

Supporting Information

Molecular Analysis of the YM-216391 Biosynthetic Gene Cluster and Improvement of the Cyclopeptide Production in Heterologous Host

Xiao-Hong Jian,[†] Hai-Xue Pan,[†] Ting-Ting Ning,[†] Yuan-Yuan Shi,[†] Yong-Sheng Chen,[†] Yan Li,[†] Xiao-Wei Zeng,[‡] Jian Xu,[‡] and Gong-Li Tang^{†,*}

[†]State Key Laboratory of Bio-organic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 345 Lingling Road, Shanghai 200032, China

[‡]Bioenergy Genome Center and Shandong Key Laboratory of Energy Genetics, Qingdao Institute of BioEnergy and BioProcess Technology, Chinese Academy of Sciences, 189 Songling Rd., Qingdao, Shandong 266101, China.

*Correspondence: Gong-Li Tang, Email: gltang@sioc.ac.cn; Tel: 086-21-54925113, Fax: 086-21-64166128.

Materials and Methods

General. A YM-216391 producer, *Streptomyces nobilis* JCM 4274, was purchased from the RIKEN BioResource Center, Japan. *E. coli* BW25113/pIJ790 and *E. coli* 5a /BT340 were used for gene knockout according the ReDirect technology. The pJTU2125 contains the *aadA* cassette used to replace the gene using PCR targeting and λ-Red-mediated recombination. The resulting mutagenised was transformed into *E. coli* S17-1 for conjugation with *S. lividans* 1326. The engineered strain was confirmed by the PCR amplified the gene using external primer outside the mutant gene. The genome DNA sequencing was accomplished by BioEnergy Genome Center, Qingdao Institute of BioEnergy and BioProcess Technology, Chinese Academy of Sciences. The open reading frames (ORFs) were deduced from the sequence by performing FramePlot 4.0beta program (<http://nocardia.nih.go.jp/fp4>) and BLAST methods.

Genomic Library Construction and Screening. A genomic library of *Streptomyces nobilis* JCM 4274 (genome DNA digested by *Sau3AI*) was constructed in the cosmid pJTU2554 (digested by *BamHI*) according to the manufacturer's protocol (Epicentre Technologies Co.). The cosmid containing the YM-216391 gene cluster was screened by PCR using primers YA1 (5'-GGGGACAAGGAAGTCAGGAGG) and YA2 (5'-GCTGCGATGCCGAAGGAGA) combining with YJ-1 (5'-CGATTCCCGCAAGTGCA) and YJ-2 (5'-CCGCTTCTCGTCGGTGTAGTGT). Three cosmids 1G10, 7A5, 4C8 was selected, and cosmid 7A5 was used to do heterologous expression and named as pTG1101.

Construction of Gene Replacement Mutants. The *aadA*-cassette PCR product was obtained by PCR amplification of the *aadA*-cassette containing in pJTU2125 with the separately prime pairs (ymR1-1: CGGTGGCGCGAGGGGTGAAAGGGAGGGACGGCGATGTCATTCCGGGATCCGTCGACC and ymR1-2: CGACATGCCGTACGCCGCCGGAAAGGCCGGCTGAAGTGTTAGGCTGGAGCTGCTTC; ymR2-1: GTCGACGACGCCGGTCCCGTACGTTGCTGGTGCTCATGCCATTCCGGGATCCGTCGACC and ymR2-2: CGCCAGCCAGCCTTGGATGCCGCCGCTCTCGTCGGTTAGGCTGGAGCTGCTTC; ymR3-1: CGAGGAGGAGCTGACGGAAAAGTCGAGGGCCGGTTGGATTCCGGGATCCGTCGACC and ymR3-2: AGCTGGACGCCGACACGATCCGGACGGATTGCCGAAGTGTTAGGCTGGAGCTGCTTC). The PCR reaction condition is: 100 pmol primers each, 50 ng template DNA, 1 x buffer, 50 mM each dNTP, 5% DMSO, 2.5 units KOD plus, 2 mM MgSO₄. The cycle conditions is: 94 °C, 2 min, 40 cycles with 94 °C, 15 s; 50 °C, 15 s; 68 °C, 90 s, 68 °C, 10 min. The PCR product was purified by Kit according to the manufacturer's protocol.

