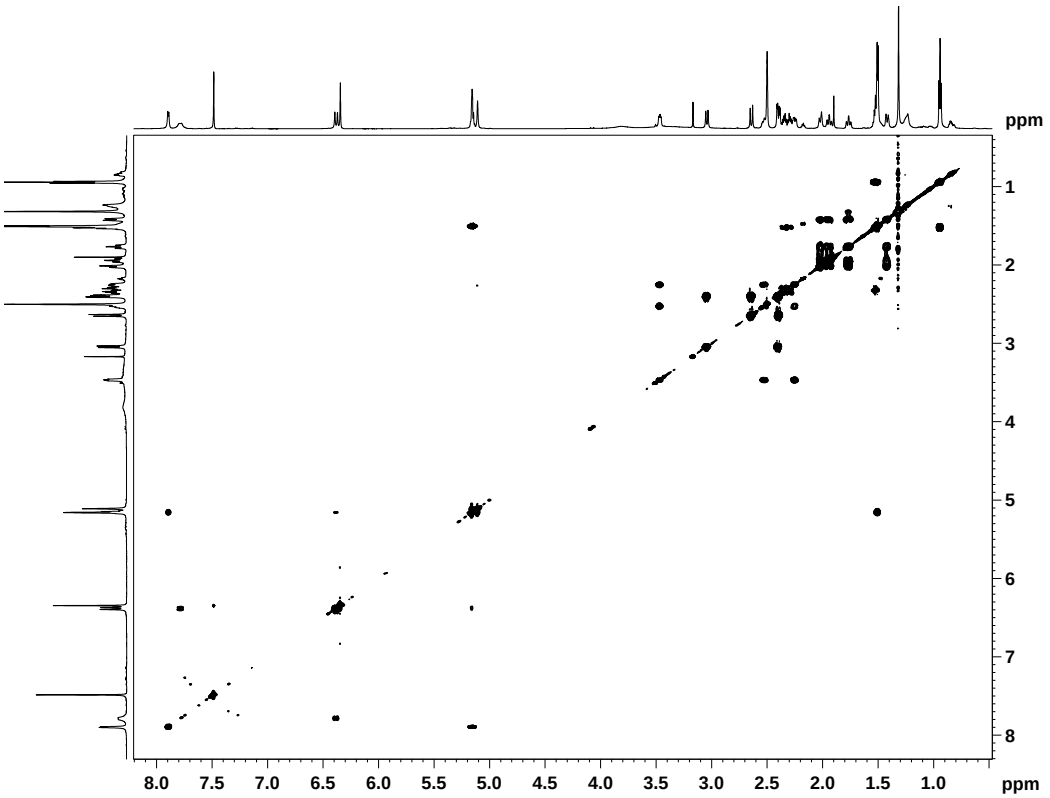


Fig. S21F. HMBC spectra of WSM A1 (DMSO-d_6)

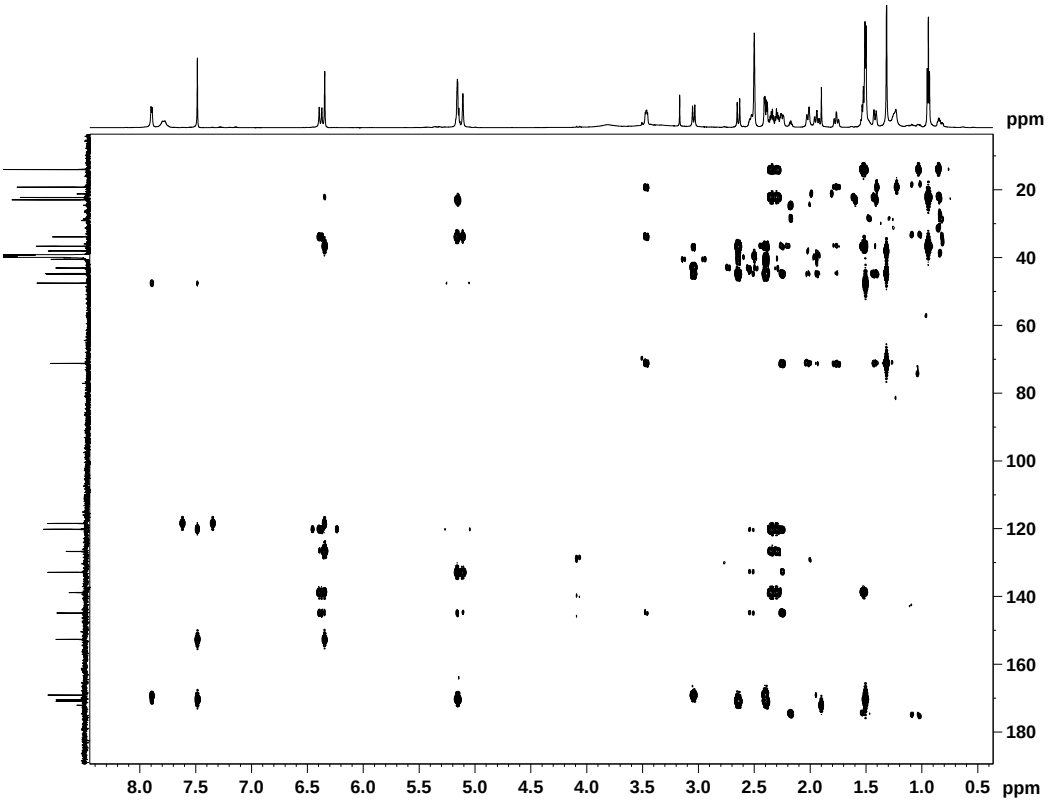


Fig. S21G. ROESY spectra of WSM A1 (DMSO-d_6)

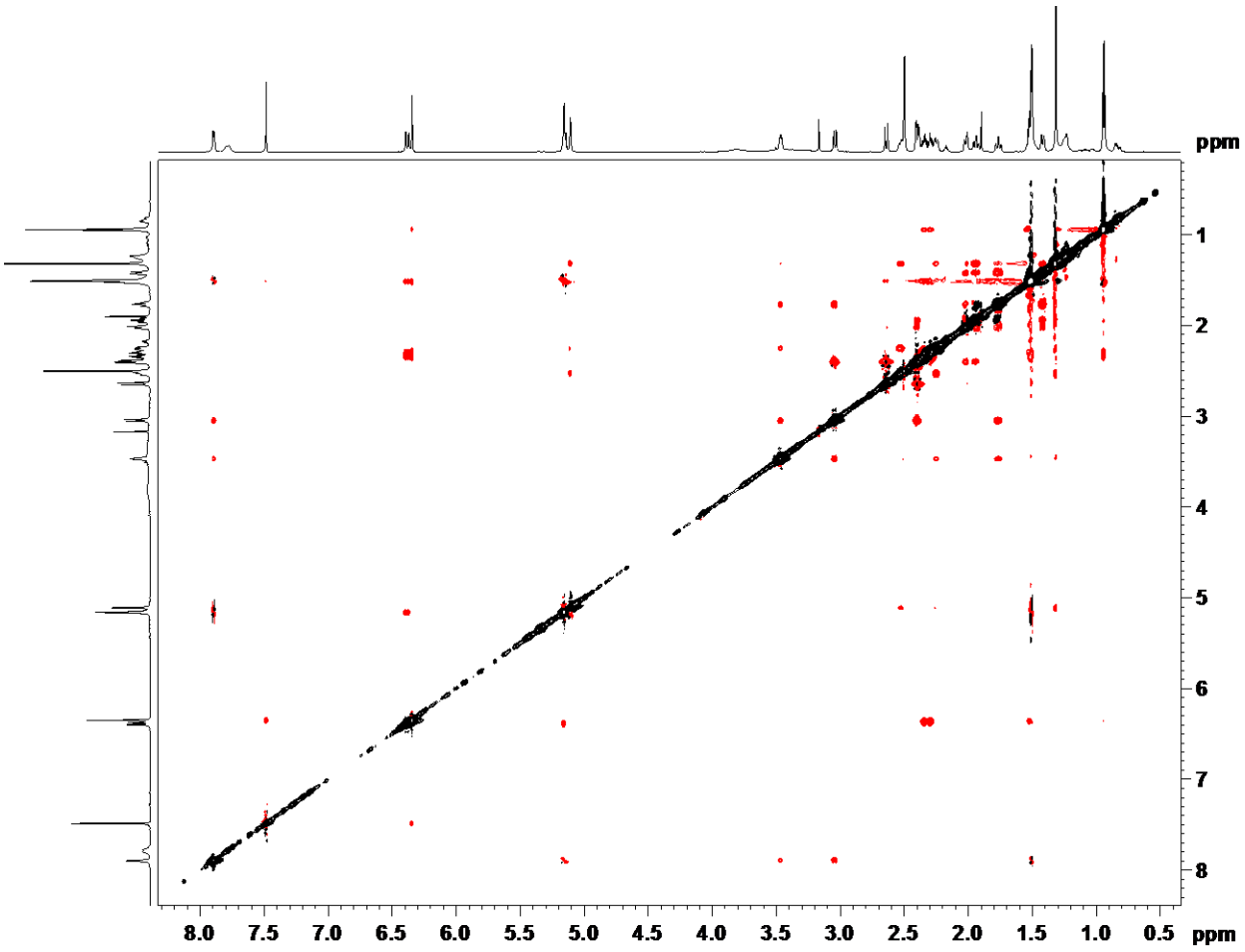


Fig. S21H. HR-MS (ESI) spectra of WSM A1

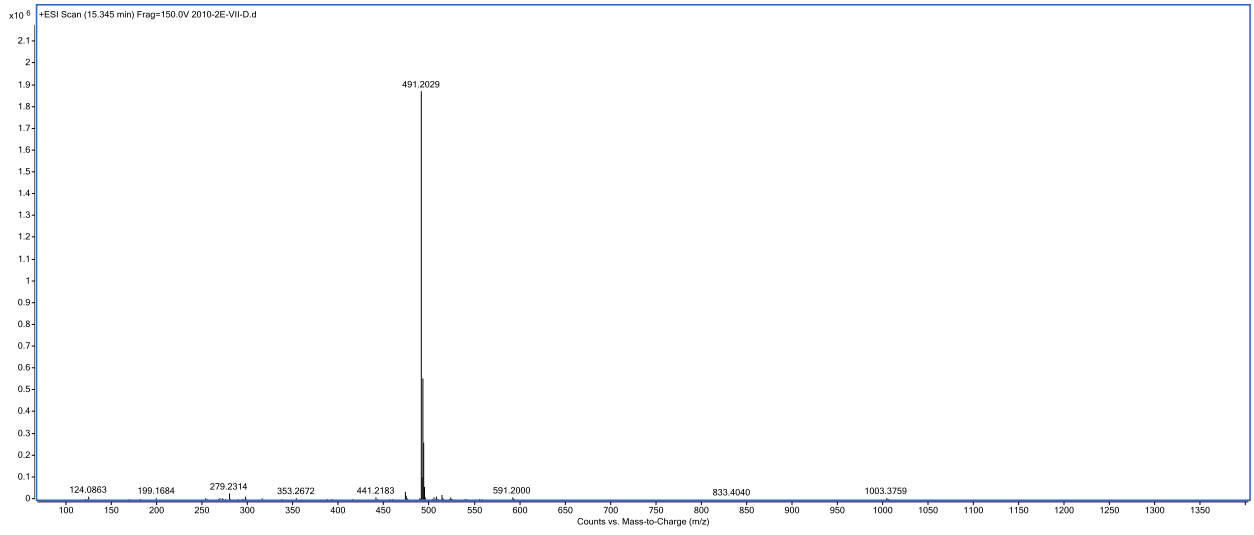


Fig. S22A. ^{1}H NMR spectra of WSM A2 (700 MHz, DMSO-d_6)

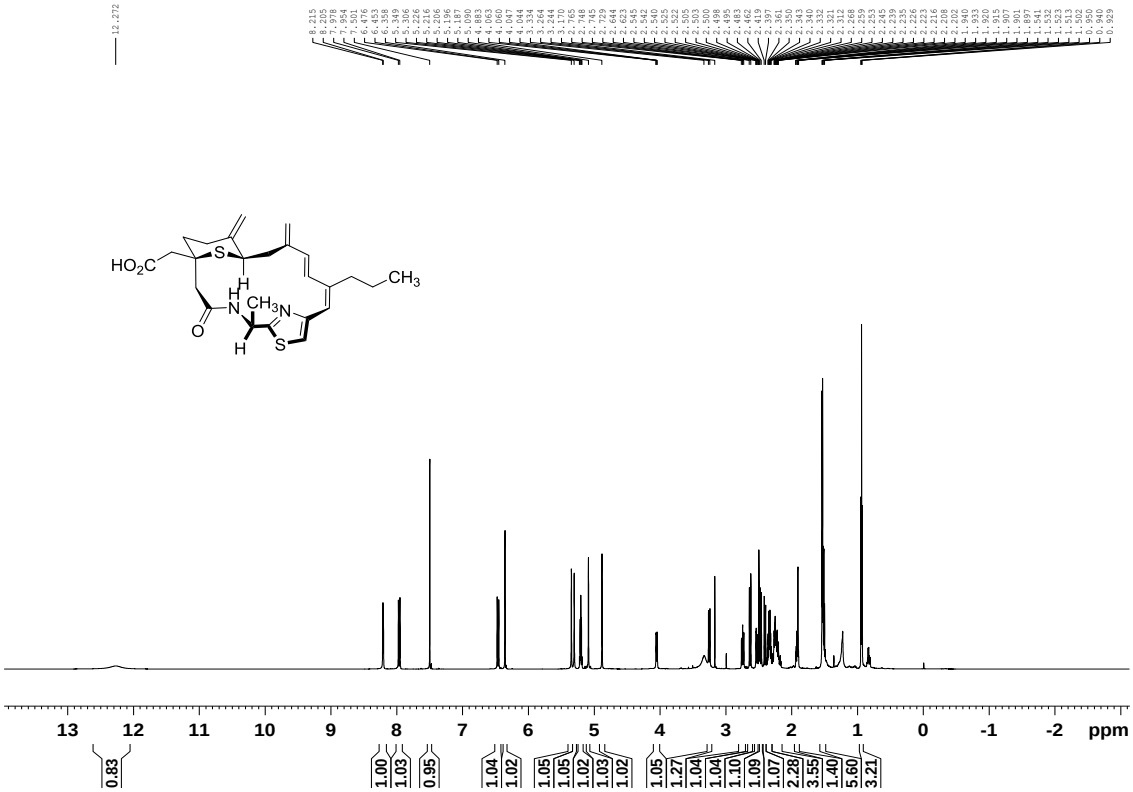


Fig. S22B. ^{13}C NMR spectra of WSM A2 (175 MHz, DMSO-d_6)

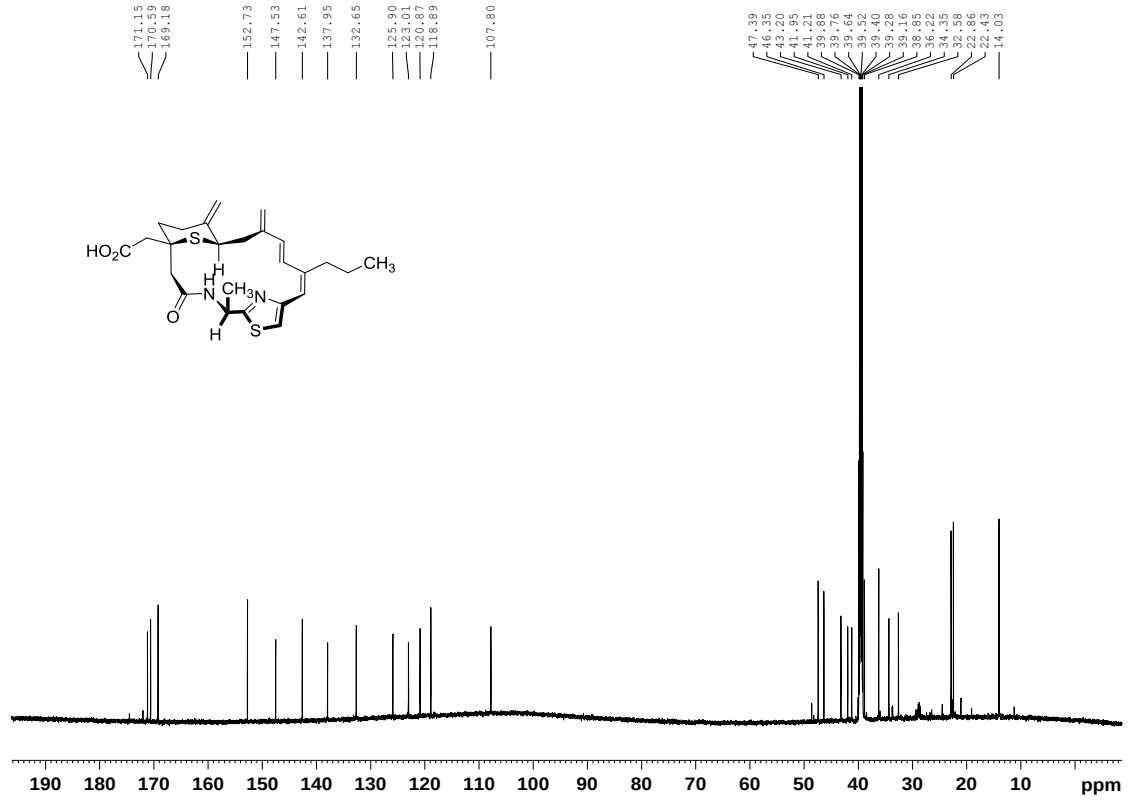


Fig. S22C. DEPT-135 spectra of WSM A2 (DMSO-d_6)

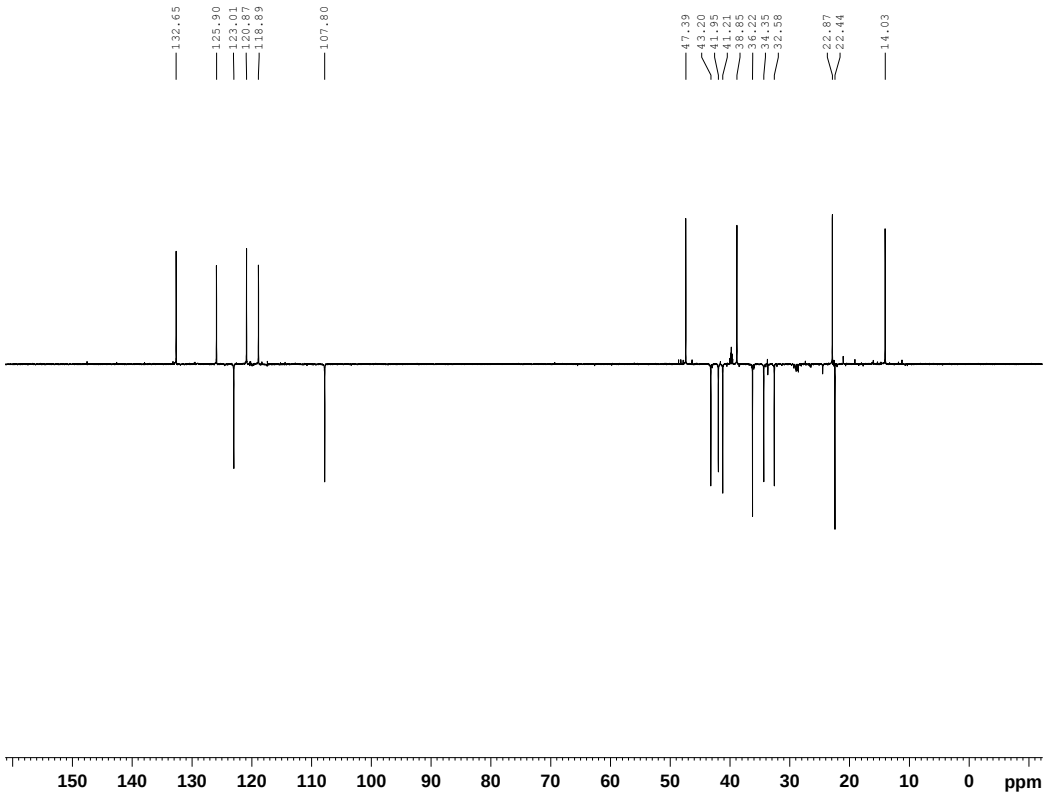


Fig. S22D. HSQC spectra of WSM A2 (DMSO-d_6)

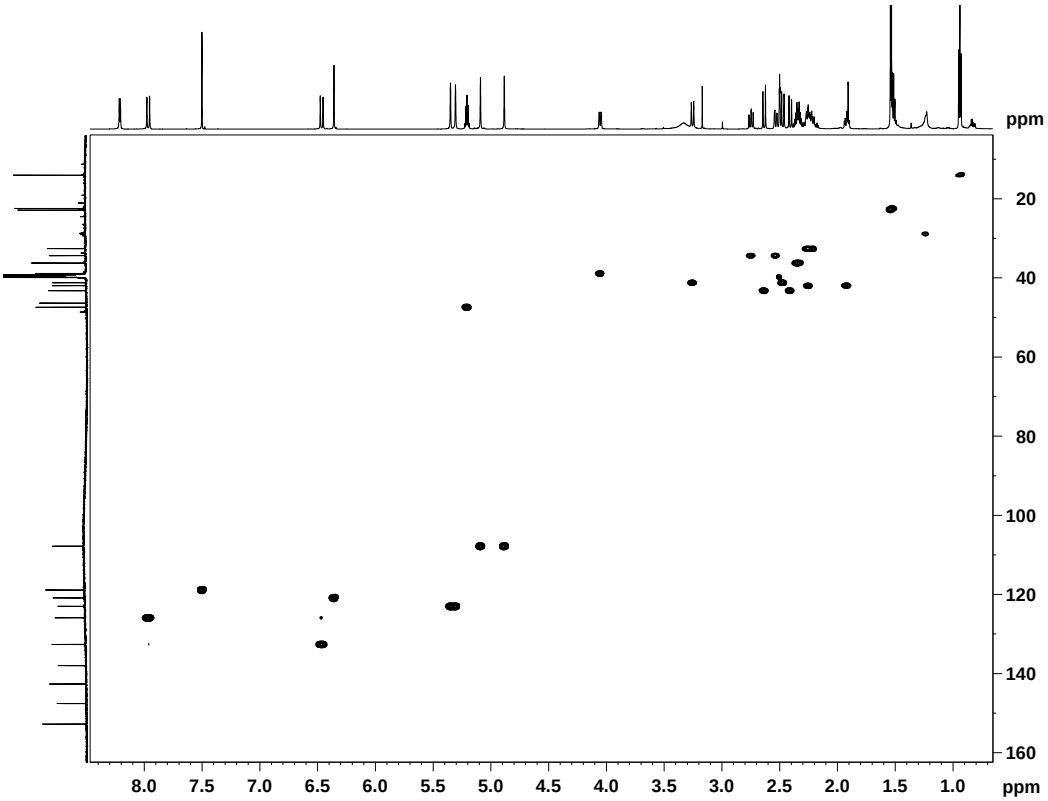


Fig. S22E. ^{1}H-^{1}H COSY spectra of WSM A2 (DMSO-d_6)

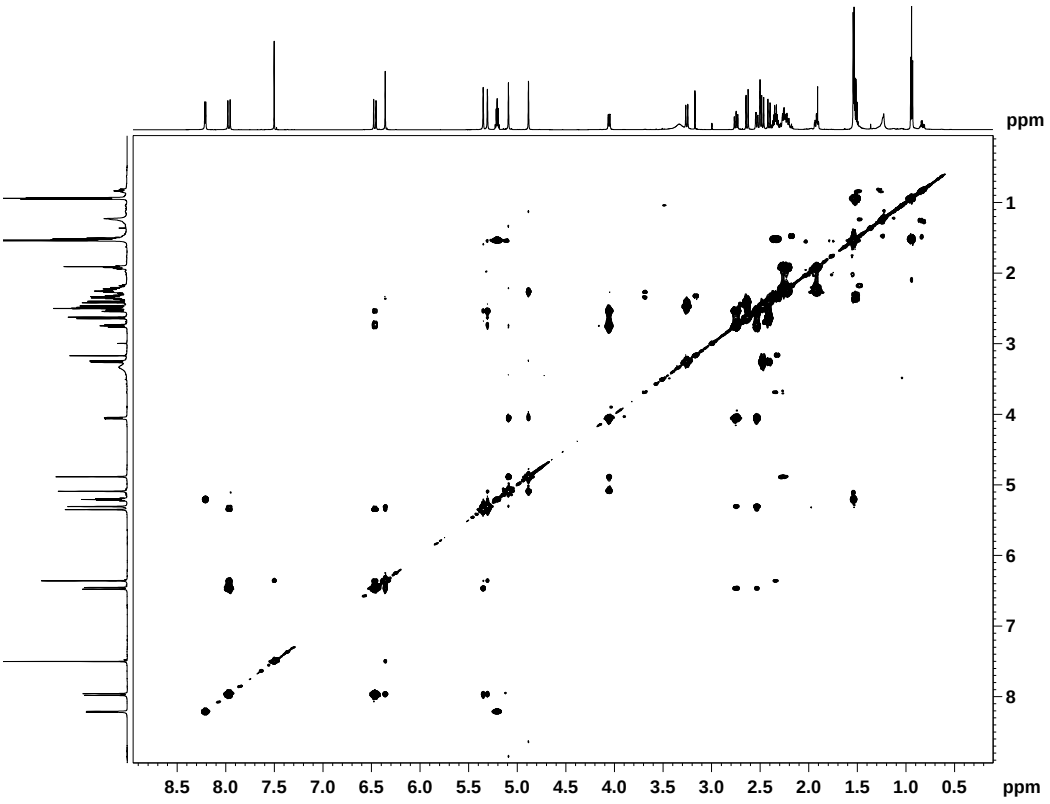


Fig. S22F. HMBC spectra of WSM A2 (DMSO-d_6)

