Supporting Information

Methyltransferases of gentamicin biosynthesis

Sicong Li[a,1], Junhong Guo[a,1], Anna Reva[b], Fanglu Huang[b], Binbin Xiong[a], Yuanzhen Liu[a], Zixin Deng[a,c], Peter F. Leadlay[b,2], and Yuhui Sun[a,2]

[a]Key Laboratory of Combinatorial Biosynthesis and Drug Discovery, Ministry of Education, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, People's Republic of China; [b]Department of Biochemistry, University of Cambridge, Cambridge CB2 1GA, United Kingdom; and [c]State Key Laboratory of Microbial Metabolism, School of Life Sciences & Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China.

[1]S.L. and J.G. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: yhsun@whu.edu.cn or pfl10@cam.ac.uk

Table of contents

Materials and Methods

Table S1. ^{1}H and ^{13}C NMR data of DAA2 and A-2, Ae (400 MHz for ^{1}H, 100 MHz for ^{13}C; in D_2O, δ in ppm) (Related to Figure 1)

Table S2. Plasmids Used in This Study

Table S3. In-frame Deletion Mutants and Complementation Strains Used in This Study

Table S4. List of Candidate Enzymes for 6'-N Methylation in *Micromonospora echinospora* ATCC15835

Table S5. Oligonucleotide Primers Used in This Study

Figure S1. Construction of Methyltransferase Gene Deletion Mutants of *M. echinospora* ATCC15835 (Related to Figure 1 and 2).
A. Construction of ΔgenD1ΔgenN and Confirmation
B. Construction of ΔgenD1ΔgenK and Confirmation
C. Construction of ΔgenNΔgenK and Confirmation
D. Construction of ΔgenD1ΔgenKΔgenN and Confirmation
E. Construction of ΔBN and Confirmation

Figure S2. LC-ESI-HRMS and MS/MS Analysis of Gentamicin Biosynthetic intermediates Isolated from Mutants of *M. echinospora* ATCC15835 (Related to Figure 1 and 2)
A. LC-ESI-HRMS and MS/MS Analysis of DAA2 (**2**)
B. LC-ESI-HRMS and MS/MS Analysis of A (**3**)
C. LC-ESI-HRMS and MS/MS Analysis of A-1 (**6**)
D. LC-ESI-HRMS and MS/MS Analysis of A-2 (**7**)
E. LC-ESI-HRMS and MS/MS Analysis of G418 (**5**)
F. LC-ESI-HRMS and MS/MS Analysis of A2e (**8**)
G. LC-ESI-HRMS and MS/MS Analysis of X2 (**4**)
H. LC-ESI-HRMS and MS/MS Analysis of Ae (**9**)
I. LC-ESI-HRMS and MS/MS Analysis of X2-1 (**10**) and X2-1-i

Figure S3. NMR Analysis of Gentamicin Biosynthetic Intermediates DAA2, A-2 and Ae (Related to Figure 1 and 2)
A. NMR Spectra of DAA2 (**2**)
i. ^{1}H-NMR Spectrum of Compound DAA2 (**2**)
ii. ^{13}C-NMR Spectrum of Compound DAA2 (**2**)
iii. ^{1}H-^{1}H COSY Spectrum of Compound DAA2 (**2**)
iv. HSQC Spectrum of Compound DAA2 (**2**)
v. HMBC Spectrum of Compound DAA2 (**2**)
B. NMR Spectra of A-2 (**7**)
i. ^{1}H-NMR Spectrum of Compound A-2 (**7**)
ii. ^{13}C-NMR Spectrum of Compound A-2 (**7**)

iii. ^{1}H-^{1}H COSY Spectrum of Compound A-2 (**7**)
iv. HSQC Spectrum of Compound A-2 (**7**)
v. HMBC Spectrum of Compound A-2 (**7**)
vi. NOESY Spectrum of Compound A-2 (**7**)

C. NMR Spectra of Ae (**9**)
i. ^{1}H-NMR Spectrum of Compound Ae (**9**)
ii. ^{13}C-NMR Spectrum of Compound Ae (**9**)
iii. ^{1}H-^{1}H COSY Spectrum of Compound Ae (**9**)
iv. HSQC Spectrum of Compound Ae (**9**)
v. HMBC Spectrum of Compound Ae (**9**)
vi. NOESY Spectrum of Compound Ae (**9**)

Figure S4. Genetic Confirmation of ΔgenL by Southern Blot and Analysis by LC-ESI-HRMS
A. Schematic representation of the in-frame deletion of *genL*
B. Confirmation of ΔgenL by Southern blot
C. Extracted ion chromatogram of gentamicin C1a, C2, C2a, C2b and C1 on LC-ESI-HRMS of wild-type, ΔgenL and ΔgenL::*genL*

Figure S5. Confirmation by PCR of Single Methyltransferase Gene Complementation of ΔBN and ΔgenL (Related to Figure 3)
A. Confirmation of ΔBN::*gmrA* by PCR
B. Confirmation of ΔBN::*gmrA+genN* by PCR
C. Confirmation of ΔBN::*gmrA+genD1* by PCR
D. Confirmation of ΔBN::*gmrA+genK* by PCR
E. Confirmation of ΔgenL::*genL* by PCR

Figure S6. LC-ESI-HRMS and MS/MS Analysis of Cell-free Assays (Related to Figure 5)
A. Cell-free Assay with C2 as Substrate
B. Cell-free Assay with C1a as Substrate

Figure S7. Antimicrobial Activity of Gentamicin-related Compounds by Agar Diffusion Assay.

Materials and Methods

Bacterial Strains, Chemicals, and Culture Conditions. *E. coli* DH10B was used as cloning host and *E. coli* ET12567/pUZ8002 for intergeneric conjugation between *E. coli* and *Micromonospora* (18, 19). *M. echinospora* ATCC15835 wild-type strain (40) was used for creation of in-frame deletion mutants and as the source of *gen* genes. Restriction endonucleases, Phusion High-Fidelity Master Mix with GC-buffer, and T4 DNA ligase were purchased from New England Biolabs. Oligonucleotide primers were synthesized by GenScript and Tsingke. DNA sequencing of PCR products was performed by GenScript or by the Department of Biochemistry DNA Sequencing Facility, University of Cambridge. DIG DNA labeling and detection kits were purchased from Roche. For use as standards, G418 was from Sigma-Aldrich, gentamicin A, gentamicin X2, gentamicin C2, gentamicin C2a, gentamicin C1a, gentamicin C2b and gentamicin C1 were from Toku-E.

M. echinospora ATCC15835 wild-type and mutants were grown in ATCC172 medium (glucose 1%, soluble starch 2%, yeast extract 0.5%, N-Z amine type A 0.5%, $CaCO_3$ 0.1%) for chromosomal DNA isolation and preparation of mycelium. *E. coli* strains were maintained in 2×TY media at 37°C with the appropriate antibiotic selection at a final concentration of 100 μg/mL ampicillin, 25 μg/mL chloramphenicol and 25 μg/mL kanamycin.

For fermentation and detection of gentamicin complex and intermediates, *M. echinospora* ATCC15835 and its mutants were cultured in two stages. A seed culture was maintained in liquid ATCC172 medium at 28°C with shaking at 220 rpm for 2 days before being inoculated into fresh liquid ATCC172 medium (5% inoculum), then incubated at 28°C with shaking at 220 rpm for 5 days. For feeding experiments in vivo, 0.005‰ $CoCl_2$ and filter-sterilized compounds (30-40 μg/mL) were added to the medium just before addition of the seed culture. For preparing intermediates on large scale, the second stage was performed in fermentation medium F50 (Soybean powder 2.0%, peptone 0.1%, glucose 0.3%, soluble starch 3.0%, $(NH_4)_2SO_4$ 0.03%, $CaCO_3$ 0.3%, KNO_3 0.03%, $CoCl_2$

0.005‰) shaken at 220 rpm and at 28°C for 5 days.

Construction of Gene Disruption Plasmids. For deletion of gentamicin biosynthesis gene cluster, two DNA fragments flanking the region from *gmrB* to *genN* were amplified from the genomic DNA of *M. echinospora* ATCC15835 by using BN-L1/L2 and BN-R1/R2 (*SI Appendix,* Table S5). The PCR products were each cloned into pUC18, then cloned together into the *Streptomyces-E. coli* shuttle vector pYH7 (41) to obtain the gene disruption plasmid pWHU39, which was verified by sequencing. pWHU2755, pWHU2751, pWHU2752, pWHU2753 and pWHU2754 used for in-frame deleting *genL, orf0086*, *orf1678*, *orf12159* and *orf3626* were constructed by the similar way.

Targeted In-Frame Gene Deletion. To create double in-frame deletion mutants ΔgenD1ΔgenN and ΔgenD1ΔgenK based on mutant ΔgenD1, the corresponding plasmids pWHU1 and pYH289 were introduced into ΔgenD1 (19) by conjugation and mutants screening was carried out using the same method described before (18, 19). The desired in-frame deletion mutants were identified by PCR using the checking primers (*SI Appendix,* Table S5) and further confirmed by Southern blot analysis. Similarly, mutants ΔgenNΔgenK, ΔgenD1ΔgenKΔgenN, ΔBN, ΔgenL, Δorf0086, Δorf1678, Δorf12159 and Δorf3626 were obtained by disruption of target gene(s) in ΔgenN, ΔgenD1ΔgenK and wild-type, respectively, through pYH289, pWHU1, pWHU39, pWHU2755, pWHU2751, pWHU2752, pWHU2753 and pWHU2754.

Gene Complementation of ΔBN and ΔgenL Mutant. For Complementation, pWHU90 was constructed as the vector by inserting *gmrA* into pWHU77 (19) between *Nde*I and *Xba*I sites. Complementation plasmids used to transported into ΔBN host were prepared by cloning *genD1*, *genN* and *genK* with P*ermE** promoter from pWHU66, pWHU68 and pWHU67, respectively (*SI Appendix,* Table S2), using primer pairs P*ermE**-gene-EP1/EP2 (*SI Appendix,* Table S5), into pWHU90 between *Nhe*I and *Ssp*I sites to give

pWHU91, pWHU92 and pWHU93. After sequence confirmation, these plasmids were introduced individually into ΔBN by conjugation. Complementation of ΔgenL mutant was done through a similar way with pWHU94, which is constructed through inserting *genL* into pWHU77 between *Nde*I and *Eco*RI sites. Complemented exconjugants were verified on A medium containing thiostrepton (25 μg/mL) and confirmed by PCR (*SI Appendix,* Fig. S5).

Extraction of Gentamicin Complex and Intermediates. Cultures of wild-type and mutant strains of *M. echinospora* ATCC15835 were adjusted to pH 2.0 with H_2SO_4 and the acidified broth was agitated for 3 hr. The clarified supernatant after centrifugation was filtered through Whatman filter paper, and agitated with DOWEX 50 WX8-200 ion-exchange resin (1 g for 30 mL broth) that was preconditioned with acetonitrile followed by Milli-Q water. After 3 hr，the resin was put in a column then washed by Milli-Q water (8 column volume) and eluted with 1 M NH_4OH (8 column volume). The eluate was freeze-dried and redissolved in Milli-Q water (0.3 mL concentrated solution was equivalent to 30 mL broth), and filtered through 0.22 μm microporous membrane before subjection to LC-ESI-HRMS analysis.

LC-ESI-HRMS Analyses of Gentamicin Complex and Intermediates. LC-ESI-HRMS analysis of extracts of wild-type and mutant strains was performed on a Thermo Electron LTQ-Orbitrap XL fitted with a Phenomenex Luna C18 column (250×4.6 mm) at a flow rate of 0.4 mL/min using a mobile phase of (A) 0.2% trifluoroacetic acid (TFA) in H_2O (adjusted to pH 2.0 with NH_4OH) and (B) 100% CH_3CN; the gradient for separation of gentamicin complex and intermediates: 0-14 min 2% B to 6% B, 14-16 min 6% B to 8% B, 16-25 min 8% B-15% B, 25-26 min 15% B to 40% B, 26-34 min 40% B, 34-35 min 40% B to 2% B, 35-45 min 2% B. MS/MS analyses were carried out in the positive ionization mode with 35% relative collision energy.

Isolation and Purification of Gentamicin Intermediates. Isolation and purification of gentamicin intermediates (except DAA2) from crude extract was performed on a Thermo Scientific HPLC (UltiMate 3000) fitted with evaporative light scattering detector (ELSD, Alltech 2000ES) and a Phenomenex Synergi C18 column (250×10 mm) at a flow rate of 3.8 mL/min using a mobile phase of (A) 0.2% trifluoroacetic acid (TFA) in H_2O and (B) 100% CH_3CN; the gradient for separation of both gentamicin complex and related intermediates: 0-8 min 2% B to 3.5% B, 8-8.5 min 3.5% B to 90% B, 8.5-16 min 90% B, 16-16.5 min 90% B to 2% B, 16.5-20 min 2% B. The temperature and gas flow of ELSD was set 108°C and 2.8 L/min. For isolation of DAA2, the column was replaced by a Phenomenex Luna C18 column (250×4.6 mm) and the flow rate was 1.0 mL/min.

NMR Characterization of Gentamicin Intermediates. The 1D (^{1}H-, ^{13}C- and DEPT) and 2D (^{1}H-^{1}H COSY, HSQC, HMBC and NOESY) NMR spectra were collected on an Agilent-NMR-vnmrs 400 spectrometer. Chemical shifts were reported in ppm using Tetramethylsilane as an internal reference, and NMR data processing was performed by using MestReNova software.

Agar diffusion assay. *Bacillus pumilus* was incubated in 20 mL 2×TY medium to A_{600} 0.6, then was centrifuged and resuspended in 200 µL 2×TY medium. 50 mL 2×TY solid medium was melted and cooled down at 55°C then mixed with the above resuspended cells and was poured into a 14 cm petri dish. 10 µL solution of each test gentamicin-related compound (0.25 mg/mL) was dropped onto filter-paper and placed on the agar surface. The plate was left at 4°C overnight then incubated at 37°C. After 12 h, the zones of inhibition were observed.

Sequencing of *M. echinospora* ATCC15835 genomic DNA. Whole genome sequencing for *M. echinospora* ATCC15835 was performed by Shanghai Southgene Technology Co., Ltd., using a combination of Illumina Hiseq data with PacBio Single

Molecule Sequencing. The DNA sequence was analyzed and annotated using Glimmer 3.02. The sequences of the gentamicin biosynthetic gene cluster (KY971520) and the gentamicin 6'-*N*-methyltransferase candidate genes (MF036116 ~ MF036138) (*SI Appendix,* Table S4) in the *M. echinospora* ATCC15835 genome have been deposited in GenBank at NCBI.

Construction of Methyltransferase Homologs Expression Plasmid. The *genL* was amplified from the genomic DNA of *M. echinospora* ATCC15835 by PCR using Q5® High-Fidelity DNA Polymerase with 30 cycles of denaturation at 98°C for 30 seconds, annealing at 58°C for 30 seconds, and extension at 72°C for 1 min plus final extension at 72°C for 10 min. The PCR product was purified by gel extraction, and inserted into vector pET28a(+) between *Nde*I and *Eco*RI to create pET28a(+)-*genL*. The resulting construct was verified by DNA sequencing, then used to transform *E. coli* BL21(DE3). The other 18 candidate genes were cloned into pET28a(+) by the same method.

