The Journal of Antibiotics (2009) 62, 325–332
© 2009 Japan Antibiotics Research Association All rights reserved 0021-8820/09 $32.00
www.nature.com/ja

ORIGINAL ARTICLE

Identification of the biosynthetic gene cluster of A-500359s in *Streptomyces griseus* SANK60196

Masanori Funabashi[1], Koichi Nonaka[1], Chieko Yada[1], Masahiko Hosobuchi[1], Nobuhisa Masuda[2], Tomoyuki Shibata[3] and Steven G Van Lanen[4]

A-500359s, produced by *Streptomyces griseus* SANK60196, are inhibitors of bacterial phospho-*N*-acetylmuramyl-pentapeptide translocase. They are composed of three distinct moieties: a 5′-carbamoyl uridine, an unsaturated hexuronic acid and an aminocaprolactam. Two contiguous cosmids covering a 65-kb region of DNA and encoding 38 open reading frames (ORFs) putatively involved in the biosynthesis of A-500359s were identified. Reverse transcriptase PCR showed that most of the 38 ORFs are highly expressed during A-500359s production, but mutants that do not produce A-500359s did not express these same ORFs. Furthermore, *orf21*, encoding a putative aminoglycoside 3′-phosphotransferase, was heterologously expressed in *Escherichia coli* and *Streptomyces albus*, yielding strains having selective resistance against A-500359B, suggesting that ORF21 phosphorylates the unsaturated hexuronic acid as a mechanism of self-resistance to A-500359s. In total, the data suggest that the cloned region is involved in the resistance, regulation and biosynthesis of A-500359s.
The Journal of Antibiotics (2009) **62,** 325–332; doi:10.1038/ja.2009.38; published online 29 May 2009

Keywords: A-500359s; biosynthetic gene cluster; resistance gene; RT-PCR; *Streptomyces griseus*

INTRODUCTION

Nucleoside antibiotics are a structurally diverse group of secondary metabolites with a broad range of biological activities, such as antibacterial, antifungal, antiviral, insecticidal, immunostimulative, immunosuppressive and antitumor activities. For example, blasticidin S[1] and mildiomycin,[2] peptidyl nucleoside antibiotics containing cytosine-derived bases, are cytotoxic to fungi by virtue of binding to the 50S ribosomal subunit. Polyoxins and nikkomycins, peptidyl nucleoside antibiotics, also exhibit antifungal activity, but in this case by the inhibition of chitin synthase,[3] and have been commercialized in the United States and Japan for agricultural application. There are a large number of uracil-containing nucleoside antibiotics that have antibacterial activities, such as mureidomycins, pacidamycins, napsamycins, liposidomycins, caprazamycins, muramycins and capuramycins.[4] Interestingly, they all inhibit bacterial phospho-*N*-acetylmuramyl-pentapeptide translocase (translocase I), which is an enzyme that catalyzes the first step in the lipid cycle of peptidoglycan biosynthesis: the transfer of phospho-*N*-acetylmuramic acid-pentapeptide to undecaprenyl phosphate to generate undecaprenyl-disphospho-*N*-acetylmuramic acid-pentapeptide, also known as lipid intermediate I. As peptidoglycan has an essential role in the vitality of bacteria, and the biosynthesis of peptidoglycan is a proven target for many antibiotics, including β-lactams, vancomycin and bacitracin, translocase I represents a valid target for the discovery and development of new antibacterial agents.

Nucleoside antibiotics not only possess potent and desirable biological activities but also are endowed with unique structural features, suggesting the occurrence of novel or unusual enzymatic transformations during their biosynthesis. To date, six complete biosynthetic gene clusters for nucleoside antibiotics have been cloned and reported: nikkomycin and polyoxin, peptidyl nucleoside antibiotics containing a uracil- or 4-formyl-imidazolin-2-one base;[5,6] puromycin, an adenine-containing aminonucleoside antibiotic;[7] streptothricins, peptidyl guanidine nucleoside antibiotics;[8] and blasticidin S[9] and toyocamycin, pyrrolopyrimidine nucleoside antibiotics.[10] Furthermore, two enzymes involved in mildiomycin biosynthesis have been functionally characterized.[11] However, there have been no reports on the cloning and characterization of a biosynthetic gene cluster for nucleoside antibiotics that target translocase I. Given that deciphering the mechanism of assembly of nucleoside antibiotics inhibiting translocase I will ultimately facilitate applications to promote the molecular diversity of natural and unnatural nucleoside antibiotics and also to enhance the

[1]Bioengineering Research Group I, Process Technology Research Laboratories, Pharmaceutical Technology Division, Daiichi Sankyo Co. Ltd, Fukushima, Japan; [2]Group IV, Biological Research Laboratories IV, Research and Development Division, Daiichi Sankyo Co. Ltd, Tokyo, Japan; [3]Group V, Exploratory Research Laboratories I, Research and Development Division, Daiichi Sankyo Co. Ltd, Tokyo, Japan and [4]Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, College of Pharmacy, Lexington, KY, USA
Correspondence: Dr K Nonaka, Process Technology Research Laboratories, Daiichi Sankyo Co. Ltd, Aza-ohtsurugi, Shimokawa, Izumi-machi, Iwaki-shi, Fukushima 971-8183, Japan.
E-mail: nonaka.koichi.k2@daiichisankyo.co.jp
Received 15 January 2009; revised 26 April 2009; accepted 7 May 2009; published online 29 May 2009

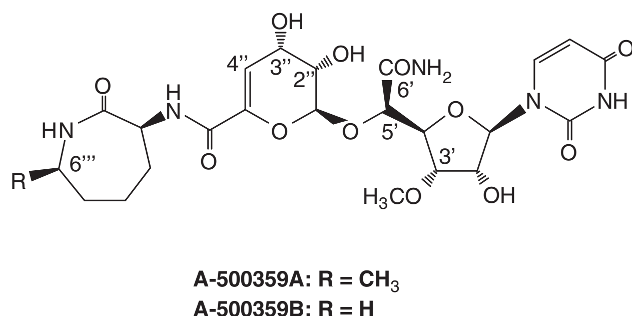

Figure 1 The structure of the two main capuramycin-related metabolites isolated from *S. griseus* SANK60196.

production of desired nucleoside antibiotics, we initiated studies to identify the biosynthetic locus for model translocase I inhibitors.