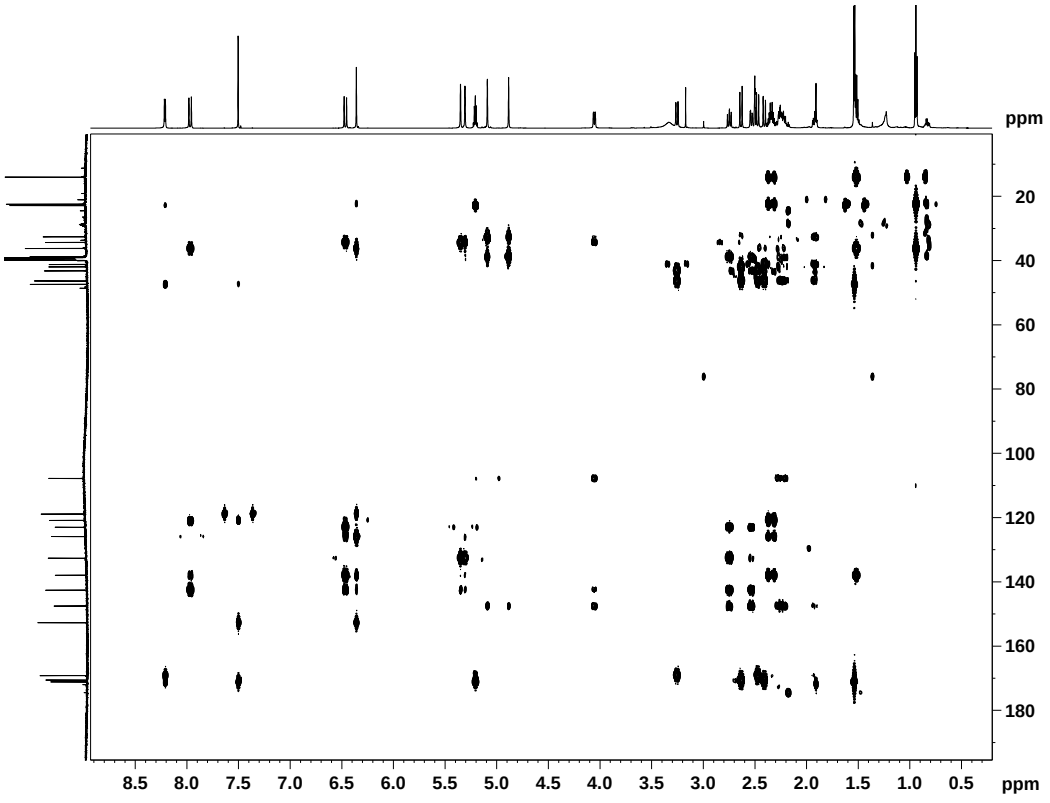


Fig. S22G. ROESY spectra of WSM A2 (DMSO-d_6)

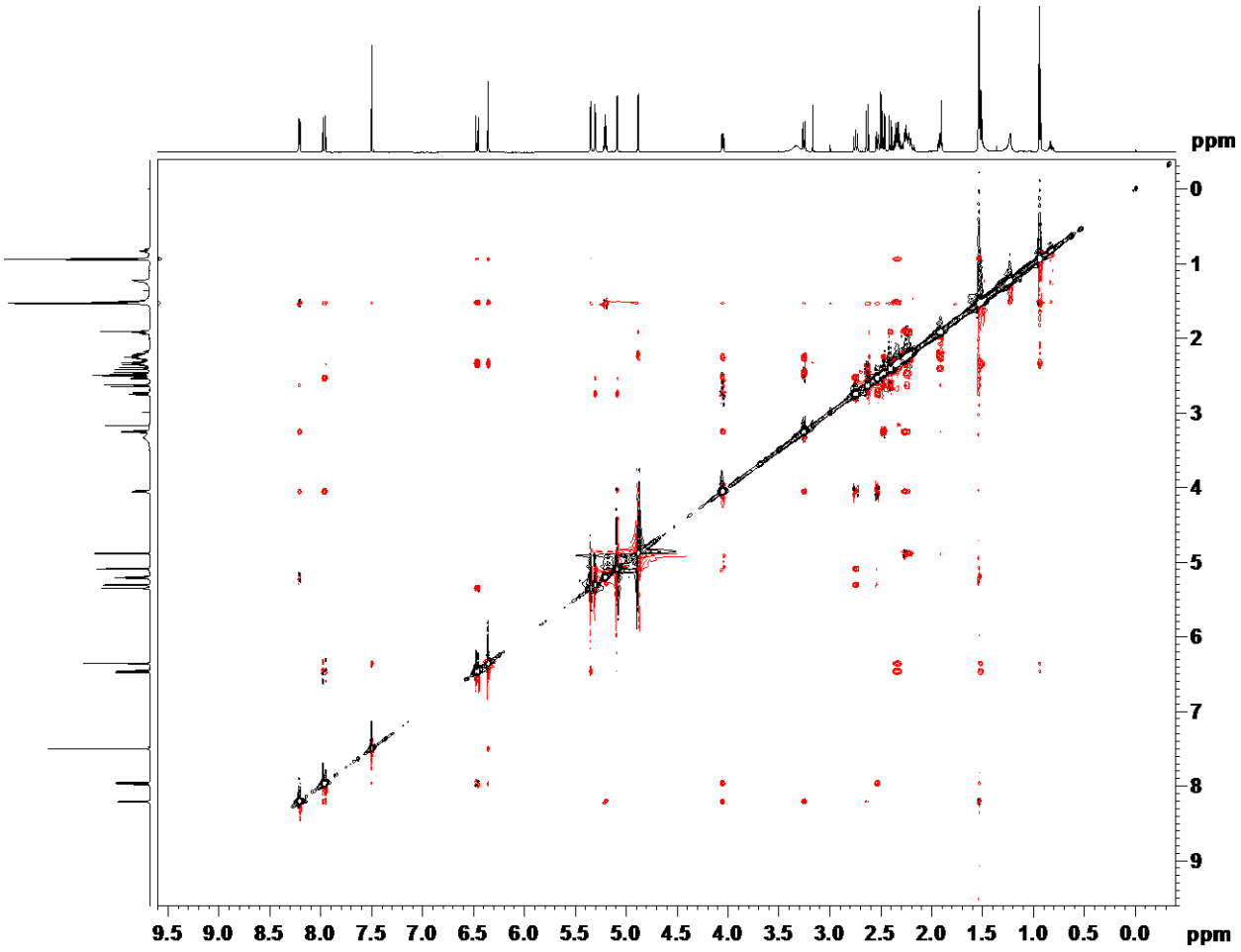


Fig. S22H. HR-MS (ESI) spectra of WSM A2

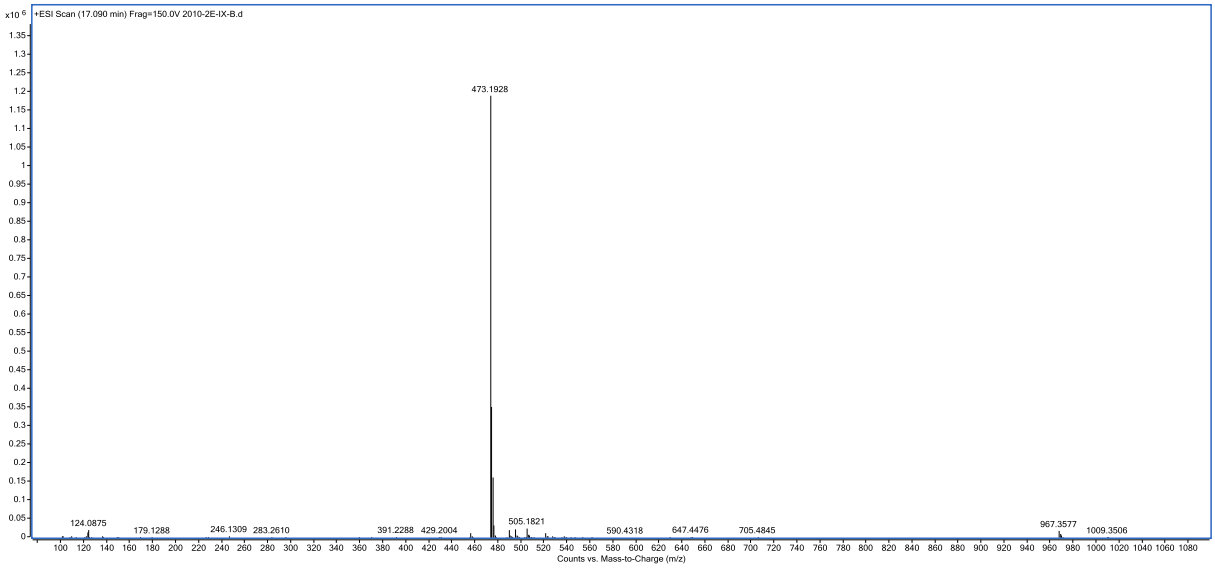


Fig. S23A. ^{1}H NMR spectra of WSM A3 (700 MHz, $CDCl_3$)

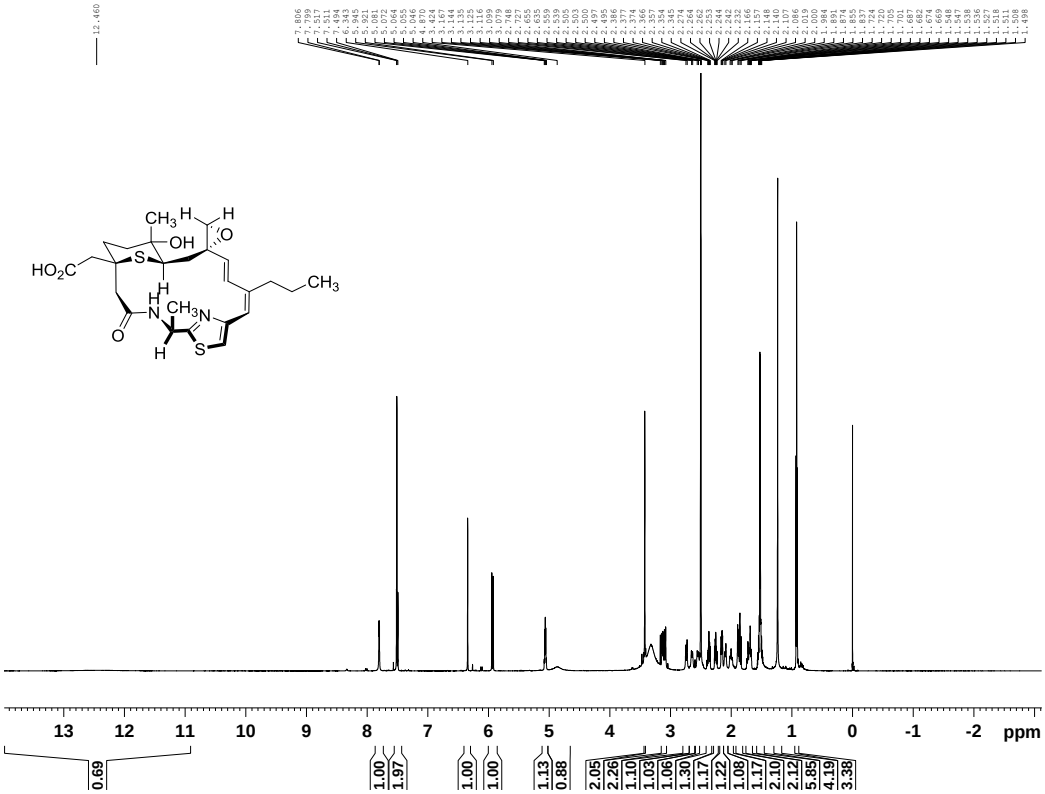


Fig. S23B. ^{13}C NMR spectra of WSM A3 (175 MHz, DMSO-d_6)

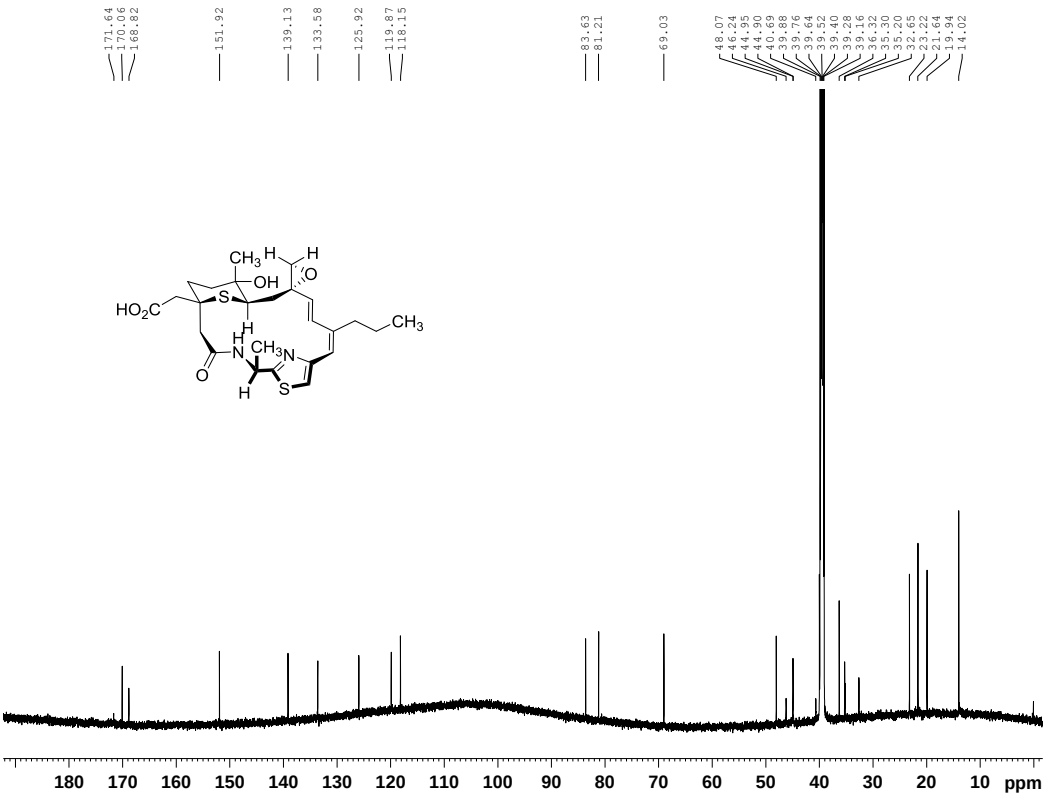


Fig. S23C. DEPT-135 spectra of WSM A3 (DMSO-d_6)

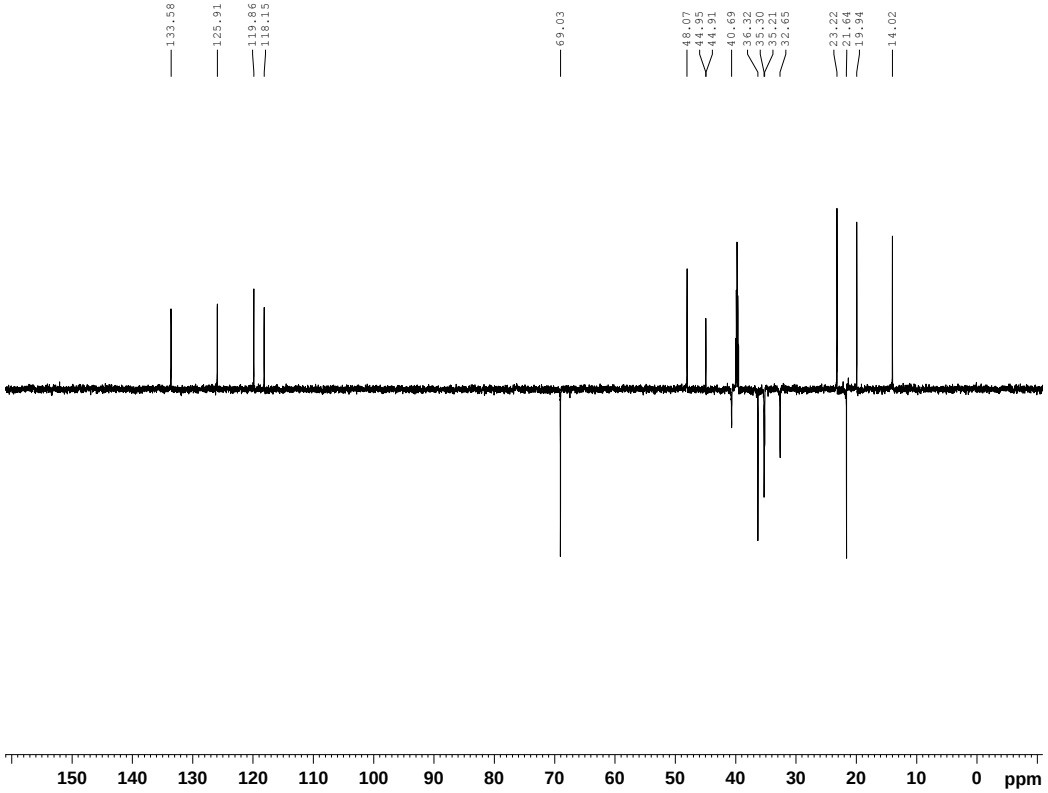


Fig. S23D. HSQC spectra of WSM A3 (DMSO-d_6)

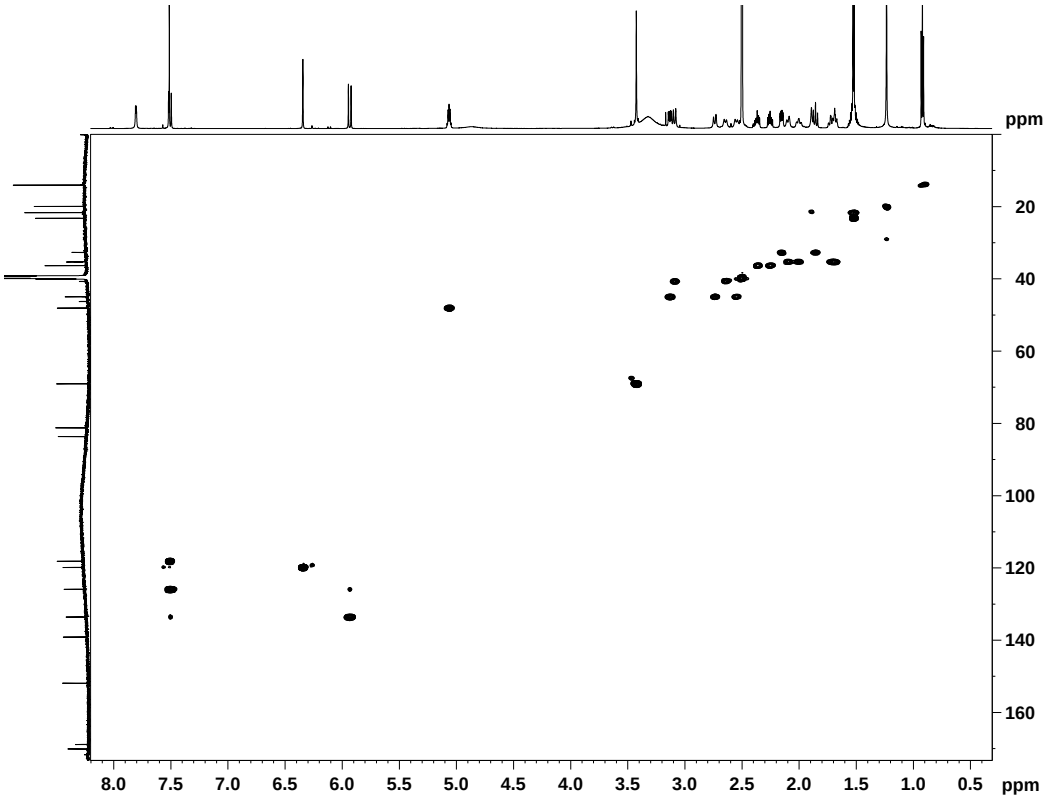


Fig. S23E. ^{1}H-^{1}H COSY spectra of WSM A3 (DMSO-d_6)

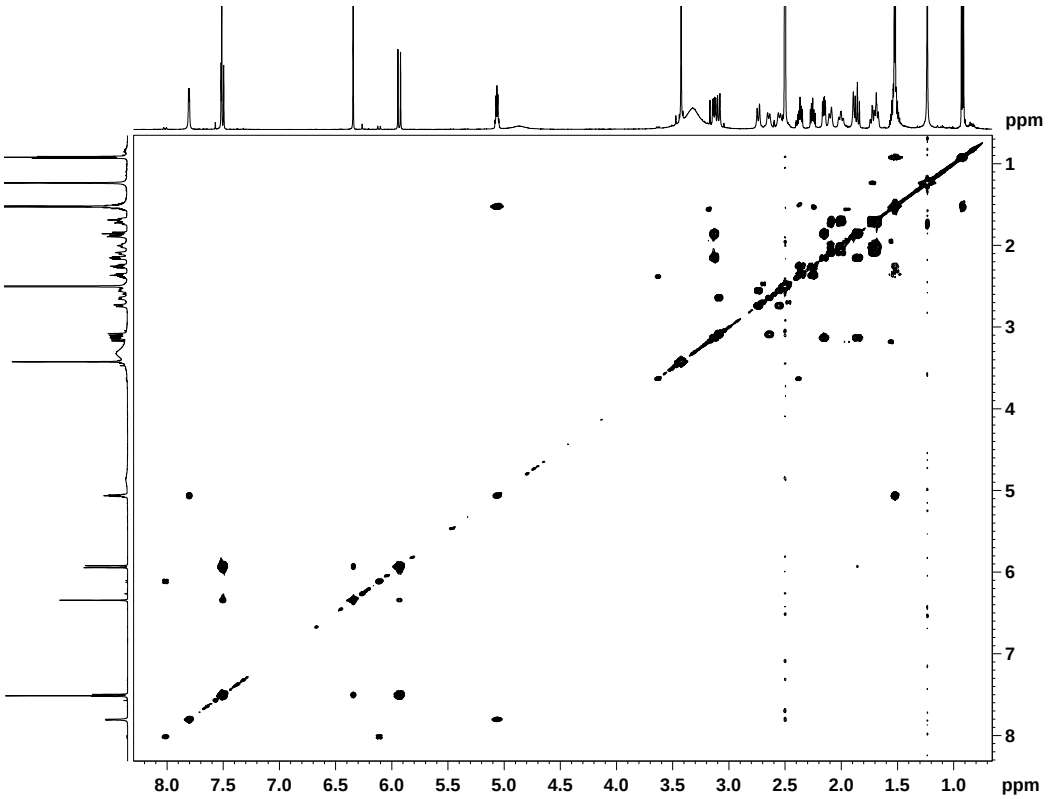


Fig. S23F. HMBC spectra of WSM A3 (DMSO-d_6)

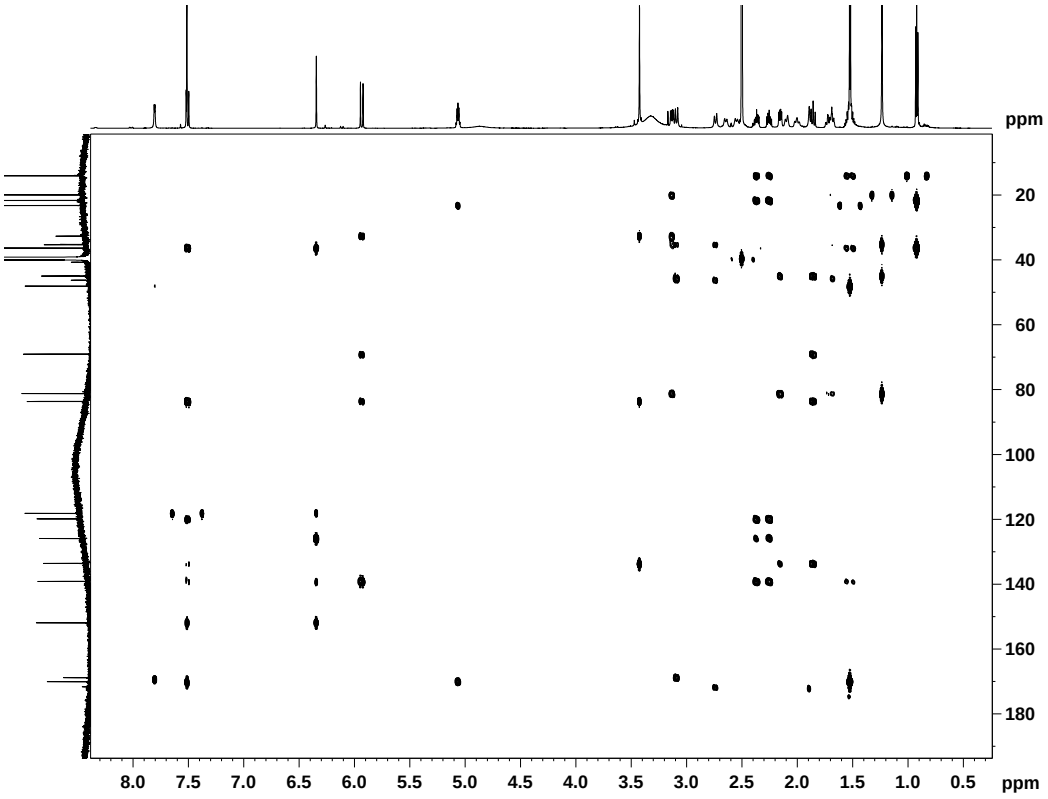


Fig. S23G. ROESY spectra of WSM A3 (DMSO-d_6)