Cell-free Assay with Gentamicin C1a and C2. BL21(DE3)/pET28a(+)-*genL* was grown in 5 mL 2×TY medium (0.5% NaCl, 1% yeast extract, 1.6% tryptone) containing kanamycin (50 mg/mL) at 37°C to absorption at 600 nm of 0.6 to 0.8, then induced by isopropylthiogalactoside (0.1 mM) at 18°C with shaking overnight. The cells were collected and suspended in 1 mL ST buffer (50 mM Tris-HCl, 200 mM NaCl, pH 8.0), then opened through ultrasonication for 3 min at 4°C. Subsequently, 0.4 mM C2 (or alternative substrate) and 2 mM SAM were added into the cell-free system and the mixture was incubated at 30°C for 6 hr. The other 19 candidate enzymes were expressed and assayed in the same way.

Overexpression and Purification of GenN and GenL. Overexpression and purification of GenN was described previously (18). For GenL expression and purification, *E. coli* BLR cells transformed with pET-28a(+)-*genL* plasmid were grown in TB broth containing

kanamycin (50 µg/mL) at 37˚C until A_{600} reached 0.7. The cells were then cooled down to 18˚C, and protein expression was induced with IPTG (0.25 mM) at 18˚C overnight. The cells were collected by centrifugation at 3600 x *g* for 10 minutes. The cell pellet was resuspended in Binding buffer (50 mM K_2HPO_4/KH_2PO_4, pH 8, 0.1 M NaCl, 10% (v/v) glycerol, 10 mM imidazole, 0.5% (v/v) Tween® 20). Emulsiflex (Avestin) was used to lyse the cells. Supernatant was separated from cell debris by centrifugation at 48000 x *g* for 1 hour at 4˚C and filtration through a 5 µm membrane. The supernatant was then passed through a 1 mL Co^{2+} His-Select resin (Sigma-Aldrich). After washing the column with 10 mL of Binding buffer followed by 10 mL of Wash buffer (50 mM K_2HPO_4/KH_2PO_4, pH 8, 0.1 M NaCl, 10% (v/v) glycerol, 40 mM imidazole, 0.5% (v/v) Tween® 20), the protein was eluted with Elution buffer (50 mM K_2HPO_4/KH_2PO_4, pH 8, 0.1 M NaCl, 10% (v/v) glycerol, 500 mM imidazole, 0.5% (v/v) Tween® 20). A CentriPure P25 desalting column was used to buffer-exchange the protein into Exchange buffer (50 mM K_2HPO_4/KH_2PO_4, pH 8, 0.1 M NaCl, 10% (v/v) glycerol, 0.5% (v/v) Tween® 20). The protein was then concentrated using a 10 kDa VivaSpin PES concentrator and stored in 10% glycerol at -20˚C.

Enzymatic Assay of GenN with as Substrates Gentamicin Intermediates Purified from *M. echinospora* Mutants. For enzyme assays of GenN, a reaction mixture containing 50 mM TrisHCl, pH 7.5, 0.4 mM aminoglycoside substrate, 2 mM SAM, and 10 µM enzyme was incubated at 30°C overnight. The reactions were stopped by addition of 0.5 volume of chloroform followed by vigorous vortex to denature the proteins. The mixture was then centrifuged at 13,000 rpm for 5 min and the supernatant was subject to LC-ESI-MS analysis. LC-ESI-MS analyses were carried out on an HP1100 high pressure liquid chromatography system (Agilent) coupled with a Finnigan LCQ (Thermo Finnigan). A 250 × 4.6 mm Prodigy 5-µ ODS3 100 Å column (Phenomenex) was used for analysis enzyme assays with a flow rate of 0.4 mL/min and the following mobile phase gradient: 0–14 min, 2–6% B; 14–16 min, 6–8% B; 16–25 min, 8–25% B; 25–26 min, 25–95% B;

26–34 min, 95% B; 34–38 min, 95–8% B; 38–45 min, 8–2% B (solvent A: 0.2% trifluoroacetic acid (TFA) in H_2O; solvent B: acetonitrile).

Enzymatic Assay of GenL with Gentamicin C Complex Components Kinetic Characterization of GenL. A typical assay for GenL activity contained 50 mM Tris-HCl, pH 7.8, 0.2 mM aminoglycoside substrate, 1 mM SAM and 20 µM enzyme. Reactions were typically incubated at 30°C for 1 hr or overnight followed by treatment with chloroform to precipitate protein as described for GenN assays. The kinetic parameters of GenL activity were obtained using a Methyltransferase Colorimetric Assay Kit (Cayman Chemical) to follow GenL-catalyzed transfer of methyl group from SAM to gentamicins C2 or C1a (29). A mixture of 1 µM recombinant GenL in 20 mM Tris-HCl, pH 7.5, and 100 mM NaCl, and various concentrations (50-800 µM) of gentamicins C2, C2a, and C1a, was pre-incubated at 30°C for 5 minutes. The reaction mixture containing 175 µM SAM, detection system enzymes and chemicals in buffer was then added to each well. Absorbance at 510 nm was recorded every 30 seconds over 30 minutes using PHERAstar® *FSX* spectrophotometer (BMG Labtech). All reactions were performed at 30°C, in triplicate.

Table S1. ^{1}H and ^{13}C NMR data of DAA2 and A-2, Ae (400 MHz for ^{1}H, 100 MHz for ^{13}C; in D_2O, δ in ppm) (Related to Figure 1)

	DAA2		A-2		Ae	
position	δ_H (*mult*)	δ_C	δ_H (*mult*)	δ_C	δ_H (*mult*)	δ_C
1	3.43 (m)	48.6/49.4[a]	3.42-3.48 (m)	48.7/49.3[a]	3.37-3.45 (m)	48.7/49.2[a]
2	2.41 (dt), 1.78 (dd)	28.0	2.40 (dt), 1.77 (dd)	27.7	2.35 (m), 1.72 (dd)	27.7
3	3.43 (m)	48.6/49.4[a]	3.42-3.48 (m)	48.7/49.3[a]	3.37-3.45 (m)	48.7/49.2[a]
4	3.75 (m)	80.5	3.72 (m)	81.8	3.68 (m)	81.7
5	3.35 (m)	69.3	3.31 (m)	69.9	3.28 (m)	69.9
6	3.62 (m)	83.6	3.61 (m)	83.4	3.60 (m)	83.4
1'	5.53 (d)	97.0	5.44 (d)	97.6	5.45 (d)	97.6
2'	3.32 (dd)	53.8	3.36 (dd)	53.9	3.32 (m)	53.9
3'	3.77 (m)	69.0	3.72 (m)	69.3	3.70 (m)	69.3
4'	3.69 (m)	73.5	3.70 (m)	73.3	3.64 (m)	73.3
5'	3.70 (m)	73.6	3.76 (m)	75.2	3.72 (m)	75.2
6'	3.60 (m), 3.79 (m)	60.2	4.12 (dq)	65.0/65.1[a]	4.07 (m)	65.0
6'-Me	/	/	1.07 (d)	14.7	1.08 (d)	14.2
1''	4.91 (d)	100.5	4.92 (d)	100.4	4.93 (d)	100.6
2''	3.78 (m)	67.9	3.79 (dd)	67.9	3.89 (m)	65.9
3''	3.27 (m)	54.9	3.24 (dd)	54.9	3.34 (m)	60.4
N-3''H- Me	/	/	/	/	2.68 (s)	29.1
4''	3.71 (m)	65.2	3.72 (m)	65.0/65.1[a]	3.90 (m)	63.1
5''	3.60 (m), 3.67(m)	62.4	3.70 (m), 3.58 (m)	62.2	3.70 (m), 3.55 (m)	62.3

[a]Assignment may be interchanged

Table S2. Plasmids Used in This Study

Plasmid	Description	Reference
pUC18	Sub-cloning vector	Joseph and David, 2001
pYH7	*Streptomyces-E. coil* shuttle vector	Sun et al., 2009
pIB139	*int, att, acc*(3)*IV,* P*ermE**	Wilkinson et al., 2002; Del Vecchio et al., 2003
pWHU77	pIB139 derivative with Tsr^R	Guo et al., 2014
pYH289	*genN* in-frame deletion construct	Huang et al., 2015
pWHU1	*genK* in-frame deletion construct	Guo et al., 2014
pWHU39	ΔBN deletion construct	This study
pWHU2755	*genL* in-frame deletion construct	This study
pWHU2751	*orf0086* in-frame deletion construct	This study
pWHU2752	*orf1678* in-frame deletion construct	This study
pWHU2753	*orf12195* in-frame deletion construct	This study
pWHU2754	*orf3626* in-frame deletion construct	This study
pWHU66	Donor of *genD1* for constructing pWHU92	Huang et al., 2015
pWHU67	Donor of *genK* for constructing pWHU93	Guo et al., 2014
pWHU68	Donor of *genN* for constructing pWHU91	Huang *et al*., 2015
pWHU90	ΔBN complementation starter construct containing *gmrA* under the control of the P*ermE** promoter	This study
pWHU91	ΔBN complementation construct containing *gmrA* and *genN* under the control of two individual P*ermE** promoters	This study

pWHU92	ΔBN complementation construct containing *gmrA* and *genD1*under the control of two individual P*ermE** promoters	This study
pWHU93	ΔBN complementation construct containing *gmrA* and *genK* under the control of two individual P*ermE** promoters	This study
pWHU94	ΔgenL complementation construct containing *genL* under the control of the P*ermE** promoters	This study
pUC18-M1-*gmrA*	Donor of *gmrA* for construct pWHU90	This study
pET28a(+)-*genN*	GenN expression construct	Huang *et al*., 2015
pET28a(+)-*genL*	GenL expression construct	This study
pET28a(+)-*orf0088*	Orf0088 expression construct	This study
pET28a(+)-*orf5598*	Orf5598 expression construct	This study
pET28a(+)-*orf0283*	Orf0283 expression construct	This study
pET28a(+)-*orf2746*	Orf2746 expression construct	This study
pET28a(+)-*orf0475*	Orf475 expression construct	This study
pET28a(+)-*orf0639*	Orf639 expression construct	This study
pET28a(+)-*orf1520*	Orf1520 expression construct	This study
pET28a(+)-*orf2298*	Orf2298 expression construct	This study
pET28a(+)-*orf3388*	Orf3388 expression construct	This study
pET28a(+)-*orf4521*	Orf4521 expression construct	This study
pET28a(+)-*orf2148*	Orf2148 expression construct	This study
pET28a(+)-*orf4864*	Orf4864 expression construct	This study
pET28a(+)-*orf5560*	Orf5560 expression construct	This study

pET28a(+)-*orf5573*	Orf5573 expression construct	This study
pET28a(+)-*orf5664*	Orf5664 expression construct	This study
pET28a(+)-*orf6656*	Orf6656 expression construct	This study
pET28a(+)-*orf5907*	Orf5907 expression construct	This study
pET28a(+)-*orf5344*	Orf15344 expression construct	This study

Table S3. In-frame Deletion Mutants and Complementation Strains Used in This Study

Gene knockout mutant or complemented strain	Parent strain	Plasmid used for deletion or complementation	Reference
ΔBN	wild-type	pWHU39	This study
ΔgenD1	wild-type	pYH287	Huang *et al.*, 2015
ΔgenN	wild-type	pYH289	Huang *et al*., 2015
ΔgenK	wild-type	pWHU1	Guo *et al*., 2014
ΔgenL	wild-type	pWHU2755	This study
ΔgenD1ΔgenK	ΔgenD1	pWHU1	This study
ΔgenNΔgenK	ΔgenN	pWHU1	This study
ΔgenD1ΔgenN	ΔgenD1	pYH289	This study
ΔgenD1ΔgenKΔgenN	ΔgenD1ΔgenK	pYH289	This study
ΔBN::*gmrA*	ΔBN	pWHU90	This study
ΔBN::*gmrA+genN*	ΔBN	pWHU91	This study
ΔBN::*gmrA+genD1*	ΔBN	pWHU92	This study
ΔBN::*gmrA+genK*	ΔBN	pWHU93	This study
ΔgenL::*genL*	ΔgenL	pWHU94	This study

Table S4. List of Candidate Enzymes for 6'-*N*-methylation in *M. echinospora* ATCC15835

IstU Homologs

ORF	Number of residues	Identity to IstU (%)	E-value	Gene ID
Orf0086	248	40	4E-37	MF036116
Orf1678	246	36	2E-32	MF036132
Orf12195	279	38	1E-24	MF036133
Orf3626	283	27	2E-05	MF036117

GenN Homologs

ORF	Number of residues	Identity to GenN (%)	E-value	Gene ID
Orf4864	493	30.60	8.00E-07	MF036134
Orf5573	246	26.67	7.00E-05	MF036118
Orf0639	215	37.31	8.00E-04	MF036119
Orf5344	264	37.50	0.008	MF036120
Orf2298	201	32.58	0.019	MF036121
Orf5365 (GenL)*	239	37.84	0.023	MF036122
Orf5664	257	40.35	0.025	MF036123
Orf1520	266	33.33	0.025	MF036135
Orf5560	394	36.21	0.06	MF036136
Orf0475	482	29.77	0.18	MF036137
Orf5598	267	51.85	0.22	MF036124
Orf2148	219	27.78	0.29	MF036125
Orf0088	263	31.43	0.41	MF036138
Orf0283	295	38.10	0.99	MF036126
Orf6656	525	39.62	1.4	MF036127
Orf2746	250	37.78	2	MF036128
Orf3388	347	35.29	3.7	MF036129
Orf5907	207	37.84	4.6	MF036130
Orf4521	326	34.04	9.1	MF036131

***, The amino acid sequence of Orf5365 (GenL) is as follows:**

VLSISDLRTDWKIFRQTMRDSTLKEALVDSAEYIRIRRHERRERFDERFGTETNGIVGLA
DIDSIGTHQEEASHYLPTRKQEFDRMMATVGELDHGEHVFVDLGCGKGRVVLLAAEKP
YKKVIGVDFSPSFISQAKENVERYTGPVATHEIELLAIDAVDFVVPPENLIVYLFSPFGPPV
FDTVMRNLVAATKKRKQKITIVYYSPDYDDVVREAGFTLVAQGKGDHWPWSVYSVGES
A

Table S5. Oligonucleotide Primers Used in This Study

Primer	Oligonucleotide sequences (5' to 3')	Restriction site
BN-L1	GCC**CATATG**AGCAAGAGTCTCGGC	*Nde*I
BN-L2	GGC**GAATTC**CTTCTGCACGGC	*Eco*RI
BN-R1	GAC**GAATTC**CTGCGGGGCTGACCCC	*Eco*RI
BN-R2	GAG**AAGCTT**GCCGCCGACTCCGACC	*Hin*dIII
BN-CP1	CCACCCTGCACAACCTGCGCTTCTACAC	
BN-CP2	TGACGACGACCGCGACGACGATGA	
genN-CP1	GGATGGGATGCCAACGACC	
genN-CP2	ACCGCGACGACGATGACG	
genD1-CP1	GAAGCTCGCCGATGCCA	
genD1-CP2	CAGGTGAAGGCGGTGGTG	
genK-CP1	CGGGCGAACCTTCGGGATA	
genK-CP2	CCGTCAGCGTTGGCAATAA	
genL-L1	GCA**CATATG**GCAGGCGTGGGTCAACAG	*Nde*I
genL-L2	CCG**TCTAGA**GTCTACTCCGTCGGCGAG	*Xba*I
genL-R1	GCG**TCTAGA**GGAGATGCTCAGCACGGT	*Xba*I
genL-R2	GCC**AAGCTT**ATCACCGAGTACGGTCGC	*Hin*dIII
genL-CP1	CGTGTCGCTGTTCTGGGTCA	
genL-CP2	CTCGTGCTCGTCATCGCCTA	
orf0086-L1	GCC**CATATG**AGTCTGCTGGTCTTCCTC	*Nde*I
orf0086-L2	GGC**AGGCCT**GTCACCGTCTACGAGAAG	*Stu*I