A-500359s, which are capuramycin derivatives classified as a family of glycosyl nucleoside antibiotics, are produced by *Streptomyces griseus* SANK60196. They consist of three primary moieties, a 5′-carbamoyl uridine, an unsaturated hexuronic acid and an aminocaprolactam (Figure 1). The biosynthesis of A-500359s was previously analyzed using isotope-feeding experiments and was reported by Ohnuki *et al.*[12] to reveal the origin of every carbon atom, including the potential precursors of each moiety as uridine, mannose and lysine, respectively. Recently, we showed that a giant linear plasmid (SGF200) in *S. griseus* SANK60196 might be required for A-500359s biosynthesis; however, the biosynthetic gene cluster was not identified in SGF200 (unpublished data). We report here the identification of a putative gene cluster involved in the biosynthesis, regulation and resistance of A-500359s. Using reverse transcriptase PCR (RT-PCR), open reading frames (ORFs) within this gene cluster are shown to be highly expressed during A-500359s production. In addition, we show that ORF21 within this locus confers selective resistance to A-500359B. The identification of the A-500359s gene cluster now sets the stage to explore the mechanism of biosynthesis of this family of nucleoside antibiotics.

MATERIALS AND METHODS

Chemicals, strains and general recombinant DNA techniques

Escherichia coli JM109, restriction enzymes and DNA-modifying enzymes were purchased from Takara Bio (Shiga, Japan). *E. coli* XL1-Blue MR was purchased from Stratagene (La Jolla, CA, USA). Media, growth conditions and general recombinant DNA techniques for *E. coli* were described by Sambrook and Russell.[13] *Streptomyces albus* J1074 and plasmid vectors pWHM3[14] and pWHM79[15] were gifts from Professor Ben Shen, University of Wisconsin at Madison. *E. coli* MG1655 (ATCC 47076) was obtained from the American Type Culture Collection (Manassas, VA, USA).

RT-PCR analysis

A loopful of mycelia of cultured *S. griseus* SANK60196 derivatives HP, 35-4, 37-3 and 37-9 was inoculated into a test tube containing 5 ml of PM-1 medium and cultured with shaking (310 r.p.m.) at 23 °C for 3 days, as described previously.[16] Then 1 ml aliquots of each culture were transferred into a 100-ml Erlenmeyer flask containing 20 ml OM-1 medium and cultivation was continued for 7 days. The cultured mycelia were harvested from the culture broth by centrifugation (20 000×*g*, 5 min, 4 °C) and were treated overnight with RNAlater (Ambion, Austin, TX, USA) at 4 °C. Total RNA was isolated from the treated mycelia using RNAqueous (Ambion), according to the manufacturer's instructions. The isolated total RNA was treated with DNaseI to digest any contaminating genomic DNA. For RT-PCR cloning of NDP-glucose dehydratase (NGDH) in *S. griseus* SANK60196, the total RNA isolated from the mycelia of strain HP cultured for 7 days was used for cDNA synthesis using TOYOBO RT-Ace (TOYOBO, Osaka, Japan), and the desired fragment was amplified by LA-Taq with GC buffer (Takara Bio). The NGDH degenerate primer pairs used for PCR (dehy-f: 5′-CSGGSGSSGCSGGSTTCATSGG-3′/ dehy-r: 5′-GGGWRCTGGYRSGGSCCGATGTTG-3′) were designed on the basis of the report by Decker *et al.*[17] RT-PCR amplification for expression analysis was carried out on a GeneAmp PCR system 9700 (Perkin-Elmer/ABI, Foster City, CA, USA) using TaKaRa One Step RNA PCR kit (AMV) (Takara Bio) and was conducted in 25 cycles. *glk*, encoding glucokinase, was used as an internal control, and unique primers (Glk-f: 5′-CGGCGGCACGAAGATC-3′/ Glk-r: 5′-GCGCAGCTTGTTGCCG-3′) were designed on the basis of the highly conserved sequences, IGGTKI and IGNKLR, corresponding to the *N*- and the *C*-terminal amino-acid sequence of glucokinase in *Streptomyces coelicolor* A3(2) (NP_626383), *Streptomyces avermitilis* MA-4680 (NP_827250) and *S. griseus* IFO13350 (YP_001826889).

Genomic library construction

S. griseus SANK60196 genomic DNA was partially digested with *Sau*3AI to give 30- to 50-kb DNA fragments. These fragments were dephosphorylated with bacterial alkaline phosphatase and ligated into *Bam*HI-digested cosmid vector SuperCos1 (Stratagene), which was dephosphorylated by bacterial alkaline phosphatase after *Xba*I digestion. The ligation products were packaged with Gigapack III Gold packaging extract (Stratagene) as described by the manufacturer, and the resulting recombinant phage was used to transfect *E. coli* XL1-Blue MR. Approximately 20 000 colonies from the obtained genomic library were screened by colony hybridization using a digoxigenin (DIG)-labeled 0.55-kb fragment, including a part of the cloned putative NGDH. Hybridization was carried out using DIG easy hyb (Roche, Indianapolis, IN, USA) at 42 °C, and the resulting filter was washed under high stringency conditions (0.1× SSC including 0.1% SDS, 68 °C). Detection was performed using CDP-Star (Roche) according to the manufacturer's procedures. The resultant positive cosmids were isolated and sequenced.

DNA sequencing

Automated DNA sequencing was carried out on an ABI PRISM 3700 DNA Analyzer (Perkin-Elmer/ABI). The DNA sequence of the isolated cosmids was determined by shotgun sequencing. The cosmid DNA was sheared using a Nebulizer Kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's procedures and the treated DNA was analyzed by agarose gel electrophoresis. The DNA fragments from 1 to 3 kb were purified using a QIAquick PCR purification kit (Qiagen, Gaithersburg, MD, USA). The recovered DNA fragments were cloned into pHSG398 (Takara Bio) and transformed into *E. coli* JM109. Approximately 400 plasmids extracted using R.E.A.L. Prep 96 (Qiagen) were sequenced and the sequence data were assembled using ATGC (Genetyx, Tokyo, Japan). Database comparison for sequence homology was performed with BLAST search tools using the National Center for Biotechnology Information (Bethesda, MD, USA). The DNA sequence has been deposited in DDBJ under the accession number AB476988.