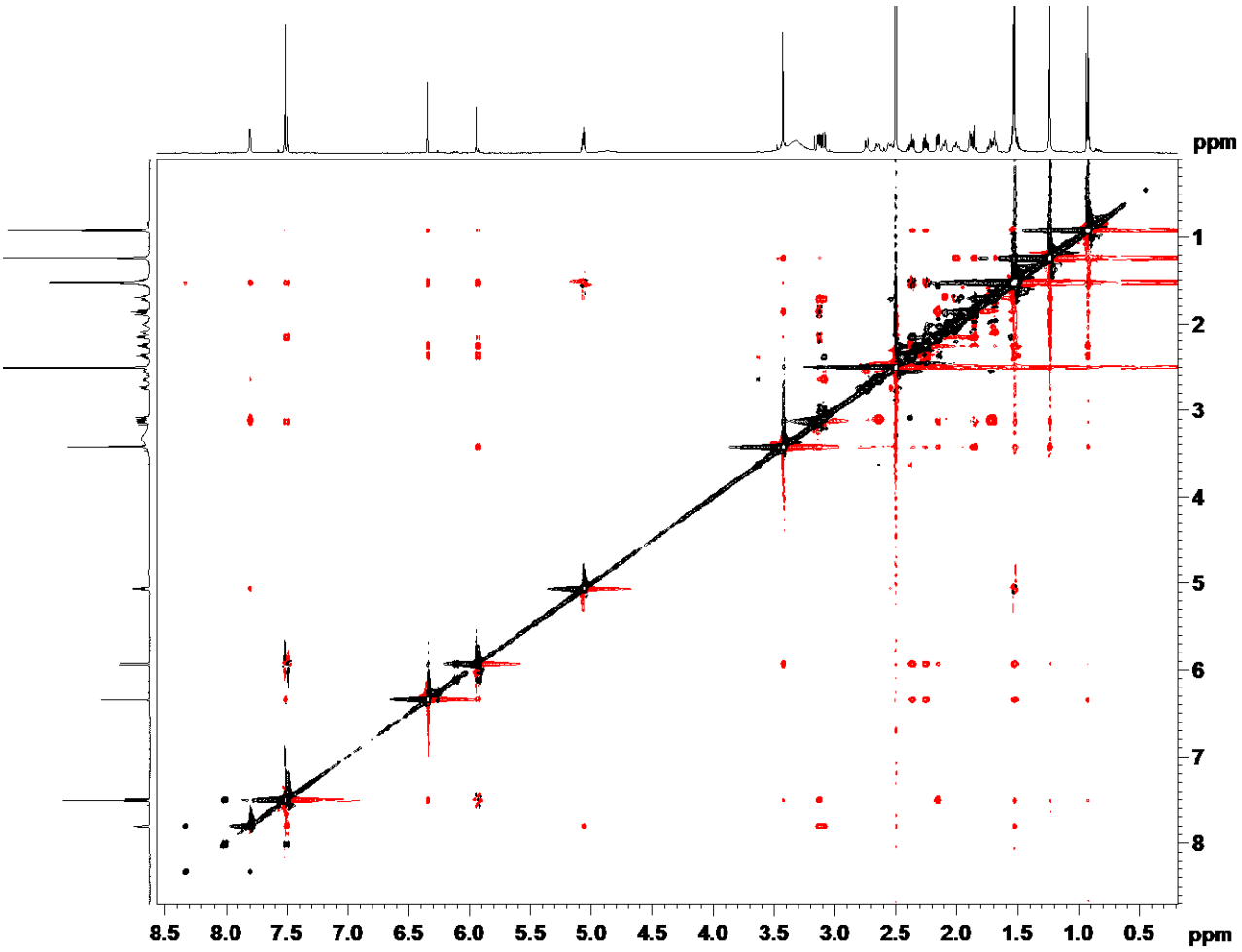


Fig. S23H. HR-MS (ESI) spectra of WSM A3

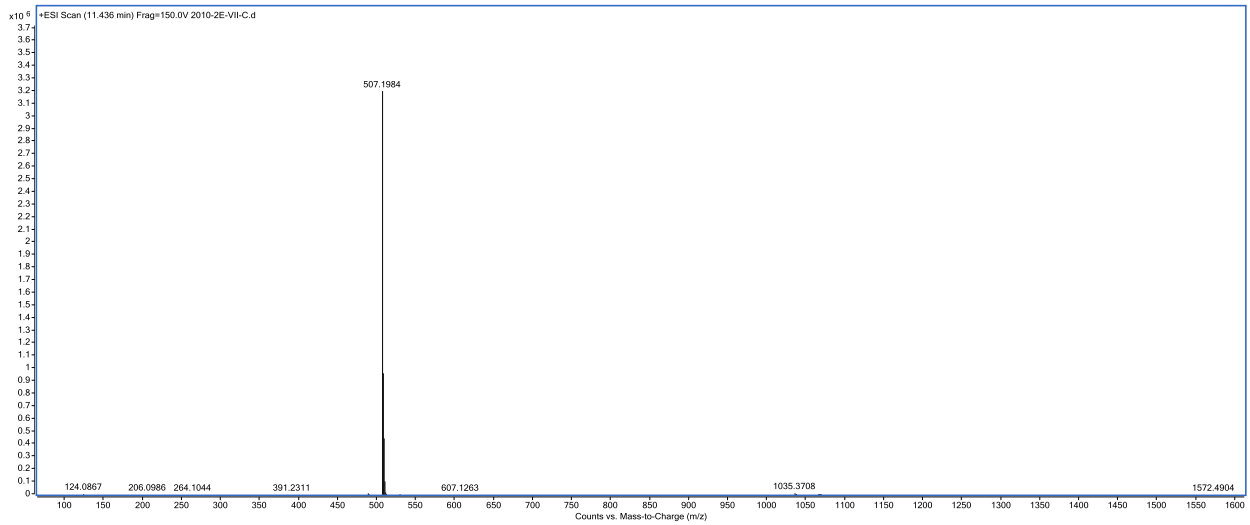


Fig. S24. Determination of the absolute configuration of WSM A2 using its (*S*)- and (*R*)-PGME derivatives. The two major conformations of (*S*)- and (*R*)-PGME-WSM A2 resembling that of GNM B2 derivatives (**Fig. S14**), which are different from that of the corresponding derivatives of LNM E2 probably due to the missing methyl group at C24 position, are shown in **A**. In the ROESY spectra (**B**), the presence of correlations between 24-H_b and 4-H_b, between 2-H_b and 4-H_a, as well as the absent correlations between 24-H_a and 2-H/4-H, support that C24-C25 and C3-S are *anti*-oriented, and 24-H_b takes a pseudo-axial position in the preferred conformation at C24. Two major conformations (**A**, conformations I and II), due to the rotation of C24-C25 single bond, could be observed based on the ROESY correlations between 25-N*H* and 24-H_a/24-H_b/2-H_b (**C** and **D**). The relatively weak correlation between 25-N*H* and 24-H_b (**C**) supports that conformation I is the favored conformation of (*R*)-PGME-WSM A2, while the relatively weak correlation between 25-N*H* and 2-H_b (**D**) supports that conformation II is the favored conformation of (*S*)-PGME-WSM A2. Although several hydrogens in the macrolactam backbone are positioned in opposite sides of PGME plane in two different conformations, 24-H_b and 2-H_a/H_b are always positioned at the two different sides of the PGME plane in both conformations I and II, hence making them diagnostic useful for determination of the absolute configuration. Based on the differences of chemical shift in ^{1}H NMR between (*R*)- and (*S*)-PGME-WSM A2 ($\Delta\delta = \delta_{(R)} - \delta_{(S)}$) as shown in **A**, the absolute configuration of C3 is elucidated to be 3*S* for WSM A2 [3*R* for (*R*)/(*S*)-PGME-WSM A2]. Additionally, since the priming amino acid during the biosynthesis has been determined biochemically as D-Alanine (**Fig. 3B**), the ROESY correlations between 24-H_a and 18-CH_3 (**B**, **E**), and between 11-H and 18-CH_3 (**F**), further support the determined absolute configuration at C3, which is opposite to that of LNM E2 (**Fig. S16**). H_a denotes one of the geminal hydrogens appearing at lower field in ^{1}H NMR, and H_b denotes one of the geminal hydrogens appearing at higher field in ^{1}H NMR.

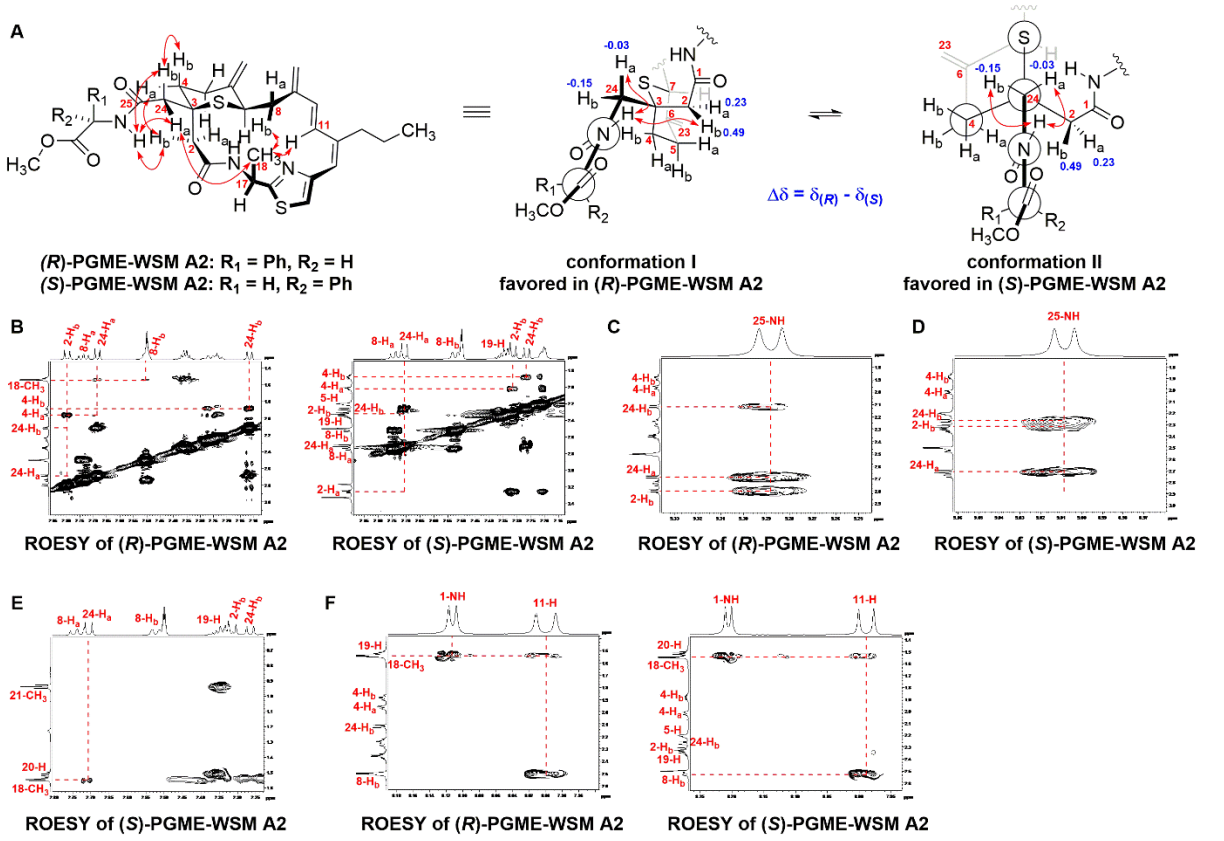


Fig. S25-SA. ^{1}H NMR spectra of (*S*)-PGME-WSM A2 (700 MHz, DMSO-d_6)

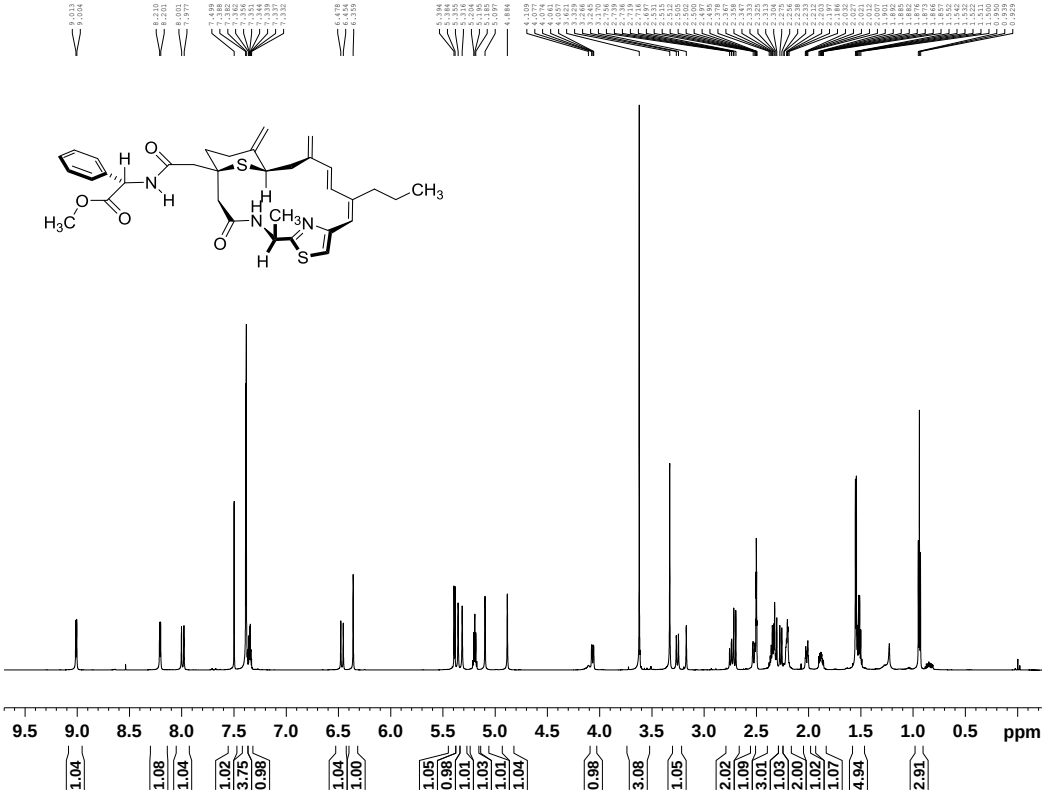


Fig. S25-SB. ^{13}C NMR spectra of (*S*)-PGME-WSM A2 (175 MHz, DMSO-d_6)

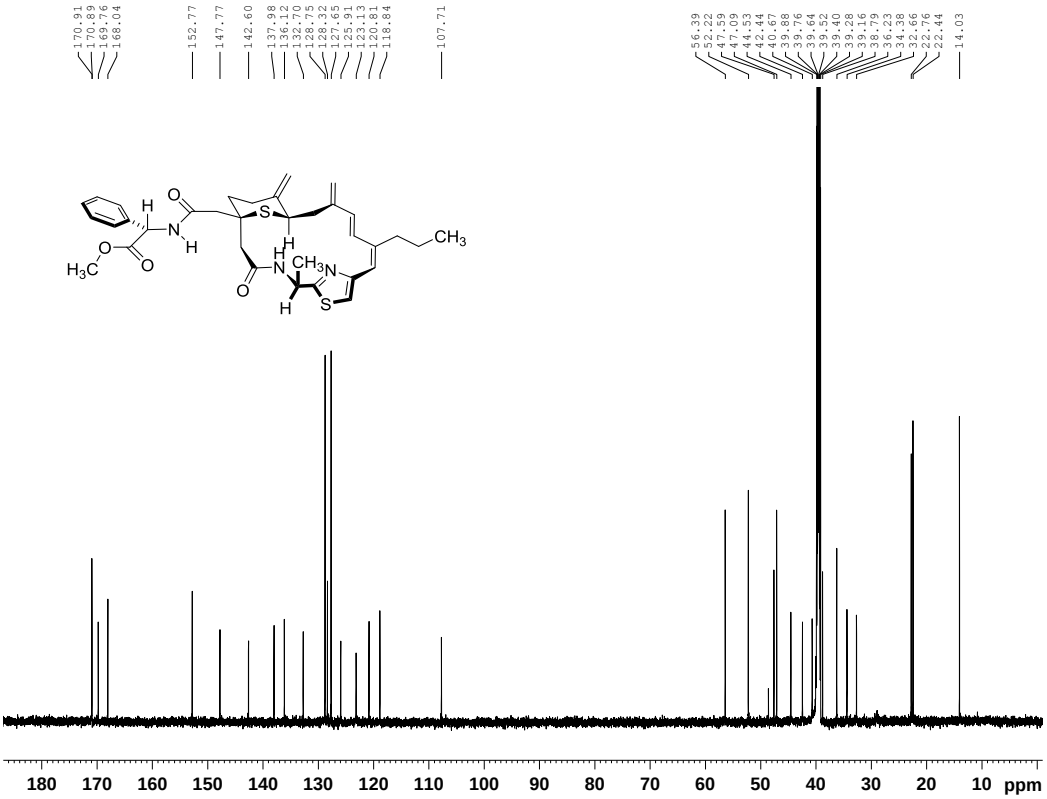


Fig. S25-SC. DEPT-135 spectra of (*S*)-PGME-WSM A2 (DMSO-d_6)

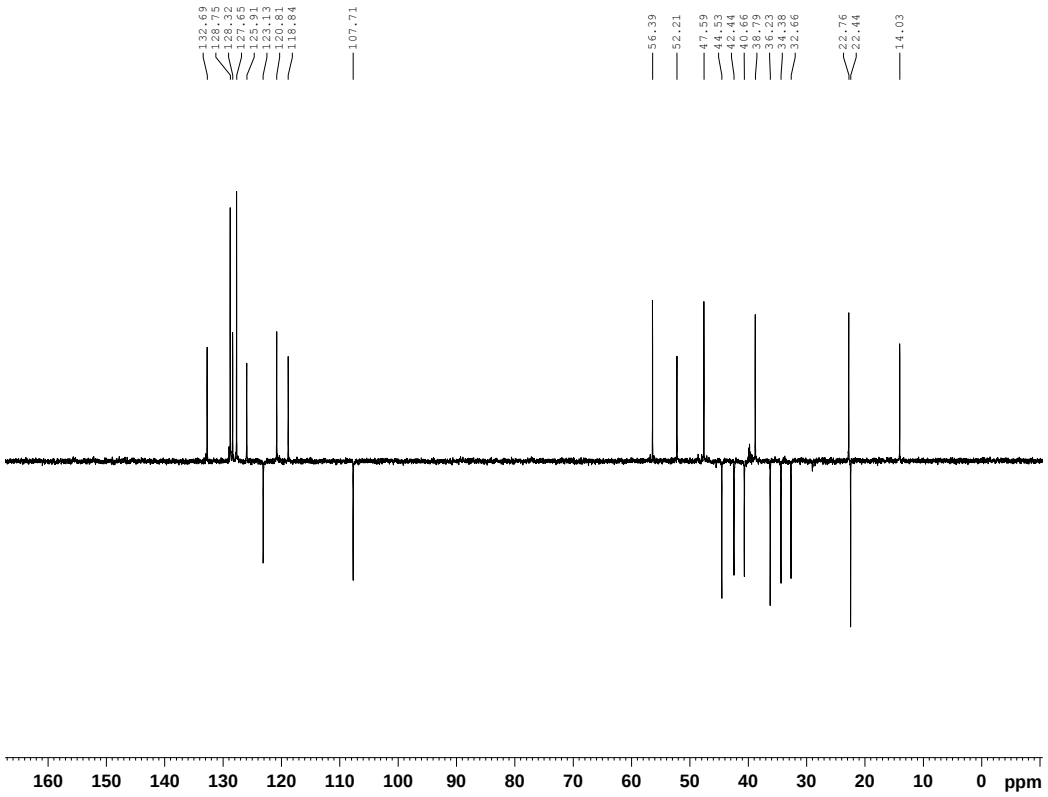


Fig. S25-SD. HSQC spectra of (*S*)-PGME-WSM A2 (DMSO-d_6)

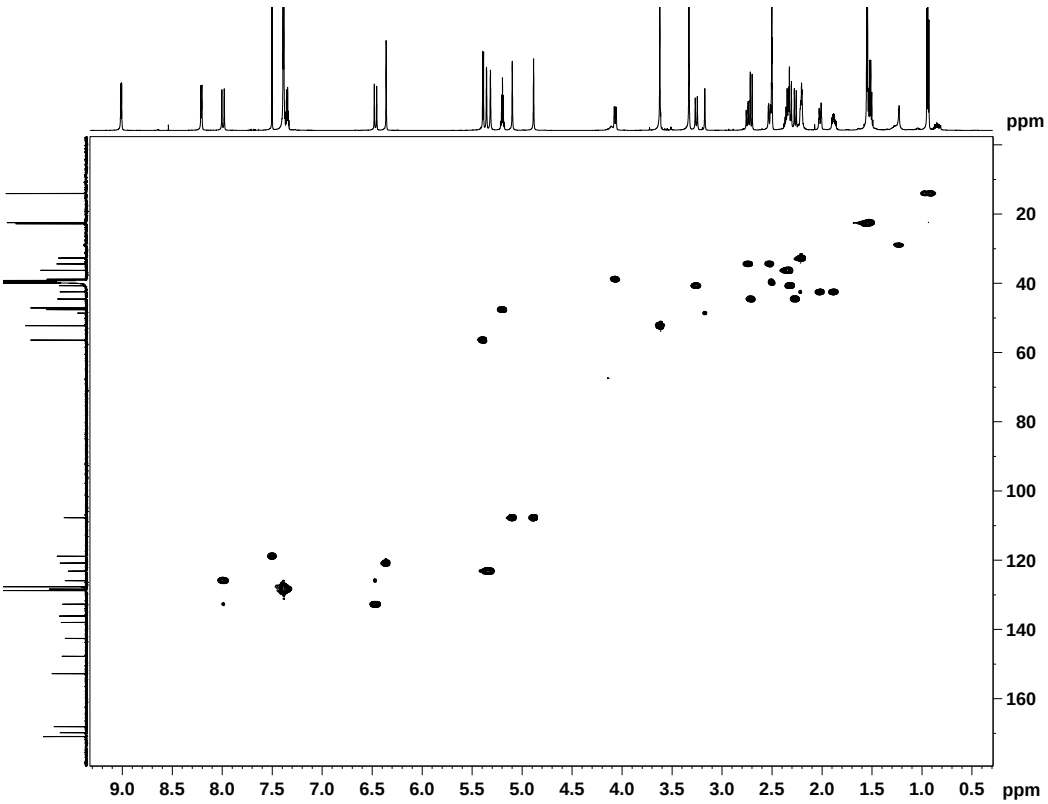


Fig. S25-SE. ^{1}H-^{1}H COSY spectra of (*S*)-PGME-WSM A2 (DMSO-d_6)

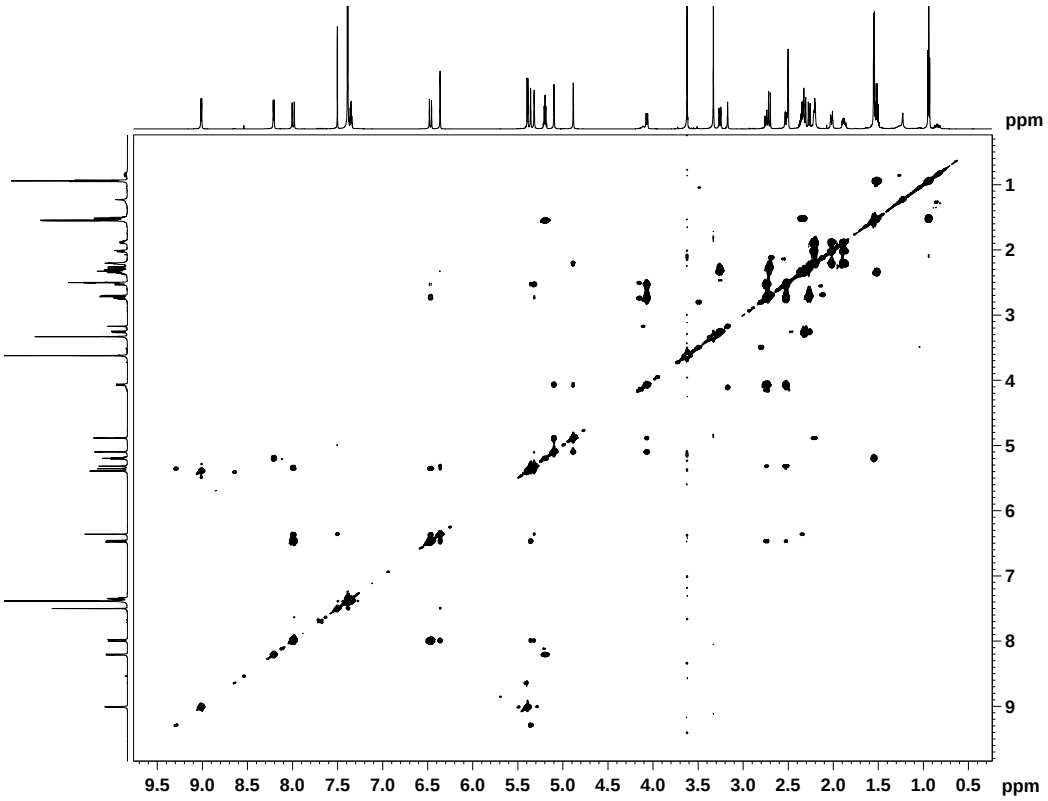


Fig. S25-SF. HMBC spectra of (*S*)-PGME-WSM A2 (DMSO-d_6)

