

# Crystal Structure of SgcJ, An NTF2-Like Superfamily Protein Involved in Biosynthesis of the 9-Membered Eneidyne Antitumor Antibiotic C-1027

Tingting Huang<sup>1,†</sup>, Chin-Yuan Chang<sup>1</sup>, Jeremy R. Lohman<sup>1,‡</sup>, Jeffrey D. Rudolf<sup>1</sup>,  
Youngchang Kim<sup>2,3,4</sup>, Changsoo Chang<sup>3,4</sup>, Dong Yang<sup>1</sup>, Ming Ma<sup>1</sup>, Xiaohui Yan<sup>1</sup>, Ivana  
Crnovcic<sup>1</sup>, Lance Bigelow<sup>4</sup>, Shonda Clancy<sup>4</sup>, Craig A. Bingman<sup>5</sup>, Ragothaman M.  
Yenamalli<sup>6,§</sup>, Gyorgy Babnigg<sup>2,4</sup>, Andrzej Joachimiak<sup>2,3,4</sup>, George N. Phillips, Jr.<sup>6</sup> and Ben  
Shen<sup>1,7,8\*</sup>

<sup>1</sup>Department of Chemistry, The Scripps Research Institute, Jupiter, FL 33458, United States;

<sup>2</sup>Center for Structural Genomics of Infectious Diseases, University of Chicago, Chicago, IL

60637, United States; <sup>3</sup>Structural Biology Center and <sup>4</sup>Midwest Center for Structural

Genomics, Biosciences Division, Argonne National Laboratory, Argonne, IL 60439, United

States; <sup>5</sup>Department of Biochemistry, University Wisconsin-Madison, Madison, WI 53705,

United States; <sup>6</sup>BioSciences at Rice and Department of Chemistry, Rice University, Houston,

TX 77251, United States; and <sup>7</sup>Department of Molecular Therapeutics and <sup>8</sup>Natural Products

Library Initiative at The Scripps Research Institute, The Scripps Research Institute, Jupiter,

FL 33458, United States

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

## Supplementary Information

<b>Table S1.</b> Bacterial strains used in this study	S2
<b>Table S2.</b> Plasmids and cosmids used in this study	S3
<b>Table S3.</b> Primers used in this study	S4
<b>Figure S1.</b> Structures of the six known 10-membered enediynes	S5
<b>Figure S2.</b> Multiple sequence alignment of SgcJ and its homologues.	S6
<b>Figure S3.</b> Construction of the $\Delta$ <i>sgcJ</i> mutant strain <i>S. globisporus</i> SB1027	S9
<b>Figure S4.</b> Size exclusion chromatography of SgcJ and NCS-Orf16	S10
<b>References</b>	S11

**Table 1.** Bacterial strains used in this study

Strain	Description	Source (Reference)
<b><i>Streptomyces globisporus</i></b>		
SB1022	<i>S. globisporus</i> $\Delta$ <i>sgcR1</i> mutant strain	1
SB1027	<i>S. globisporus</i> $\Delta$ <i>sgcR1</i> / $\Delta$ <i>sgcJ</i> mutant strain	This work
SB1028	$\Delta$ <i>sgcJ</i> mutant complemented by pBS1146	This work
SB1029	$\Delta$ <i>sgcJ</i> mutant complemented by pBS1147	This work
SB1030	$\Delta$ <i>sgcJ</i> mutant complemented by pBS1148	This work
SB1031	$\Delta$ <i>sgcJ</i> mutant complemented by pBS1149	This work
SB1032	$\Delta$ <i>sgcJ</i> mutant complemented by pBS1150	This work
SB1033	$\Delta$ <i>sgcJ</i> mutant complemented by pBS1151	This work
SB1034	$\Delta$ <i>sgcJ</i> mutant complemented by pBS1152	This work
SB1035	$\Delta$ <i>sgcJ</i> mutant complemented by pBS1153	This work
<b><i>Escherichia coli</i></b>		
<i>E. coli</i> DH5 $\alpha$	General subcloning host	commercial
<i>E. coli</i> BR513	Bioassay test strain for BIA	2
<b><i>Micrococcus luteus</i></b>		
<i>M. luteus</i> ATCC 9431	Test strain for testing of the antibacterial activity of C-1027	ATCC

