ChemistryOPEN

Supporting Information

Genome Mining-Mediated Discovery of a New Avermipeptin Analogue in *Streptomyces actuosus* ATCC 25421

Weiying Liu,[a] Fengxian Sun,[b] and Yang Hu*[a]

open_201800130_sm_miscellaneous_information.pdf

Table of Contents

Table S1 Features of *S. actuosus* ATCC 25421 chromosome.

Component of chromosome	Property
Genome size	8,145,579 bp
GC content	72.53%
Gene Average Length (bp)	958
Number of Contigs	1
rRNAs	18
tRNAs	70
sRNAs	16
Protein-coding genes	7,536

Table S2 Secondary metabolites in *S. actuosus* ATCC 25421.

Cluster	Type	From	To	Most similar known biosynthetic gene cluster (percent of similarity, %)	MIBiG BGC-ID
1	PKS-NRPS	35338	113331	Antimycin (86)	BGC0000958_c1
2	β-lactam	117855	139684	Spinosad (8)	BGC0000148_c1
3	Other	158144	182801	Herboxidiene (2)	BGC0001065_c1
4	Other	373066	398132	Nocathiacin (4)	BGC0000609_c1
5	Melanin	608692	619063	Melanin (71)	BGC0000908_c1
6	Terpene	683296	708900	Carotenoid (63)	BGC0000633_c1
7	Other	963361	977997	Macrotetrolide (33)	BGC0000244_c1
8	Other	1188024	1210424	Kirromycin (3)	BGC0001070_c1
9	Saccharide	1640023	1661884	-	-
10	Other	1755360	1768611	Calicheamicin (2)	BGC0000033_c1
11	Ectoine	1826893	1837297	Ectoine (100)	BGC0000853_c1
12	Other	2208379	2242570	Nataxazole (7)	BGC0001213_c1
13	Indole	2457525	2478628	Lactonamycin (5)	BGC0000238_c1
14	PKS	2483340	2536604	Rabelomycin (41)	BGC0000262_c1
15	Fatty acid	2597423	2618424	Tetracenomycin (16)	BGC0000275_c1
16	Other	2691170	2706906	Salinilactam (8)	BGC0000142_c1
17	Saccharide	2820518	2845659	-	-
18	Saccharide-Melanin-Butyrolactone-fatty acid-PKS	2963331	3053214	RK-682 (45)	BGC0000140_c1
19	Siderophore	3115427	3127199	Desferrioxamine B (100)	BGC0000940_c1
20	Other	3481224	3498212	Echosides (11)	BGC0000340_c1
21	Other	3850963	3859392	Streptolydigin (5)	BGC0001046_c1
22	Fatty acid-Butyrolactone	4185302	4230096	Simocyclinone (6)	BGC0001072_c1

23	PKS	4337231	4379746	Spore pigment (83)	BGC0000271_c1
24	Arylpolyene	4436423	4477583	Lomaiviticin (3)	BGC0000240_c1
25	Other	4845352	4863025	Coelimycin (8)	BGC0000038_c1
26	Thiopeptide	4903675	4941503	Nosiheptide (80)	BGC0000610_c1
27	NRPS	5308792	5393699	Telomycin (17)	BGC0001406_c1
28	Other	5439042	5458140	Meilingmycin (2)	BGC0000093_c1
29	Terpene	5562219	5583316	Albaflavenone (100)	BGC0000660_c1
30	Fatty acid	5671893	5693677	Fluostatin (3)	BGC0000223_c1
31	Other	6376197	6386714	JBIR-34, JBIR-35 (12)	BGC0000376_c1
32	Bacteriocin	6437806	6449149	-	-
33	Terpene-Butyrolactone	6459628	6486085	γ-butyrolactone (100)	BGC0000850_c1
34	Saccharide	6593964	6628360	-	-
35	Bacteriocin-PKS-Siderophore	6635180	6742967	A-503083 (7)	BGC0000288_c1
36	Other	6789738	6806927	Kanamycin (20)	BGC0000704_c1
37	Other	6807005	6817184	Kanamycin (3)	BGC0000703_c1
38	Other	6817191	6823998	Kanamycin (1)	BGC0000703_c1
39	Fatty acid	6872844	6893791	-	-
40	Other	6934432	6950718	Kinamycin (5)	BGC0000236_c1
41	Other	7001939	7008198	GE81112 (7)	BGC0000360_c1
42	Terpene	7054188	7080941	Hopene (92)	BGC0000663_c1
43	Saccharide	7120847	7155983	-	-
44	Other	7230369	7241721	Lysolipin (4)	BGC0000242_c1

45	Other	7584174	7592752	Oxazolomycin (6)	BGC0001106_c1
46	Lantipeptide	7634678	7664630	Informatipeptin (85)	BGC0000518_c1
47	Lassopeptide	7907890	7930577	-	-
48	NRPS	7974127	8021481	Leinamycin (2)	BGC0001101_c1
49	Other	8071675	8077451	9-methylstreptimidone (9)	BGC0000171_c1

Table S3 Deduced functions of ORFs in the actuospeptin biosynthetic gene cluster.