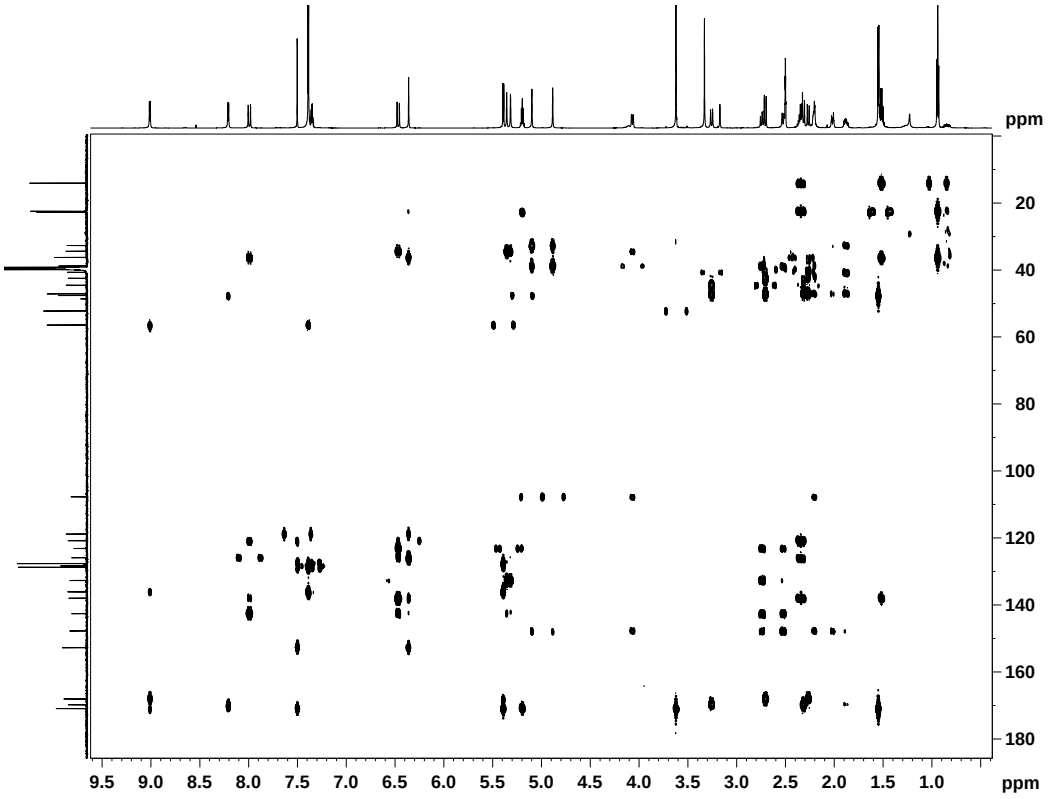


Fig. S25-SG. ROESY spectra of (*S*)-PGME-WSM A2 (DMSO-d_6)

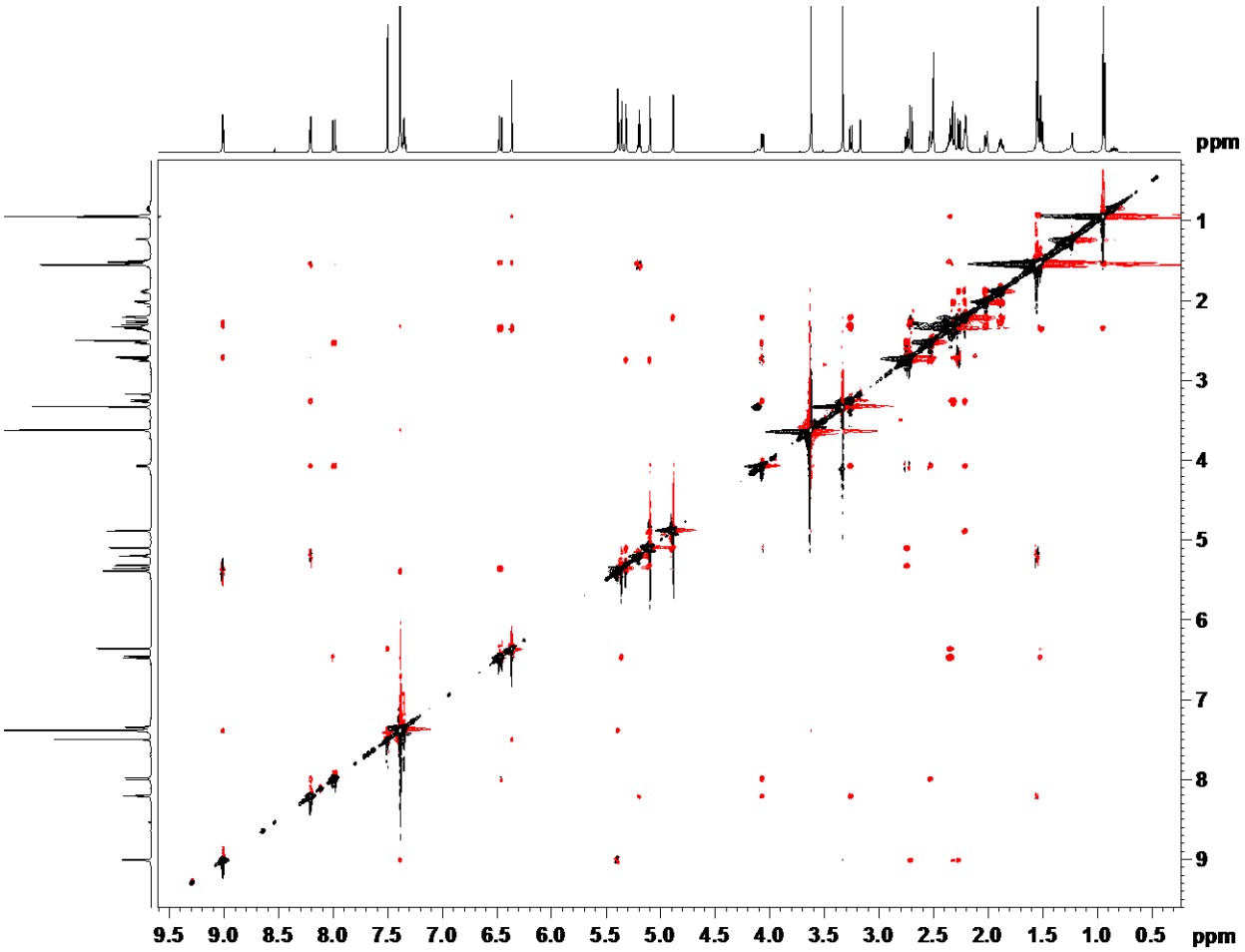


Fig. S25-SH. HR-MS (ESI) spectra of (*S*)-PGME-WSM A2

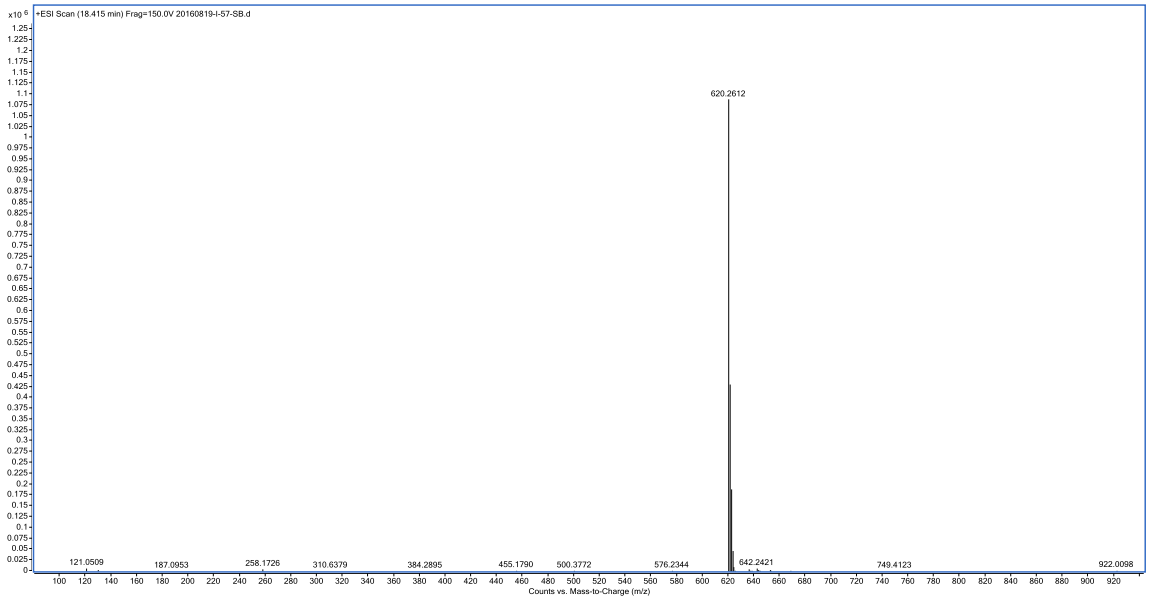


Fig. S25-RA. ^{1}H NMR spectra of (*R*)-PGME-WSM A2 (700 MHz, DMSO-d_6)

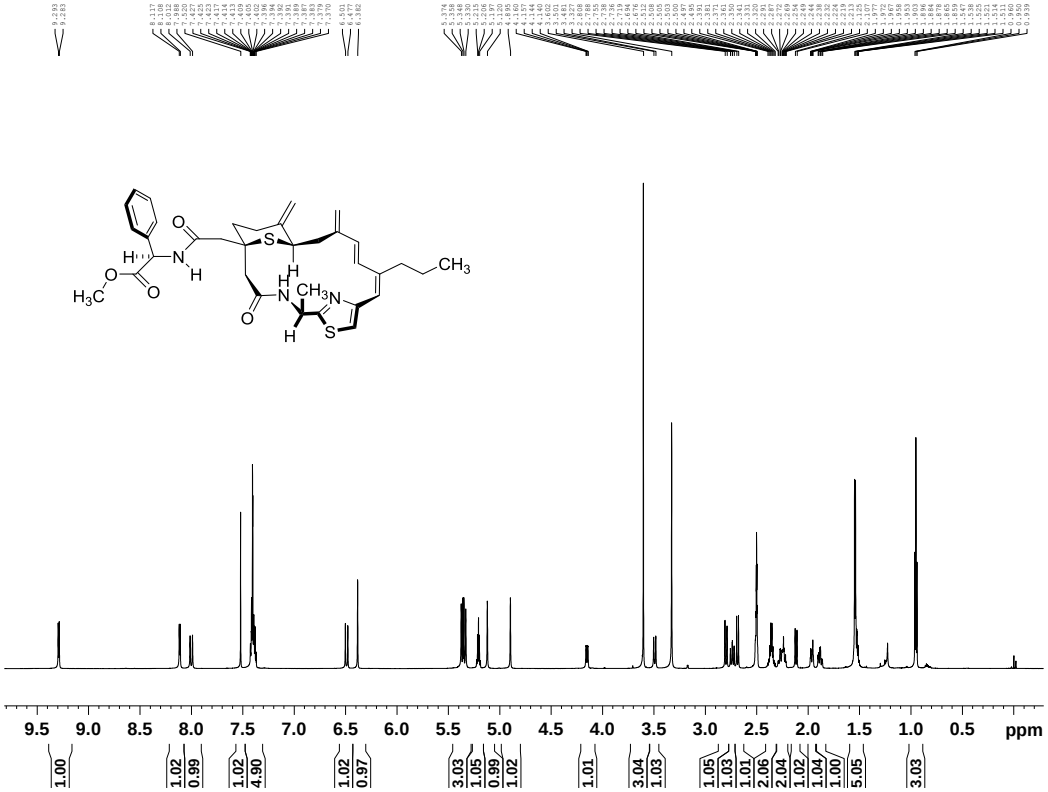


Fig. S25-RB. ^{13}C NMR spectra of (*R*)-PGME-WSM A2 (175 MHz, DMSO-d_6)

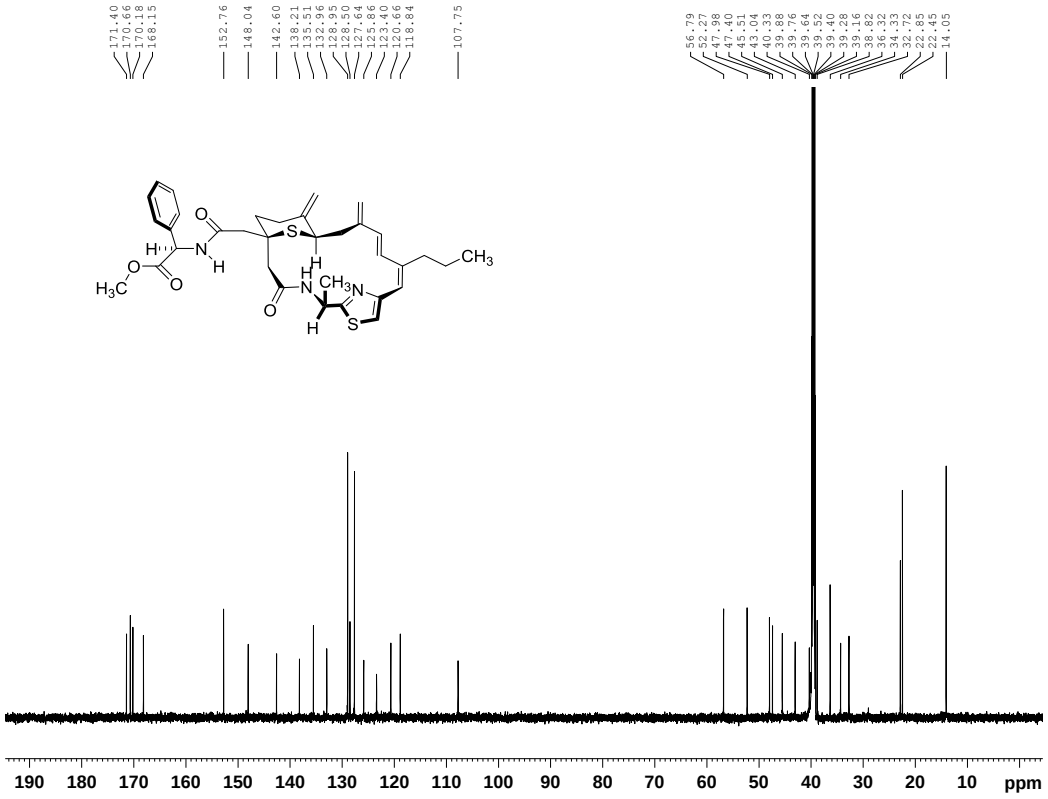


Fig. S25-RC. DEPT-135 spectra of (*R*)-PGME-WSM A2 (DMSO-d_6)

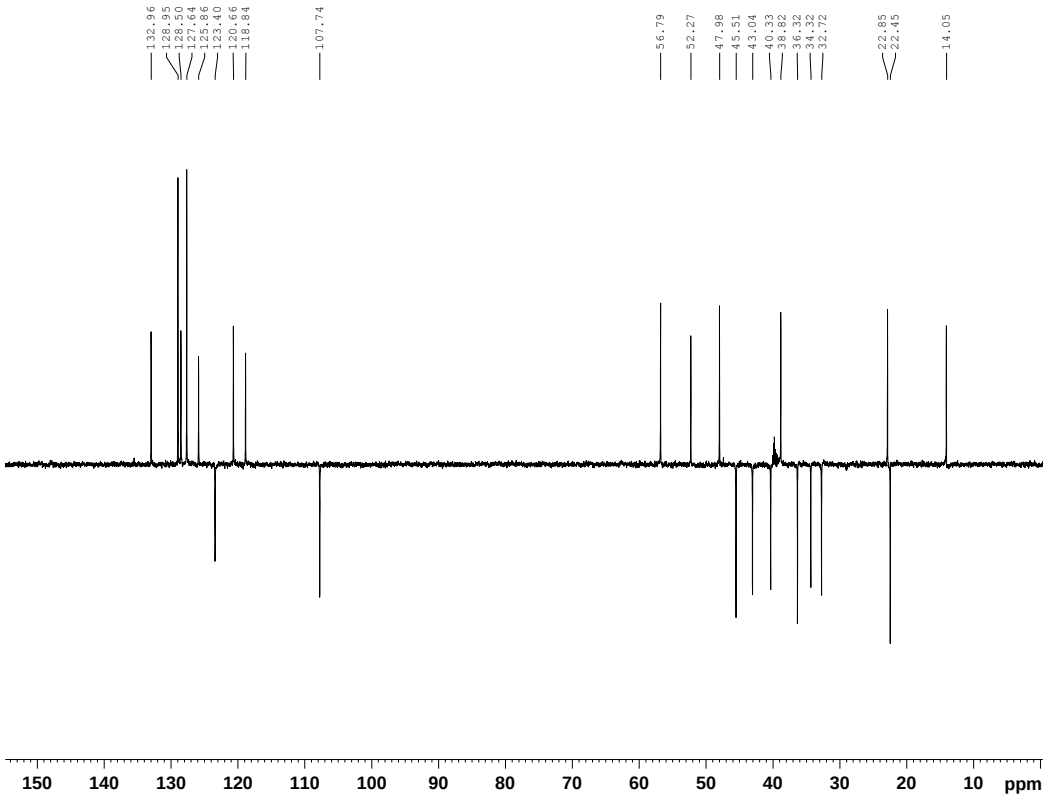


Fig. S25-RD. HSQC spectra of (*R*)-PGME-WSM A2 (DMSO-d_6)

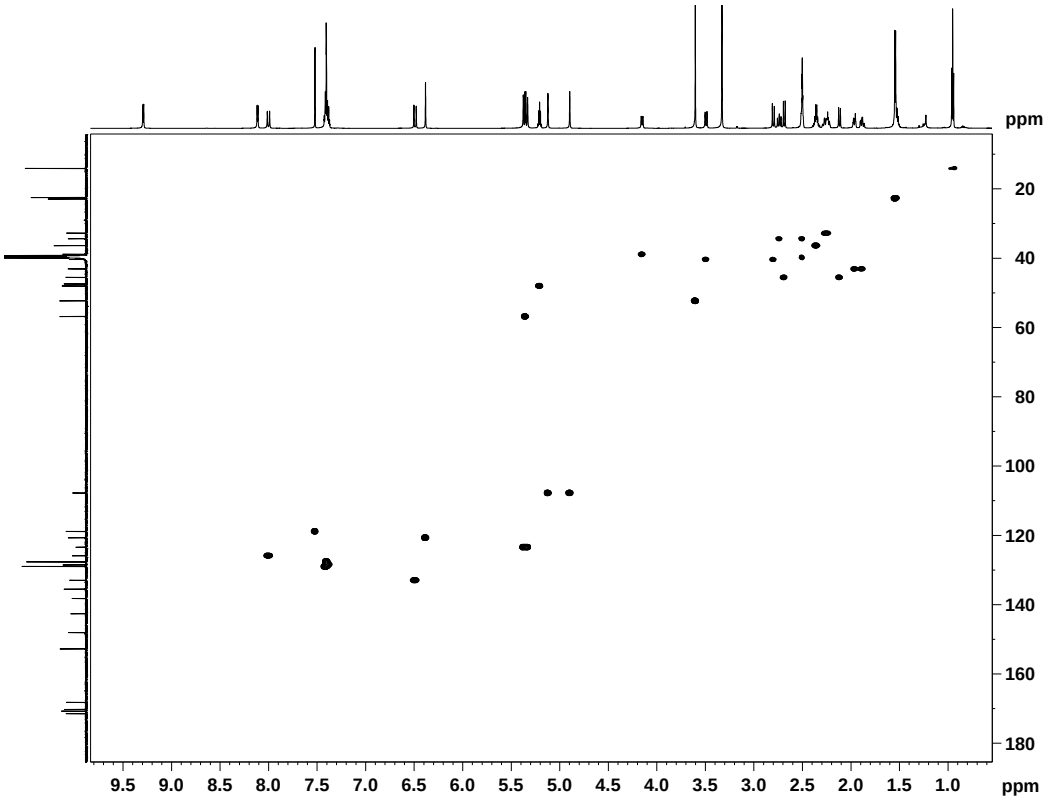


Fig. S25-RE. ^{1}H-^{1}H COSY spectra of (*R*)-PGME-WSM A2 (DMSO-d_6)

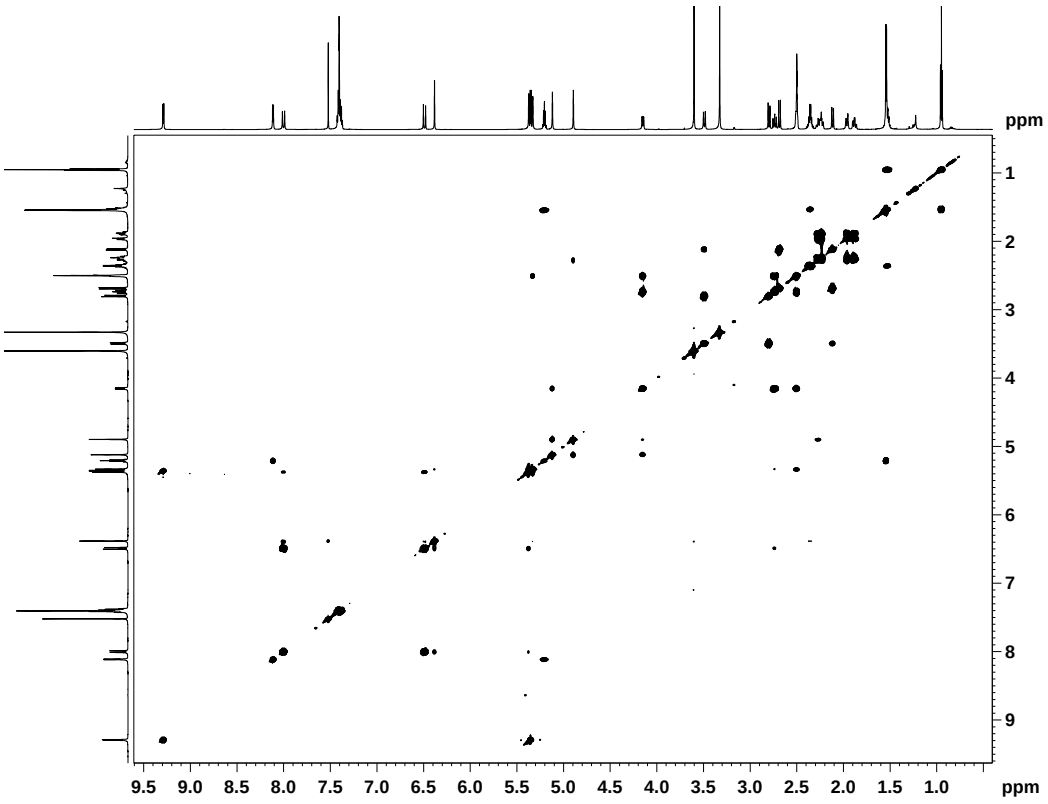


Fig. S25-RF. HMBC spectra of (*R*)-PGME-WSM A2 (DMSO-d_6)

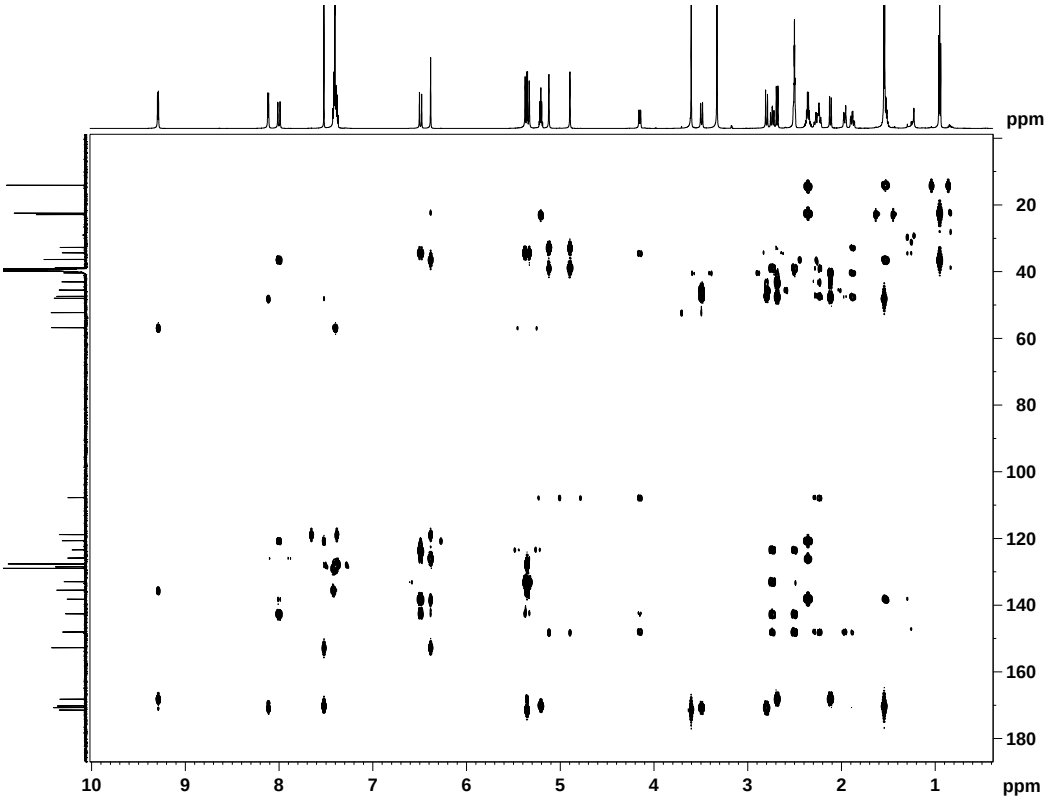


Fig. S25-RG. ROESY spectra of *R*)-PGME-WSM A2 (DMSO-d_6)