Disruption of *tolC* in *E. coli* MG1655 for a test of resistance ability against A-500359B

An in-frame deletion of the *tolC* gene (AC_000091) in *E. coli* MG1655 was carried out using the pKO3-derived plasmid carrying an in-frame fusion of the 5′ and 3′ flanking regions of *tolC* reported previously.[18] The pKO3-derived plasmid for the deletion of *ΔtolC* was introduced into *E. coli* MG1655-competent cells by electroporation, and integrants, which contained the plasmid in the chromosome, were selected using Luria-Bertani (LB) medium containing chloramphenicol at 43 °C. One of the integrants was grown in LB liquid medium without selection pressure for 9 h and the serially diluted culture broth was plated and incubated on LB agar supplemented with 10% (wt/vol) sucrose at 30 °C for 24 h. Chloramphenicol-susceptible and sucrose-resistant colonies were isolated and subjected to PCR for the confirmation of gene deletion using the following pairs of primers I: 5′-AAGGAAAAAAGC GGCCGCTGCTAAACAGTATCGCAACCAGTC-3′ and II: 5′-CGCACGCATGT CGACTCGTATAGTGACGTTGGCGTATC-3′. The resulting clone, named

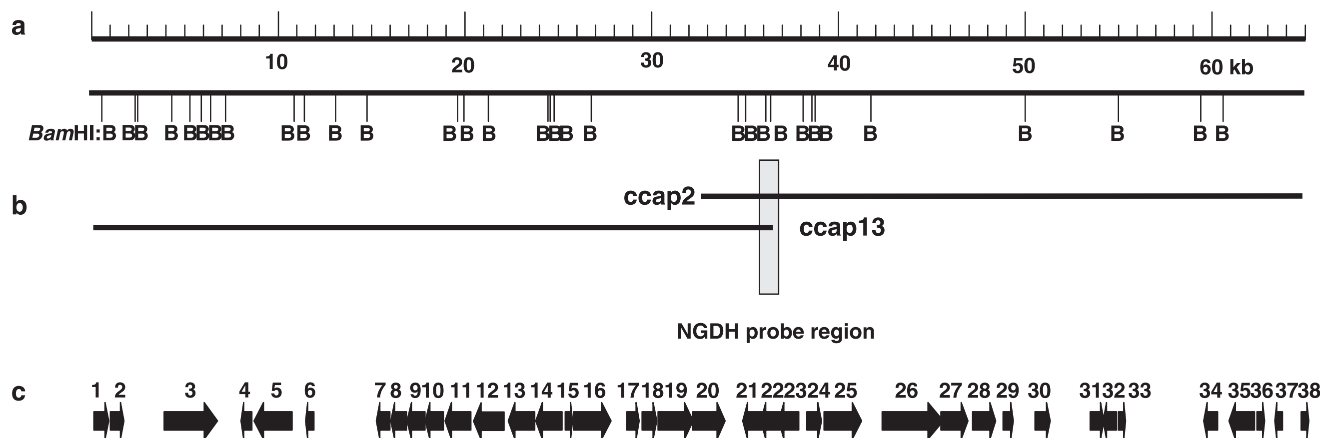

Figure 2 Restriction map of *Bam*HI and genetic organization of the A-500359s biosynthetic gene cluster. (**a**) *Bam*HI cleavage sites are illustrated. (**b**) Overlapping cosmids, ccap2 and ccap13, which contain the NGDH homolog. (**c**) Genetic organization of the fully sequenced ccap13 and ccap2 inserts encoding the A-500359s biosynthetic gene cluster.

E. coli ΔtolC, was sensitive against A-500359s and was used to test resistance against A-500359B.

Construction of *orf21* expression vector in *E. coli ΔtolC*

An *orf21* expression construct in *E. coli* was generated as follows: *orf21* was amplified using primers III: 5′-GCGA***AGC****TT*GGTGGCAGCGGACGGG-3′ (the *Hin*dIII site is shown in italics and the Ser residue of LacZ in bold) and IV: 5′-GCG*GAATTC*TCAGGTTCGAGTCGCG-3′ (the *Eco*RI site is shown in italics). The resulting 1-kb amplified fragment was cloned into pSC-B using a StrataClone PCR Cloning Kit (Stratagene) to yield strata-ec-*orf21*. After sequencing to confirm PCR fidelity, the *Hin*dIII–*Eco*RI fragment was excised from strata-ec-*orf21* and introduced into pUC19 (Takara Bio) at *Hin*dIII–*Eco*RI sites to yield pUC19-*orf21*. The desired plasmid was transformed into *E. coli ΔtolC*-competent cells to confirm its resistance against A-500359B. *E. coli ΔtolC* harboring pUC19 was used as a control.

Construction of *orf21* expression vector in *Streptomyces*

An *orf21* expression vector in *Streptomyces* was constructed as follows: a 450-bp *Eco*RI–*Bam*HI fragment that harbored the *ermE** promoter from pWHM79 was ligated at *Eco*RI–*Bam*HI sites in pWHM3 to yield pWHM3-*Ep*. *orf21* was amplified using primers V: 5′-GCG*CTGCAG*GTGGCAGCGGACGGG-3′ (the *Pst*I site is shown in italics) and VI: 5′-GCG*AAGCTT*TCAGGTTC-GAGTCGCG-3′ (the *Hin*dIII site is shown in italics) and the resulting 1-kb fragment was cloned into pSC-B using a StrataClone PCR Cloning Kit (Stratagene) to yield strata-st-*orf21*. After sequencing to confirm PCR fidelity, the *Pst*I–*Hin*dIII fragment was excised from strata-st-*orf21* and ligated into pWHM3-*Ep* at *Pst*I–*Hin*dIII sites to yield pWHM3-*Ep-orf21*. The desired plasmid was introduced into *S. albus* J1074 by polyethylene glycol-mediated protoplast transformation[19] to confirm its resistance against A-500359B. *S. albus* J1074, harboring pWHM3, was used as control.

Test of resistance to A-500359B in *Streptomyces* and *E. coli*

S. albus J1074, harboring the appropriate plasmid, was inoculated into 5 ml of Trypto Soy Broth (TSB) medium containing 50 μg ml^{-1} of thiostrepton and grown at 28 °C for 3 days. After homogenization, the culture broth was diluted with TSB medium to OD_{600}=3.0. Ten microliters of 100- and 1000-fold dilution prepared from the resulting culture broth was spotted on the ISP-2 agar containing 0, 100, 200, 500 or 1000 μg ml^{-1} of A-500359B, and the spotted plate was incubated at 27 °C for 4 days.