Gene	Size[a]	Protein homologue[b] and origin	Identity, %	Proposed function
*ave*A complement (7648041-7649948)	635	ABC transporter ATP-binding protein [Streptomyces sp. Ag82_G6-1] WP_097222466.1	72	transporter
*ave*B complement (7650014-7651747)	577	ABC transporter ATP-binding protein [Streptomyces sp. CdTB01] WP_058921347.1	78	transporter
*ave*S complement (7651849-7651965)	38	SapB/AmfS family lantipeptide [Streptomyces sp. 142MFCol3.1] WP_028798592.1	100	precursor peptide
*ave*T complement (7652048-7654630)	860	lantipeptide synthetase [Streptomyces sp. Ru71] WP_103782009.1	85	lantipeptide synthetase

[a]Numbers are in amino acids. [b]NCBI accession numbers are given in parentheses.

Table S4 Minimum inhibitory concentrations (MICs) of avermipeptin B.

Organism	MIC (μg/mL)
Staphylococcus aureus subsp. aureus ATCC43300	0.0016
Enterococcus faecalis ATCC29212	0.0032
Bacillus subtilis ATCC6633	0.0032

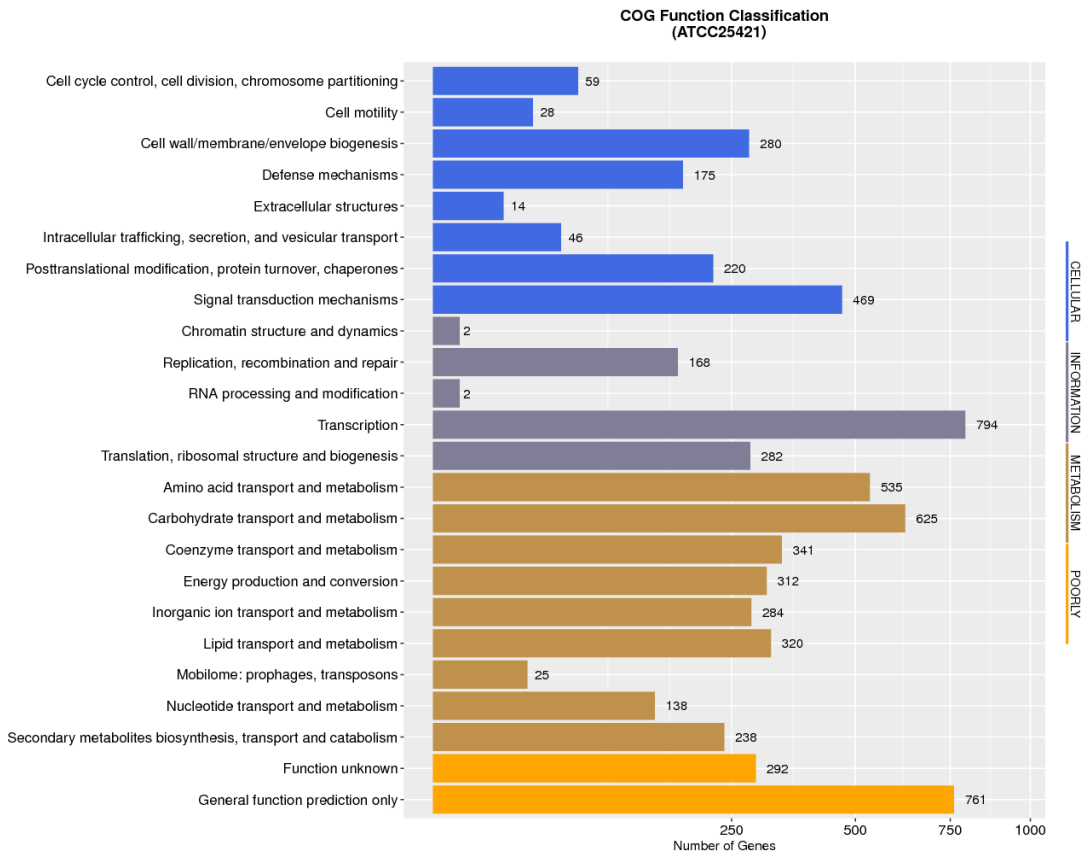

Figure S1. COG-based functional classification of genes located on *S. actuosus* ATCC 25421 genome. Genes might be classified into more than one category.