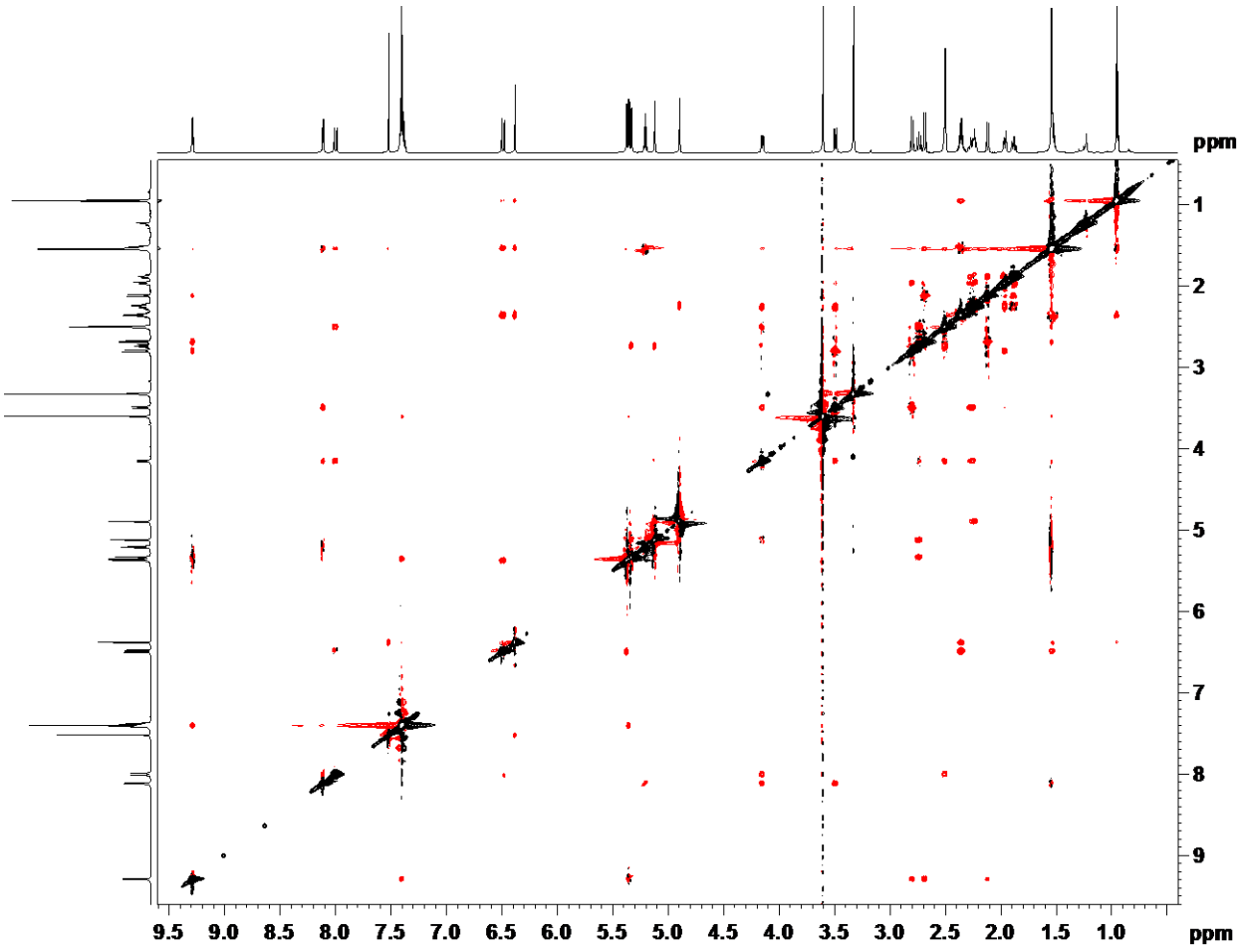


Fig. S25-RH. HR-MS (ESI) spectra of (*R*)-PGME-WSM A2

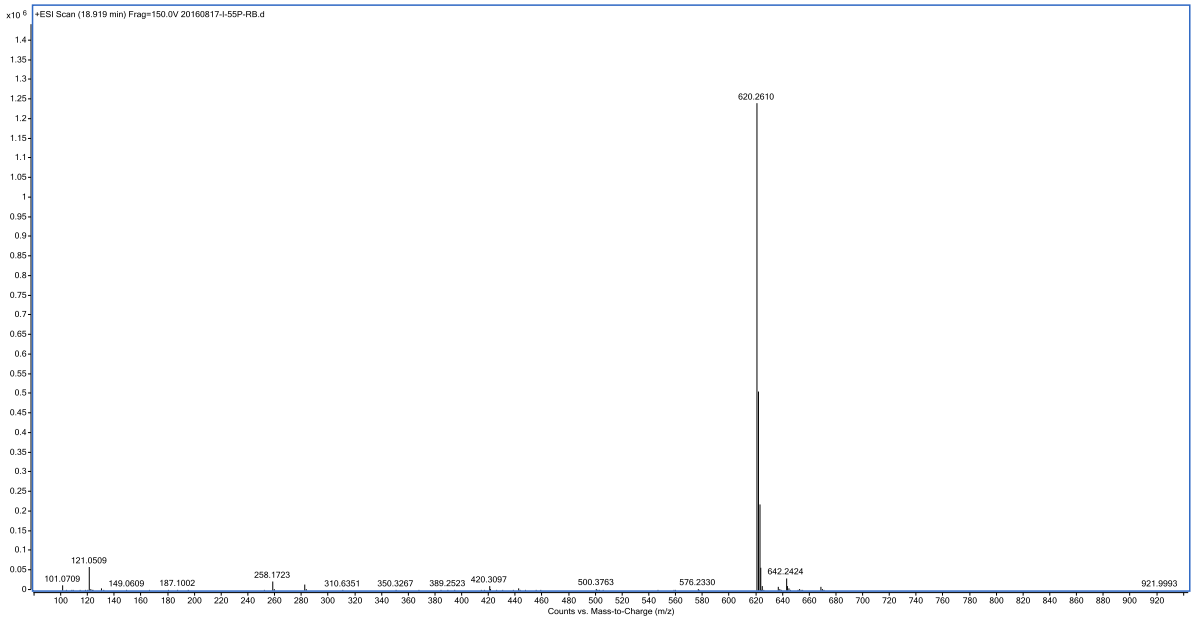


Fig. S26. Construction of the in-frame deletion mutant strains SB22002 (i.e. Δ*wsmZ3*), and SB22004 (i.e. Δ*wsmZ4*), and confirmation of their genotypes by Southern analysis. (**A**) Schematic representation for the deletion of *wsmZ3* in strain CB02120-2 by homologous recombination. The probe (526 bp) for Southern blot was amplified with primers orfZ3-SBlot-5 and orfZ3-SBlot-3 using genomic DNA of strain CB02120-2 as the template. Apr^S, apramycin sensitive; Apr^R, apramycin resistance; *acc(3)IV*, apramycin resistance gene. (**B**) Southern blot verification of strain CB02120-2 wild type (1819 bp) and mutant strain SB22002 (1025 bp). Genomic DNAs were digested with *Apa*LI and then used for Southern analysis. Lane 1, DNA marker VII, DIG-labeled (Roche); lane 2, strain CB02120-2 wild type; lane 3, SB22002. (**C**) Schematic representation for the deletion of *wsmZ4* in strain CB02120-2 by homologous recombination. The probe (492 bp) for Southern blot was amplified with primers orfZ4-SBlot-5 and orfZ4-SBlot-3 using genomic DNA of strain CB02120-2 as the template. (**D**) Southern blot verification of strain CB02120-2 wild-type (2265 bp) and mutant strain SB22004 (1626 bp). Genomic DNAs were digested with *Kpn*I and then used for Southern analysis. Lane 4, DNA marker VII, DIG-labeled (Roche); lane 5, strain CB02120-2 wild type; lane 6, SB22004.

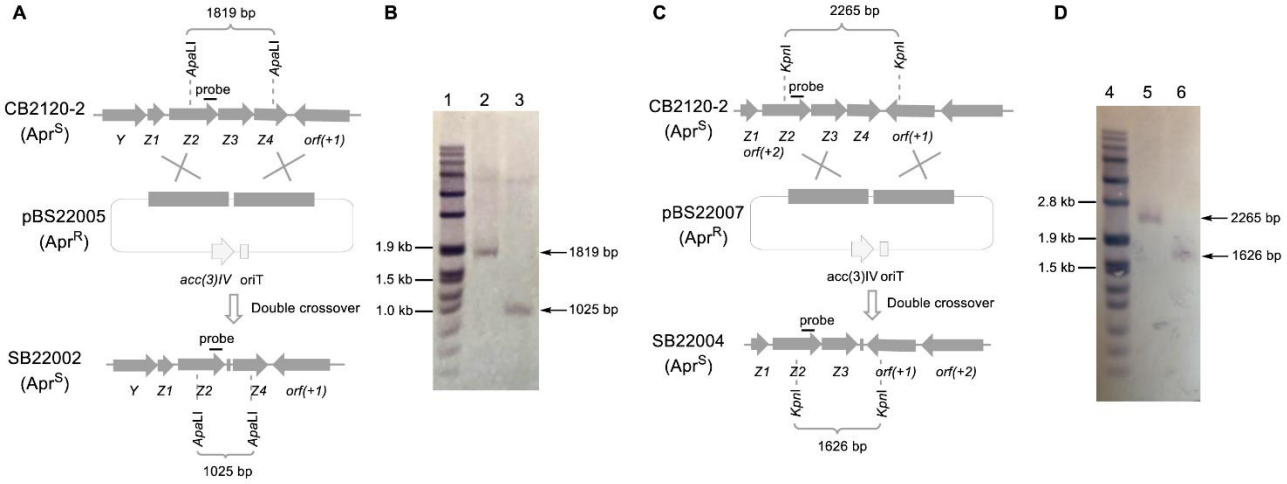


Fig. S27. Discovery of *S.* sp. CB01635 as an alternative LNM E1 producer. (**A**) Comparison of *lnm* gene cluster from *S. atroolivaceus* S-140 and *lnm*-type gene cluster from strain CB01635 revealing almost identical genetic organization. (**B**) The metabolite profile of strain CB01635 (I) in comparison with that of *S. atroolivaceus* S-140 that produces both LNM and LNM E1 (II). While LNM E1 production was reliably observed for both CB01635 and *S. atroolivaceus* S-140, LNM production was not detected in CB01635. (**C**) Confirmation of the identity of LNM E1 produced by strain CB01635 by HR-ESI-MS analysis: *m/z* 449.1562 $[M + H]^+$ (calcd for $C_{22}H_{29}N_2O_4S_2$, 449.1563).

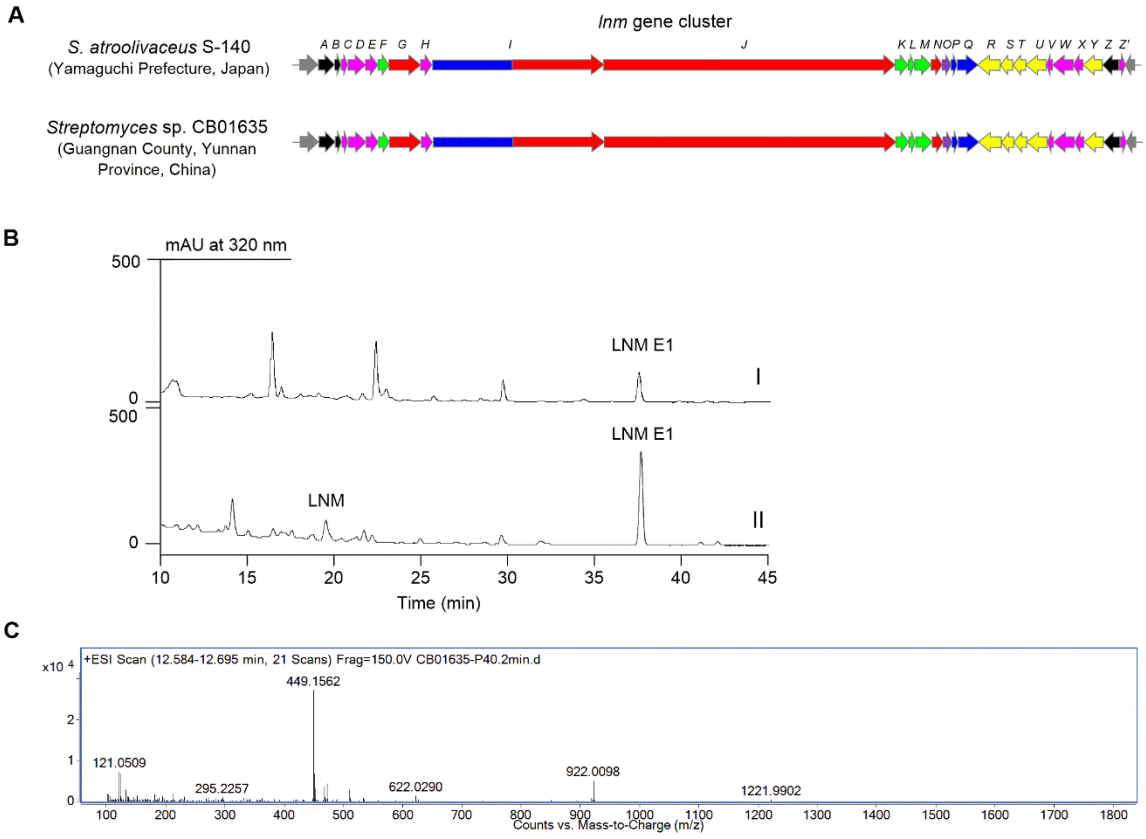


Fig. S28. Discovery of *Streptomyces* sp. NRRL F-5630 and *S. aureofaciens* NRRL B-2183 as alternative producers of GNM A and GNM B. (**A**) The *gnm*-like gene clusters from strains *Streptomyces* sp. NRRL F-5630 and *S. aureofaciens* NRRL B-2183 in comparison with the *gnm* gene cluster from strain CB01883. Geographical locations of the three strains (where they were originally isolated) were shown in the parentheses. (**B**) HPLC analysis of the metabolite profiles of CB01883 wild type (III), NRRL F-5630 (IV) and NRRL B-2183 (V) in comparison with the GNM A (I) and GNM B (II) standards.

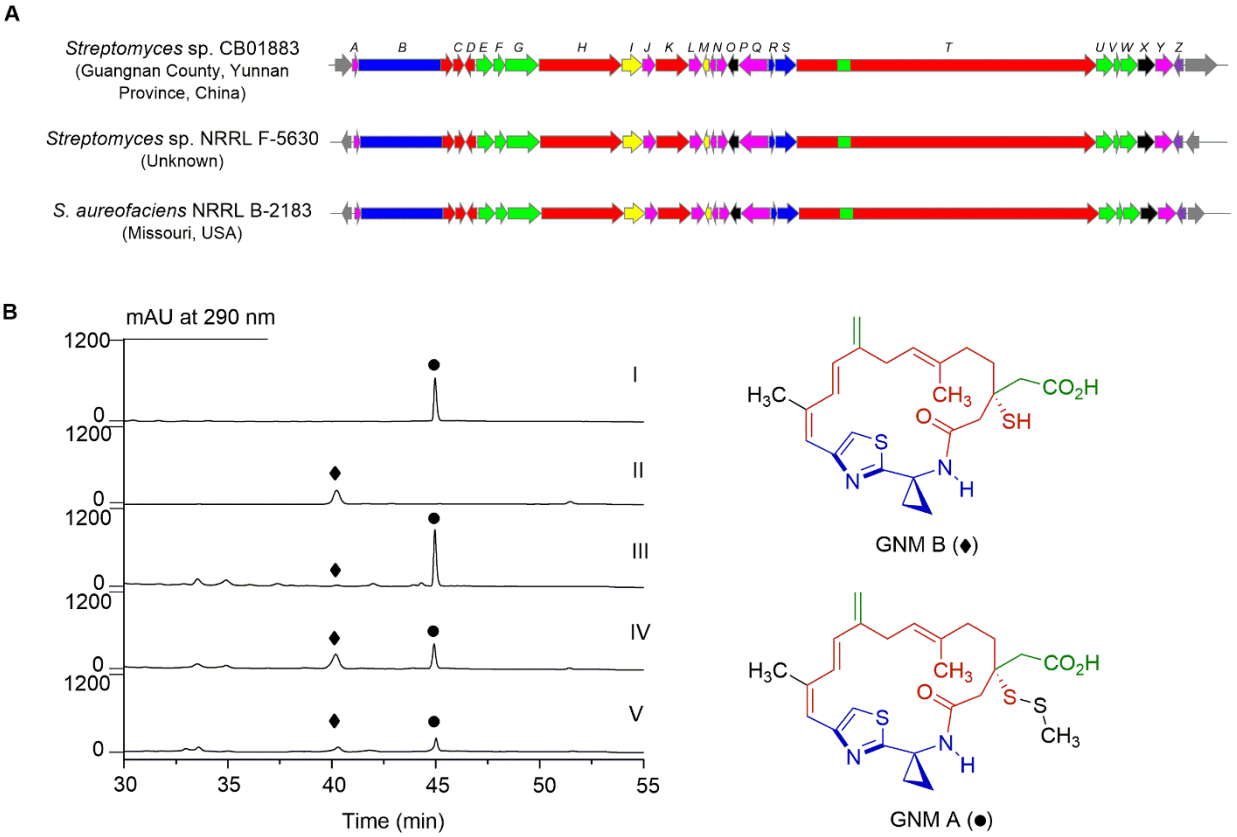


Fig. S29. Functional characterization of representative adenylation proteins from newly discovered LNM-type biosynthetic machineries. (**A**) SDS-PAGE analysis of the purified adenylation proteins. Lane 1, GnmS (59.8 kD); Lane 2, WsmQ (57.5 kD); Lane 3, Maur_X (A domain, 58.4 kD); Lane 4, CB01373_Q (A domain, 59.0 kD); Lane 5, CB02613_Z17 (A domain including an extra N-terminal with ~200 amino acid residues, 81.7 kD); Lane 6, CB02613_M (A domain including an extra N-terminal with ~200 residues, 80.6 kD); M, protein ladder. (**B**) The mechanism of hydroxylamine-trapping assay used in this study for probing the substrate specificities of the adenylation proteins (15). (**C**) Substrate specificities of the representative adenylation proteins from different LNM-type pathways. Under tested conditions, GnmS, WsmQ, Maur_X and CB01373_Q showed highest activities towards ACC (73% substrate conversion), D-Ala (32% substrate conversion), L-Thr (19% substrate conversion) and L-Thr (22% substrate conversion), respectively, while CB02613_Z17 and CB02613_M did not show obvious activity towards all the tested substrates (< 2% substrate conversion). The percentage of the substrate conversion was estimated based on the assumption that the UV absorbance of 3 mM DL-alanine hydroxamate complexed with $FeCl_3$ corresponded to the 100% substrate conversion. The relative activities of GnmS, WsmQ, Maur_X and CB01373_Q towards their best substrates were normalized to 100%, and the relative activities of CB02613_Z17 and CB02613_M towards different amino acids were calculated based on the activity of CB01373_Q towards its best substrate L-Thr.

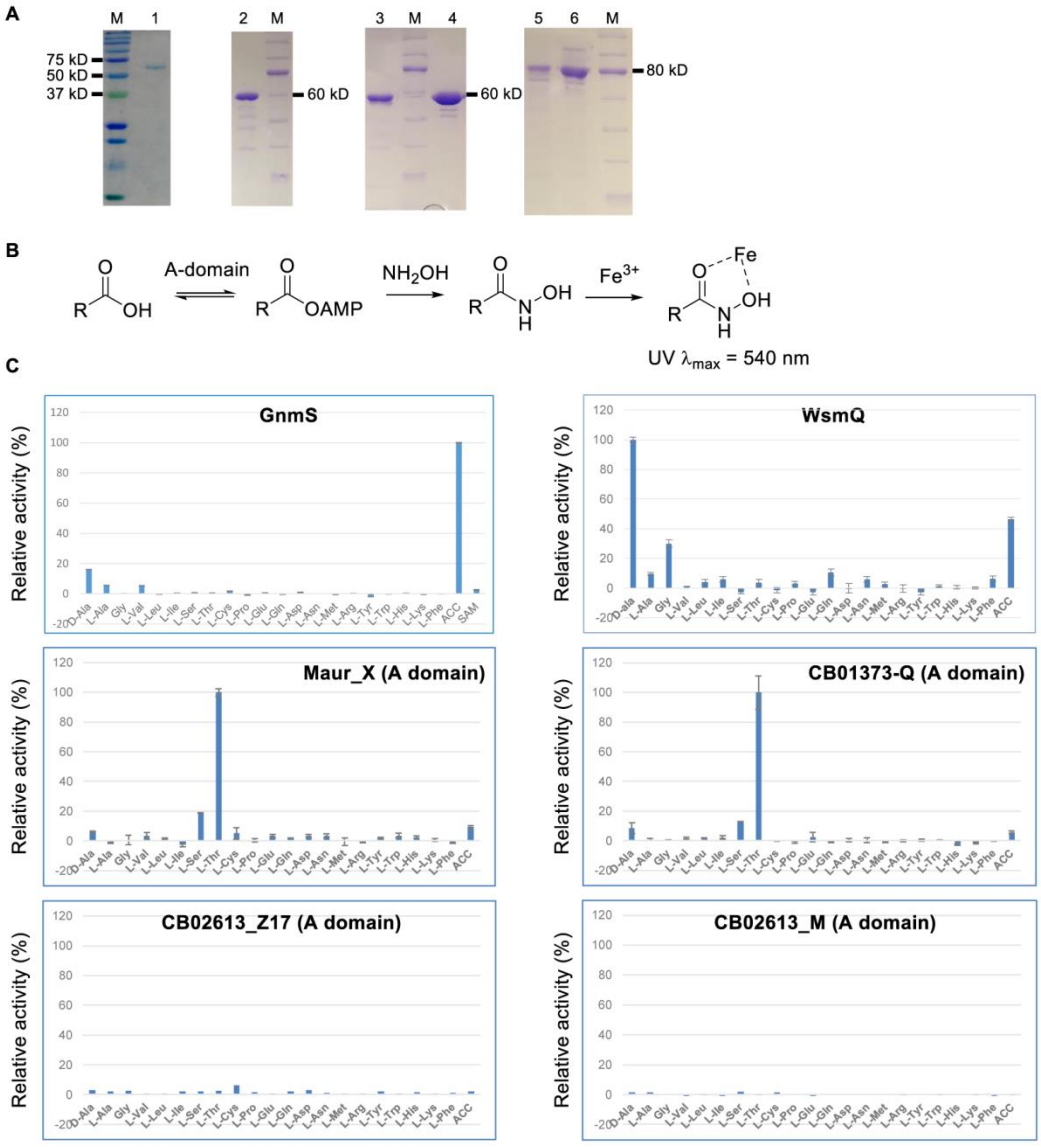

A
M
1
2
M
3
M
4
5
6
M
75 kD
50 kD
37 kD
60 kD
60 kD
80 kD
B
A-domain
OAMP
NH2OH
Fe3+
UV λmax = 540 nm
C
GnmS
WsmQ
Maur_X (A domain)
CB01373-Q (A domain)
CB02613_Z17 (A domain)
CB02613_M (A domain)
Relative activity (%)

Fig. S30. Construction of the in-frame deletion mutant strain SB21003 (i.e. Δ*gnmO*) and confirmation of its genotype by Southern analysis. (**A**) Schematic representation for the deletion of *gnmO* in strain CB01883 by homologous recombination. The probe (542 bp) for Southern blot was amplified with primers orf132-SBlot-5nn and orf132-SBlot-3nn using genomic DNA of strain CB01883 as the template. Apr^S, apramycin sensitive; Apr^R, apramycin resistance; *acc(3)IV*, apramycin resistance gene. (**B**) Southern blot verification of strain CB01883 wild type (1.7 kb) and mutant strain SB21003 (3.3 kb). Genomic DNAs were digested with *Aat*II and then used for Southern analysis. Lane 1, DNA marker VII, DIG-labeled (Roche); lane 2, strain CB01883 wild type; lane 3, SB21003.