E. coli ΔtolC, harboring the appropriate plasmid, was inoculated into 2 ml of LB medium containing 100 μg ml^{-1} of ampicillin sodium salt and cultured at 37 °C for 1 day. The culture broth was diluted with LB medium to OD_{600}=1.0. Twenty microliters of the resulting culture broth was spotted on the LB agar containing 0, 10, 100 or 1000 μg ml^{-1} of A-500359B with 1 mM isopropyl-β-D-thiogalactopyranoside and the spotted plate was incubated at 37 °C for 1 day.

RESULTS

Identification of A-500359s biosynthetic gene cluster

Using RT-PCR, it was observed that a gene for a putative NGDH was abundantly transcribed during the production of A-500359s in a high-producing strain (strain HP). In contrast, this same gene was not expressed in a mutant strain that does not produce A-500359s. A 550-bp fragment of the NGDH gene was subsequently cloned and sequenced, and the deduced amino-acid sequence of the cloned region resembled that of other NGDHs cloned to date.

The cloned NGDH fragment was used as a probe leading to the entire A-500359s biosynthetic gene cluster. Approximately 20 000 clones from the genomic library were screened using the DIG-labeled NGDH fragment. The resulting two contiguous cosmids, named ccap2 and ccap13, were isolated and sequenced by the shotgun method. Consequently, the sequenced DNA covered approximately 65 kb, revealing 38 *orf*s (Figure 2), including *orf*22, the NGDH gene. The sequence was analyzed by a comparison with the database, and the predicted functions of 38 ORFs are summarized in Table 1.

The deduced functions of *orfs 9, 11, 12, 13, 22* and *23* are those of 3-ketoreductase, NDP-4-keto-6-deoxy-glucose-2,3-dehydratase, UDP-glucose-4-epimerase, glycosyltransferase, UDP-glucose-4,6-dehydratase and glucose-1-phosphate thymidylyltransferase, respectively. They have been frequently identified in natural product biosynthesis containing a sugar moiety such as that observed in the structure of A-500359s. The deduced function of *orf7* is that of a non-heme, iron-dependent oxygenase. The *orf10* gene encodes a putative clavaminic synthase (also a non-heme, iron-dependent oxygenase), which is a key enzyme in clavulanic acid biosynthesis. The *orf8* gene encodes a putative truncated carbamoyltransferase, and the presumed function of *orf14* is that of a serine hydroxymethyltransferase (SHMT). Although it is well known that SHMT has an important role in amino-acid catabolism, the deduced amino-acid sequence of ORF14 is significantly different from that of SHMTs conserved in *Streptomyces*. The *orf16, -17* and *-18* genes encode a putative CO dehydrogenase large subunit, small subunit and medium subunit, respectively. The *orf21* gene encodes a putative aminoglycoside 3′-phosphotransferase, which is commonly observed as a mechanism of self-resistance and is often found in aminoglycoside-producing strains. The *orf26* and *orf27* genes were deduced to encode non-ribosomal peptide synthetase,

Table 1 Deduced function of ORFs in the A-500359s biosynthetic gene cluster

Protein	*Size*[a]	*Proposed function*	*Sequence similarity (protein, accession no., origin)*	*Identity %/ similarity %*
ORF1	>296	Unknown	Mflv_2879 (YP _001134144) *Mycobacterium gilvum* PYR-GCK	60/70
ORF2	245	Deoxyribonuclease	Mflv_2878 (YP _001134143) *Mycobacterium gilvum* PYR-GCK	58/74
ORF3	897	Endonuclease	SGR_4678 (YP_001826190) *Streptomyces griseus* NBRC 13350	46/61
ORF4	185	Terminal protein	TpgA1 (NP_828744) *Streptomyces avermitilis* MA-4680	87/95
ORF5	744	Telomere-associated protein	TapA1 (NP_828743) *Streptomyces avermitilis* MA-4680	83/86
ORF6	148	Unknown	SCO0007 (NP_624368) *Streptomyces coelicolor* A3(2)	89/92
ORF7	293	Dioxygenase	Aave_3719 (YP_972040) *Acidovorax avenae* subsp. *citrulli* AAC00-1	35/51
ORF8	302	Carbamoyl transferase	Bphy_7715 (YP_001863651) *Burkholderia phymatum* STM815	47/62
ORF9	316	3-Ketoreductase	ChlC4 (AAZ77681) *Streptomyces antibioticus*	30/48
ORF10	324	Clavaminic synthase	CAS (CAA58905) *Streptomyces clavuligerus*	57/70
ORF11	465	NDP-4-keto-6-deoxy-Glc-2,3-dehydratase	Sim20 (AAL15606) *Streptomyces antibioticus*	42/59
ORF12	313	UDP-Glc-4-epimerase	GalE3 (YP_134444) *Haloarcula marismortui* ATCC 43049	32/44
ORF13	383	Glycosyltransferase	SACE_6476 (YP_001108570) *Saccharopolyspora erythraea* NRRL 2338	25/40
ORF14	461	Serine hydroxymethyltransferase	Orf(-4) (AAN85510) *Streptomyces atroolivaceus*	41/59
ORF15	87	Pyrophosphatase	SAV_326 (NP_821500) *Streptomyces avermitilis* MA-4680	52/72
ORF16	759	Carbonmonoxide dehydrogenase	SACE_1162 (YP_001103415) *Saccharopolyspora erythraea* NRRL 2338	58/68
ORF17	168	Carbonmonoxide dehydrogenase	SACE_0536 (YP_001102808) *Saccharopolyspora erythraea* NRRL 2338	72/83
ORF18	289	Carbonmonoxide dehydrogenase	SACE_0538 (YP_001102810) *Saccharopolyspora erythraea* NRRL 2338	56/69
ORF19	606	ABC transporter	SGR_2091 (YP_001823603) *Streptomyces griseus* NBRC 13350	68/78
ORF20	649	ABC transporter	SGR_2092 (YP_001823604) *Streptomyces griseus* NBRC 13350	69/78
ORF21	334	Aminoglycoside phosphotransferase	Strop_0209 (YP_001157072) *Salinispora tropica* CNB-440	64/76
ORF22	321	UDP-Glc-4,6-dehydratase	StaB (BAC55206) *Streptomyces* sp. TP-A0274	70/80
ORF23	356	Glc-1-phosphate thymidylyltransferase	StaA (BAC55207) *Streptomyces* sp. TP-A0274	71/84
ORF24	255	Methyltransferase	MAV_4317 (YP_883454) *Mycobacterium avium* 104	44/56
ORF25	645	*C*-Methyltransferase	PctJ (BAF92592) *Streptomyces pactum*	47/63
ORF26	1087	NRPS	Npun_F2463 (YP_001865967) *Nostoc punctiforme* PCC 73102	29/48
ORF27	469	NRPS	OciB (ABI26078) *Planktothrix agardhii* NIVA-CYA 116	29/45
ORF28	402	β-Lactamase	Pjdr2DRAFT_1210 (ZP_02846103) *Paenibacillus* sp. JDR-2	44/56
ORF29	145	DNA ligase	SSAG_00834 (EDX21043) *Streptomyces* sp. Mg1	54/65
ORF30	258	ABC transporter	SSAG_06370 (EDX26579)	85/90