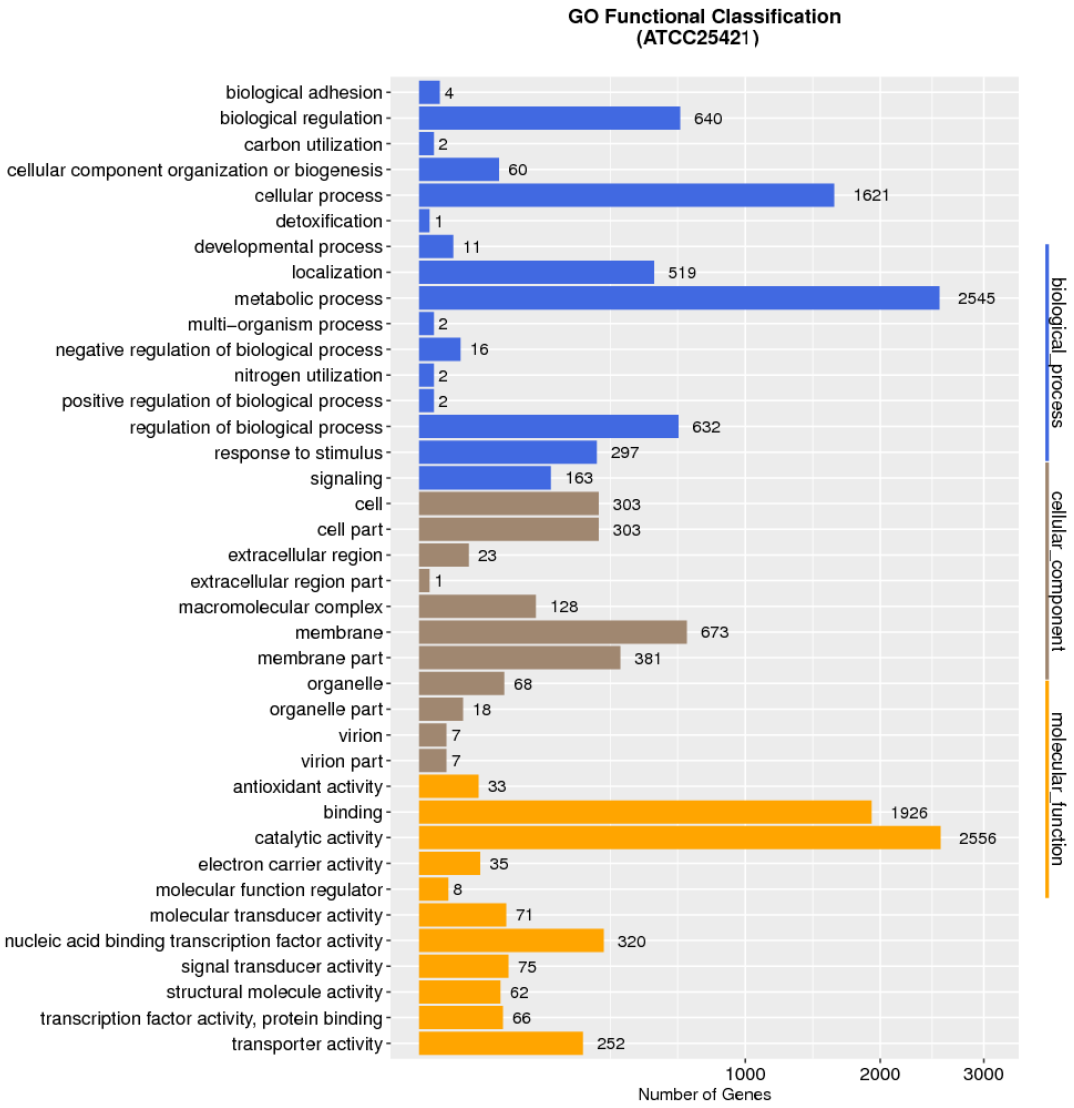

Figure S2. GO-based functional classification of genes located on *S. actuosus* ATCC 25421 genome. Genes might be classified into more than one category.