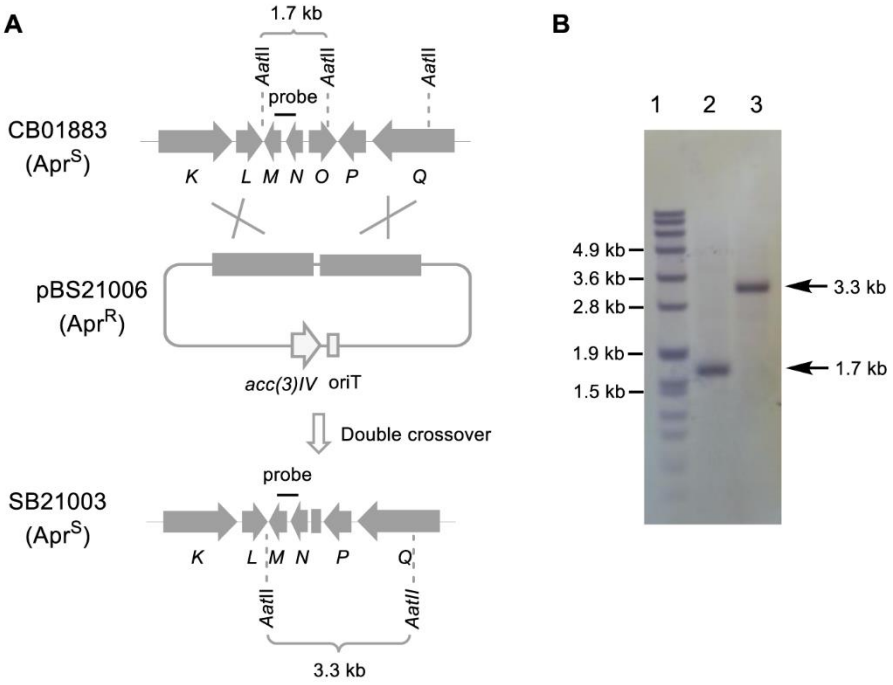


Fig. S31A. ^{1}H NMR spectra of GNM B1 (700 MHz, DMSO-d_6)

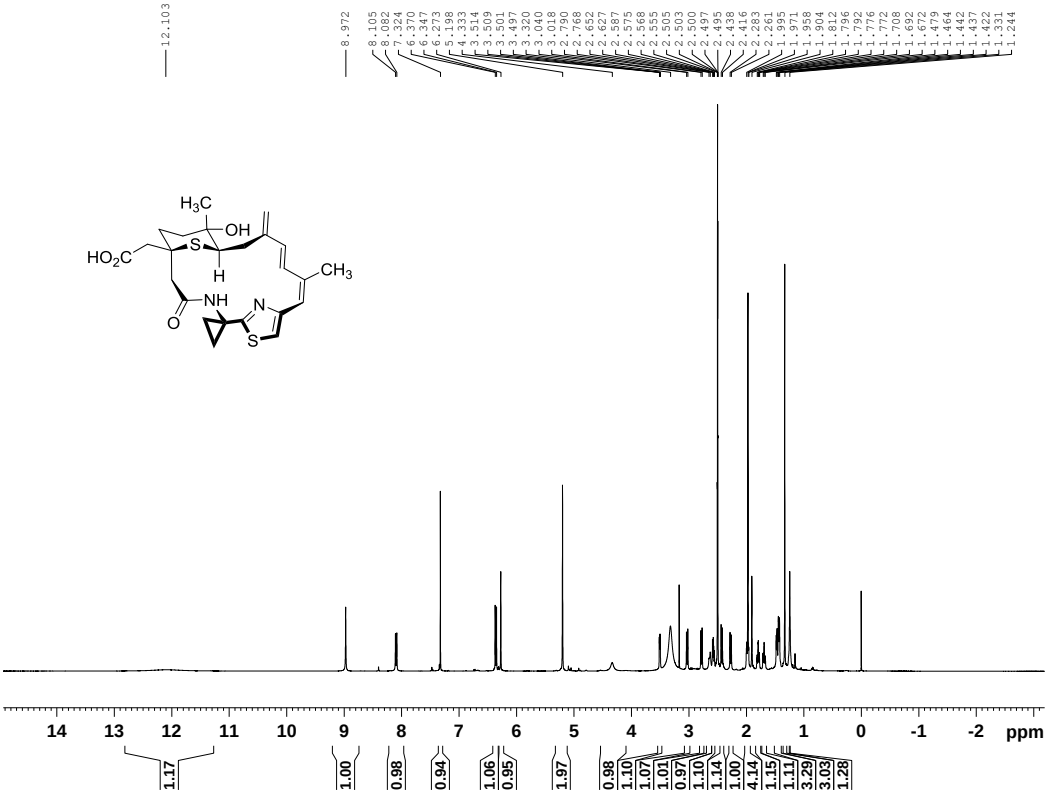


Fig. S31B. ^{13}C NMR spectra of GNM B1 (175 MHz, DMSO-d_6)

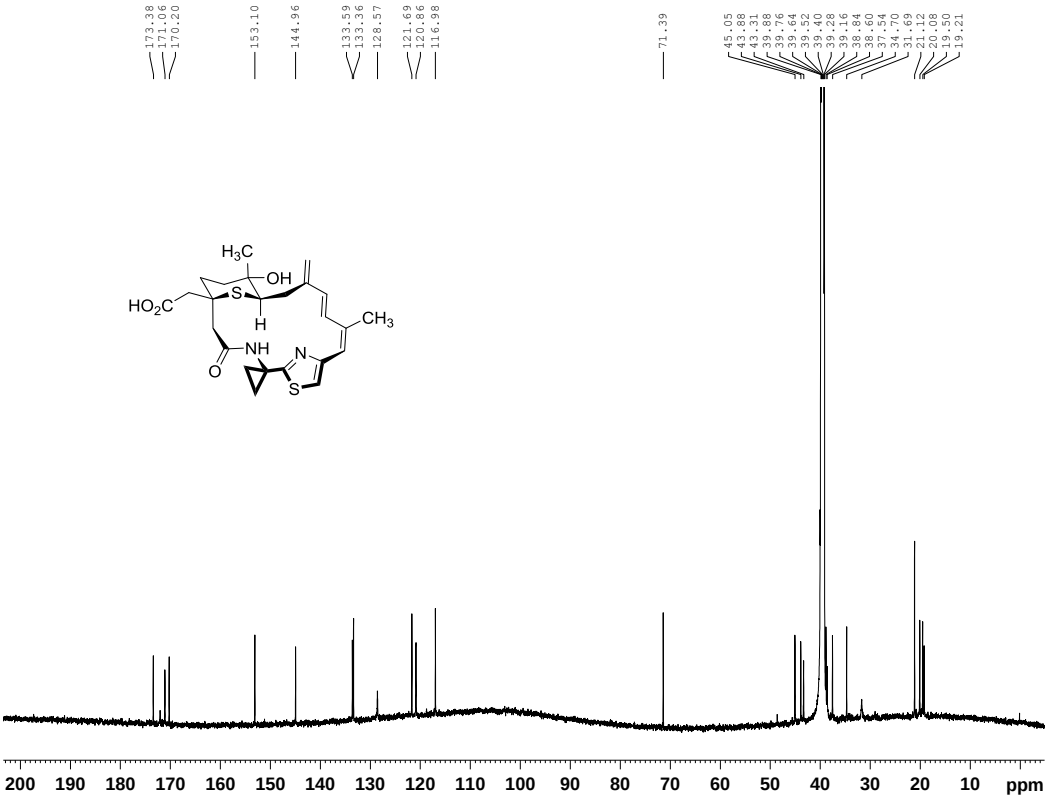


Fig. S31C. DEPT-135 spectra of GNM B1 (DMSO-d_6)

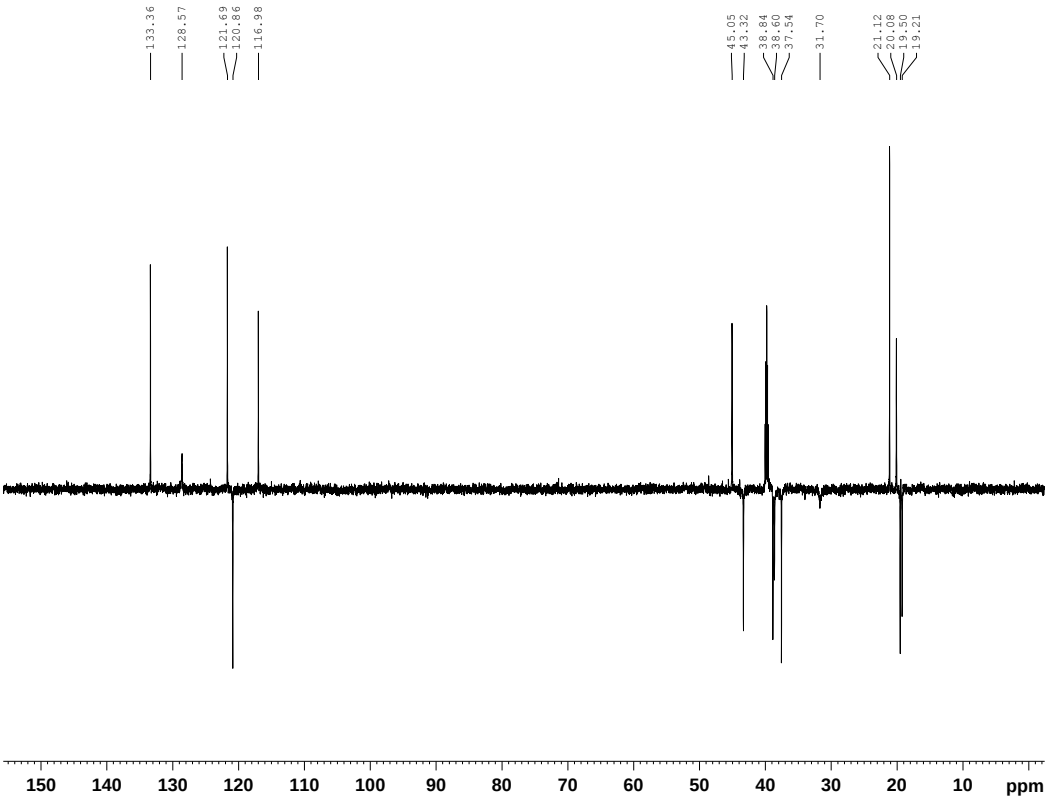


Fig. S31D. HSQC spectra of GNM B1 (DMSO-d_6)

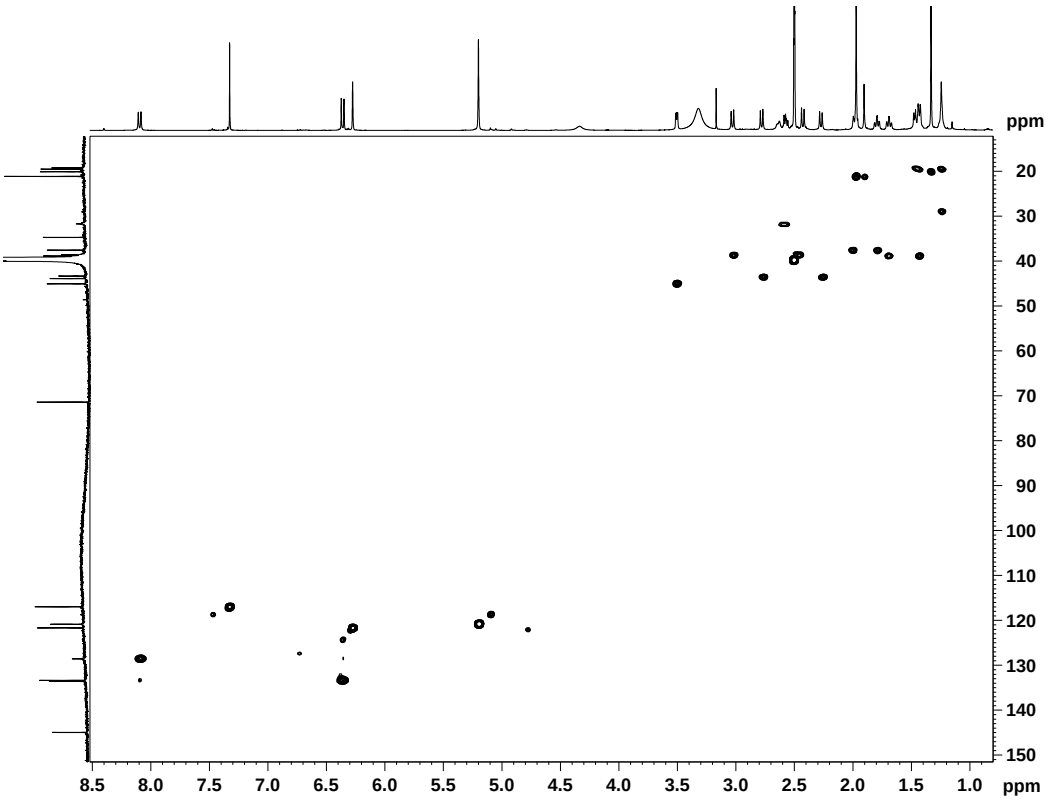


Fig. S31E. ^{1}H-^{1}H COSY spectra of GNM B1 (DMSO-d_6)

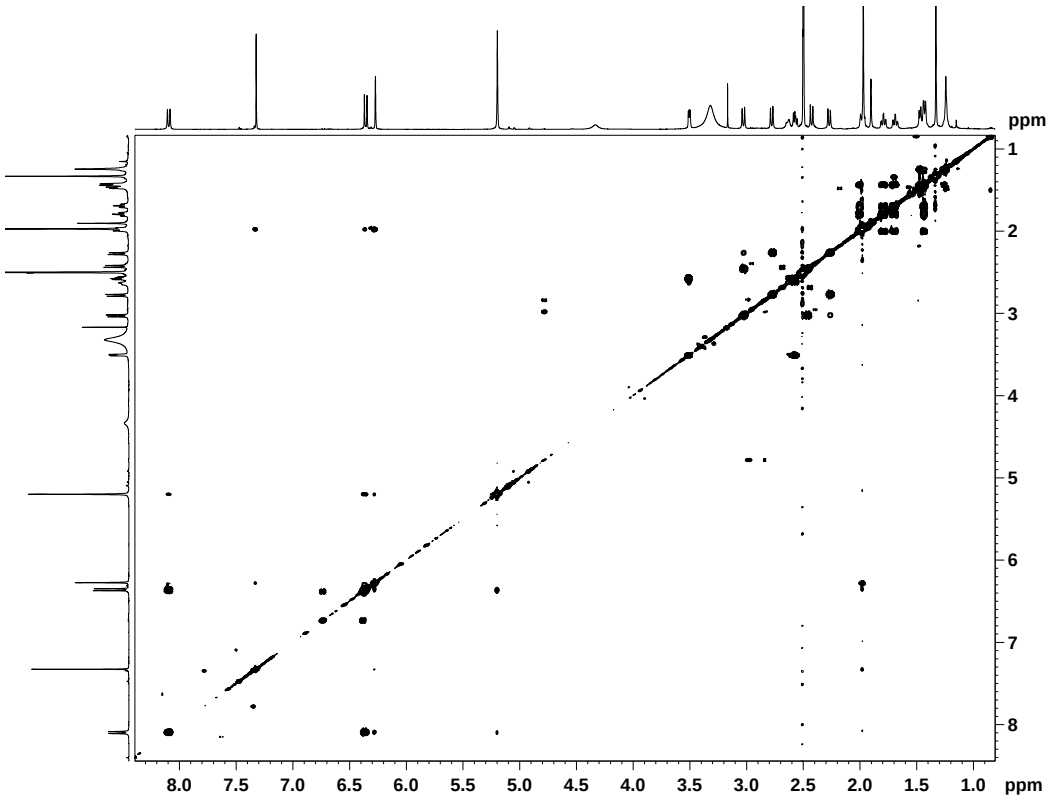


Fig. S31F. HMBC spectra of GNM B1 (DMSO-d_6)

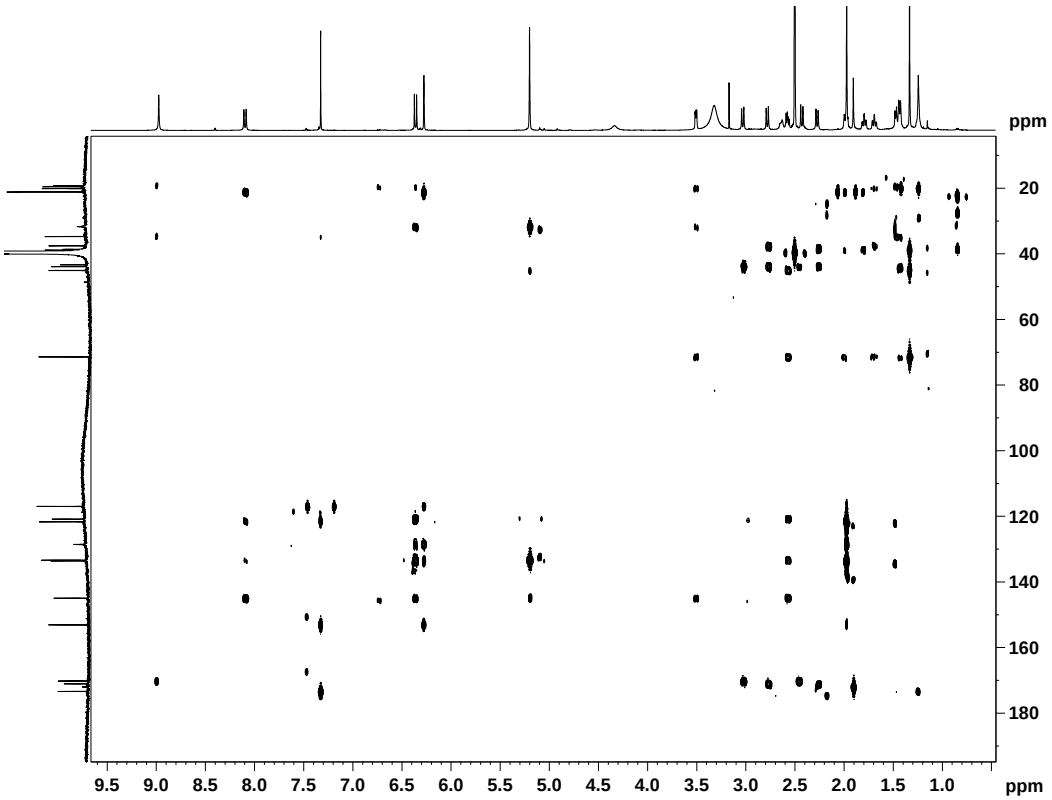


Fig. S31G. ROESY spectra of GNM B1 (DMSO-d_6)

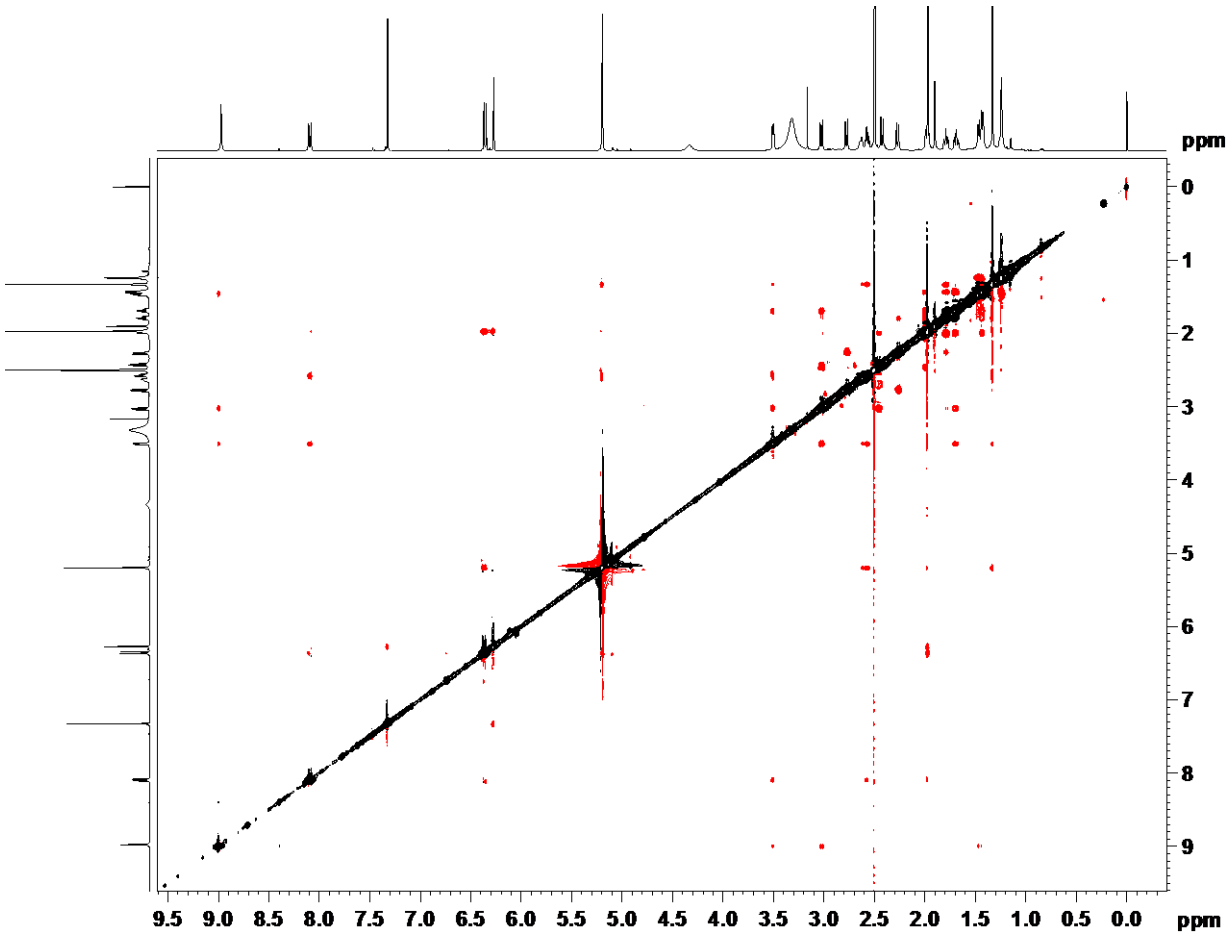


Fig. S31H. HR-MS (ESI) spectra of GNM B1

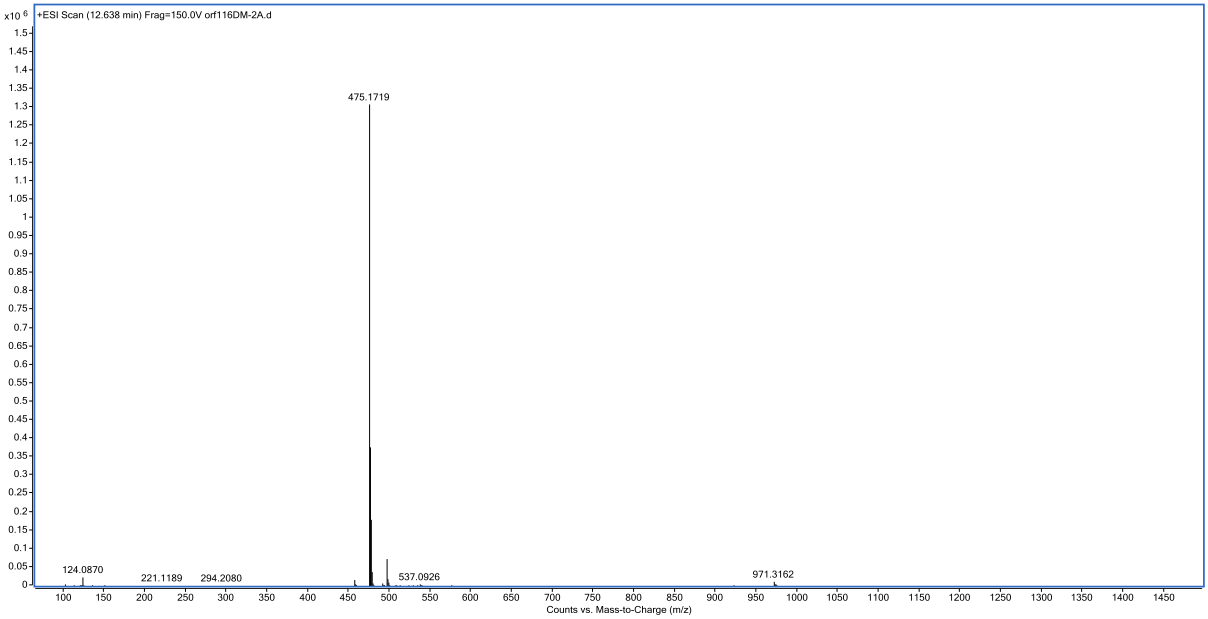


Fig. S32A. ^{1}H NMR spectra of GNM B2 (700 MHz, DMSO-d_6)