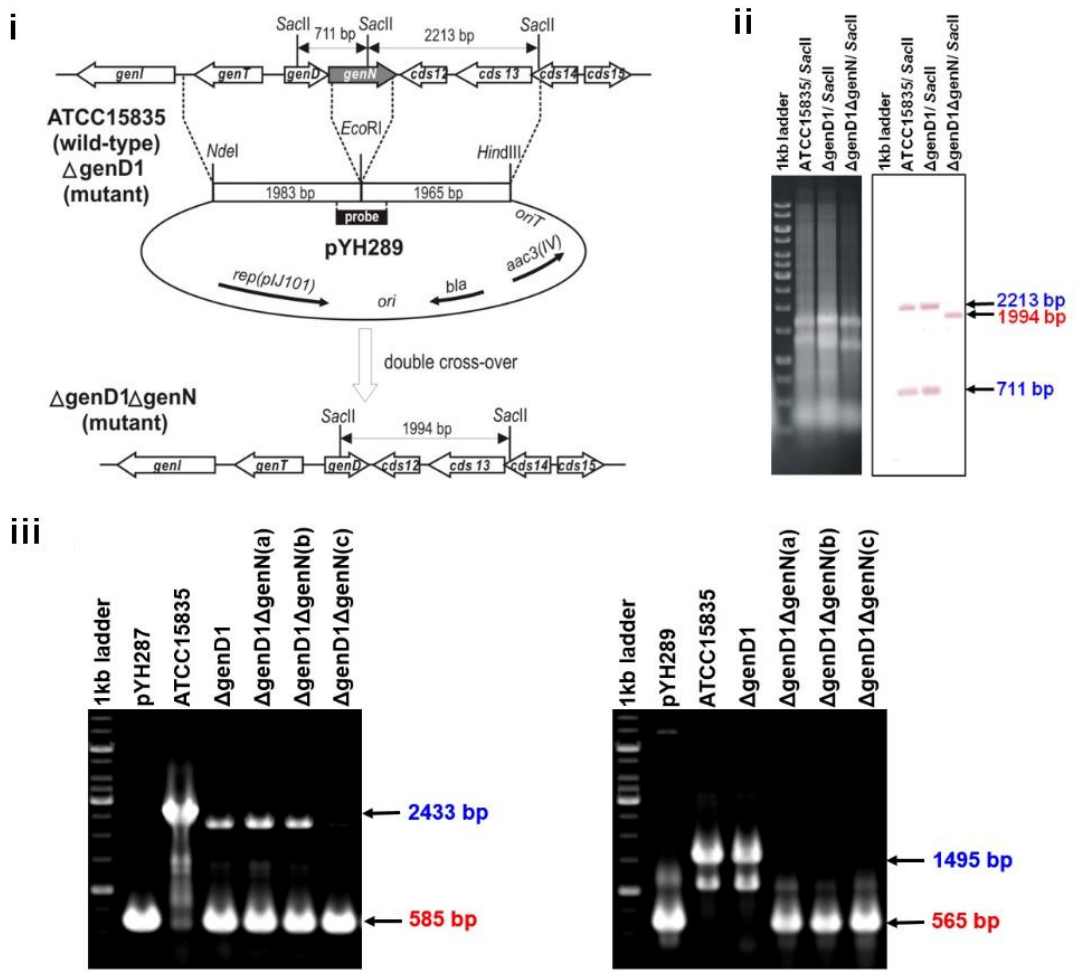
orf0086-R1	GGT**AGGCCT**TCCGTAGTTCCTCATGGT	*Stu*I
orf0086-R2	GGG**AAGCTT**TGCCAGTCACGCTAAGAC	*Hin*dIII
orf0086-CP1	CGCTGTCTTTCCAGTCGCTC	
orf0086-CP2	GGCTACGGCATGACCTTCCT	
orf1678-L1	GGT**CATATG**GCTGACCTCGGCGATGCG	*Nde*I
orf1678-L2	GCT**AGGCCT**GAAGGACAACGGCACTCCC	*Stu*I
orf1678-R1	CAG**AGGCCT**TCGGTCTACCGACGGGAGG	*Stu*I
orf1678-R2	GCG**AAGCTT**TGGTGGCGCTGTTCCTCTC	*Hin*dIII
orf1678-CP1	TCATCACGTCTCCTCCAACCG	
orf1678-CP2	TGAGCACCCCGAGCACCTTTC	
orf12195-L1	CCG**CATATGT**CCCTTTCTGCAACTCCCG	*Nde*I
orf12195-L2	GGA**AGGCCT**GGGTGACGACCTCGTCCA	*Stu*I
orf12195-R1	CCA**AGGCCT**CTCTTGTCGCATACCGTCG	*Stu*I
orf12195-R2	GAG**AAGCTT**GTTGGTGGCGTAGGTCGGG	*Hin*dIII
orf12195-CP1	CGGTGGTGAGCCTGGTCGTGC	
orf12195-CP2	CGGACTGGCCCTGATGGTGCC	
orf3626-L1	CCG**CATATG**GTCGATTCATCGCTGTGC	*Nde*I
orf3626-L2	GTC**TCTAGA**GCCGCCTGACCCGTCAGA	*Xba*I
orf3626-R1	CCC**TCTAGA**GCACTCGGCGTCGTAGACG	*Xba*I
orf3626-R2	GCA**AAGCTT**ACGGGAACATCACCAGCA	*Hin*dIII
orf3626-CP1	GCGTTGAACGACCCGCTCTCTGACG	
orf3626-CP2	GCCACTTCCCACCACGATCACCTGA	
P*ermE**-gene-EP1	CGAC**GCTAGC**ATGCATGCGAGTGTCCGTTC	*Nhe*I
P*ermE**-gene-EP2	AAAC**AATATT**GACATGATTACGAATTCGATAT	*Ssp*I
P*ermE**-gene-CP1	AAGGGAATAAGGGCGACACGG	
P*ermE**-gene-CP2	GAAGATCGAAGAGAAGCAGGACGAG	
gmrA-EP/CP1	AAC**CATATG**ACTAGTCAGGATCTTTC	*Nde*I

gmrA-EP/CP2	GCTT**TCTAGA**GCCGAAAAGCTAT	*Xba*I
orf0088-1	TTC**GAATTC**AGCAGGCACAGCCGGTGTC	*Eco*RI
orf0088-2	CCT**CATATG**CGTCCGCCGCCGTACGAAT	*Nde*I
orf5598-1	CGG**GAATTC**CGGTAGTGCCTCGCGGTCA	*Eco*RI
orf5598-2	GTC**CATATG**GGAACGGGCTTCAGCGGCG	*Nde*I
orf0283-1	CGA**GAATTC**GAAGACCACGGTGGACGAG	*Eco*RI
orf0283-2	ACA**CATATG**TTGGAGAACCTCCGCACCC	*Nde*I
orf2746-1	TCC**GAATTC**CTCGTCGCCCACCACCTG	*Eco*RI
orf2746-2	CCG**CATATG**CTGCCGCCGAACGATCC	*Nde*I
orf0475-1	TGT**GAATTC**ACGAGCTGCTGTGGGATGT	*Eco*RI
orf0475-2	GCG**CATATG**GTTGGCTGGGAAGCTGTGA	*Nde*I
orf0639-1	ACG**CATATG**ACCGGGGACGGTTTC	*Nde*I
orf0639-2	GGC**GAATTC**CTCATCGCTGCGCAC	*Eco*RI
orf1520-1	TGA**GAATTC**CGCGCTTGGTCTTGGTCCT	*Eco*RI
orf1520-2	CCC**CATATG**GGTGTGATGGCGGGCAACG	*Nde*I
orf2298-1	CCA**GAATTC**CCACGTCCACGTATCCCAC	*Eco*RI
orf2298-2	GCG**CATATG**ACCGGCGACCACTACTTCA	*Nde*I
orf3388-1	CGG**CATATG**CGGACGCCCCCGGACGT	*Nde*I
orf3388-2	GCC**GAATTC**CCAGCTCCCCCCGACCG	*Eco*RI
orf4521-1	ACG**GAATTC**ACCTCTGCGGACATGCC	*Eco*RI
orf4521-2	GAC**CATATG**GAGCTGAGCTACGCCTTTC	*Nde*I
orf2148-1	GGC**GAATTC**CGTCCTTTAGAGCTGCCCG	*Eco*RI
orf2148-2	CGA**CATATG**ATGGACGCGGGGACCAGGG	*Nde*I
orf4864-1	GTC**GAATTC**ACCGCCCGCATCAGG	*Eco*RI
orf4864-2	GGC**CATATG**GACGAACACGACATG	*Nde*I
orf5560-1	GCC**GAATTC**ACCCTGTTGCGCCCGTGAC	*Eco*RI
orf5560-2	GAG**CATATG**CCCCGGCACAACGTCCTTG	*Nde*I

orf5573-1	ACC**CATATG**AACCGACTCGGCCTC	*Nde*I
orf5573-2	GCG**GAATTC**GGAACGGGACGGTCC	*Eco*RI
orf5664-1	GCC**GAATTC**CTGACCGACACCCGGTAGG	*Eco*RI
orf5664-2	GGC**CATATG**AGTGGACTCGCGGACCTGA	*Nde*I
orf6656-1	TGC**CATATG**GACGACAGGTACGACGTG	*Nde*I
orf6656-2	GGC**GAATTC**CGTCAAGCGTCTTCACCG	*Eco*RI
orf5907-1	TTG**CATATG**GAGTTCGACGGGGAGTACT	*Nde*I
orf5907-2	GAC**GAATTC**CATAACGCGACCATGAC	*Eco*RI
orf5344-1	TTG**CATATG**TGGGATCCGAGCAGC	*Nde*I
orf5344-2	CGA**GAATTC**AGGGCAGGTCGGTCA	*Eco*RI
genL-U	TGA**CATATG**CTGAGCATCTCCGATCTAC	*Nde*I
genL-D	AGT**GAATTC**ACGAGGTGGTCTACGCGAT	*Eco*RI

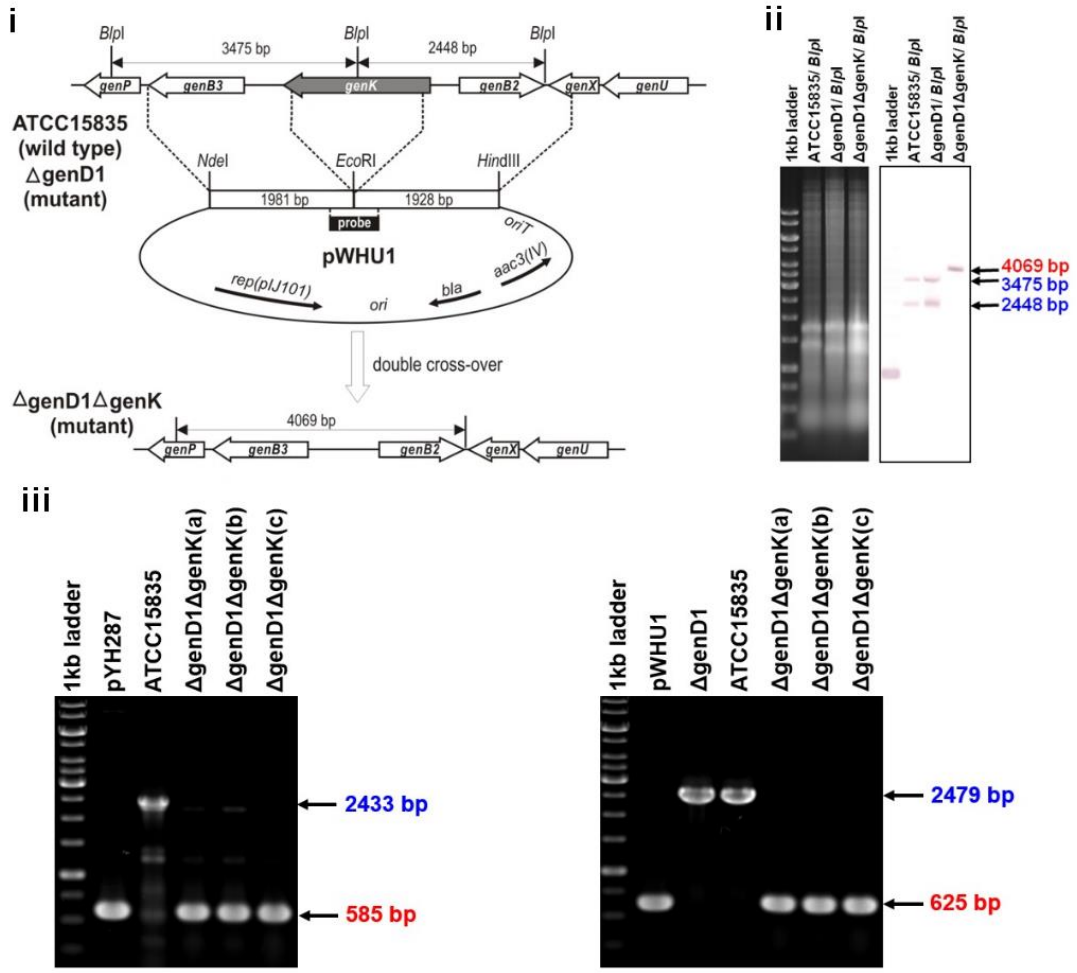
Figure S1. Construction of Methyltransferase Gene Deletion Mutants of *M. echinospora* ATCC15835 (Related to Figure 1 and 2).