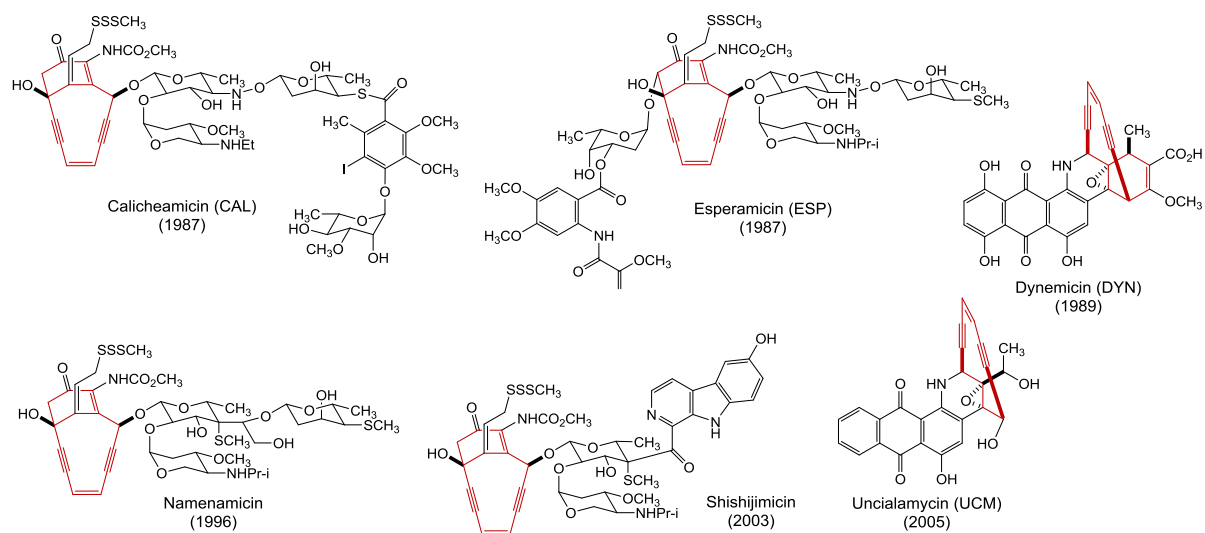
**Table S2.** Plasmids and cosmids used in this study

Plasmid	Description	Source (Reference)
pSET152	Integrative <i>Streptomyces</i> vector	commercial
pSET151	Integrative <i>Streptomyces</i> vector	commercial
pJTU4659	Vector containing kanamycin ( <i>kan</i> ) resistance cassette	This work
pUWL201pw	<i>E. coli-Streptomyces</i> expression shuttle vector	3
pBS1005	Cosmid containing partial <i>C-1027</i> gene cluster	4
pBS5007	Cosmid containing partial <i>ncs</i> gene cluster	5
pMCSG57	<i>E. coli</i> expression vector	6
pBS1143	pBS1005 in which <i>sgcJ</i> has been inactivated by inserting a kanamycin ( <i>kan</i> ) cassette	This work
pBS1144	pSET151 harboring <i>kan</i> cassette (cloned into the <i>XbaI</i> site)	This work
pBS1145	pUWL201pw with <i>oriT</i> fragment inserted into the <i>KpnI</i> site	This work
pBS1146	pBS1145 harboring <i>sgcJ</i> gene (cloned into the <i>NdeI</i> and <i>EcoRI</i> sites)	This work
pBS1147	pBS1145 harboring <i>ncs-orf16</i> gene (cloned into the <i>NdeI</i> and <i>HindIII</i> sites)	This work
pBS1148	SgcJ mutant (W29A) generated by site-directed mutagenesis of pBS1146	This work
pBS1149	SgcJ mutant (F37A) generated by site-directed mutagenesis of pBS1146	This work
pBS1150	SgcJ mutant (Y72A) generated by site-directed mutagenesis of pBS1146	This work
pBS1151	SgcJ mutant (D111A) generated by site-directed mutagenesis of pBS1146	This work
pBS1152	SgcJ mutant (W118A) generated by site-directed mutagenesis of pBS1146	This work
pBS1153	SgcJ mutant (Y132A) generated by site-directed mutagenesis of pBS1146	This work
pBS1154 (APC109082)	pMCSG57 containing <i>sgcJ</i> gene for protein expression	This work
pBS1155 (APC109139)	pMCSG57 containing <i>ncs-orf16</i> gene for protein expression	This work