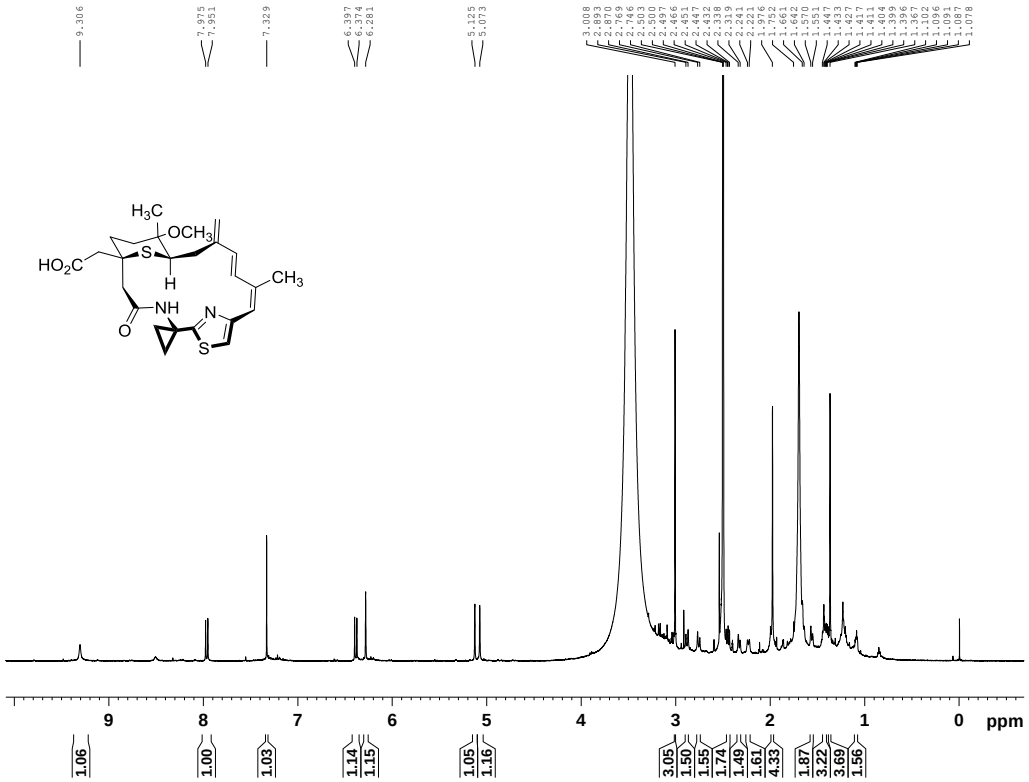


Fig. S32B. ^{13}C NMR spectra of GNM B2 (175 MHz, DMSO-d_6)

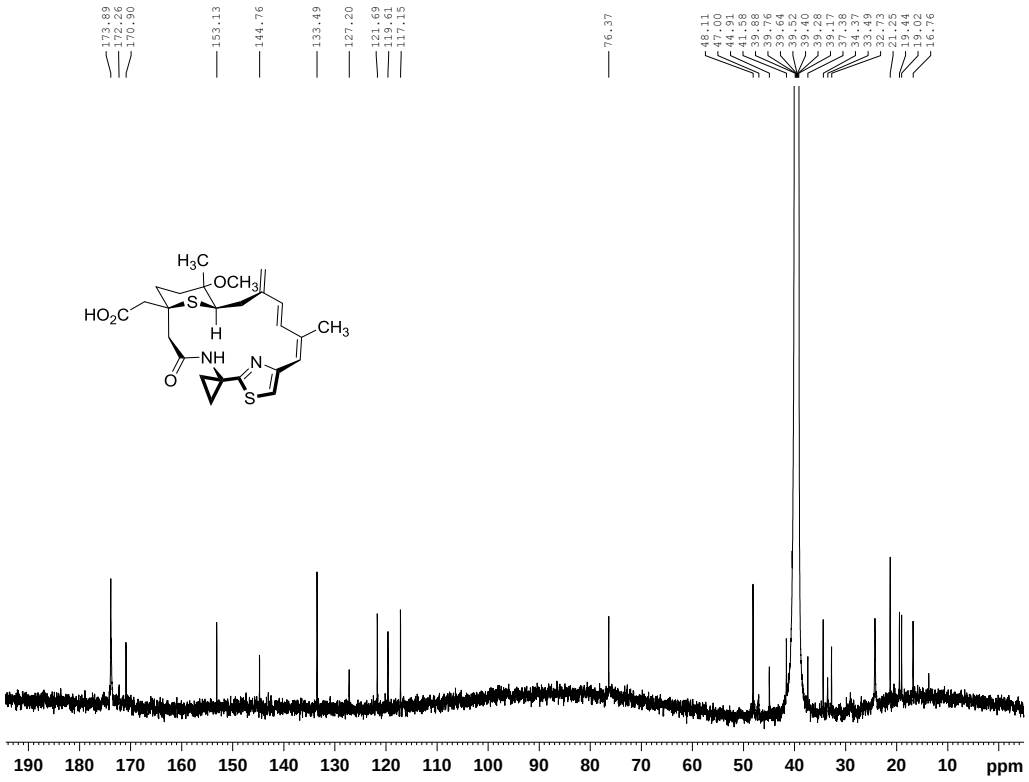


Fig. S32C. DEPT-135 spectra of GNM B2 (DMSO-d_6)

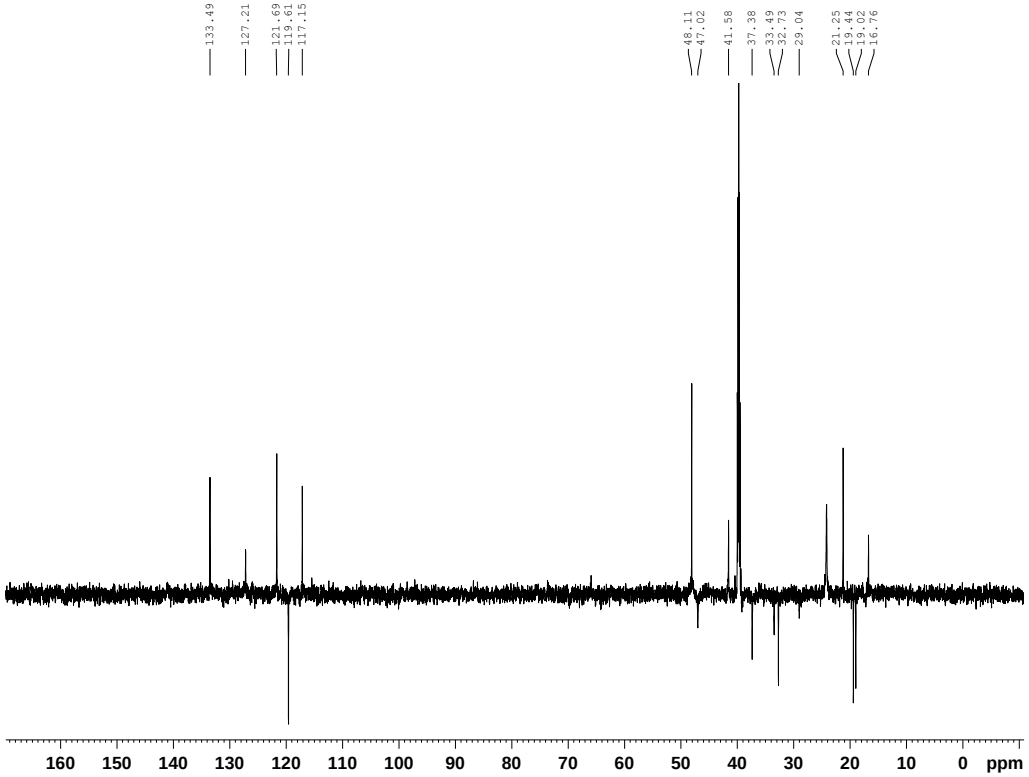


Fig. S32D. HSQC spectra of GNM B2 (DMSO-d_6)

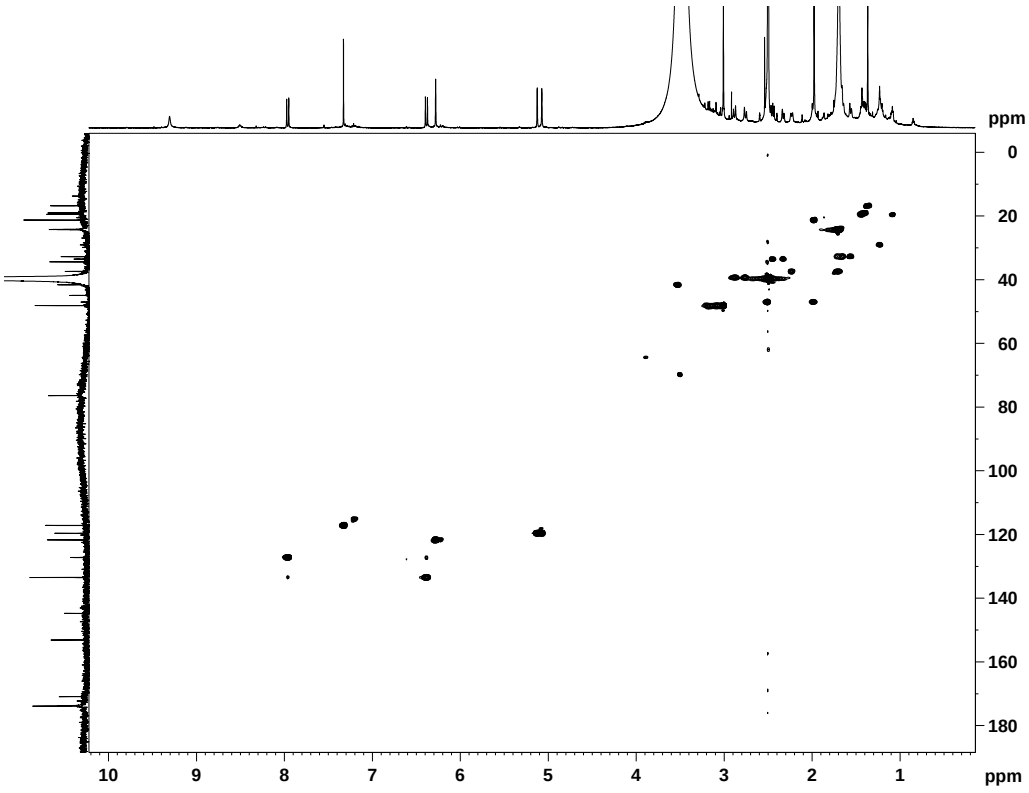


Fig. S32E. ^{1}H-^{1}H COSY spectra of GNM B2 (DMSO-d_6)

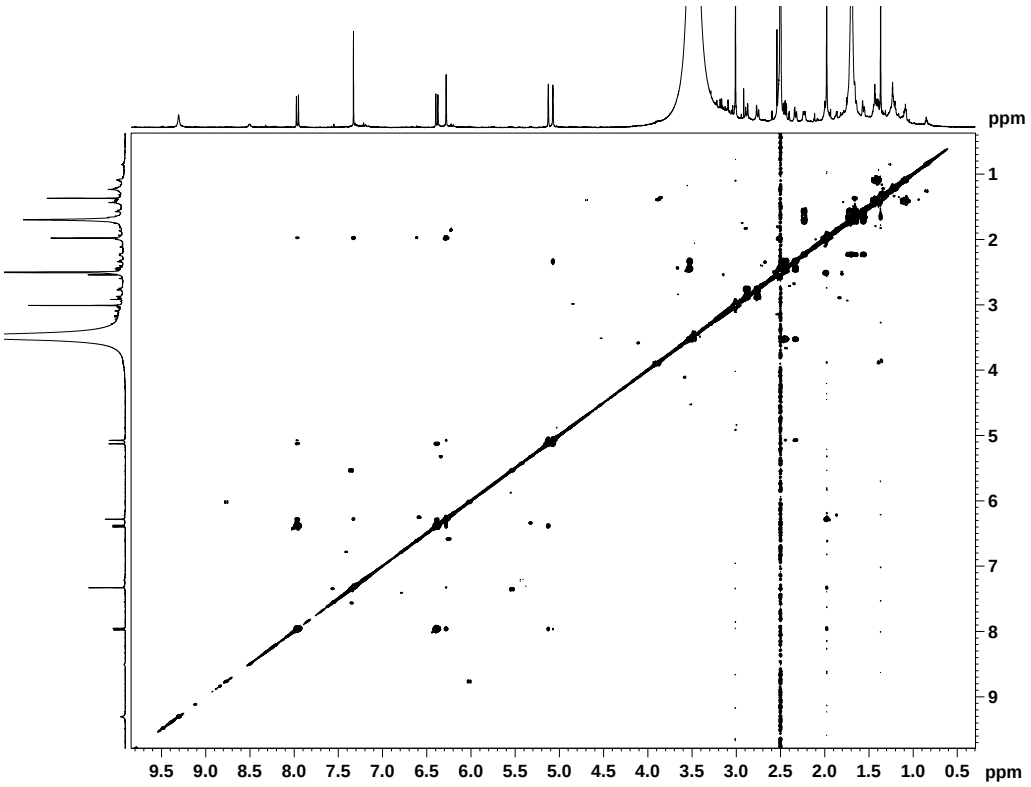


Fig. S32F. HMBC spectra of GNM B2 (DMSO-d_6)

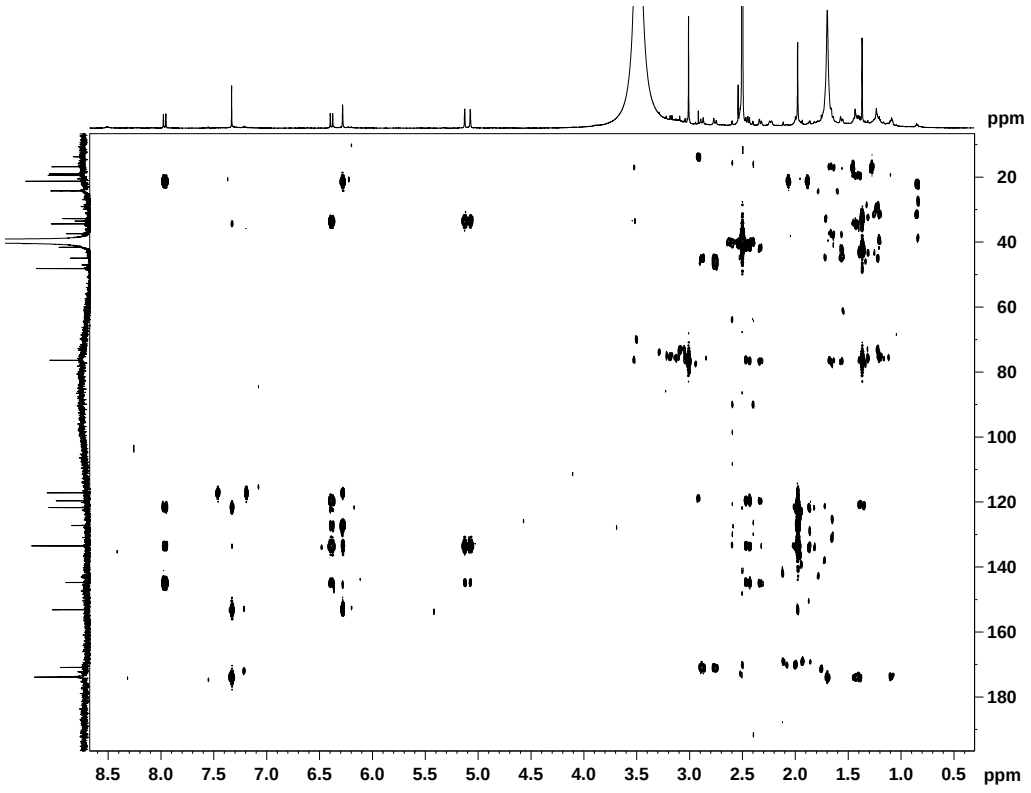


Fig. S32G. ROESY spectra of GNM B2 (DMSO-d_6)

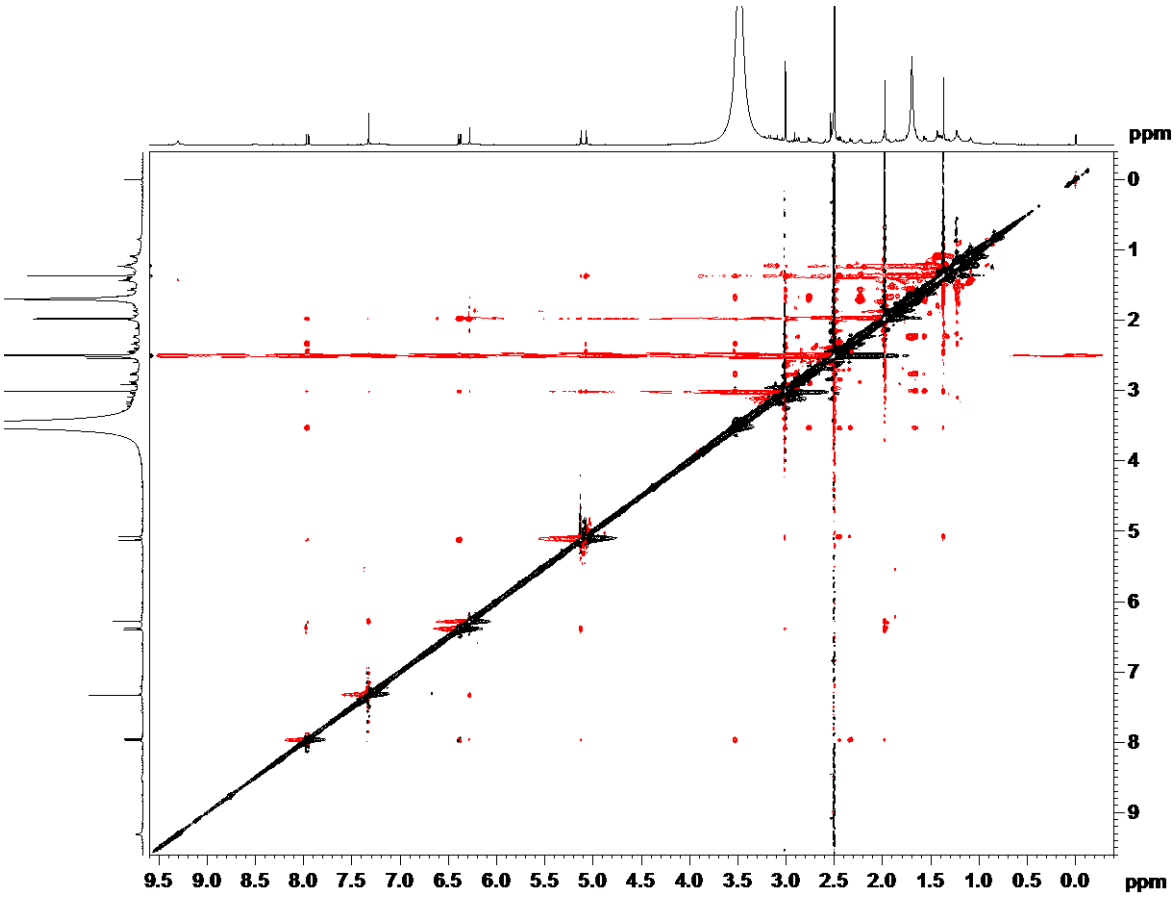


Fig. S32H. HR-MS (ESI) spectra of GNM B2

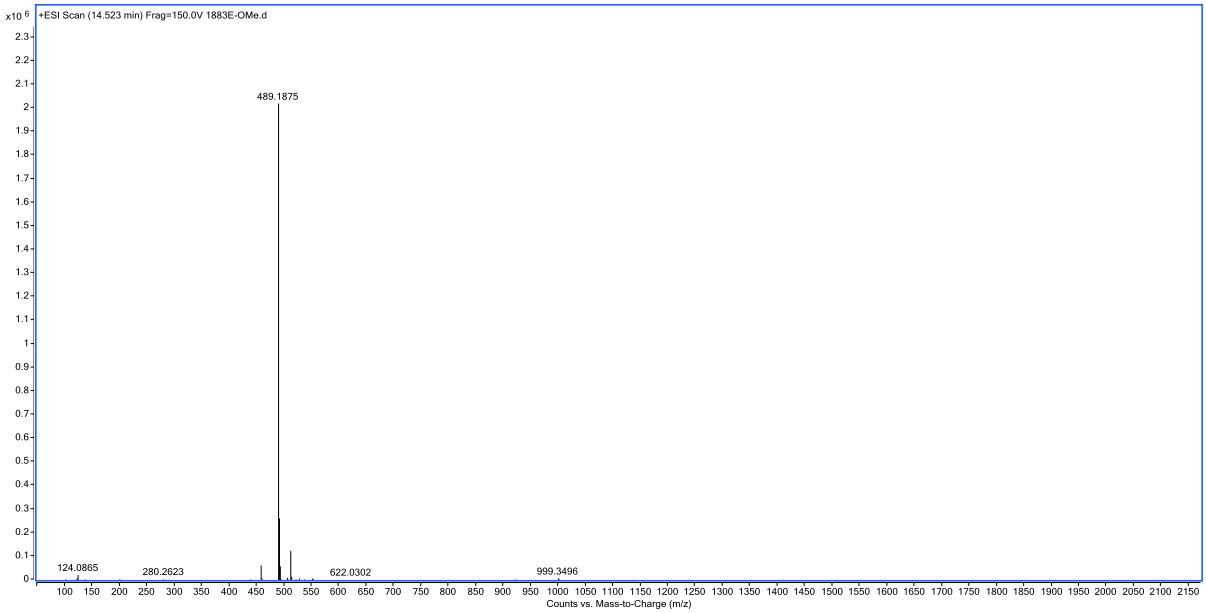


Fig. S33A. ^{1}H NMR spectra of GNM B3 (700 MHz, DMSO-d_6)

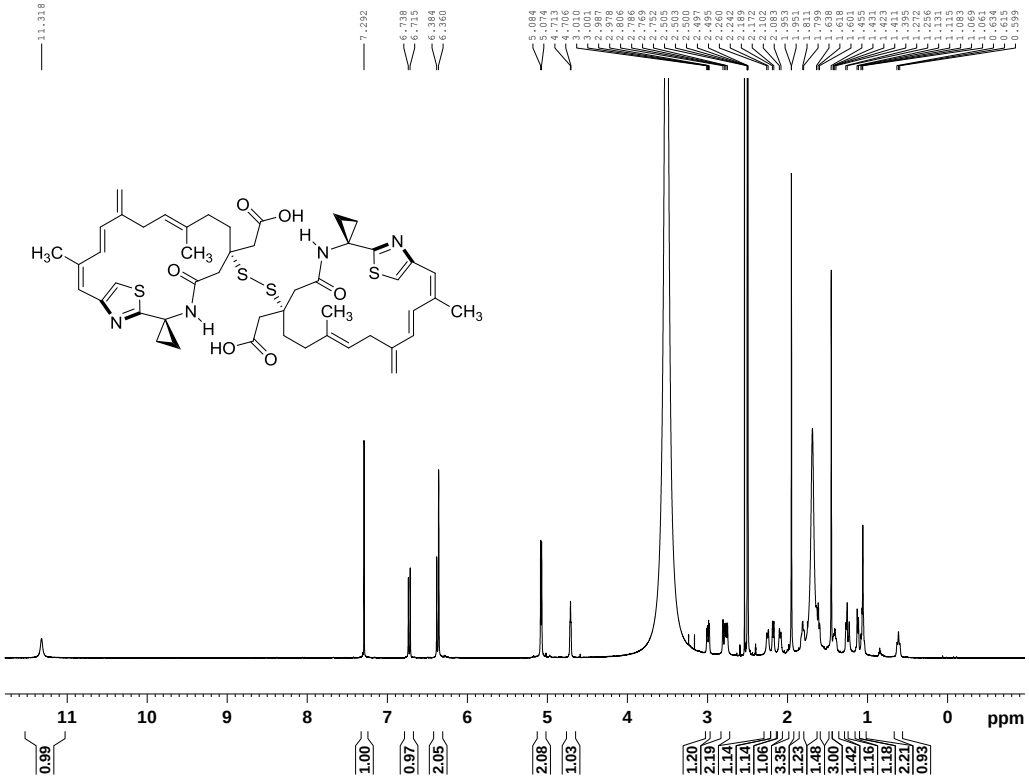


Fig. S33B. ^{13}C NMR spectra of GNM B3 (175 MHz, DMSO-d_6)

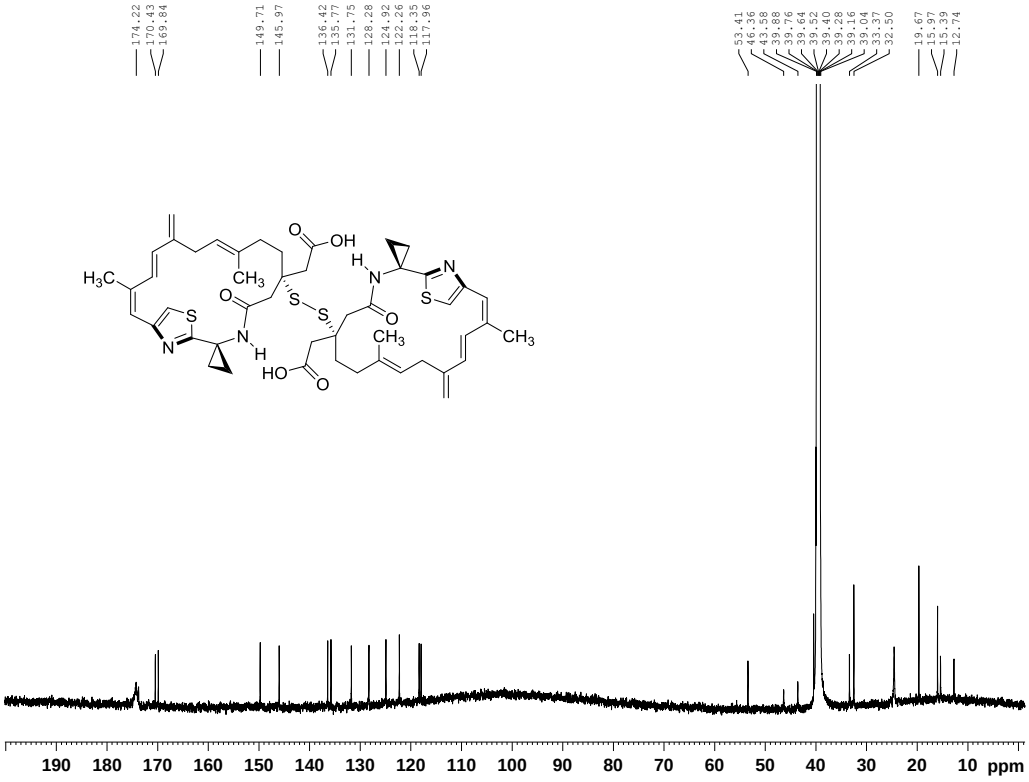


Fig. S33C. DEPT-135 spectra of GNM B3 (DMSO-d_6)