Table 1 Continued

Protein	*Size*[a]	*Proposed function*	*Sequence similarity (protein, accession no., origin)*	*Identity %/ similarity %*
			Streptomyces sp. Mg1	
ORF31	232	Transposase	SCP1.214 (NP_639820)	74/77
			Streptomyces coelicolor A3(2)	
ORF32	251	Transposase	SGR_6970 (YP_001828482)	94/96
			Streptomyces griseus NBRC13350	
ORF33	161	Unknown	SGR_6969 (YP_001828481)	98/100
			Streptomyces griseus NBRC13350	
ORF34	219	Endonuclease	SGR_6967 (YP_001828479)	100/100
			Streptomyces griseus NBRC13350	
ORF35	428	Transporter	SGR_6966 (YP_001828478)	99/100
			Streptomyces griseus NBRC13350	
ORF36	155	Transcriptional regulator	SGR_6965 (YP_001828477)	100/100
			Streptomyces griseus NBRC13350	
ORF37	132	Regulatory protein	SGR_6963 (YP_001828475)	100/100
			Streptomyces griseus NBRC13350	
ORF38	>182	Alkaline serine protease	SGR_6962 (YP_001828474)	100/100
			Streptomyces griseus NBRC13350	

Abbreviation: ORFs, open reading frames.
[a]Numbers are in amino acids.

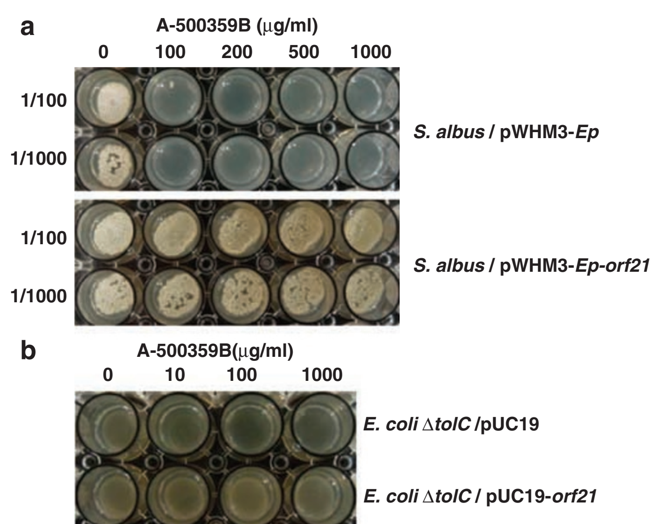

Figure 3 Overexpression of ORF21 in *S. albus* J1074 and *E. coli ΔtolC*. (**a**) *S. albus* J1074, harboring pWHM3-*Ep* and pWHM3-*Ep-orf21,* was incubated on ISP-2 agar containing 0, 100, 200, 500 and 1000 μg ml^{-1} of A-500359B. 1/100 and 1/1000 indicate the dilution ratio of cultured mycelium spotted on the agar plate. (**b**) *E. coli ΔtolC,* harboring pUC19 or pUC19-*orf21,* was incubated on LB agar containing 0, 10, 100 and 1000 μg ml^{-1} of A-500359B with 1 mM IPTG.

which catalyzes peptide bond formation, with the *orf26* gene product consisting of a condensation (C), adenylation (A) and a peptidyl-carrier protein (PCP) domain, and the *orf27* gene product consisting of a C domain. The *orf24* gene encodes a putative *S*-adenosylmethionine (SAM)-dependent methyltransferase, which is well known as a tailoring enzyme in secondary metabolite biosynthesis. The *orf25* gene product belongs to the radical SAM superfamily, which is a group of enzymes that catalyze a wide range of reactions such as protein radicals, sulfur insertion, isomerization, ring formation, oxidation, dehydrogenation and unusual methylation included in various biosyntheses of secondary metabolites.[20] In particular, the function of *orf25* was speculated to be that of *C*-methyltransferase by BLAST analysis. The deduced function of *orf28* gene is that of a β-lactamase, which is often involved in the self-resistance to β-lactam compounds. The *orf36* and *orf37* genes encode putative regulatory factors for A-500359s biosynthesis.

Functional analysis of ORF21

The *orf21* gene product, consisting of 334 amino acids, was deduced to belong to the aminoglycoside 3′-phosphotransferase family, which catalyzes the phosphorylation of aminoglycoside antibiotics and confers resistance. The gene for the aminoglycoside 3′-phosphotransferase is typically found in or near the biosynthetic gene cluster as observed for the aminoglycosides neomycin, ribostamycin, streptomycin and gentamicin, among others.[21] The 3′-hydroxy group of A-500359s is essential for translocase I inhibitory activity (unpublished data), thus consistent with ORF21 catalyzing the phosphorylation of the hexuronic acid moiety of A-500359s as a mechanism for self-resistance.

An efficient transformation system in *S. griseus* SANK60196 has not yet been developed; thus, the functional analysis of *orf21* was heterologously conducted in *S. albus* J1074 and *E. coli ΔtolC*. *S. albus* J1074 does not grow on ISP-2 agar containing more than 100 μg ml^{-1} of A-500359B. On the other hand, *E. coli* MG1655 is resistant to A-500359B at all concentrations tested. Therefore, the gene, *tolC*, encoding a multifunctional outer-membrane channel,[22] was disrupted in *E. coli* MG1655 to yield the mutant strain *E. coli ΔtolC,* which was sensitive to A-500359B and does not grow on LB agar containing 100 μg ml^{-1} of A-500359B. *S. albus* J1074 and *E. coli ΔtolC* strains were thus utilized as test organisms.