The cosmid pTG1101 was introduced into *E. coli* BW25113/pIJ790 by electroporation. Then *E. coli* BW25113/pIJ790 containing pTG1101 is electro-transformed with the spectinomycin resistance cassette. The Apra^R and Spec^R transformants were selected. The mutagenised cosmid was transformed into *E. coli* DH5 α

containing the temperature sensitive FLP-recombination plasmid BT340. The Apra^R and spec^R transformers was incubated at 30 °C for 1 days. A few colonies is streaked on LB agar without antibiotics at 42 °C overnight to induce express of the FLP recombinase followed by the loss of plasmid BT340, leaving behind a 81-bp "scar" sequence. Then Apra^R spec^S colones containing in-frame deletion cosmid was selected, characterized by restriction and PCR analysis, which resulting cosmid pTG1102 (Δ ymR1), pTG1103 (Δ ymR2), and pTG1104 (Δ ymR3). The genotype of these gene replacement cosmids were characterized by restriction analysis and PCR with primer pair (ymR1-3: CGCAGAACCTCCAACCGTC and ymR1-4: CATGGGTCCGTCCTGTCG; ymR2-3: GACAGGACGGACCCATGACG and ymR2-4: TGGATTGCCGCCTACCCT; ymR3-3: TCGGTTGACACTGCACCCCTCC and ymR3-4: CGTCGTCGTTGATGCCCTTGA) and the results were shown in Figure S2. These cosmids were introduced into *E. coli* S17-1 and then transferred to *Streptomyces lividans* 1326 by conjugation. Apra^R conjugants was selected and conformed by PCR.

Production and analysis of YM-216391. *S. lividans* containing plasmid was grown on TSB for 36 hours at 28 °C. Then 12-mL culture liquid was transferred into 120-mL fermentation medium (yeast 0.4%, malt extract 1%, glucose 0.4%) and incubated at 28 °C for 7 days. The fermentation broth was extracted with Me₂CO, After Me₂CO was removed by concentrating *in vacuo*, the solution was extracted twice with EtOAc, The organic layer was concentrated *in vacuo* and then the residue was dissolved in MeOH for anlaysis. The identity of compound was confirmed by liquid chromatography-mass spectrometry(LC-MS) analysis performed on LCMS-2010 A (Liquid Chromatpgraph Mass Spectrometer, SHIMADZU, JP) carried out on a GraceSmart RP 18 5u 4.6 x 250 mm). The column was equilibrated with 90% A (H₂O) and 10% B (CH₃CN), and developed with the following program: 0-5 min 90% A/10% B; 5-25 min, 90% A/10% B to 10% A/95% B; 25-28 min, 10% A/95% B to 90% A/10% B; 28-30 min, 90% A/10% B. The flowing rate is 1 mL min⁻¹, and UV detection is at 287 nm.

For further isolation the compound, the MeOH solution was subjected to SephadextTM LH-20 gel flash column chromatography eluted with MeOH. The active fraction was concentrated *in vacuo*. Further purification was achieved by preparative HPLC (venusil XBP C18 10 x 250 mm, Agela Technologies) with following program: 0-15 min a linear gradient from 0-15 min 50% A/50% B to 20% A/80% B; 15-18 min 20% A/80% B to 50% A/50% B, 18-20 min 50% A/50% B. NMR spectra were measured at 400 M Hz for ¹H and ¹³C in DMSO-*d*₆.