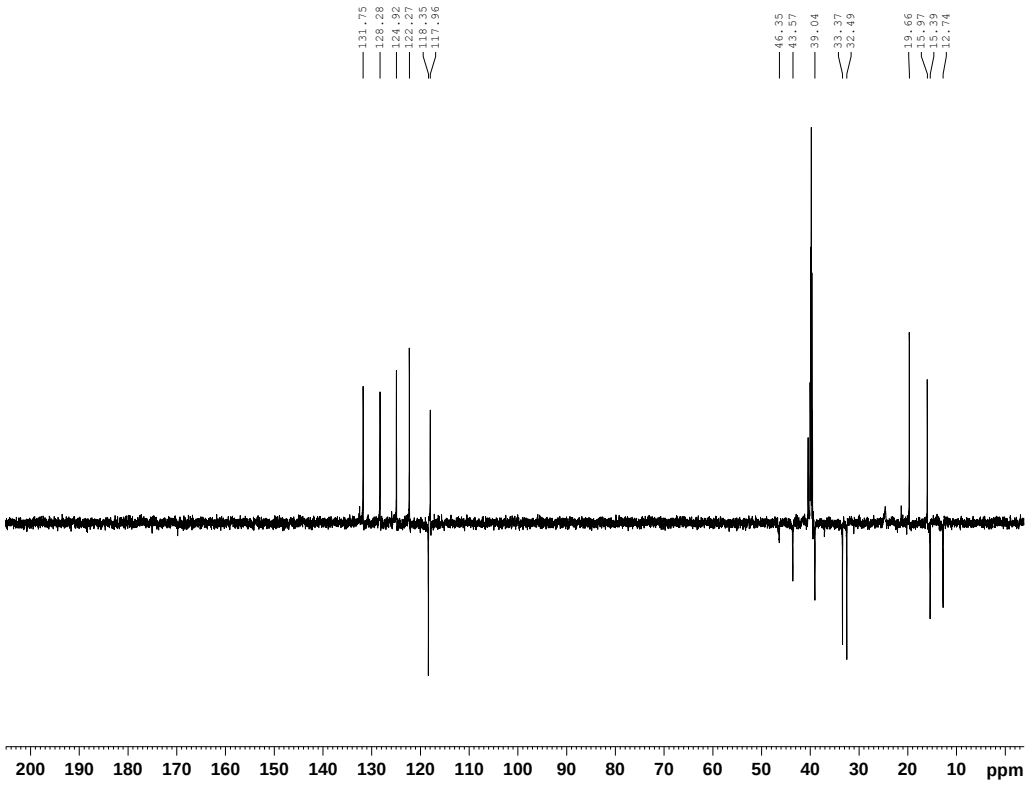


Fig. S33D. HSQC spectra of GNM B3 (DMSO-d_6)

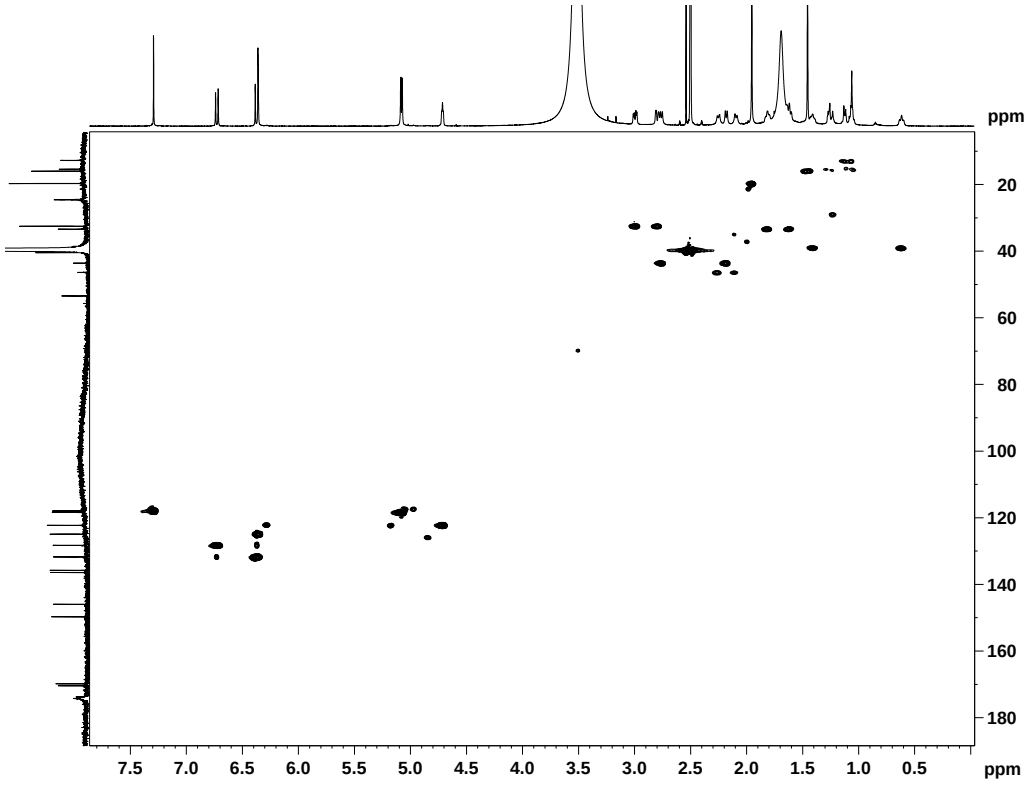


Fig. S33E. ^{1}H-^{1}H COSY spectra of GNM B3 (DMSO-d_6)

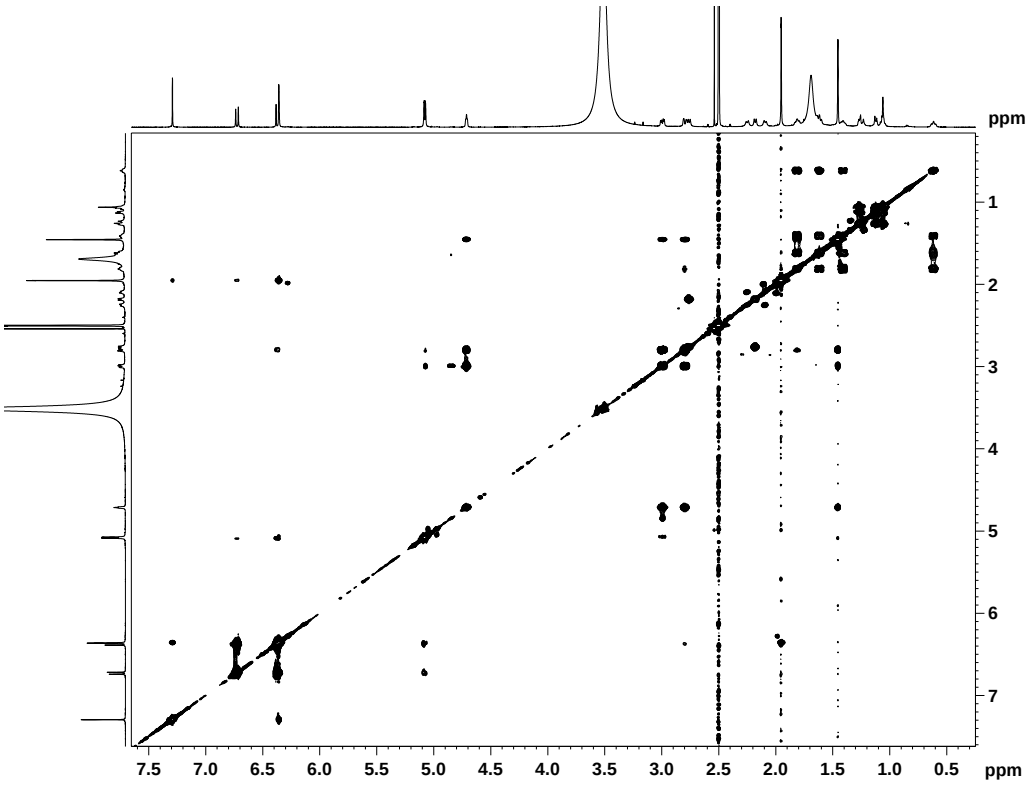


Fig. S33F. HMBC spectra of GNM B3 (DMSO-d_6)

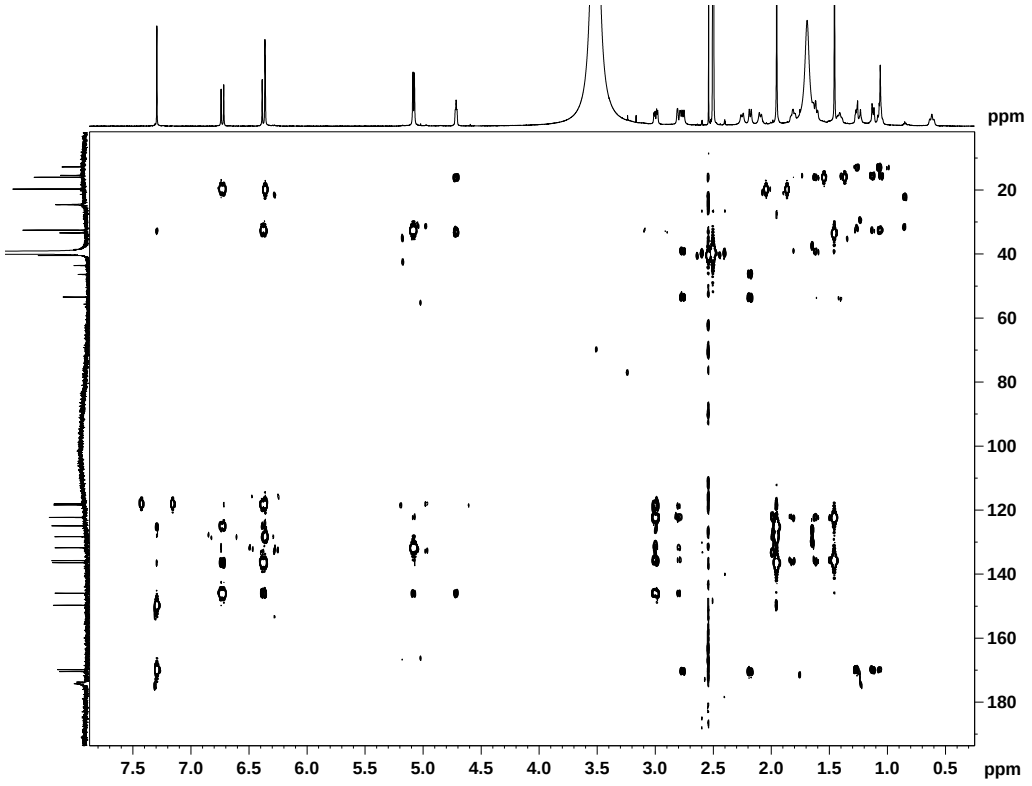


Fig. S33G. ROESY spectra of GNM B3 (DMSO-d_6)

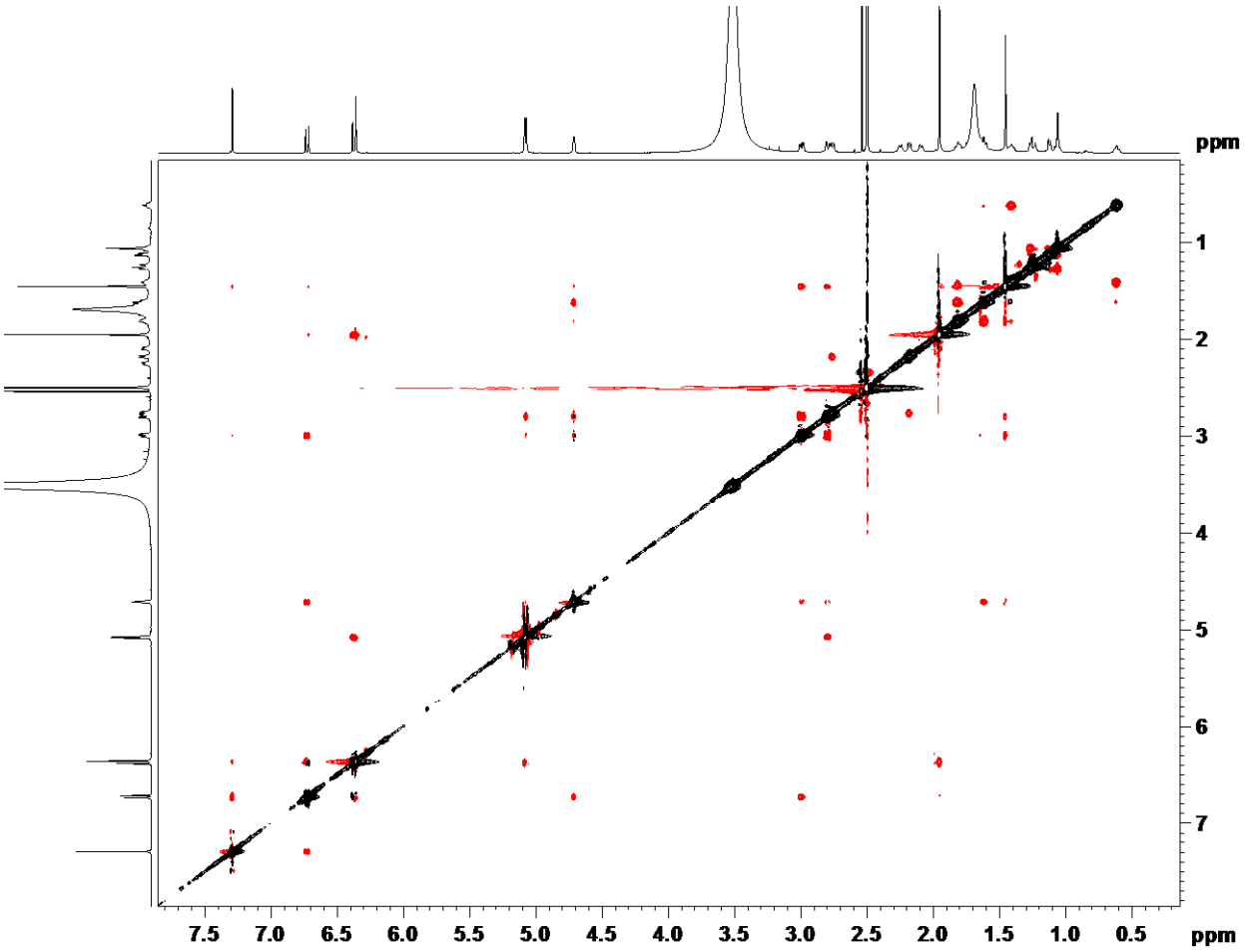


Fig. S33H. HR-MS (ESI) spectra of GNM B3

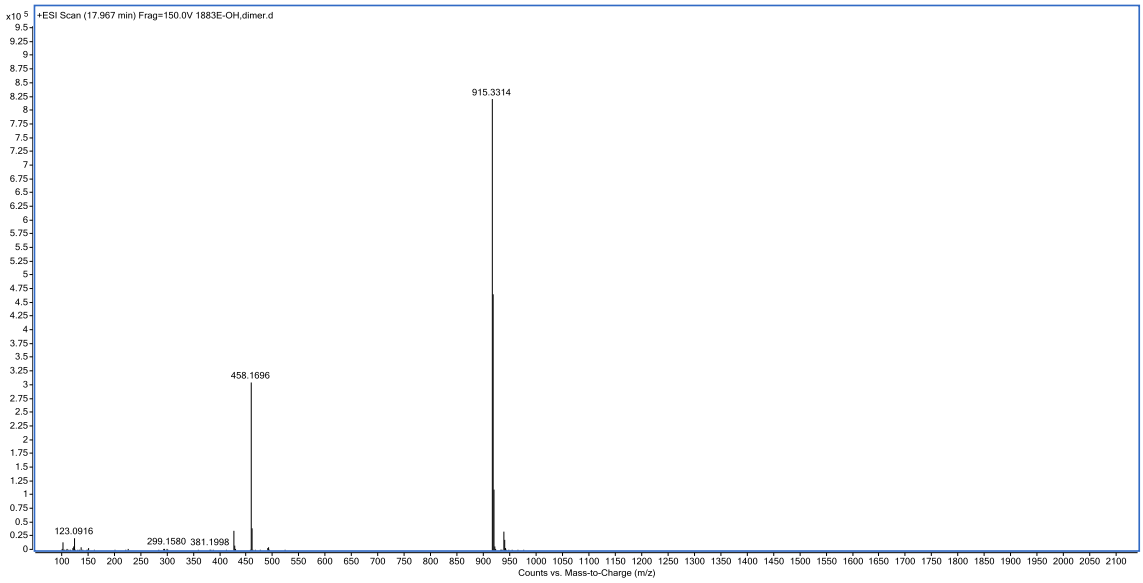
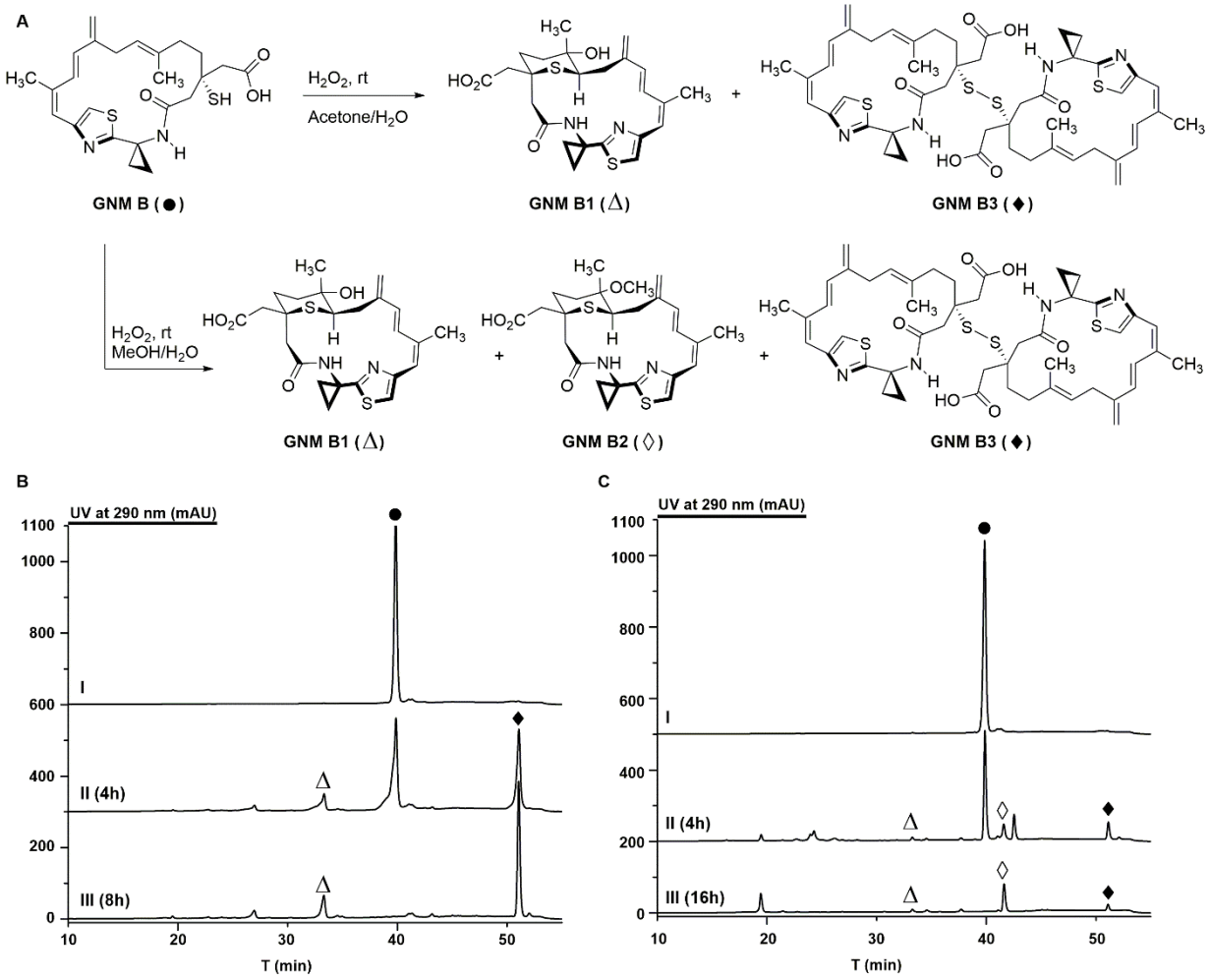


Fig. S34. Chemical transformation of GNM B to GNM B1, GNM B2 and GNM B3 (24). (**A**) Oxidation of GNM B with H_2O_2. Conditions: A solution of GNM B (0.2 mg) in acetone/H_2O (440 µL/15 µL) or in methanol/H_2O (490 µL/5 µL) was added 5 µL H_2O_2 (30% in water), and the reaction mixture was stirred at room temperature for certain time as indicated. The reaction was followed by HPLC and LC-MS using the same conditions as that for the analysis of GNM A. (**B**) Oxidation of GNM B in acetone solution. I: GNM B standard; II: reaction for 4 h; III: reaction for 8 h. (**C**) Oxidation of GNM B in methanol solution. I: GNM B standard; II: reaction for 4 h; III: reaction for 16 h. GNM B (●),GNM B1 (Δ); GNM B2 (◊);GNM B3 (♦).



References

1. Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, NY.
2. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical Streptomyces genetics. The John Innes Foundation, Norwich.
3. Larkin MA, et al. (2007) ClustalW and ClustalX version 2. *Bioinformatics* 23:2947-2948.
4. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729.
5. Weber T, et al. (2015) antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res* 43:W237-W243.
6. Rudolf JD, Yan X, Shen B (2016) Genome neighborhood network reveals insights into enediyne biosynthesis and facilitates prediction and prioritization for discovery. *J Ind Microbiol Biotechnol* 43(2-3):261-276.
7. Hindra, et al. (2014) Strain prioritization for natural product discovery by a high-throughput real-time PCR method. *J Nat Prod* 77:2296–2303.
8. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W (2011) Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–579.
9. Boetzer M, Pirovano W (2012) Toward almost closed genomes with GapFiller. *Genome Biol* 13:R56.
10. Bierman M, Logan R, O'Brien K, Seno ET, Rao RN, Schoner BE (1992) Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to *Streptomyces* spp. *Gene* 116:43-49.
11. Li L, et al. (2017) Engineered jadomycin analogues with altered sugar moieties revealing JadS as a substrate flexible O-glycosyltransferase. *Appl Microbiol Biotechnol* 101:5291-5300.
12. Wang W, Li X, Wang J, Xiang S, Feng X, Yang K (2013) An engineered strong promoter for streptomycetes. *Appl Environ Microbiol* 79:4484–4492.
13. Cheng YQ, Tang GL, Shen B (2002) Identification and localization of the gene cluster encoding biosynthesis of the antitumor macrolactam leinamycin in *Streptomyces atroolivaceus* S-140. *J Bacteriol* 184(24):7013–7024.
14. Yabuuchi T, Kusumi T (2000) Phenylglycine methyl ester, a useful tool for absolute configuration determination of various chiral carboxylic acids. *J Org Chem* 65:397-404.
15. Lohman JR, Bingman CA, Phillips GN, Jr, Shen B (2013) Structure of the bifunctional acyltransferase/decarboxylase LnmK from the leinamycin biosynthetic pathway revealing novel activity for a double-hot-dog fold. *Biochemistry* 52(5):902–911.
16. Aslanidis C, de Jong PJ (1990) Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res* 18:6069-6074.
17. Kadi N, Challis GL (2009) Siderophore biosynthesis: a substrate specificity assay for nonribosomal peptide synthetase-independent siderophore synthatases involving trapping of acyl-adenylate intermediates with hydroxylamine. *Methods Enzymol* 458:431-457.
18. Röttig M, Medema MH, Blin K, Weber T, Rausch C, Kohlbacher O (2011) NRPSpredictor2-a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res* 39:W362-W367.
19. Rausch C, Weber T, Kohlbacher O, Wohlleben W, Huson DH (2005) Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Res* 33(18):5799-5808.
20. Stachelhaus T, Mootz HD, Marahiel MA (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol* 6(8):493-505.
21. González-de-Castro Á, Broughton H, Martínez-Pérez JA, Espinosa JF (2015) Conformational features of secondary N-cyclopropyl amides. *J Org Chem* 80(8):3914-3920.

22. Hu DX, Grice P, Ley SV (2012) Rotamers or diastereomers? An overlooked NMR solution. *J Org Chem* 77:5198-5202.
23. Nagai Y, Kusumi T (1995) New chiral anisotropic reagents for determining the absolute configuration of carboxylic acids. *Tetrahedron Lett* 36(11):1853-1856.
24. Huang SX, et al. (2015) Leinamycin E1 acting as an anticancer prodrug activated by reactive oxygen species. *Proc Natl Acad Sci USA* 112(27):8278-8283.