S. albus J1074, harboring pWHM3-*Ep* or pWHM3-*Ep-orf21,* was initially analyzed for resistance to A-500359B. *S. albus* J1074/pWHM3-*Ep* did not grow on a TSB agar supplemented with A-500359B. On the other hand, *S. albus* J1074/pWHM3-*Ep-orf21* acquired resistance against A-500359B and grew on ISP-2 agar at concentrations of A-500359B greater than 1000 μg ml^{-1}, as shown in Figure 3a. The *orf*21 gene was next expressed under the control of *lac* promoter in *E. coli ΔtolC*, and *E. coli ΔtolC,* harboring pUC19 or pUC19-*orf21,* was isolated and analyzed for resistance to A-500359B. *E. coli ΔtolC/*

Table 2 Primers for RT-PCR

Primers	*Sequence (5′–3′)*	*Description*
ORF1-RT-f	AAACCACCACCCGATCACG	Forward primer for *orf1*
ORF1-RT-r	AGTGGACCGTTGCGCAGG	Reverse primer for *orf1*
ORF2-RT-f	GCTCGGCAGACGCCCTGG	Forward primer for *orf2*
ORF2-RT-r	GGCGAGGTGAACAATGACG	Reverse primer for *orf2*
ORF3-RT-f	CCCAGGTCGAGCAGGAGC	Forward primer for *orf3*
ORF3-RT-r	GACCGTGCCGCAGGAACC	Reverse primer for *orf3*
ORF4-RT-f	ACGCCGCGGTGCACAAGG	Forward primer for *orf4*
ORF4-RT-r	TCGAACTCAAGGTGCTCG	Reverse primer for *orf4*
ORF5-RT-f	TGGACTGGACGCTCAAGG	Forward primer for *orf5*
ORF5-RT-r	CTCCGAGGTGCGTTTGCC	Reverse primer for *orf5*
ORF6-RT-f	CAGGCTCCGGCACGGCCC	Forward primer for *orf6*
ORF6-RT-r	CAACGTCTCGCCGGCACC	Reverse primer for *orf6*
ORF7-RT-f	CCGAGTGGGAGTTCGTCC	Forward primer for *orf7*
ORF7-RT-r	AGAGAAGGGCTTCCGCTGC	Reverse primer for *orf7*
ORF8-RT-f	CCGGAAGTCCGGCCCGACG	Forward primer for *orf8*
ORF8-RT-r	GTAGCCGGCTCAGTGCTTG	Reverse primer for *orf8*
ORF9-RT-f	GCGGAGGCCACCAACTACGC	Forward primer for *orf9*
ORF9-RT-r	GGTAGGCAGTCGTGAAGCCG	Reverse primer for *orf9*
ORF10-RT-f	GGCTATCTGCTCCTTCGAGG	Forward primer for *orf10*
ORF10-RT-r	GTCGATGATCAGCAGGTCGC	Reverse primer for *orf10*
ORF11-RT-f	ATCTGGTCCAGTACGCCGCG	Forward primer for *orf11*
ORF11-RT-r	GCCTGGACGAGGAAGTGCAG	Reverse primer for *orf11*
ORF12-RT-f	CGCTGGTGATCGACCTCTGC	Forward primer for *orf12*
ORF12-RT-r	CGACGTTGACCGTTGCAGGC	Reverse primer for *orf12*
ORF13-RT-f	ATGACCGACCAACTCATCG	Forward primer for *orf13*
ORF13-RT-r	CCAGGGTCGAGGACCGCAC	Reverse primer for *orf13*
ORF14-RT-f	GCGGAAAGCGGCCACCGC	Forward primer for *orf14*
ORF14-RT-r	GTGCTGGGAGAGTACTCC	Reverse primer for *orf14*
ORF15-RT-f	GTTCCCTGCAGCGACTCG	Forward primer for *orf15*
ORF15-RT-r	ACTCGGATTAGCCGCCG	Reverse primer for *orf15*
ORF16-RT-f	TGCTCGACGACACCCTCC	Forward primer for *orf16*
ORF16-RT-r	GATCCATGCCGATCTCGG	Reverse primer for *orf16*
ORF17-RT-f	TGCGTAAACGGCACGACG	Forward primer for *orf17*
ORF17-RT-r	TCATGTACACGCCTGGCC	Reverse primer for *orf17*
ORF18-RT-f	TGCTTGTCGACATCAACC	Forward primer for *orf18*
ORF18-RT-r	TCGGCGTGGTGCCCTCG	Reverse primer for *orf18*
ORF19-RT-f	ACGGGCACACTGGTGGCG	Forward primer for *orf19*
ORF19-RT-r	AGCCCCTGTTCGCCGACC	Reverse primer for *orf19*
ORF20-RT-f	GCTCCATGCTCGCCTACC	Forward primer for *orf20*
ORF20-RT-r	CGAGCGTCAGGCTGAAGC	Reverse primer for *orf20*
ORF21-RT-f	GCAGAAGCGTACGGTCGCG	Forward primer for *orf21*
ORF21-RT-r	GGGCTGATGCAGGGCGGTG	Reverse primer for *orf21*
ORF22-RT-f	GTCTCCGGTGGCTCCCCGGC	Forward primer for *orf22*
ORF22-RT-r	GGTGACGTGGTAGGAGCGTGC	Reverse primer for *orf22*
ORF23-RT-f	GTCCTCGCAGGAGGTTCCG	Forward primer for *orf23*
ORF23-RT-r	CGAACGGACGTGGATCGGC	Reverse primer for *orf23*
ORF24-RT-f	CCTGAACAGGTCGCCAGAG	Forward primer for *orf24*
ORF24-RT-r	TGGTCGGCGCTCCTTCTCC	Reverse primer for *orf24*
ORF25-RT-f	CTTCAGCGAGGAAGTACCGG	Forward primer for *orf25*
ORF25-RT-r	TGGTGAAGTATTCGTCGCCG	Reverse primer for *orf25*
ORF26-RT-f	AACAGGCTCCCCTGGCAGGC	Forward primer for *orf26*
ORF26-RT-r	TGTCGGCGTTCTCGTAGACG	Reverse primer for *orf26*
ORF27-RT-f	CTGCAAGAACGGCACGAGGC	Forward primer for *orf27*
ORF27-RT-r	GGGTGAATTCTCCCTGCTGG	Reverse primer for *orf27*
ORF28-RT-f	TCTCGCCTTCGAGGGAGC	Forward primer for *orf28*
ORF28-RT-r	CGAGCGGTTCGCAGATCC	Reverse primer for *orf28*
ORF29-RT-f	GTGGGACGGGTATCGGG	Forward primer for *orf29*
ORF29-RT-r	TCACGTCTCCAGCCACTC	Reverse primer for *orf29*
ORF30-RT-f	TCTCTTTCGCACTCTGGC	Forward primer for *orf30*
ORF30-RT-r	CTGCGGTGGCGTACTTGG	Reverse primer for *orf30*
ORF31-RT-f	GGCGGGATCACCGGCAGG	Forward primer for *orf31*
ORF31-RT-r	GGCCCCGGTCCGCAGCCG	Reverse primer for *orf31*
ORF32-RT-f	ACTCTCGCGCCGGGTACC	Forward primer for *orf32*
ORF32-RT-r	TCTTCAACCAGGCCAAGC	Reverse primer for *orf32*
ORF33-RT-f	GTCCACGAGGCCAGGACG	Forward primer for *orf33*
ORF33-RT-r	CAGGCGTCTTCCTCGTCC	Reverse primer for *orf33*
ORF34-RT-f	ACTCGGATTAGCCGCCG	Forward primer for *orf34*
ORF34-RT-r	GCATGTGCAGCTGCTGGG	Reverse primer for *orf34*
ORF35-RT-f	TTGTCTCGATCGGCCAGC	Forward primer for *orf35*
ORF35-RT-r	GCAGCGACAGCCCGCTCC	Reverse primer for *orf35*
ORF36-RT-f	CGAGAACCACGCCTGGACCG	Forward primer for *orf36*
ORF36-RT-r	CTGTGAGGGGGGCGCCATCG	Reverse primer for *orf36*
ORF37-RT-f	GCACCGTCAGGGCTGAGAGC	Forward primer for *orf37*
ORF37-RT-r	GGAGTGGGCGAAGAGTGCCC	Reverse primer for *orf37*
ORF38-RT-f	GTAGGCCATATCCCCGAC	Forward primer for *orf38*
ORF38-RT-r	TCCGGTCGAGGCCCCAGG	Reverse primer for *orf38*
Glk-f	CGGCGGCACGAAGATC	Forward primer for glucokinase
Glk-r	GCGCAGCTTGTTGCCG	Reverse primer for glucokinase