Table S1. ^1H (400 MHz) and ^{13}C (100 MHz) NMR data for YM-216391 in $(\text{CD}_3)_2\text{SO}$

Position	^{13}C	^1H
1	160.3	
2		8.22 1H(d, 4.0)
3	57.1	4.80 1H(dd; 8.0,4.0)
4		2.09 1H(m)
5	14.8	
6	25.7	1.66 1H (m) 1.09 1H (m)
7	12.2	
8	170.2	
9		8.57 1H(d ; 12.0)
10	57.5	4.58 1H(dd; 4.0,8.0)
11	31.6	2.14 1H(m)
12	17.5	
13	19.8	
14	170.8	
15		8.68 1H(d; 8.0)
16	35.2	5.05 1H(dd; 16.0,8.0) 4.18 1H(d; 16.0)
17	163.0	
19	139.8	8.89 1H(s)
20	129.1	
22	155.6	
24	139.3	8.99 1H(s)
25	129.9	
27	155.0	
29	139.5	9.09 1H(s)
30	135.6	
32	157.5	
34	122.4	8.65 1H(s)
35	141.5	
37	154.2	
39	150.7	
40	126.7	
41/45	127.5	8.35 1H(d; 4.0)
42/44	128.7	7.57 1H(t; 8.0)
43	130.0	7.52 1H(t; 8.0)
46	130.8	

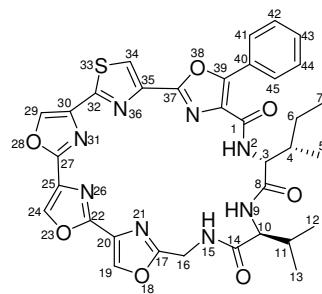


Figure S1. HR-MS analysis of the YM-216391 production by fermentation of *S. lividans* TG1101.