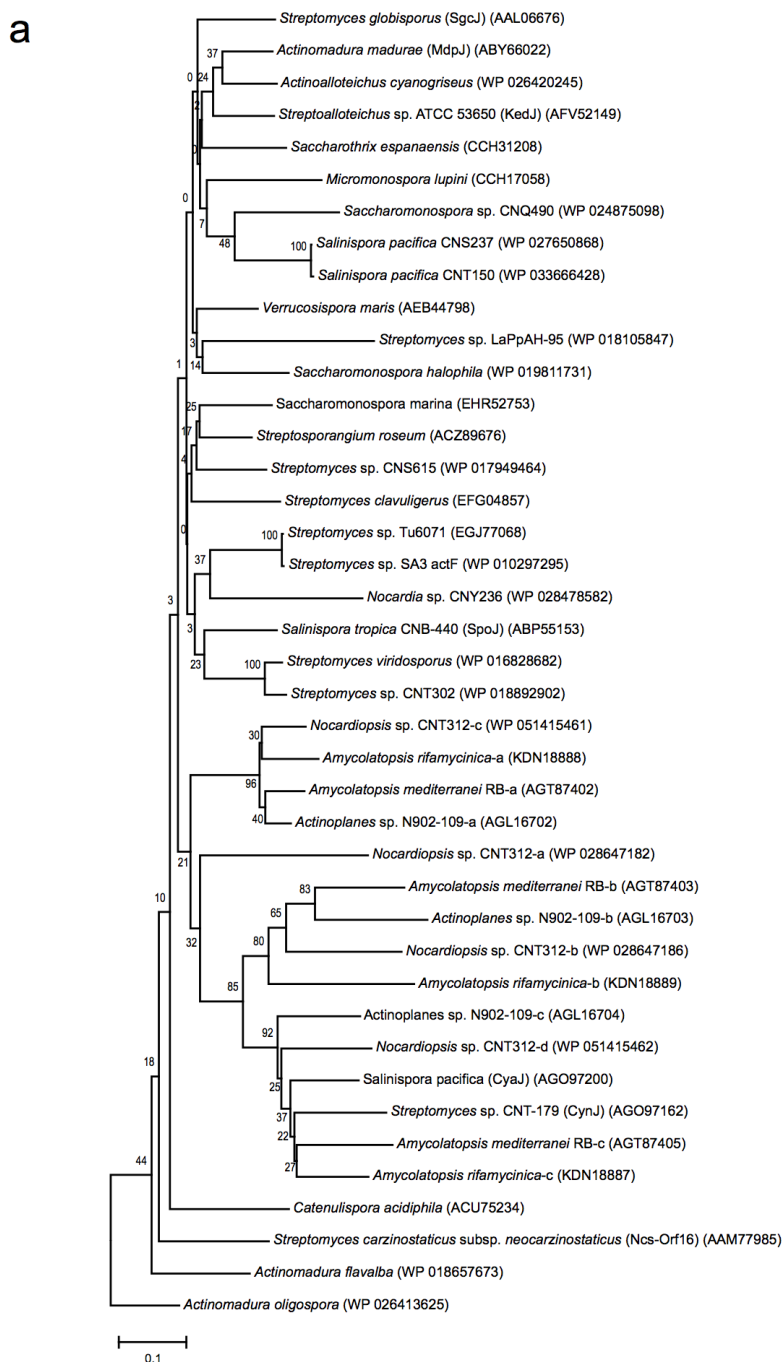
**Table S3.** Primers used in this study

Primer	Nucleotide Sequence	Function
sgcJtgtF	5'-ATGACCAGCACCGACTCGACCACCTCCGCTCCCGACGCCG ATTCCGGGGATCCGTCGACC-3'	PCR targeting for replacement of <i>sgcJ</i>
sgcJtgtR	5'-TCAGTCGTTGCCGCGCGGGCTGTTCTGGTAAGCGGCCAGA CTGTAGGCTGGAGCTGCTTC-3'	PCR targeting for replacement of <i>sgcJ</i>
oriT152F	5'-CGGGTACCAGTATGCAGGTGCAGCGGATCT-3' ( <i>KpnI</i> )	<i>oriT</i> amplification
oriT152R	5'-CCGGTACCCTTGCTCGTCGGTGATGTAC-3' ( <i>KpnI</i> )	<i>oriT</i> amplification
sgcJ201NdeIF	5'-GCAATTCATATGACCAGCACCGACTCGA-3' ( <i>NdeI</i> )	<i>sgcJ</i> amplification
sgcJ201EcoRIR	5'-GGAATTCAGTCGTTGCCGCGCGG-3' ( <i>EcoRI</i> )	<i>sgcJ</i> amplification
ncs16NdeIF	5'-GCAATTCATATGGGAGTGAAGATGAGTTC-3' ( <i>NdeI</i> )	<i>ncs-orf16</i> amplification
ncs16HindIIIR	5'-CCCAAGCTTTGATGGCAGACGGACGGG-3' ( <i>HindIII</i> )	<i>ncs-orf16</i> amplification
sgcJW29A-F	5'-ATCGTGGCCGCCGCGGCCGACCACGACGCCGAC-3'	Site-directed mutagenesis for W29A of SgcJ
sgcJW29A-R	5'-GTCGTGGTCGGCCGCGGCCACGATGCGCGC-3'	Site-directed mutagenesis for W29A of SgcJ
sgcJF37A-F	5'-GACGCCGACCGGCCGCCGACGTCTTCGCCGAG-3'	Site-directed mutagenesis for F37A of SgcJ
sgcJF37A-R	5'-GAAGACGTCGGCGGCCCGGTCGGCGTCGTGGTC-3'	Site-directed mutagenesis for F37A of SgcJ