A. Construction of ΔgenD1ΔgenN and Confirmation



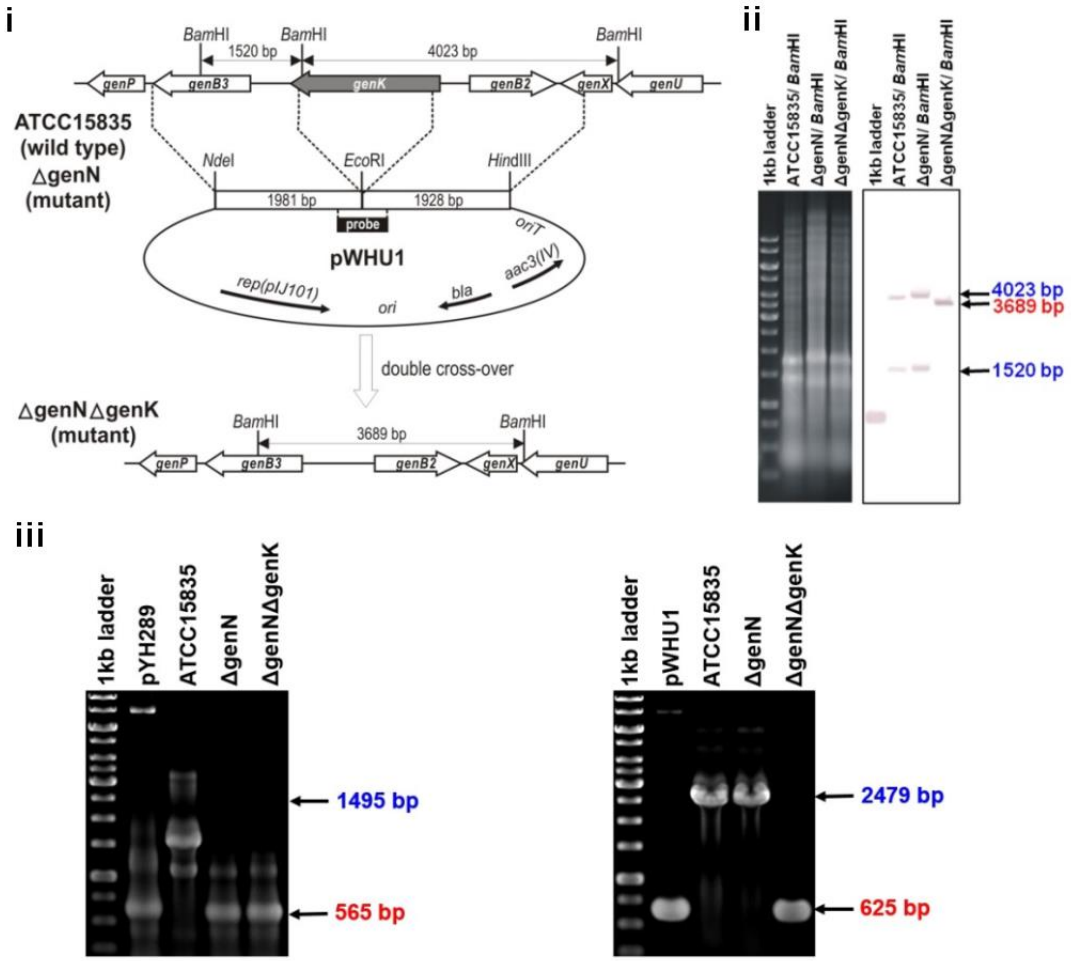
(i) Schematic representation of the in-frame deletion; (ii) Confirmation by Southern blot; (iii) Confirmation by PCR. The arrows indicate the expected size of the PCR fragments in the wild-type and mutants. The dash arrows show the location of PCR primers for check mutants. The probe for Southern blot was obtained by PCR-amplified from pYH289 using primers *genN*-CP1 and *genN*-CP2.

B. Construction of ΔgenD1ΔgenK and Confirmation



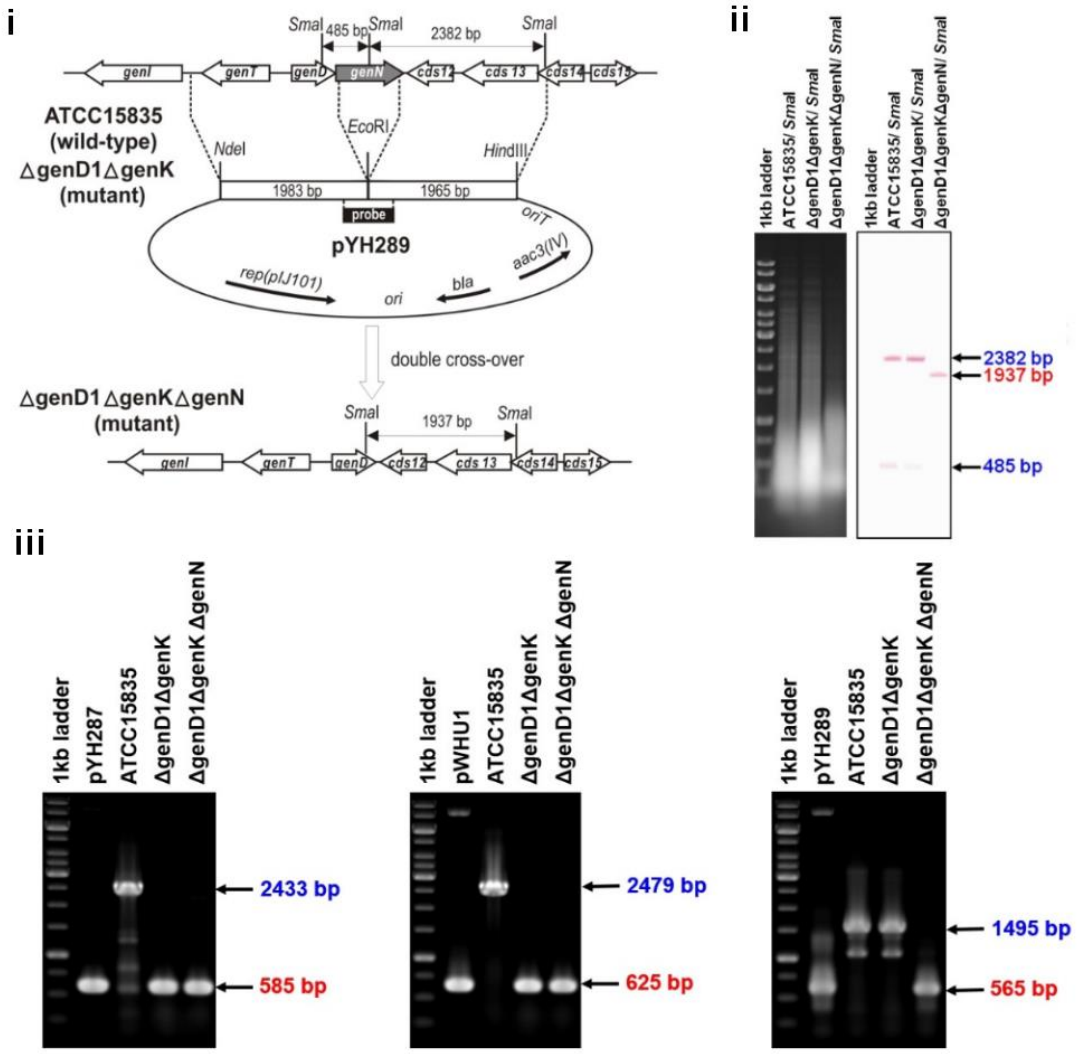
(i) Schematic representation of the in-frame deletion; (ii) Confirmation by Southern blot; (iii) Confirmation by PCR. The arrows indicate the expected size of the PCR fragments in the wild-type and mutants. The dash arrows show the location of PCR primers for check mutants. The probe for Southern blot was obtained by PCR-amplified from pWHU1 using primers *genK*-CP1 and *genK*-CP2.

C. Construction of ΔgenNΔgenK and Confirmation



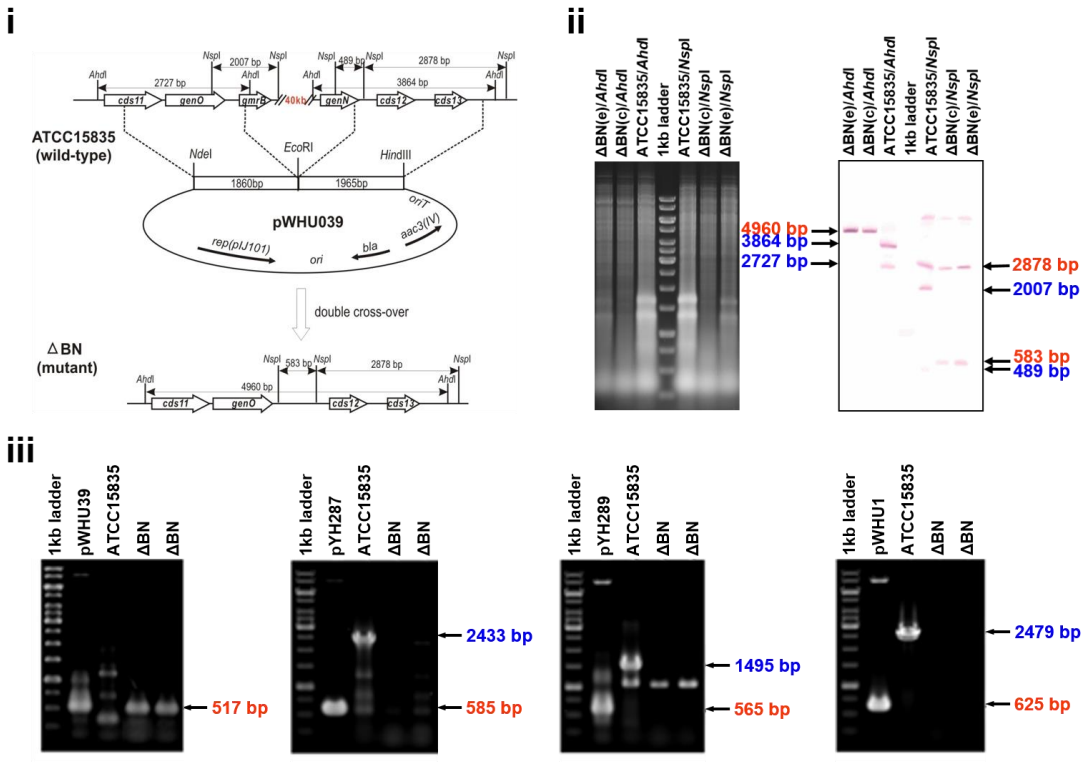
(i) Schematic representation of the in-frame deletion; (ii) Confirmation by Southern blot; (iii) Confirmation by PCR. The arrows indicate the expected size of the PCR fragments in the wild-type and mutants. The dash arrows show the location of PCR primers for check mutants. The probe for Southern blot was obtained by PCR-amplified from pWHU1 using primers *genK*-CP1 and *genK*-CP2.

D. Construction of ΔgenD1ΔgenKΔgenN and Confirmation



(i) Schematic representation of the in-frame deletion; (ii) Confirmation by Southern blot; (iii) Confirmation by PCR. The arrows indicate the expected size of the PCR fragments in the wild-type and mutants. The dash arrows show the location of PCR primers for check mutants. The probe for Southern blot was obtained by PCR-amplified from pYH289 using primers *genN*-CP1 and *genN*-CP2.

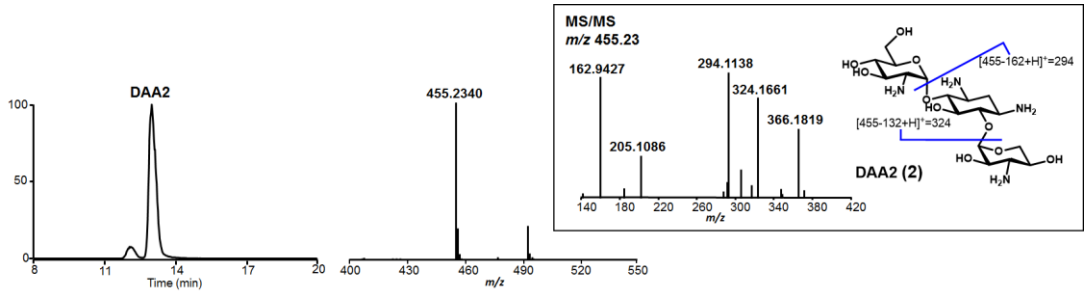
E. Construction of ΔBN and Confirmation



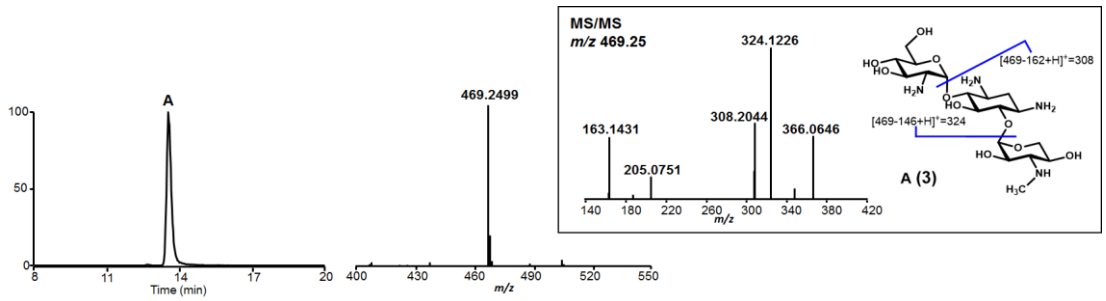
(i) Schematic representation of the in-frame deletion; (ii) Confirmation by Southern blot; (iii) Confirmation by PCR of the deletion of the large fragment using BN-CP1, BN-CP2, *genD1*-CP1, *genD1*-CP2, *genN*-CP1, *genN*-CP2, *genK*-CP1 and *genK*-CP2. The arrows indicate the expected size of the PCR fragments in the wild-type and mutant. The dash arrows show the location of PCR primers for check mutants. The probe for Southern blot was obtained by PCR-amplified from pWHU39 using primers BN-CP1 and BN-CP2.

Figure S2. LC-ESI-HRMS and MS/MS Analysis of Gentamicin Biosynthetic intermediates Isolated from Mutants of *M. echinospora* ATCC15835 (Related to Figure 1 and 2)

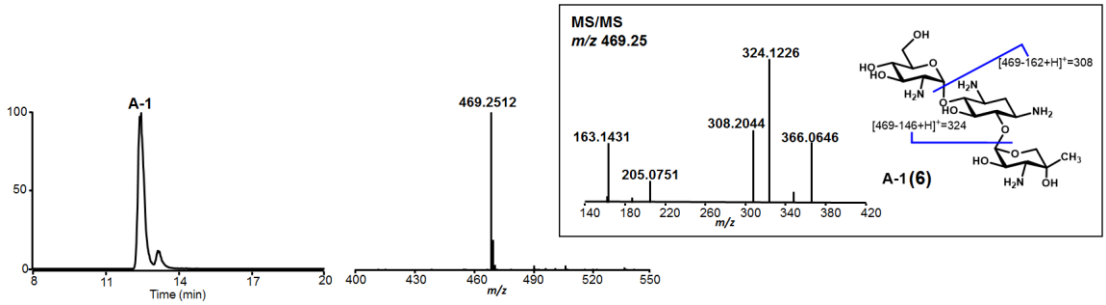
A. LC-ESI-HRMS and MS/MS Analysis of DAA2 (**2**)



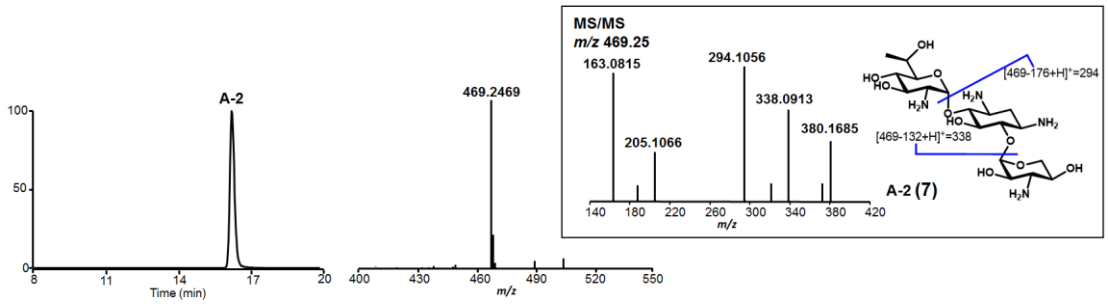
B. LC-ESI-HRMS and MS/MS Analysis of A (**3**)