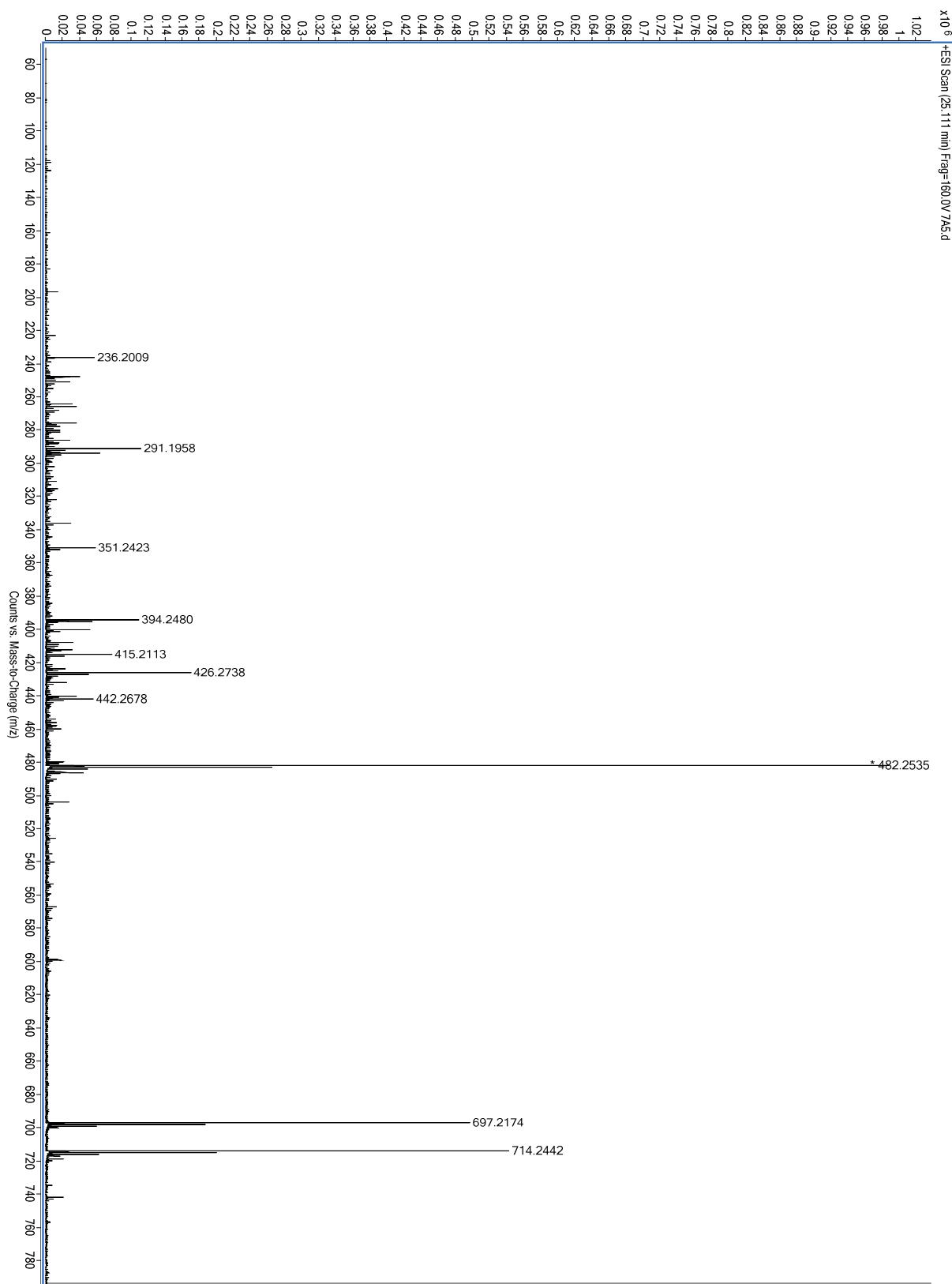
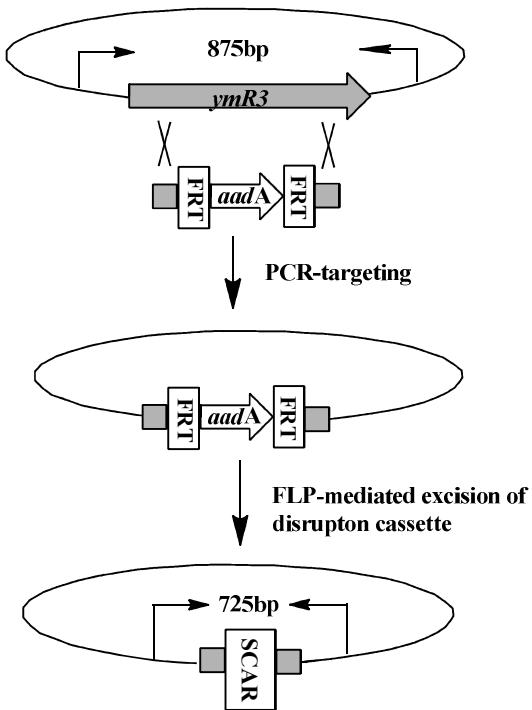


Figure S2. Genotype analysis of mutant cosmid pTG1102 ($\Delta ymR1$), pTG1103 ($\Delta ymR2$), and pTG1104 ($\Delta ymR3$)

A



B

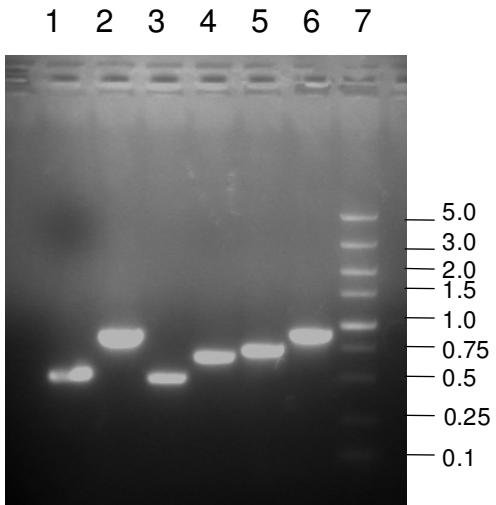


Figure S3. ^1H NMR (400 MHz) spectrum of YM-216391 isolated from *S. lividans* TG1104.