sgcJY72A-F	5'-TTCGCGGGCCCGCCAAGGGCACCCGTGTCATC-3'	Site-directed mutagenesis for Y72A of SgcJ
sgcJY72A-R	5'-ACGGGTGCCCTTGCCGGGCCCGCAAGGCGGC-3'	Site-directed mutagenesis for Y72A of SgcJ
sgcJD111A-F	5'-GAGGCGAGCGCGCCGGCGCCGTCCGGGCCTCC-3'	Site-directed mutagenesis for D111A of SgcJ
sgcJD111A-R	5'-CCGGACGGCGCCGGCGCCGCTCGCCTCGGTCTC-3'	Site-directed mutagenesis for D111A of SgcJ
sgcJW118A-F	5'-GTCCGGGCCTCCGCGCTGGCCGTCGAACAGGAC-3'	Site-directed mutagenesis for W118A of SgcJ
sgcJW118A-R	5'-TTCGACGGCCAGCGCGGAGGCCCGGACGGCGCC-3'	Site-directed mutagenesis for W118A of SgcJ
sgcJY132A-F	5'-CGTCTGGCCGCTGCCAGAACAGCCCGCGCGGC-3'	Site-directed mutagenesis for Y132A of SgcJ
sgcJY132A-R	5'-CGGGCTGTTCTGGGCAGCGGCCAGACGCCACTG-3'	Site-directed mutagenesis for Y132A of SgcJ
sgcJ-F	5'-TACTTCCAATCCAATGCCATGACCAGCACCGACTCGACC-3'	<i>sgcJ</i> amplification
sgcJ-R	5'-TTATCCACTTCCAATGTTAGTCGTTGCCGCGCGGG-3'	<i>sgcJ</i> amplification
ncs-orf16-F	5'-TACTTCCAATCCAATGCCATGGGAGTGAAGATGAGTTCTG GC-3'	<i>ncs-orf16</i> amplification
ncs-orf16-R	5'-TTATCCACTTCCAATGTTATGGCAGACGGACGGGGCT-3'	<i>ncs-orf16</i> amplification
sgcJidF	5'-ATGACCAGCACCGACTCGA-3'	PCR check of SB1027
sgcJidR	5'-GCTTACCAGAACAGCCCGC-3'	PCR check of SB1027
sgcJ-Sprobe-F	5'-TCCGTTGGTTGGCGGTTTC-3'	Southern hybridization for SB1027
sgcJ-Sprobe-R	5'-CGATGTGCTCGTAGGTGGG-3'	Southern hybridization for SB1027