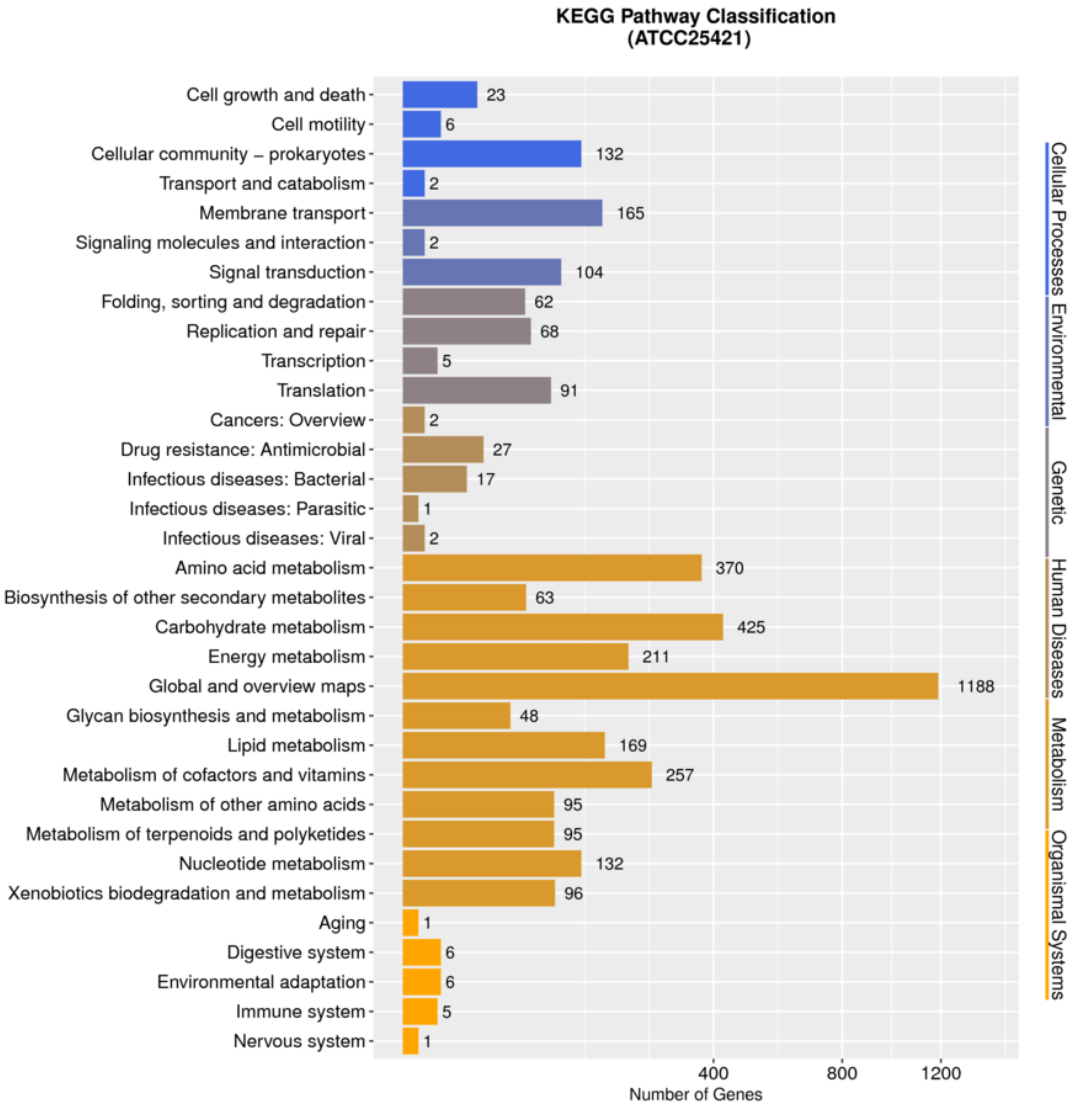

Figure S3. KEGG-based functional classification of genes located on *S. actuosus* ATCC 25421 genome. Genes might be classified into more than one category.