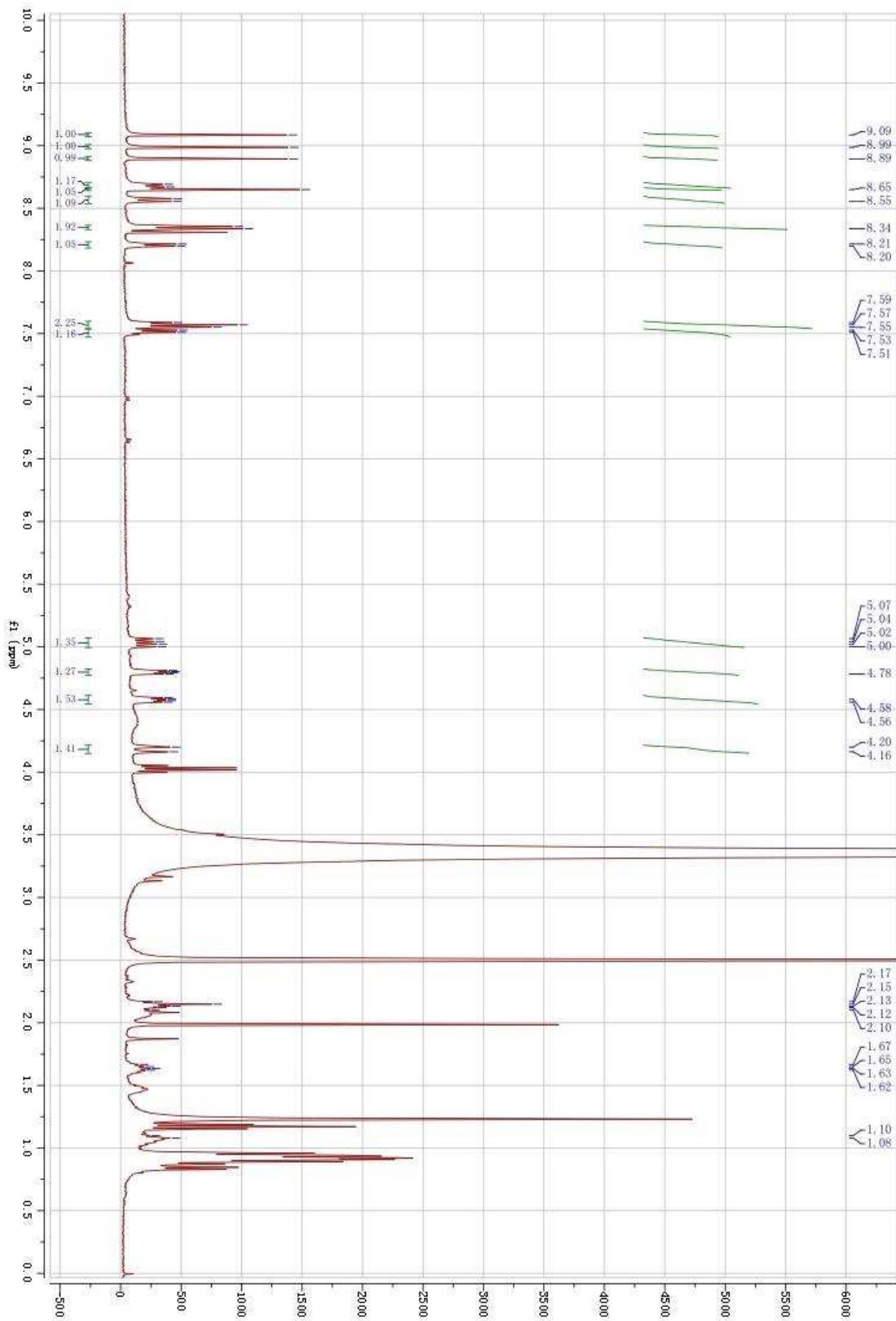


Figure S4. ^{13}C NMR (100 MHz) spectrum of YM-216391 isolated from *S. lividans* TG1104.