**Figure S1.** Structures of the six known 10-membered enediyne natural products with their enediyne core highlighted in red. Given in parentheses are the years when each of the enediyne structures was established.



**Figure S2.** Sequence analysis of SgcJ and its homologues from the seven known and 34 putative 9-membered enediyne biosynthetic gene clusters (given in parenthesis are accession numbers). (a) Phylogenetic analysis of SgcJ and its homologues. The phylogenetic tree was constructed with neighbor-joining method. Numbers of each node indicated the percentage of bootstrapping of a 1000 replications. The scale bar indicates branch length. (b) Multiple sequence alignment of SgcJ and its homologues. Aligned residues are colored on the basis of the level of conservation (red background with white character shows strict identity, red character similarity and blue frame similarity across groups).



b

			TT	α1	α2	α3
<i>Streptomyces globisporus</i> (SgcJ) (AAL06676)	1	.....	.....	.....	.....	.....
<i>Streptomyces globisporus</i> (SgcJ) (AAL06676)	1	.....	.....	.....	.....	.....
<i>Actinonadura madurae</i> (MdpJ) (ABY66022)	1	.....	.....	.....	.....	.....
<i>Actinonadurella cyanogriseus</i> (WP 026420245)	1	.....	.....	.....	.....	.....
<i>Streptoactinotrichus</i> sp. ATCC 53650 (KedJ) (AFV52149)	1	.....	.....	.....	.....	.....
<i>Saccharothrix espanaensis</i> (CCH31208)	1	.....	.....	.....	.....	.....
<i>Micromonospora lupini</i> (CCH17058)	1	.....	.....	.....	.....	.....
<i>Saccharomonospora</i> sp. CNQ490 (WP 024875098)	1	.....	.....	.....	.....	.....
<i>Salinispora pacifica</i> CNS237 (WP 027650868)	1	.....	.....	.....	.....	.....
<i>Salinispora pacifica</i> CNT150 (WP 033666428)	1	.....	.....	.....	.....	.....
<i>Verrucosipora maris</i> (AEB44798)	1	.....	.....	.....	.....	.....
<i>Streptomyces</i> sp. LaPaAH-95 (WP 018105847)	1	.....	.....	.....	.....	.....
<i>Saccharomonospora halophila</i> (WP 019811731)	1	.....	.....	.....	.....	.....
<i>Saccharomonospora marina</i> (EHR52753)	1	.....	.....	.....	.....	.....
<i>Streptosporangium roseum</i> (ACZ89676)	1	.....	.....	.....	.....	.....
<i>Streptomyces</i> sp. CNS615 (WP 017949464)	1	.....	.....	.....	.....	.....
<i>Streptomyces clavuligerus</i> (EFG04857)	1	.....	.....	.....	.....	.....
<i>Streptomyces</i> sp. Tu6071 (EGJ77068)	1	.....	.....	.....	.....	.....
<i>Streptomyces</i> sp. SA3 actf (WP 010297295)	1	.....	.....	.....	.....	.....
<i>Nocardia</i> sp. CNY236 (WP 028478582)	1	.....	.....	.....	.....	.....
<i>Salinispora tropica</i> CNB-440 (SpoJ) (ABP55153)	1	.....	.....	.....	.....	.....
<i>Streptomyces viridosporus</i> (WP 016828682)	1	.....	.....	.....	.....	.....
<i>Streptomyces</i> sp. CNT302 (WP 018892902)	1	.....	.....	.....	.....	.....
<i>Nocardiosis</i> sp. CNT312-c (WP 051415461)	1	.....	.....	.....	.....	.....
<i>Amycolatopsis rifamycinica</i> -a (KDN18888)	1	.....	.....	.....	.....	.....
<i>Amycolatopsis mediterranei</i> RB-a (AGT87402)	1	.....	.....	.....	.....	.....
<i>Actinoplanes</i> sp. N902-109-a (AGL16703)	1	.....	.....	.....	.....	.....
<i>Nocardiosis</i> sp. CNT312-a (WP 028647182)	1	.....	.....	.....	.....	.....
<i>Amycolatopsis mediterranei</i> RB-b (AGT87403)	1	.....	.....	.....	.....	.....
<i>Actinoplanes</i> sp. N902-109-b (AGL16703)	1	.....	.....	.....	.....	.....
<i>Nocardiosis</i> sp. CNT312-b (WP 028647186)	1	.....	.....	.....	.....	.....
<i>Amycolatopsis rifamycinica</i> -b (KDN18889)	1	.....	.....	.....	.....	.....
<i>Actinoplanes</i> sp. N902-109-c (AGL16704)	1	.....	.....	.....	.....	.....
<i>Nocardiosis</i> sp. CNT312-d (WP 051415462)	1	.....	.....	.....	.....	.....
<i>Salinispora pacifica</i> (CysJ) (AGO97200)	1	.....	.....	.....	.....	.....
<i>Nocardiosis</i> sp. CNT-179 (CynJ) (AGO97162)	1	.....	.....	.....	.....	.....
<i>Amycolatopsis mediterranei</i> RB-c (AGT87405)	1	.....	.....	.....	.....	.....
<i>Amycolatopsis rifamycinica</i> -c (KDN18887)	1	.....	.....	.....	.....	.....
<i>Catenulispora acidiphila</i> (ACU75234)	1	.....	.....	.....	.....	.....
<i>Streptomyces carzinostaticus</i> subsp. <i>neocarzinostaticus</i> (Ncs-Orf16) (AAM77985)	1	.....	.....	.....	.....	.....
<i>Actinonadura flavaiba</i> (WP 018657673)	1	.....	.....	.....	.....	.....
<i>Actinonadura oligospora</i> (WP 026413625)	1	.....	.....	.....	.....	.....