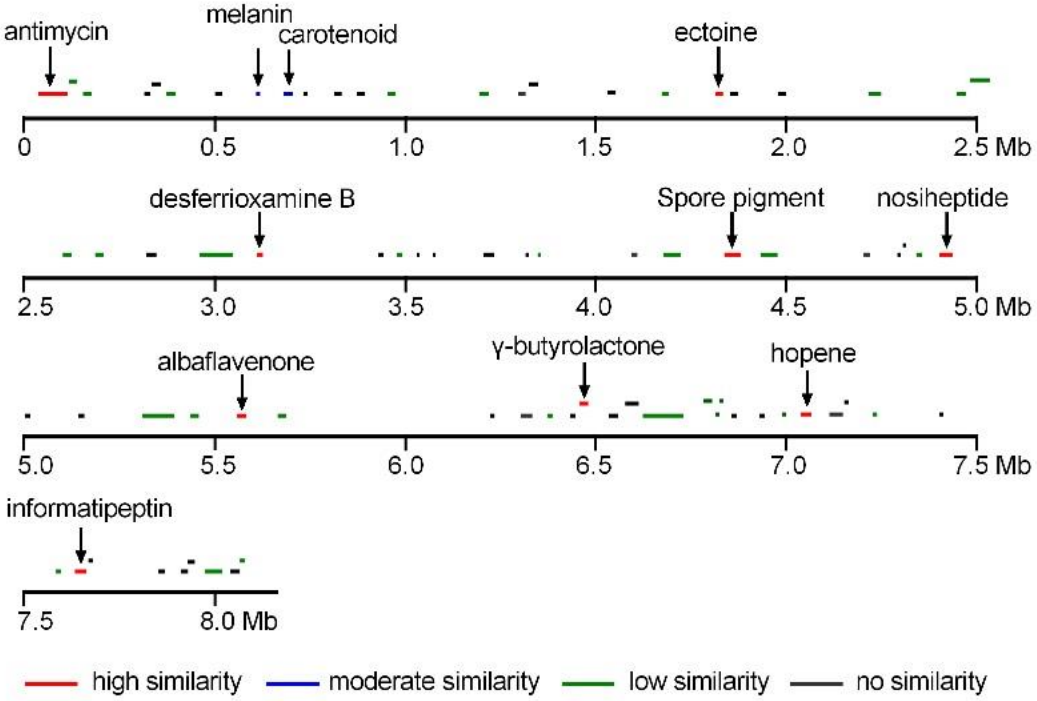

Figure S4. Distribution of secondary metabolite gene clusters in *S. actuosus* ATCC 25421. The degree of similarity with known gene clusters from other microorganisms was shown with different color of the bars.

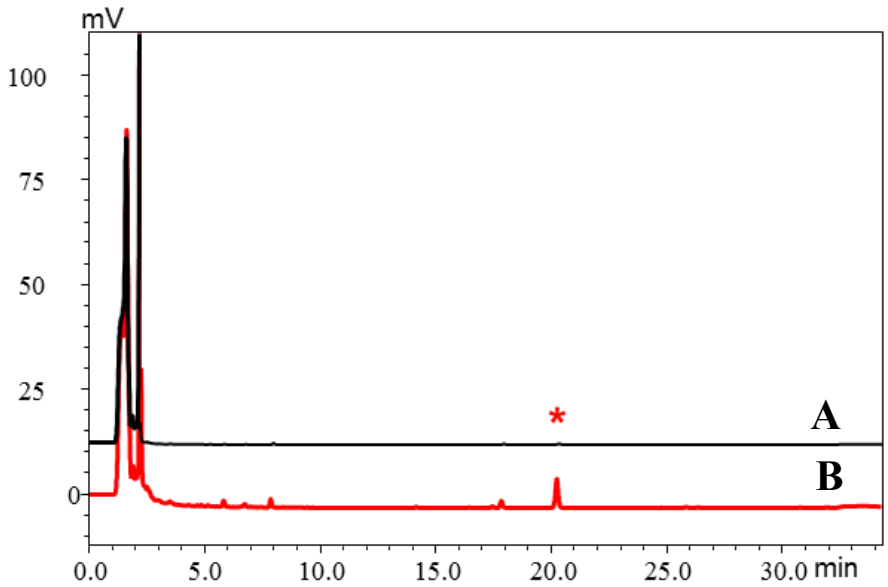

Figure S5. HPLC analysis of the extracts from (A) *S. lividans* TK24 and (B) *S. lividans* TK24/pSET152-ave.