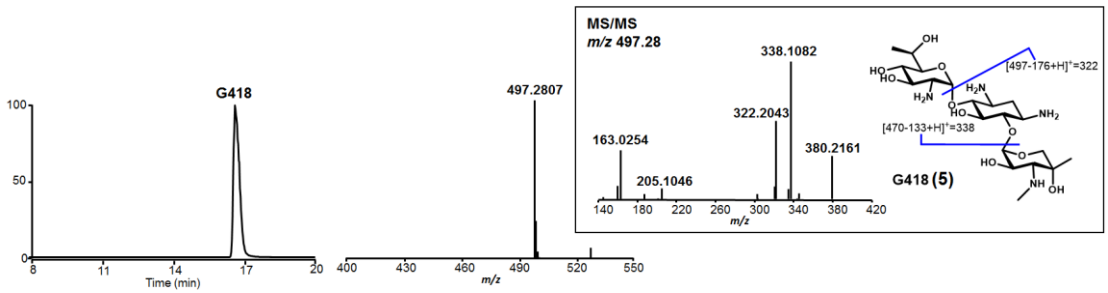
C. LC-ESI-HRMS and MS/MS Analysis of A-1 (**6**)



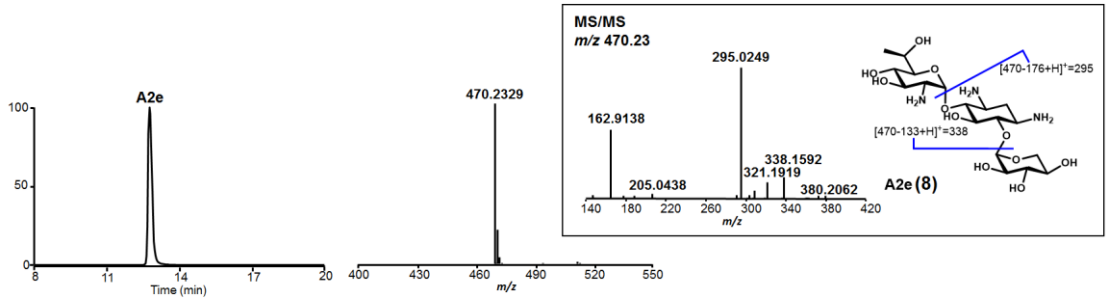
D. LC-ESI-HRMS and MS/MS Analysis of A-2 (**7**)



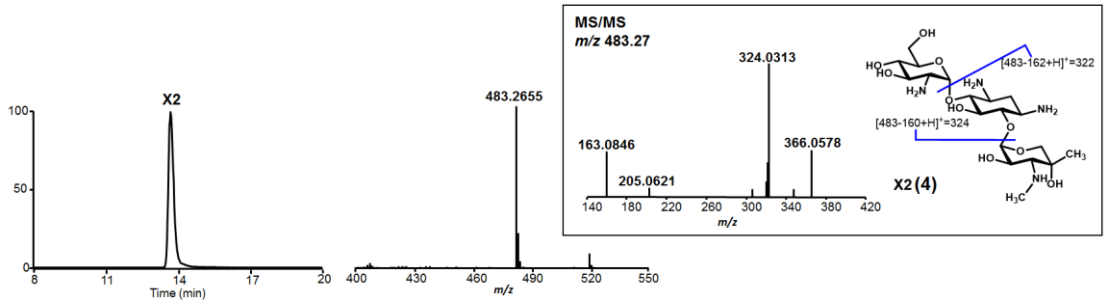
E. LC-ESI-HRMS and MS/MS Analysis of G418 (**5**)



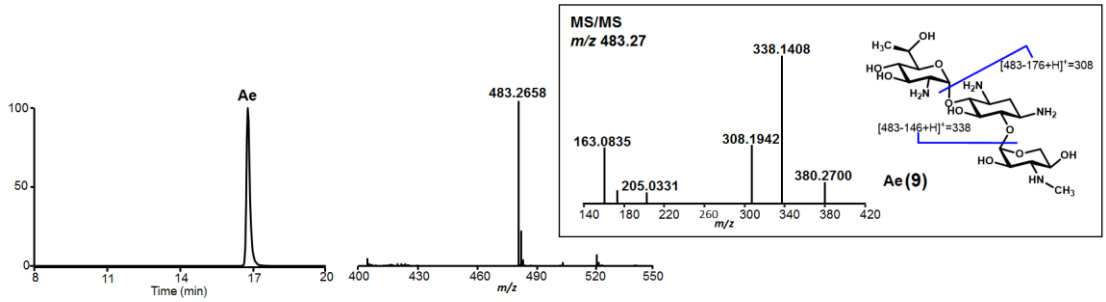
F. LC-ESI-HRMS and MS/MS Analysis of A2e (**8**)



G. LC-ESI-HRMS and MS/MS Analysis of X2 (**4**)



H. LC-ESI-HRMS and MS/MS Analysis of Ae (**9**)



I. LC-ESI-HRMS and MS/MS Analysis of X2-1(**10**) and X2-1-i

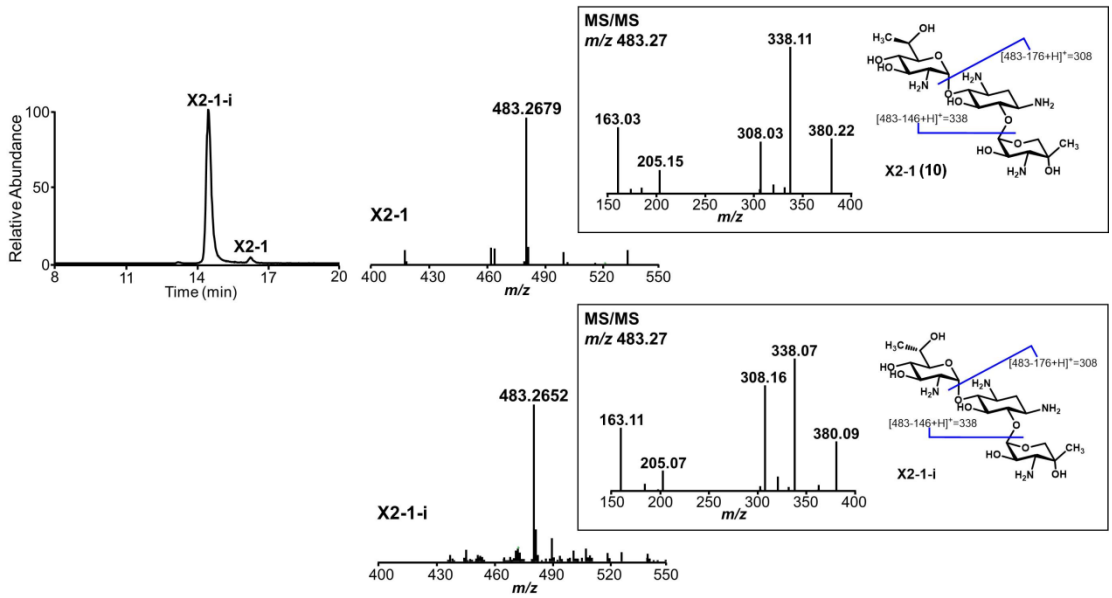
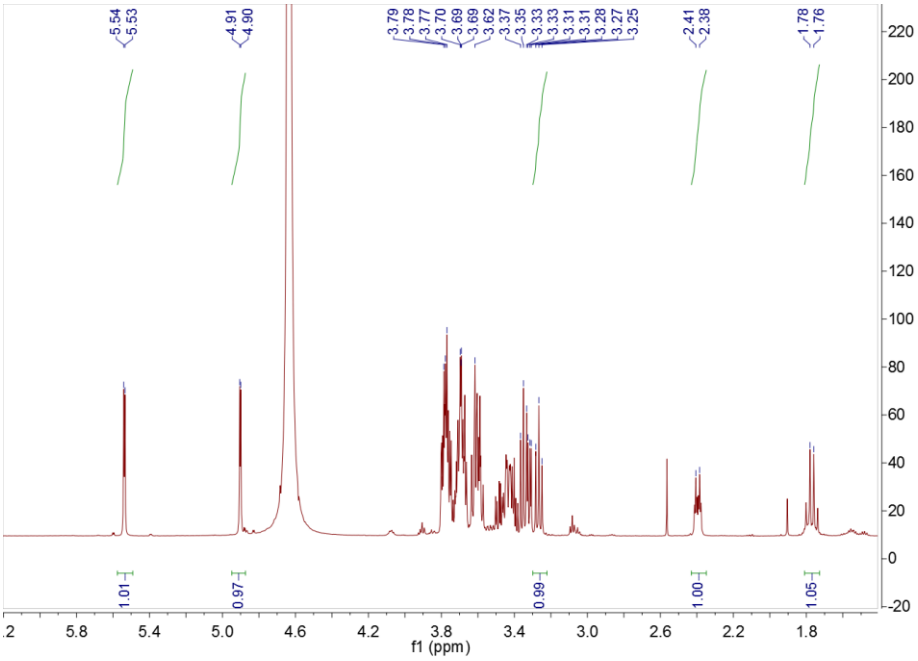


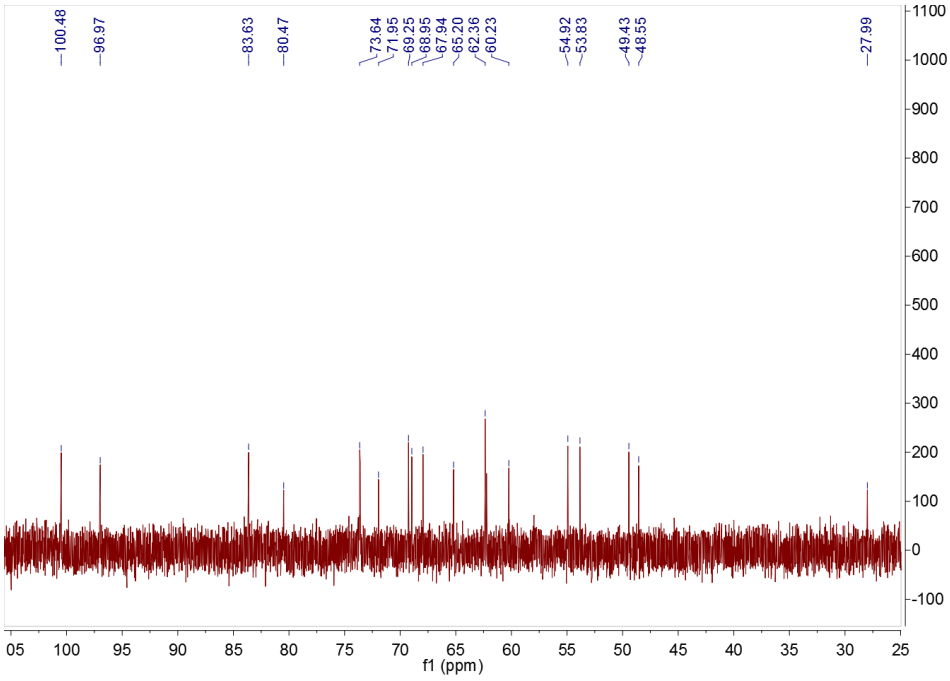
Figure S3. NMR Analysis of Gentamicin Biosynthetic Intermediates DAA2, A-2 and Ae (Related to Figure 1 and 2)

A. NMR spectra of DAA2 (2)

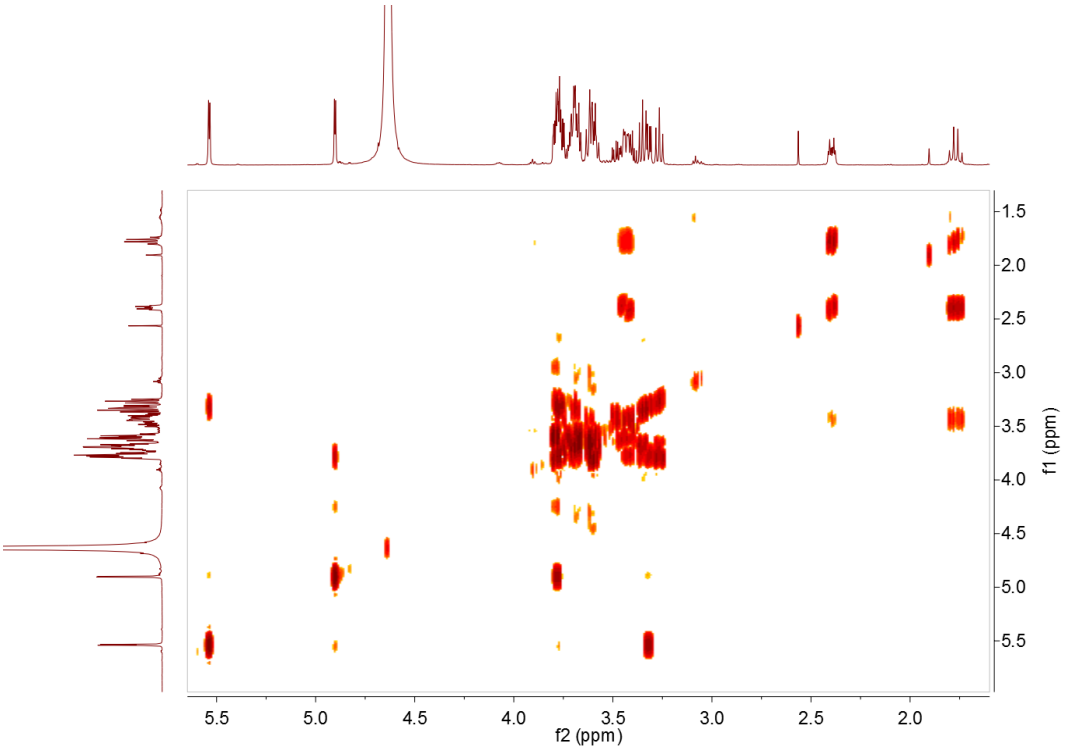
i. ^{1}H-NMR Spectrum of DAA2 (**2**)



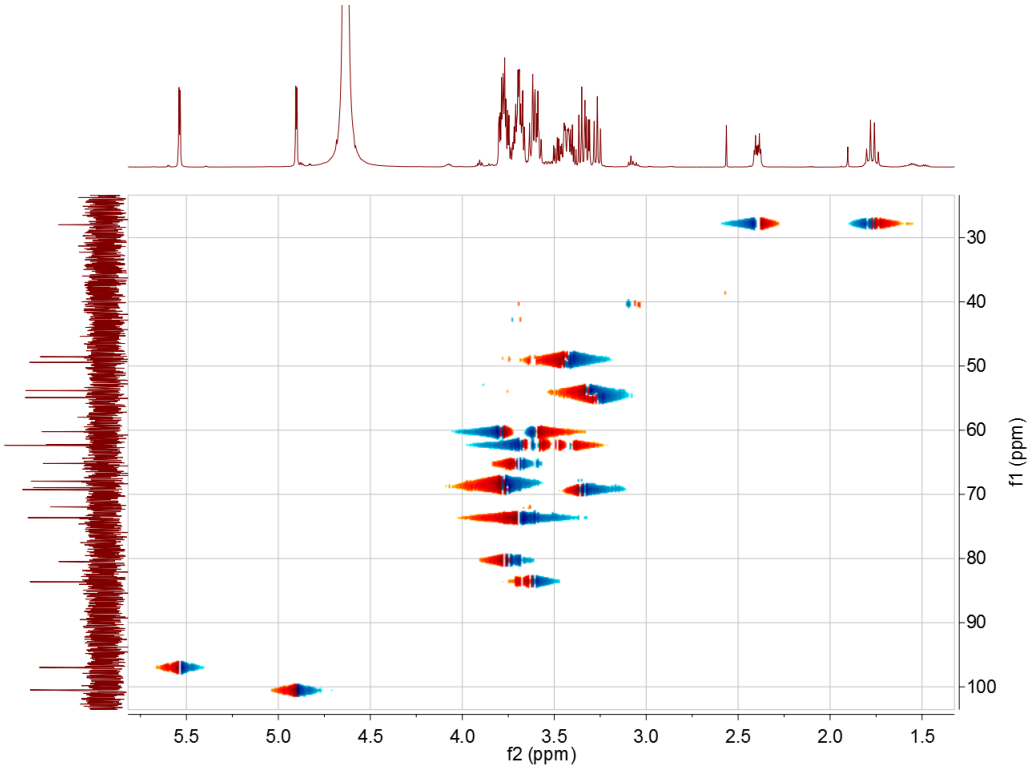
ii. ^{13}C-NMR Spectrum of DAA2 (**2**)