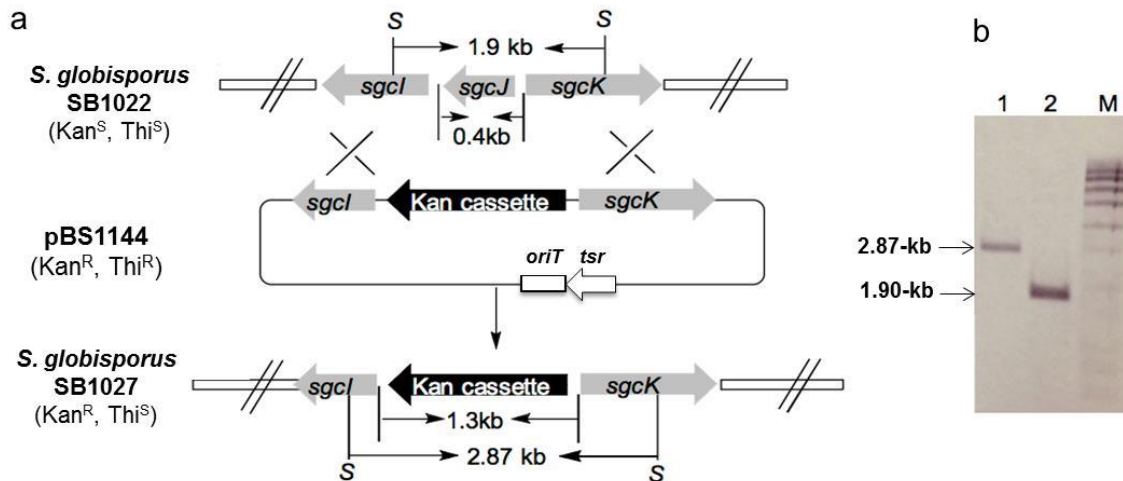
		α4	β3	TT	β4
<i>Streptomyces globisporus</i> (SgcJ) (AAL06676)	37	FADVFEDDGTMLLPG	LFRKGRNERTTHMAAFAGSYKGRVIGSTFDARLIGDG	..	IALLITIG
<i>Actinonadura madurae</i> (MdpJ) (ABY66022)	36	FAGVFIDDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	TALLITIQ
<i>Actinonadurella cyanogriseus</i> (WP 026420245)	35	FADVFEDDGTMLLPG	IYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	TALLITIQ
<i>Streptoactinotrichus</i> sp. ATCC 53650 (KedJ) (AFV52149)	37	FAGVFIDDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	SGVLLITQ
<i>Saccharothrix espanaensis</i> (CCH31208)	40	IADLFTEDDGTMLLPG	LYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	FGILVITQ
<i>Micromonospora lupini</i> (CCH17058)	37	FADLFAEDDGTMLLPG	VYVYRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	AVALLITQ
<i>Saccharomonospora</i> sp. CNQ490 (WP 024875098)	33	FADLFAEDDGTMLLPG	VYVYRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	VVVLLITQ
<i>Salinispora pacifica</i> CNS237 (WP 027650868)	37	FADLFAEDDGTMLLPG	IYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	AAILLSYQ
<i>Salinispora pacifica</i> CNT150 (WP 033666428)	37	FADLFAEDDGTMLLPG	IYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	AAILLSYQ
<i>Verrucosipora maris</i> (AEB44798)	37	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	FALLITIG
<i>Streptomyces</i> sp. LaPaAH-95 (WP 018105847)	40	FADLFAEDDGTMLLPG	AALKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	VVLLITQ
<i>Saccharomonospora halophila</i> (WP 019811731)	42	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	TAVVITQ
<i>Saccharomonospora marina</i> (EHR52753)	40	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	SAVLLITQ
<i>Streptosporangium roseum</i> (ACZ89676)	40	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	AVGLITQ
<i>Streptomyces</i> sp. CNS615 (WP 017949464)	39	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	SGILLITQ
<i>Streptomyces clavuligerus</i> (EFG04857)	35	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	SALLITIG
<i>Streptomyces</i> sp. Tu6071 (EGJ77068)	34	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	SAVLLITQ
<i>Streptomyces</i> sp. SA3 actf (WP 010297295)	34	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	SAILVITQ
<i>Nocardia</i> sp. CNY236 (WP 028478582)	26	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	VVLLITQ
<i>Salinispora tropica</i> CNB-440 (SpoJ) (ABP55153)	36	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	TVLLITQ
<i>Streptomyces viridosporus</i> (WP 016828682)	38	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	TAVVITQ
<i>Streptomyces</i> sp. CNT302 (WP 018892902)	42	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	AVVITQ
<i>Nocardiosis</i> sp. CNT312-c (WP 051415461)	50	VADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	VCLVITQ
<i>Amycolatopsis rifamycinica</i> -a (KDN18888)	35	VAKAFVIEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	VALIRITQ
<i>Amycolatopsis mediterranei</i> RB-a (AGT87402)	43	VAKAFVIEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	VALLITQ
<i>Actinoplanes</i> sp. N902-109-a (AGL16703)	50	VAGVFIEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	VALIRITQ
<i>Nocardiosis</i> sp. CNT312-a (WP 028647182)	33	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	VVLLITQ
<i>Amycolatopsis mediterranei</i> RB-b (AGT87403)	49	LADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	IADVITQ
<i>Actinoplanes</i> sp. N902-109-b (AGL16703)	48	LADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	IADVITQ
<i>Nocardiosis</i> sp. CNT312-b (WP 028647186)	49	LADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	IADVITQ
<i>Amycolatopsis rifamycinica</i> -b (KDN18889)	49	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	IVVLLITQ
<i>Actinoplanes</i> sp. N902-109-c (AGL16704)	54	FAGVFIEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	VAMVITQ
<i>Nocardiosis</i> sp. CNT312-d (WP 051415462)	66	FAGVFIEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	VAMVITQ
<i>Salinispora pacifica</i> (CysJ) (AGO97200)	57	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	VVLLITQ
<i>Streptomyces</i> sp. CNT-179 (CynJ) (AGO97162)	53	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	VAMVITQ
<i>Amycolatopsis mediterranei</i> RB-c (AGT87405)	54	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	TAMVITQ
<i>Amycolatopsis rifamycinica</i> -c (KDN18887)	55	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	VAMVITQ
<i>Catenulispora acidiphila</i> (ACU75234)	34	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	FAMVITQ
<i>Streptomyces carzinostaticus</i> subsp. <i>neocarzinostaticus</i> (Ncs-Orf16) (AAM77985)	40	FAGLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	TAILITQ
<i>Actinonadura flavaiba</i> (WP 018657673)	34	FAGLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	TAMVITQ
<i>Actinonadura oligospora</i> (WP 026413625)	27	FADLFAEDDGTMLLPG	VYRKRGRDERTAYMDAFAAGSYKGRVIGSTFDARLIGDG	..	AVVLLITQ

b (continued)