Experimental Section

Genome sequencing and assembly

The *S. actuosus* ATCC 25421 genome was sequenced using a PacBio RS II platform and Illumina HiSeq 4000 platform at the Beijing Genomics Institute (BGI, Shenzhen, China). Four SMRT cells Zero-Mode Waveguide arrays of sequencing, were used by the PacBio platform to generate the subreads set. PacBio subreads (length < 1 kb) were removed. The program Pbdagcon (https://github.com/PacificBiosciences/pbdagcon) was used for selfcorrection. Draft genomic unitigs, which are uncontested groups of fragments, were assembled using the Celera Assembler against a highquality corrected circular consensus sequence subreads set. To improve the accuracy of the genome sequences, GATK (https://www.broadinstitute.org/gatk/) and SOAP tool packages (SOAP2, SOAPsnp, SOAPindel) were used to make single-base corrections.To trace the presence of any plasmid, the filtered Illumina reads were mapped using SOAP to the bacterial plasmid database (http://www.ebi.ac.uk/genomes/plasmid.html, last accessed July 8, 2016).

Genome component prediction

Gene prediction was performed on the *S. actuosus* ATCC 25421 genome assembly by glimmer3 (http://www.cbcb.umd.edu/software/glimmer/) with Hidden Markov models. tRNA, rRNA and sRNAs recognition made use of tRNAscan-SE (Lowe and Eddy, 1997).

Gene annotation and protein classification

The best hit abstracted using Blast alignment tool for function annotation. Seven databases which are KEGG (Kyoto Encyclopedia of Genes and Genomes), COG (Clusters of Orthologous Groups), NR (Non-Redundant Protein Database databases), Swiss-Prot, and GO (Gene Ontology), TrEMBL,

EggNOG are used for general function annotation.

Cloning of the *act* gene cluster into the expression vector pSET152

PCR method was used for amplification of the eryth gene cluster from the genome of *S. actuosus* ATCC 25421 by using the primers P1/P2 (P1: 5'-CATATGCCCCGGACG GTCACACGCGTCAACAATGGG-3'; P2; 5-TCTAGAGGCGTCGGCGAGGCGTT CCAGACTGGCGCG-3'). Then, the recovered fragments were cloned to pMD-18T simple vector and confirmed by DNA sequenceing. The recombinant plasmid was digested with the restriction enzymes *Nde*I and *Xba*I, and then ligated into the vector pSET152 under the control of a promoter *PermE** to give a recombinant plasmid pSET152-ave. The recombinant plasmid pSET152-ave was transformed in competent *E. coli* DH5a cells for amplification and verified by restriction enzymes *Nde*I and *Xba*I and DNA sequencing.

***E. coli-Streptomyces* conjugation**

The recombinant plasmid pSET152-act was transformed by electroporation into *E. coli* ET12567/pUZ8002 according to a standard protocol. The pSET152-ave was introduced into *S. lividans* TK24 by *E. coli-Streptomyces* conjugation as described previously[1].

Production and Isolation of type III lantibiotics

S. lividans TK24 and *S. lividans* TK24/pSET152-ave were cultivated on ISP2 agar plates (4 g yeast extracts, 10 g malt extract, 4 g D -glucose, 18 g agar, and 1000 ml water) containing apramycin (25 mg/ml) at 28℃. The plates were incubated for 10 d at 28 °C. The agar ~ 1 cm^2 from each plate was extracted with acetonitrile (500 µl) and centrifugated (14000 rpm, 5 min), the solution was collected for further experiments.

HPLC-ESI-MS/MS analysis of microbial extracts

Lantibiotic-containing samples were analyzed by HPLC-ESI-MS/MS experiments according to the methods described previously[2].

Isolation and extraction of avermipeptin B

Isolation and extraction of avermipeptin B from *S. lividans* TK24/pSET152-ave were performed according to the methods described previously [2].

Assay of antibacterial activity

Avermipeptin B was dissolved in DMSO to produce a stock solution (1000 μg/mL), which was serially diluted into Mueller-Hinton broth (50 μL) in a 96-well microtiter plate. For the test of *Staphylococcus aureus*, *Bacillus subtilis* and *Enterococcus faecali*, the final concentration was ranging from 0.0256 to 0 μg /mL. After the compounds were diluted, 50 μl aliquots of the test strains (about 5×10^{-5} CFU/mL) were added to the appropriate wells of the microtiter plate. After incubation at 37℃ 150 rpm for 18-24 h, the MIC was determined as the lowest concentration of compound that inhibits visible bacterial growth. All assays were repeated for 3 times.

References

[1] T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater, D. A. Hopwood, **2000**, *3*.

[2] G. H. Völler, J. M. Krawczyk, A. Pesic, B. Krawczyk, J. Nachtigall, R. D. Süssmuth, **2012**, *13*, 1174-1183.