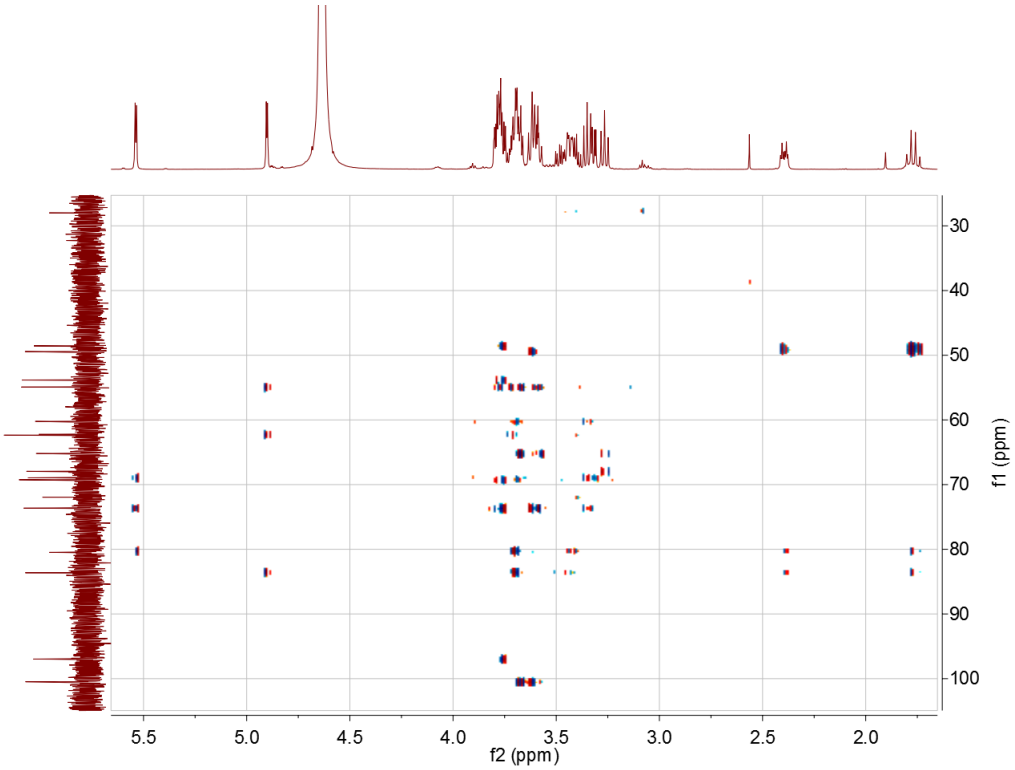
iii. ^{1}H-^{1}H COSY Spectrum of DAA2 (**2**)



iv. HSQC Spectrum of DAA2 (**2**)

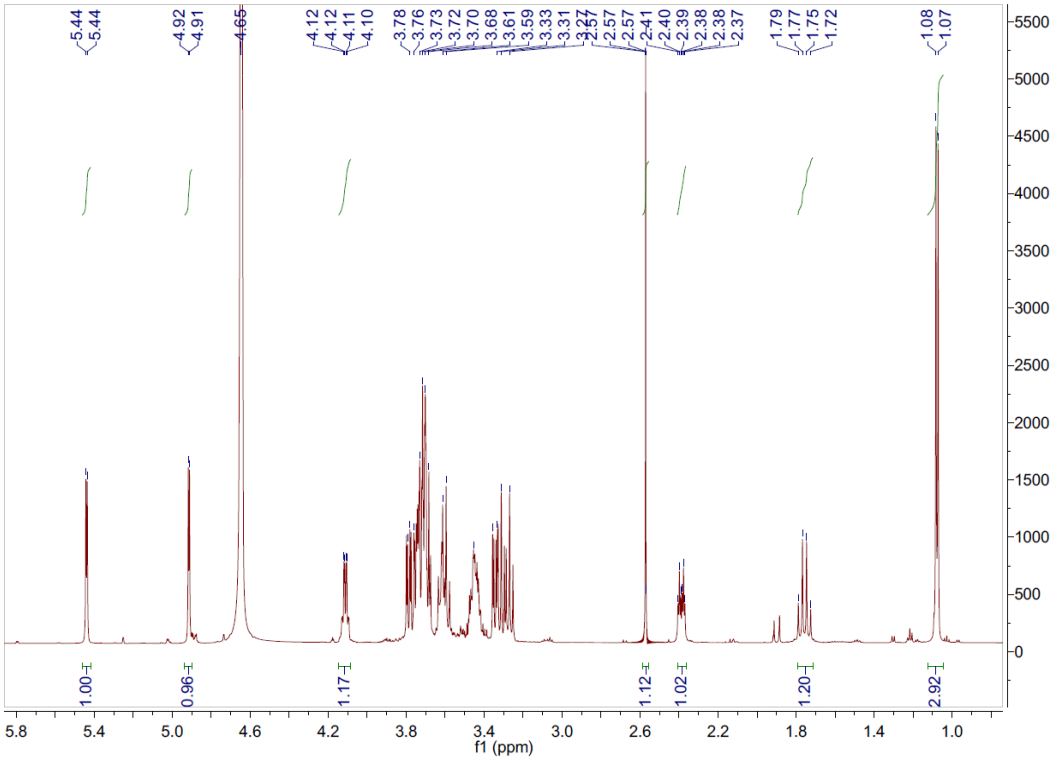


v. HMBC Spectrum of DAA2 (**2**)

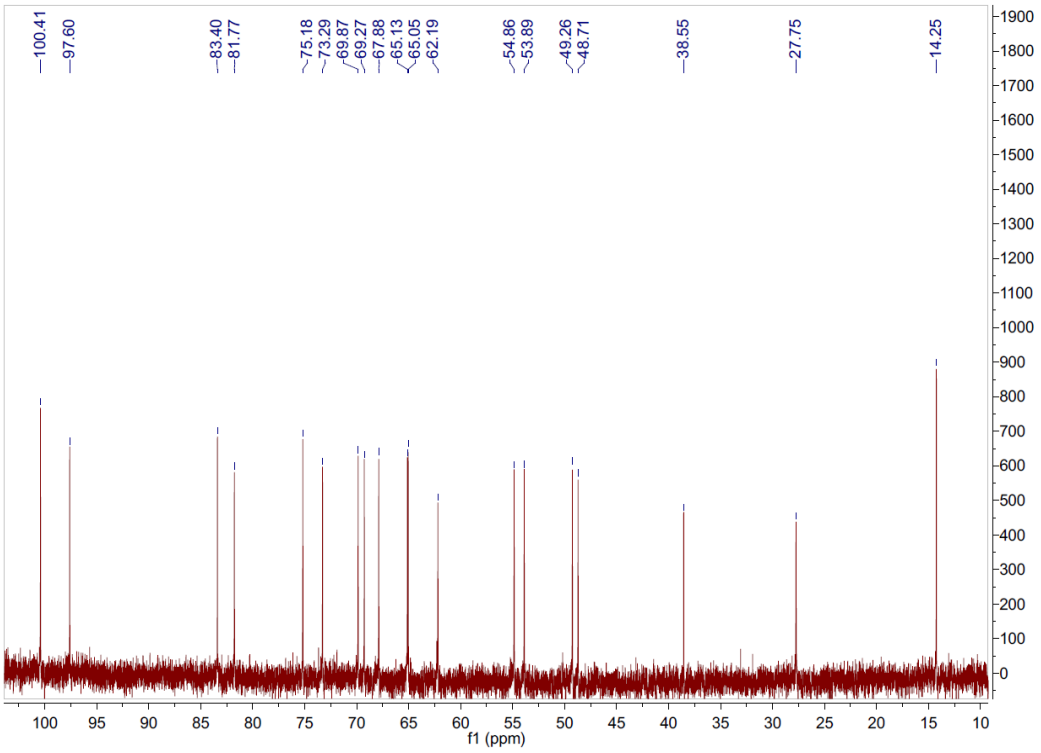


B. NMR spectra of A-2 (7)

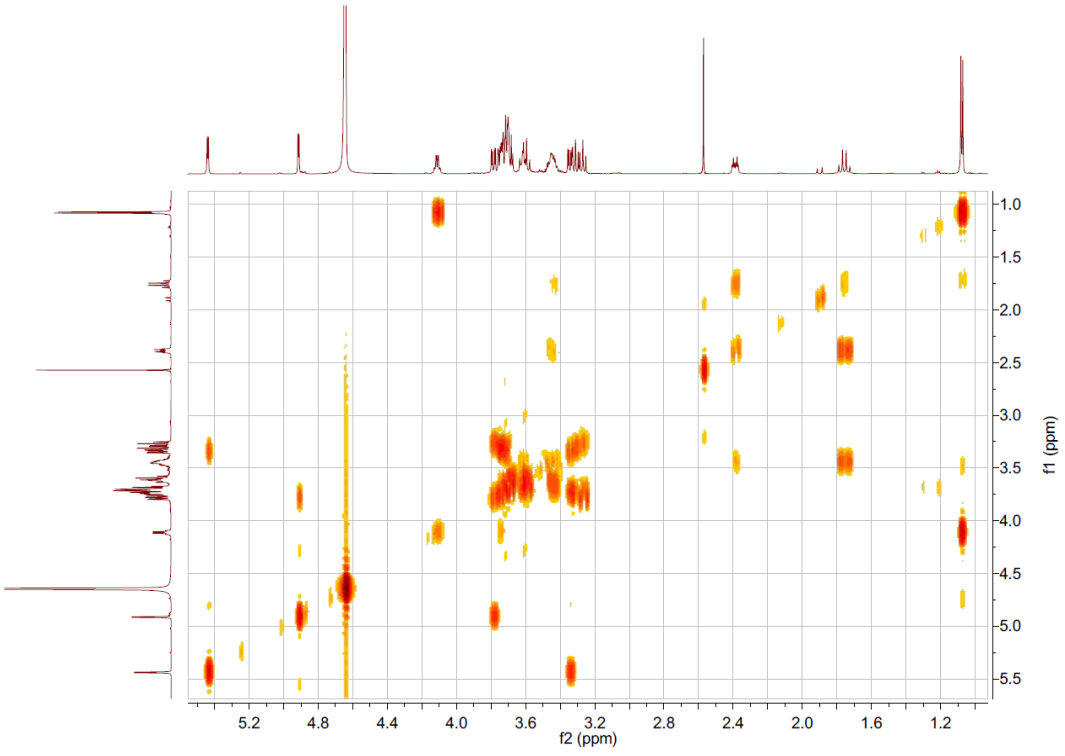
i. ^{1}H-NMR Spectrum of A-2 (**7**)



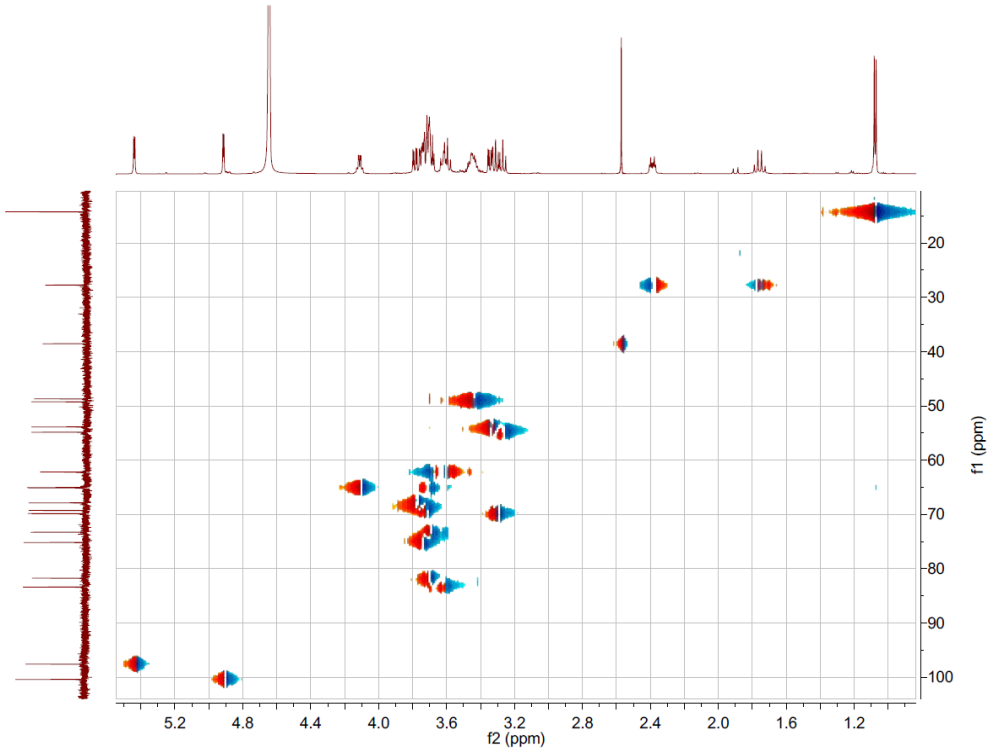
ii. ^{13}C-NMR Spectrum of A-2 (**7**)