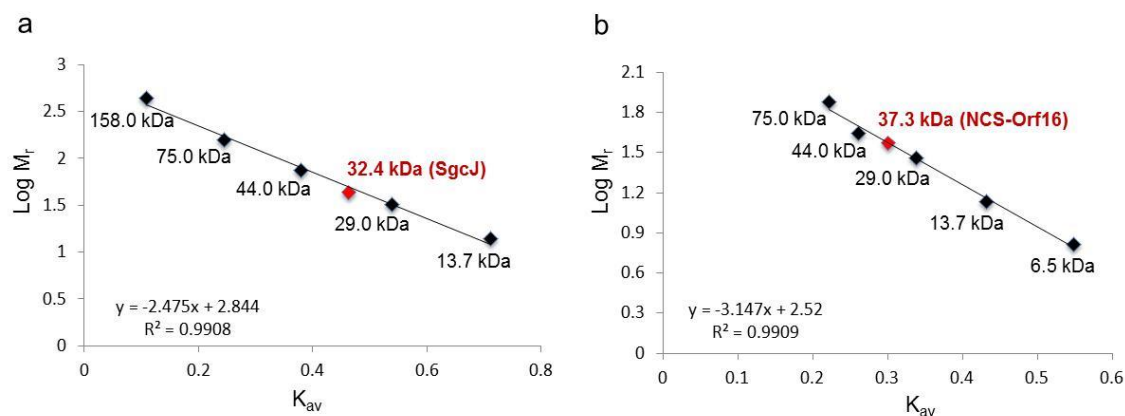
		→	T	T		η1	β5		T	T		β6			
						Q	Q								
<i>Streptomyces globisporus</i> (SgcJ) (AAL06676)			G	I	L	A	P	G	E	T	E	A	S	G	E
<i>Streptomyces globisporus</i> (SgcJ) (AAL06676)	99		G	I	L	A	P	G	E	T	E	A	S	G	E
<i>Actinomadura madurae</i> (MdpJ) (ABY66022)	98		G	V	L	A	N	G	E	S	E	V	S	D	E
<i>Actinoalloteichus cyanogriseus</i> (WP 026420245)	97		G	V	L	A	E	G	E	S	E	V	S	D	E
<i>Streptoalloteichus</i> sp. ATCC 53650 (KedJ) (AFV52149)	99		G	V	L	K	P	G	T	T	E	V	T	E	S
<i>Saccharothrix espanaensis</i> (CCH31208)	102		G	V	L	A	P	G	D	T	E	V	T	A	E
<i>Micromonospora lupini</i> (CCH17058)	99		G	V	I	A	A	G	S	T	E	L	A	A	D
<i>Saccharomonospora</i> sp. CNQ490 (WP 024875098)	95		G	V	M	E	E	G	Q	T	E	V	S	P	E
<i>Salinispora pacifica</i> CNS237 (WP 027650868)	99		G	V	L	E	S	G	E	S	E	V	S	S	K
<i>Salinispora pacifica</i> CNT150 (WP 033666428)	99		G	V	L	E	S	G	E	S	E	V	S	S	K
<i>Verrucosipora maris</i> (AEB44798)	99		G	V	L	A	P	G	E	T	E	V	A	A	E
<i>Streptomyces</i> sp. LaPpAH-95 (WP 018105847)	104		G	A	Y	P	P	D	A	T	E	V	P	T	E
<i>Saccharomonospora halophila</i> (WP 019811731)	104		G	V	L	T	P	G	E	T	E	P	A	A	E
<i>Saccharomonospora marina</i> (EHR52753)	102		G	I	L	A	P	G	E	T	E	V	S	D	E
<i>Streptosporangium roseum</i> (ACZ89676)	102		G	V	L	A	P	G	E	T	E	V	S	D	E
<i>Streptomyces</i> sp. CNS615 (WP 017949464)	101		G	I	L	A	P	G	E	S	E	P	S	A	D
<i>Streptomyces clavuligerus</i> (EFG04857)	97		G	V	L	E	G	E	E	V	A	A	E	R	A
<i>Streptomyces</i> sp. Tu6071 (EGJ77068)	96		G	V	L	L	P	G	E	K	K	V	S	A	E
<i>Streptomyces</i> sp. SA3 actF (WP 010297295)	96		G	V	L	L	P	G	E	K	K	V	S	A	E
<i>Nocardia</i> sp. CNY236 (WP 028478582)	88		G	I	R	P	T	A	D	S	L	P	D	G	A
<i>Salinispora tropica</i> CNB-440 (SpoJ) (ABP55153)	98		G	V	L	A	P	G	E	T	E	P	A	A	E
<i>Streptomyces viridosporus</i> (WP 016828682)	100		G	V	L	A	P	G	E	S	E	P	A	A	E
<i>Streptomyces</i> sp. CNT302 (WP 018892902)	104		G	V	L	T	P	G	E	T	E	P	A	A	E
<i>Nocardopsis</i> sp. CNT312-c (WP 051415461)	113		G	I	L	A	P	G	E	T	E	I	A	P	E
<i>Amycolatopsis rifamycinica</i> -a (KDN18888)	98		G	I	L	A	P	G	E	T	E	I	D	E	L
<i>Amycolatopsis mediterranei</i> RB-a (AGT87402)	106		G	I	L	A	A	G	E	T	E	I	D	E	L
<i>Actinoplanes</i> sp. N902-109-a (AGL16702)	113		G	I	L	A	P	G	Q	T	E	I	E	P	E
<i>Nocardopsis</i> sp. CNT312-a (WP 028647182)	94		R	V	E	I	P	G	V	P	A	E	T	P	L
<i>Amycolatopsis mediterranei</i> RB-b (AGT87403)	111		G	I	V	L	A	G	E	T	E	V	A	P	E
<i>Actinoplanes</i> sp. N902-109-b (AGL16703)	110		A	I	V	P	A	G	Q	D	E	P	A	A	
<i>Nocardopsis</i> sp. CNT312-b (WP 028647186)	111		G	V	A	R	S	G	E	S	P	A	E	E	Y
<i>Amycolatopsis rifamycinica</i> -b (KDN18889)	111		G	V	L	Y	D	G	E	T	T	V	P	A	E
<i>Actinoplanes</i> sp. N902-109-c (AGL16704)	116		G	I	V	R	P	G	G	H	E	V	A	P	A
<i>Nocardopsis</i> sp. CNT312-d (WP 051415462)	128		G	I	V	M	P	G	E	D	A	V	A	P	E
<i>Salinispora pacifica</i> (CyaJ) (AGO97200)	119		G	I	R	L	P	G	E	T	T	X	R	E	Q
<i>Streptomyces</i> sp. CNT-179 (CynJ) (AGO97162)	115		G	I	L	D	G	E	S	E	I	A	P	E	R
<i>Amycolatopsis mediterranei</i> RB-c (AGT87405)	116		G	I	L	D	G	E	S	E	I	A	P	E	R
<i>Amycolatopsis rifamycinica</i> -c (KDN18887)	117		G	I	L	A	P	G	E	S	E	I	A	P	E
<i>Catenulispora acidiphila</i> (ACU15234)	97		G	V	M	A	P	G	E	T	E	V	A	P	E
<i>Streptomyces carzinostaticus</i> subsp. <i>neocarzinostaticus</i> (Ncs-Orf16) (AAM77985)	103		G	V	M	A	P	G	E	H	S	V	A	P	D
<i>Actinomadura flavaiba</i> (WP 018657673)	97		G	V	L	L	P	G	E	T	E	V	S	A	P
<i>Actinomadura oligospora</i> (WP 026413625)	90		G	V	V	A	A	G	A	K	E	P	S	A	E