Abbreviation: RT-PCR, reverse transcriptase PCR.

pUC19 did not grow on LB agar supplemented with A-500359B at concentrations ranging from 10 to 1000 μg ml^{-1}. On the other hand, *E. coli ΔtolC*/pUC19-*orf21* was resistant to A-500359B within the same concentration range, as shown in Figure 3b. Importantly, both *S. albus* J1074/pWHM3-*Ep-orf21* and *E. coli ΔtolC*/pUC19-*orf21* were not resistant to other tested aminoglycoside antibiotics, neomycin, kanamycin, G418, apramycin, gentamicin and streptomycin (data not shown). Therefore, it was concluded that ORF21 confers self-resistance to A-500359B and supports the fact that the A-500359s biosynthetic gene cluster has been cloned.

Expression analysis of A-500359s biosynthetic genes

A series of *S. griseus* SANK60196 mutants: an A-500359s high-producing strain (strain HP), an A-500359s low producer (strain 37-3) and A-500359s non-producers (strains 35-4 and 37-9) (unpublished data) were utilized to test the expression levels of *orfs* of the cloned region at various time points during A-500359s production. The production of A-500359A from strain HP was about 100-fold higher than that of strain 37-3 in a 7-day culture broth, and the gene expression pattern of A-500359s biosynthetic genes was compared between these two strains. The total RNA was prepared from 4- and 7-day-cultured mycelia and was used as a template for RT-PCR analysis. Specific oligonucleotide primers were designed to amplify all 38 *orfs*, as shown in Table 2. *glk*, used as a control, was expressed in all the tested strains and no differences in the *glk* expression were detected. Most of the genes that are likely to be involved in A-500359s biosynthesis, such as *orf7–14*, *-17*, *-18*, *-21–28*, *-30* and *-37*, were expressed in strains HP and 37-3. However, their expression levels in strain HP were significantly higher than those in strain 37-3. In addition, the expression of these same *orfs* could not be detected in the non-producing strains 35-4 and 37-9 (Figure 4).

DISCUSSION

We identified an NGDH gene, *orf*22, which is expressed in the A-500359s high-producing strain HP and not expressed in a mutant strain devoid of A-500359s production. Using this cloned NGDH gene, two contiguous cosmids (ccap2 and ccap13) were isolated from around 20 000 clones of a genomic library using the DIG

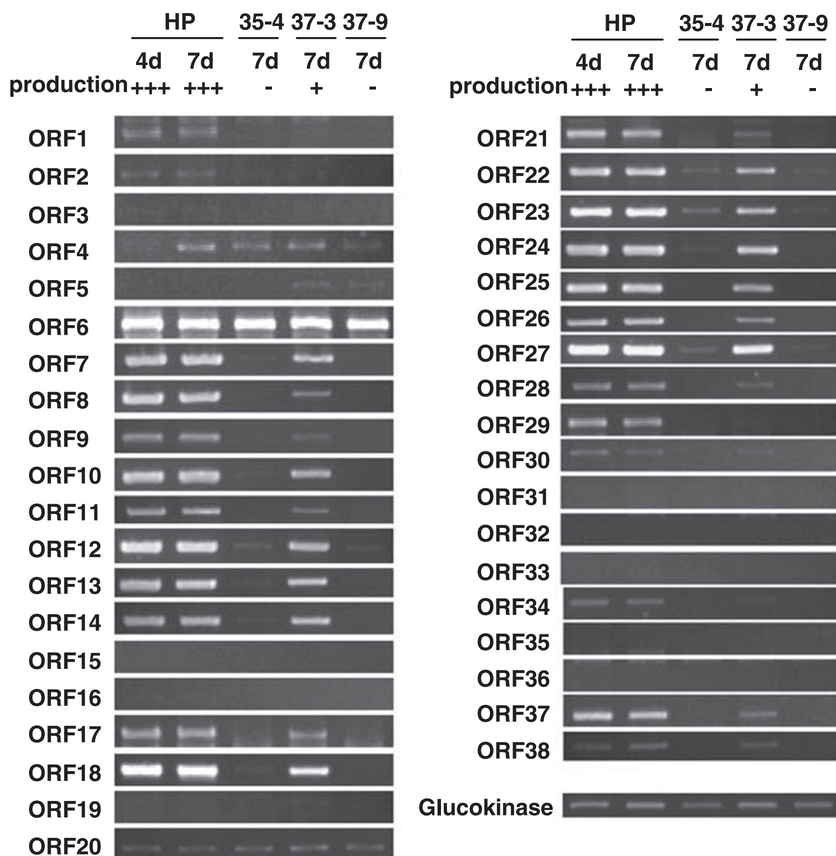

Figure 4 Gene expression analysis of 38 *orfs* in strains HP, 35-4, 37-3 and 37-9 by RT-PCR. HP, 35-4, 37-3 and 37-9 indicate the derivatives of the A-500359s producer. 4d and 7d are total RNA samples isolated from 4- and 7-day-cultured mycelia of *S. griseus* SANK60196 derivatives, respectively.

system. Sequencing analysis by the shotgun method identified 38 ORFs within the span of the 65-kb region, as shown in Figure 2 and Table 1.