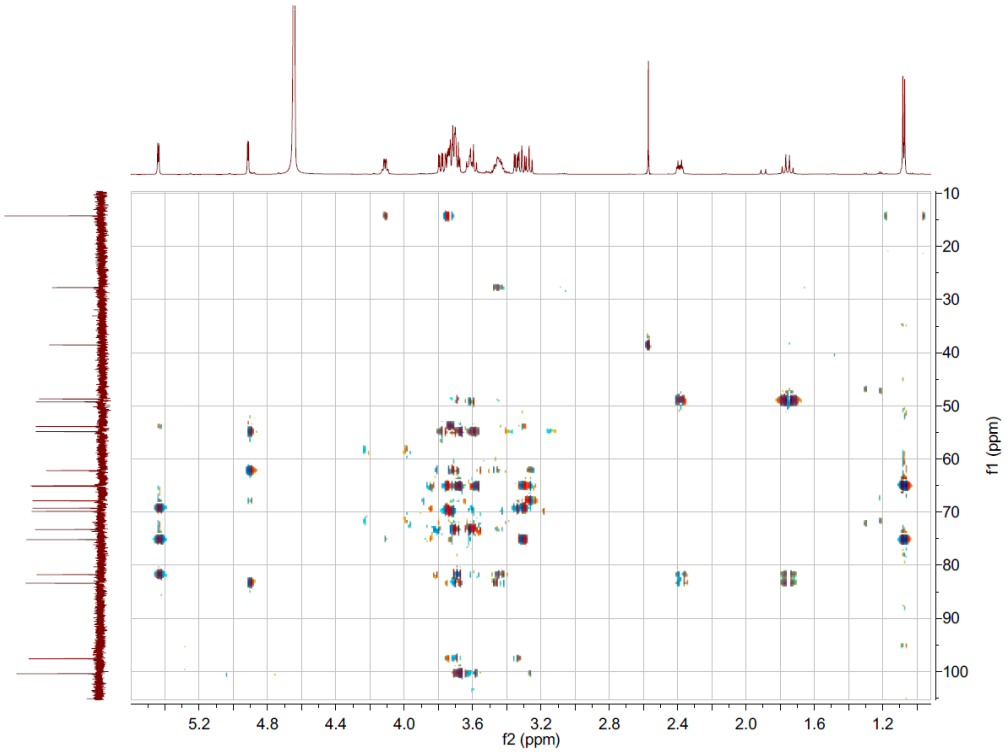
iii. ^{1}H-^{1}H COSY Spectrum of A-2 (**7**)



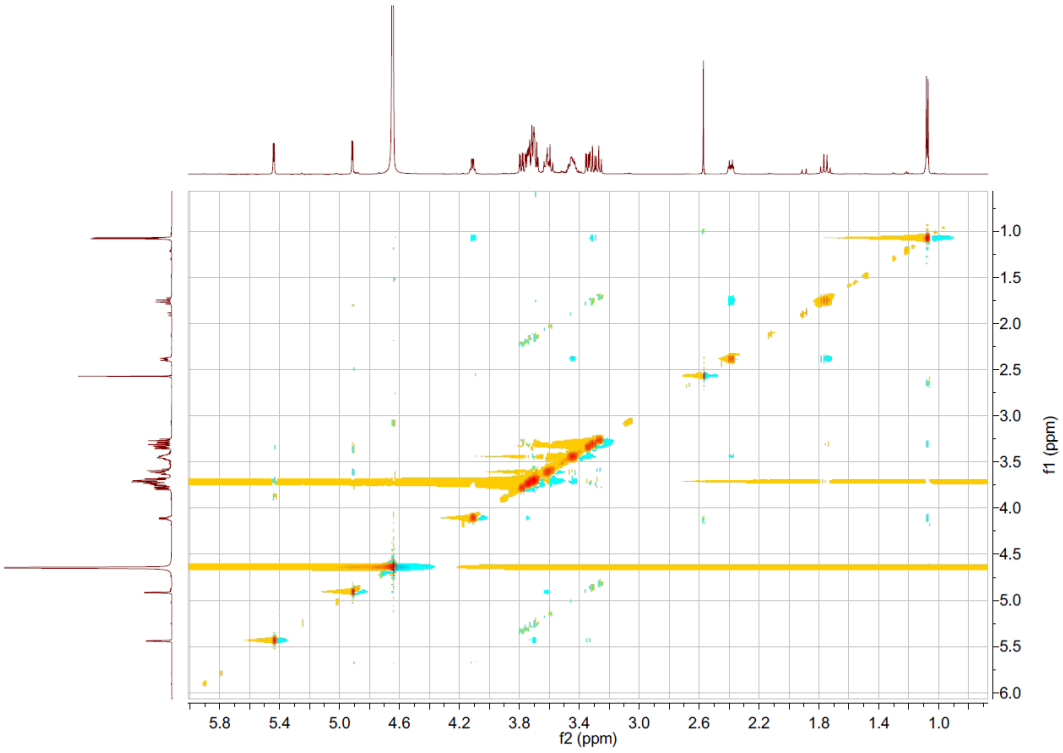
iv. HSQC Spectrum of A-2 (**7**)



v. HMBC Spectrum of A-2 (**7**)

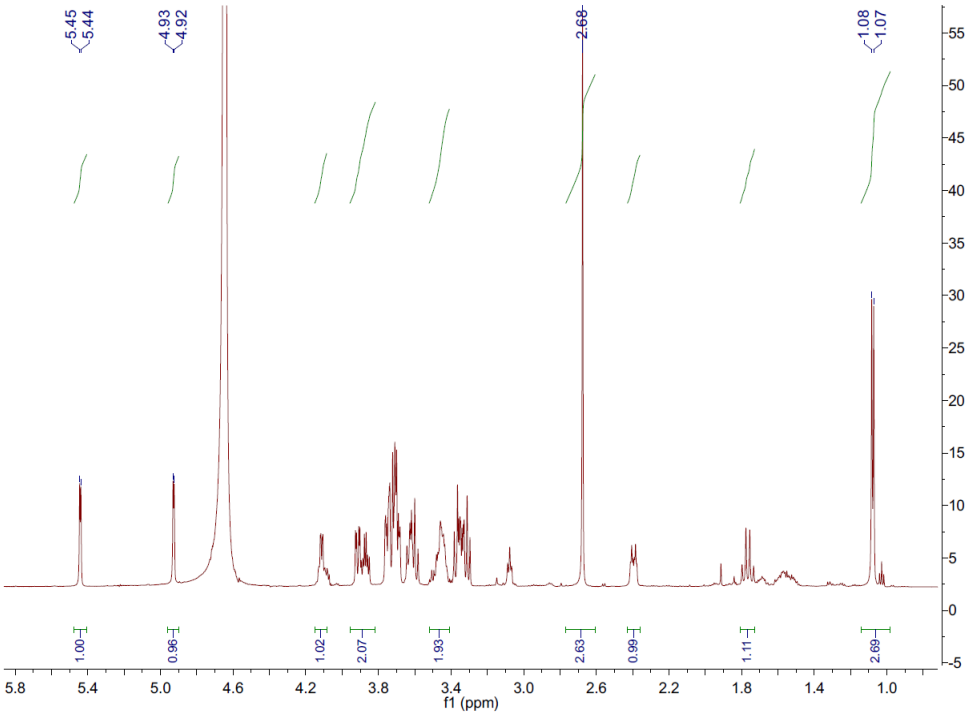


vi. NOESY Spectrum of A-2 (**7**)

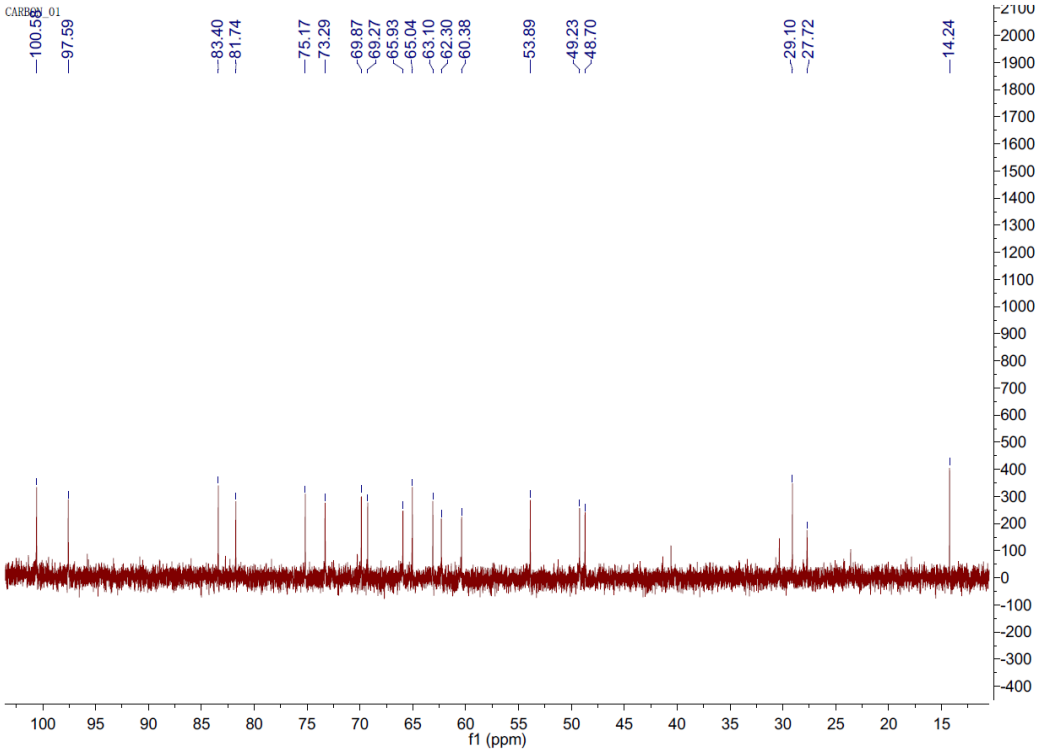


C. NMR spectra of Ae (**9**)

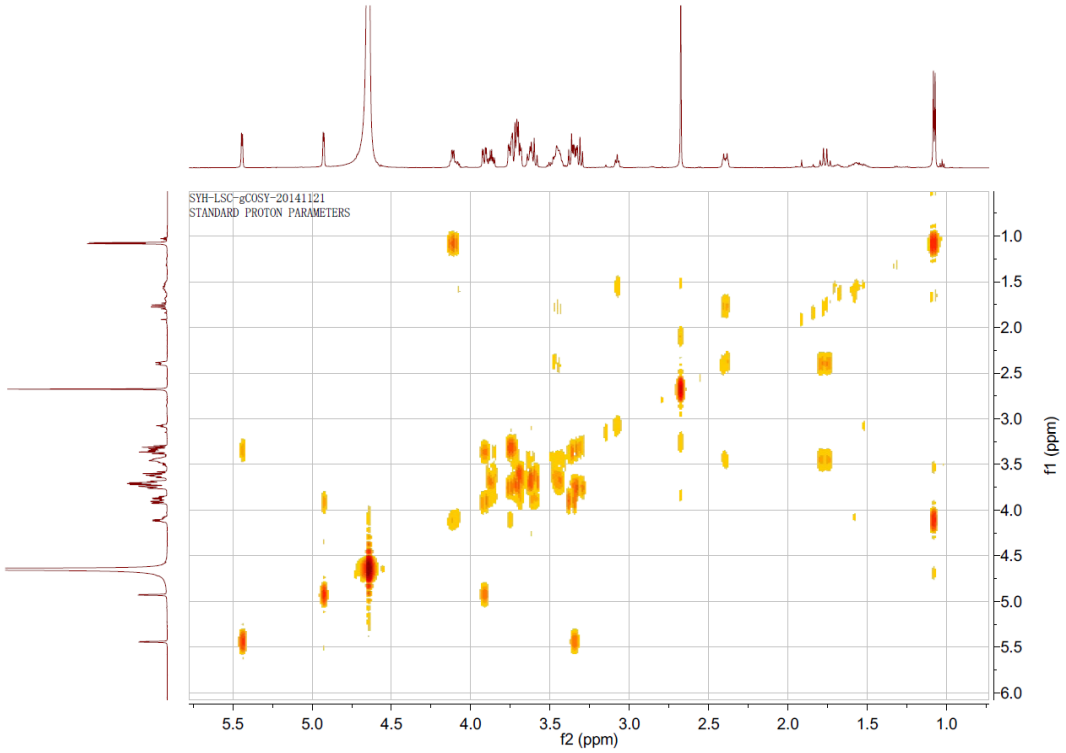
i. ^{1}H-NMR Spectrum of Ae (**9**)



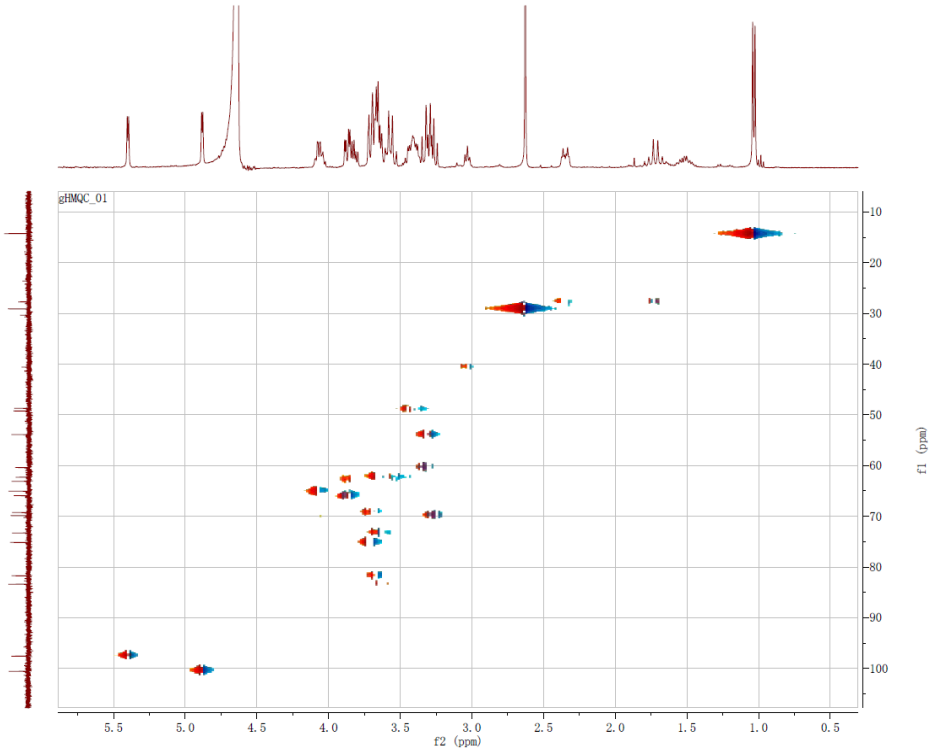
ii. ^{13}C-NMR Spectrum of Ae (**9**)



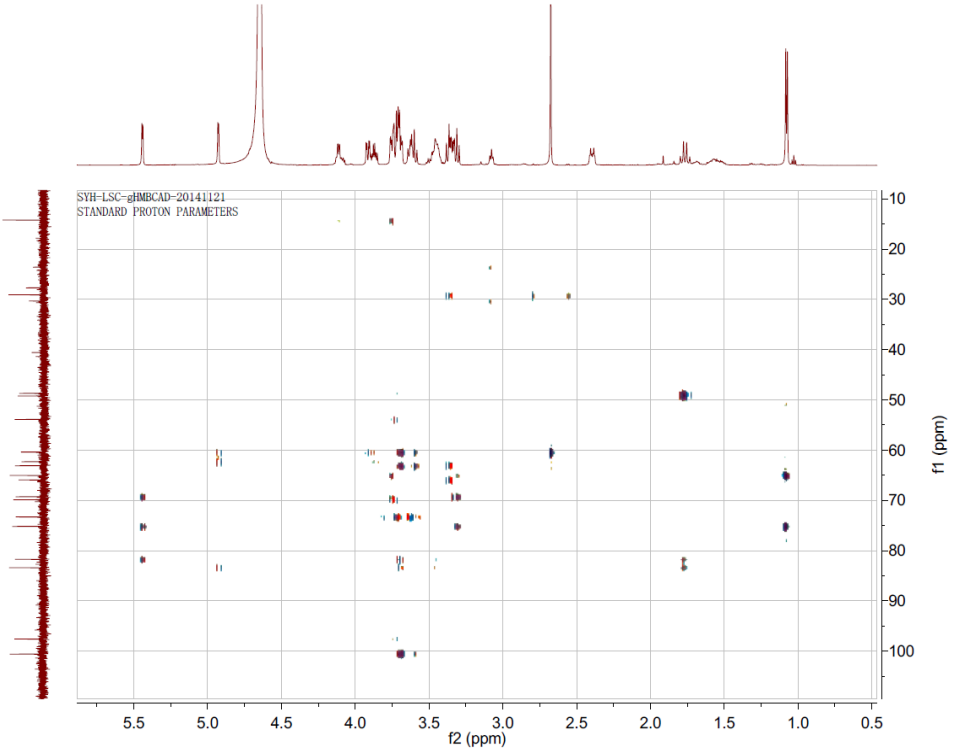
iii. ^{1}H-^{1}H COSY Spectrum of Ae (**9**)



iv. HSQC Spectrum of Ae (**9**)



v. HMBC Spectrum of Ae (**9**)



vi. NOESY Spectrum of Ae (**9**)