**Figure S3.** Construction and confirmation of the  $\Delta$ *sgcJ* mutant strain *S. globisporus* SB1027. (a) Construction of the  $\Delta$ *sgcJ* mutant strain *S. globisporus* SB1027. S, *StuI*; Kan<sup>R</sup>, kanamycin resistant; Kan<sup>S</sup>, kanamycin sensitive, Thi<sup>R</sup>, thiostrepton resistant; Thi<sup>S</sup>, thiostrepton sensitive. (b) Southern analysis confirming the genotype of SB1027. M, DNA ladder; lane 1, SB1027; lane 2, SB1022.



**Figure S4.** Molecular weight estimation of SgcJ and NCS-Orf16 in solution by size exclusion chromatography. (a) SgcJ and (b) NCS-Orf16 estimated on a Superdex 200 16/600 column calibrated against molecular weight standards. The apparent molecule weights of SgcJ and NCS-Orf16 were estimated to be 32.4 kDa and 37.3 kDa, suggesting that both SgcJ (calculated molecular weight of 14.6 KDa) and NCS-Orf16 (calculated molecular weight of 15.3 KDa) are dimers.



## References

- 1 Chen, Y., Yin, M., Horsman, G. P. & Shen, B. Improvement of the enediyne antitumor antibiotic C-1027 production by manipulating its biosynthetic pathway regulation in *Streptomyces globisporus*. *J. Nat. Prod.* **74**, 420-424 (2011).
- 2 Elespuru, R. K. & White, R. J. Biochemical prophage induction assay: a rapid test for antitumor agents that interact with DNA. *Cancer Res.* **43**, 2819-2830 (1983).
- 3 Doumith, M. *et al.* Analysis of genes involved in 6-deoxyhexose biosynthesis and transfer in *Saccharopolyspora erythraea*. *Mol. Gen. Genet.* **264**, 477-485 (2000).
- 4 Liu, W. & Shen, B. Genes for production of the enediyne antitumor antibiotic C-1027 in *Streptomyces globisporus* are clustered with the *cagA* gene that encodes the C-1027 apoprotein. *Antimicrob. Agents Chemother.* **44**, 382-392 (2000).
- 5 Liu, W. *et al.* The neocarzinostatin biosynthetic gene cluster from *Streptomyces carzinostaticus* ATCC 15944 involving two iterative type I polyketide synthases. *Chem. Biol.* **12**, 293-302 (2005).
- 6 Eschenfeldt, W. H., Lucy, S., Millard, C. S., Joachimiak, A. & Mark, I. D. A family of LIC vectors for high-throughput cloning and purification of proteins. *Methods Mol. Biol.* **498**, 105-115 (2009).