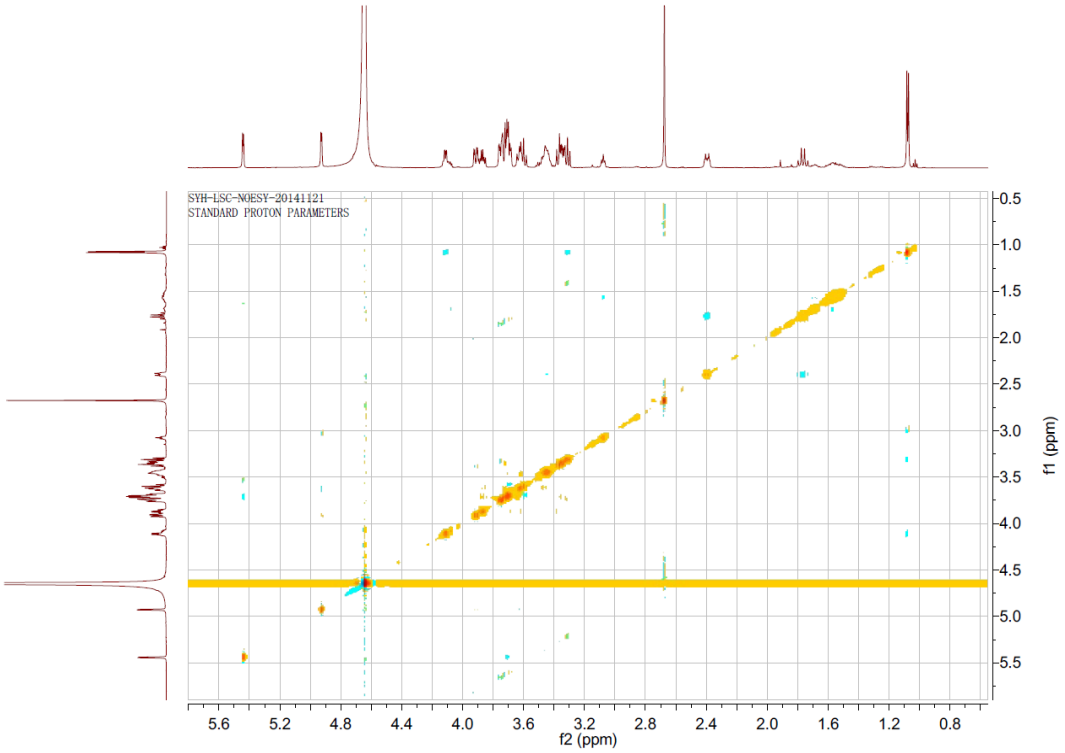
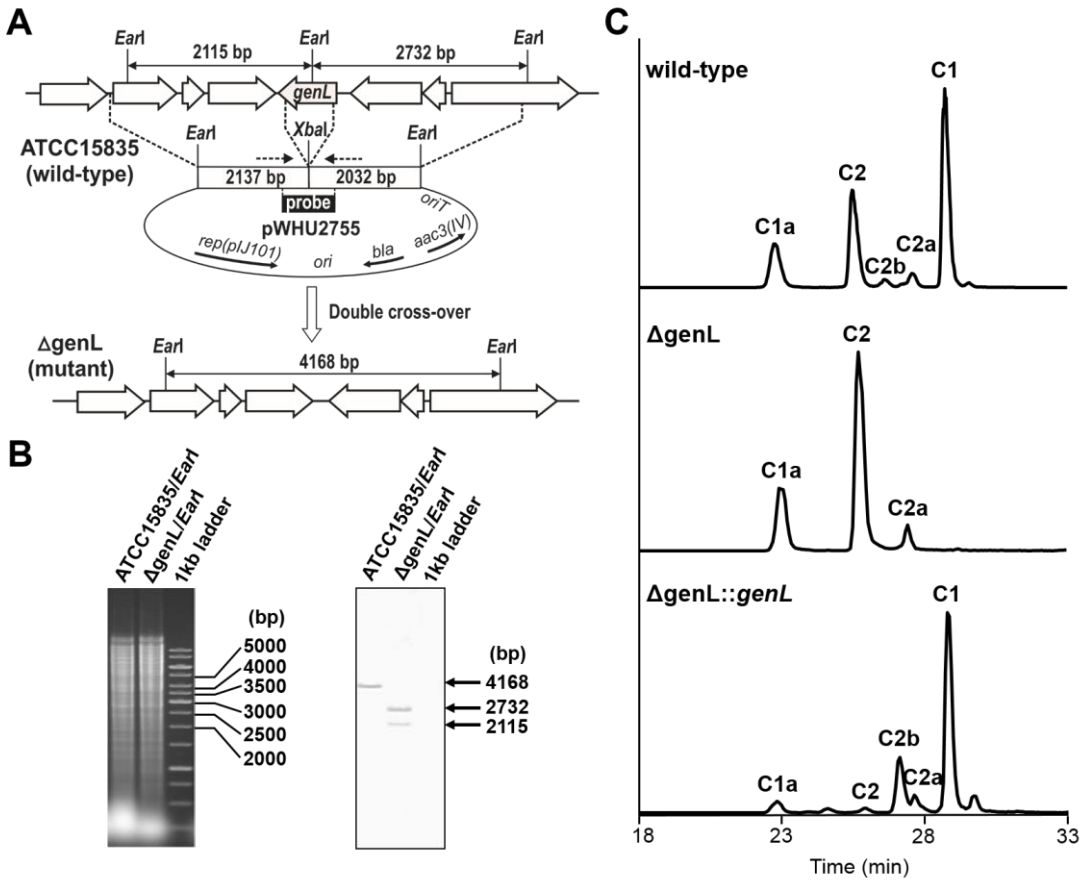
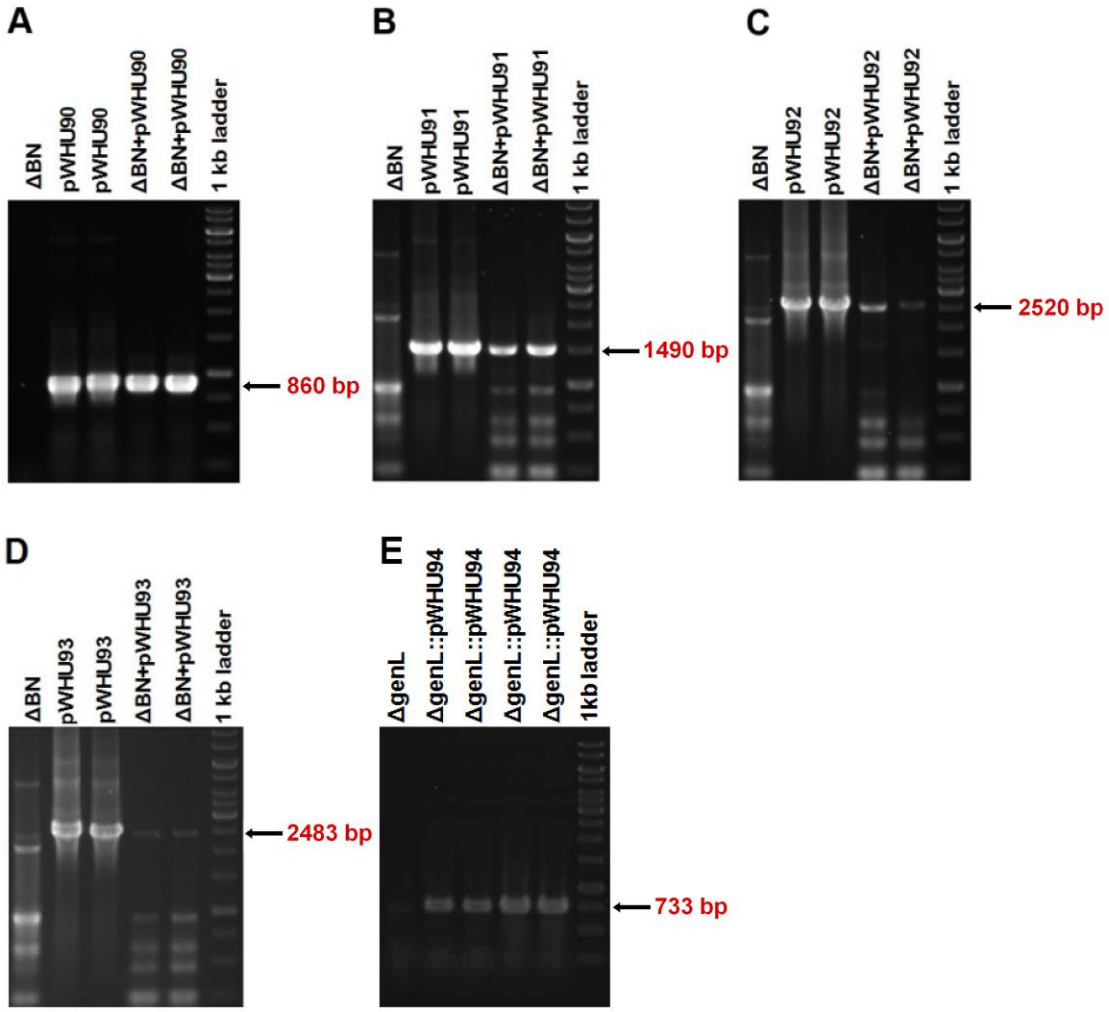


Figure S4. Genetic Confirmation of ΔgenL by Southern Blot and Analysis by LC-ESI-HRMS



(A) Schematic representation of the in-frame deletion of *genL*. (B) Confirmation of ΔgenL by Southern blot. The probe for Southern blot was obtained by PCR-amplified from pWHU2755 using primers *genL*-CP1 and *genL*-CP2. (C) Extracted ion chromatogram of gentamicin C1a, C2, C2a, C2b and C1 on LC-ESI-HRMS of wild-type, ΔgenL and ΔgenL::*genL*.

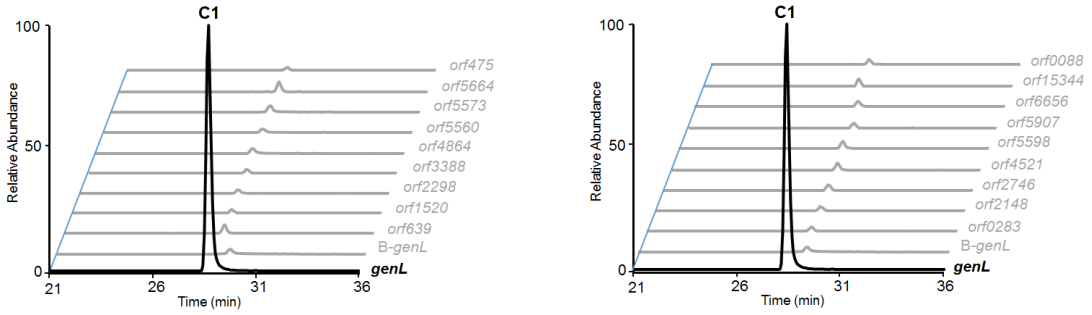
Figure S5. Confirmation by PCR of Single Methyltransferase Gene Complementation of ΔBN and ΔgenL (Related to Figure 3)



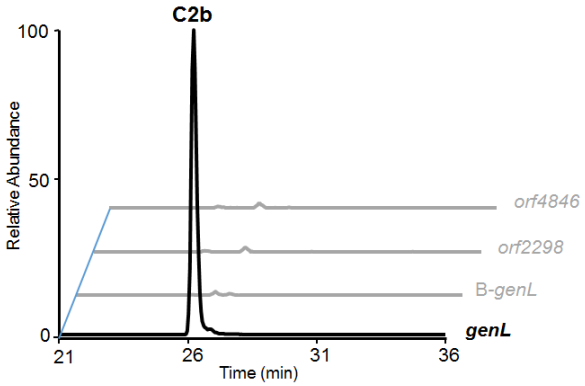
(A) ΔBN::*gmrA*; (B) ΔBN::*gmrA+genN*; (C) ΔBN::*gmrA+genD1*; (D) ΔBN::*gmrA+genK*; (E) ΔgenL::*genL*

Figure S6. LC-ESI-HRMS and MS/MS Analysis of Cell-free Assays (Related to Figure 5)

A. Cell-free Assay with C2 as Substrate

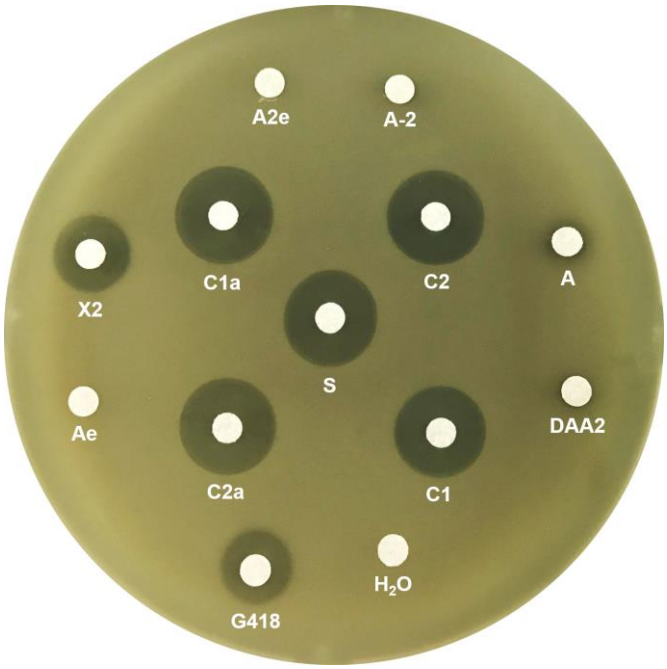


B. Cell-free Assay with C1a as Substrate



(A) Gentamicin C2 as substrate; (B) Gentamicin C1a as substrate. The black or gray lines indicate the signal of C1 (in A) or C2b (in B) in LC-ESI-HRMS detection of cell-free assays with BL21 (DE3) containing each candidate gene inserted in pET28a(+). B-*genL* is a control, in which the cell lysate of BL21(DE3)/pET28a(+)-*genL* was inactivated at 95℃ for 10 min.

Figure S7. Antimicrobial Activity of Gentamicin-related Compounds by Agar Diffusion Assay.



The indicator strains used is *Bacillus pumilus*. 2.5 μg compound was added to each filter paper. Gentamicin C complex (S) and H_2O were used as positive and negative control, respectively.