To confirm the relationship between the cloned region and A-500359s biosynthesis, *orf21*, encoding a predicted aminoglycoside 3′-phosphotransferase, was heterologously expressed in *S. albus* J1074 and in *E. coli ΔtolC* to confirm its function as a mechanism for A-500359s resistance. Genes that confer self-resistance are frequently found in regions clustered with biosynthetic and regulatory genes, and this includes many examples of natural products the producing organisms of which utilize a phosphotransferase for self-resistance. For example, the neomycin resistance gene, *neo1* (AAA26699), from *Streptomyces fradiae*,[23] the ribostamycin resistance gene, *rph* (AJ748131), from *Streptomyces ribosidificus*,[24] the streptomycin resistance gene, *aphD* (AJ862840), from *S. griseus*[25] and the hygromycin resistance gene, *hyg21* (DQ314862), from *Streptomyces hygroscopicus*[26] are representative genes encoding phosphotransferases that confer selective resistance to the respective aminoglycoside. In addition to these aminoglycosides, the biosynthetic gene cluster of certain cyclic peptides also contain a gene encoding a phosphotransferase that confers resistance, including the viomycin resistance gene, *vph* (AY263398), from *Streptomyces vinaceus*[27] and the capreomycin resistance gene, *cph* (U13078), from *Streptomyces capreolus*.[28] As expected, the overexpression of *orf21* under a strong promoter permitted *S. albus* J1074 and *E. coli ΔtolC* to grow on agar plates and liquid media containing high levels of A-500359B (Figure 3), and significantly, the resistance was selective for A-500359B. This result strongly suggests that the 65-kb region including *orf21* is responsible for A-500359s biosynthesis and production.

To provide additional evidence that the A-500359s gene cluster was cloned, RT-PCR analysis was used to show that *orfs 7–14*, *17*, *18*, *21–30*, *34*, *36* and *37* were expressed in strains HP and 37-3, which were A-500359s high and low producers, respectively, but were not expressed in strains 35-4 and 37-9, which were A-500359s non-producers (Figure 4). Thus, the data support the fact that these *orf*s are required for A-500359s production. The minimal genes required for A-500359s biosynthesis were also deduced using RT-PCR and bioinformatics analyses. The A-500359s gene cluster is proposed to be contained within *orfs 7–30* and, in addition, *orf34* has an unknown function (a putative endonuclease), and *orf*36 and *orf*37 act as regulatory factors. Thus, it is proposed that the A-500359s gene cluster consists minimally of 26 *orf*s, with 18 *orf*s involved in biosynthesis, 6 *orf*s involved in resistance, regulation and transport (*orf*s 19–21, 30, 36 and 37), and 2 *orf*s of unclear function (*orf*s 29 and 34).

In conclusion, we have identified and cloned a gene cluster involved in the resistance and likely biosynthesis of A-500359s. Although the locus was identified using a probe for NGDH, it remains to be seen what role *orf*22 plays in A-500359s assembly. Despite this unknown function, we have shown that expression of the ORFs within this genetic locus is highly correlated with the production of A-500359s, and a gene (*orf*21) located within the cloned locus has been shown to confer selective resistance to A-500359s. Cloning of the A-500359s

gene cluster now permits a thorough functional characterization of the genes involved in A-500359s biosynthesis, which will ultimately facilitate combinatorial biosynthetic methods to prepare novel compounds and expand the molecular diversity of both natural and unnatural nucleoside antibiotics.

ACKNOWLEDGEMENTS
We thank Dr Ben Shen for his kind gifts in this work.

1 Takeuchi, S., Hirayama, K., Ueda, K., Sakai, H. & Yonehara, H. Blasticidin S, a new antibiotic. *J. Antibiot.* **11,** 1–5 (1958).
2 Iwasa, T., Suetomi, K. & Kusaka, T. Taxonomic study and fermentation of producing organism and antimicrobial activity of mildiomycin. *J. Antibiot.* **31,** 511–518 (1978).
3 Worthington, P. A. Antibiotics with antifungal and antibacterial activity against plant diseases. *Nat. Prod. Rep.* **5,** 47–66 (1988).
4 Kimura, K. & Bugg, D. H. Recent advances in antimicrobial nucleoside antibiotics targeting cell wall biosynthesis. *Nat. Prod. Rep.* **20,** 252–273 (2003).
5 Bormann, C., Möhrle, V. & Bruntner, C. Cloning and heterologous expression of the entire set of structural genes for nikkomycin synthesis from *Streptomyces tendae* Tü901 in *Streptomyces lividans*. *J. Bacteriol.* **178,** 1216–1218 (1996).
6 Chen, W. *et al.* Characterization of the polyoxin biosynthetic gene cluster from *Streptomyces cacaoi* and engineered production of polyoxin H. *J. Biol. Chem.* **284,** 10627–10638 (2009).
7 Lacalle, R. A., Tercero, J. A. & Jimenez, A. Cloning of the complete biosynthetic gene cluster for an aminonucleoside antibiotic, puromycin, and its regulated expression in heterologous hosts. *EMBO J.* **11,** 785–792 (1992).
8 Fernández-Moreno, M. A., Vallín, C. & Malpartida, F. Streptothricin biosynthesis is catalyzed by enzymes related to nonribosomal peptide bond formation. *J. Bacteriol.* **179,** 6929–6936 (1997).
9 Cone, M. C., Yin, X., Grochowski, L. L., Parker, M. R. & Zabriskie, T. M. The blasticidin S biosynthesis gene cluster from *Streptomyces griseochromogenes*: sequence analysis, organization, and initial characterization. *Chembiochem* **4,** 821–828 (2003).
10 McCarty, R. M. & Bandarian, V. Deciphering deazapurine biosynthesis: pathway for pyrrolopyrimidine nucleosides toyocamycin and sangivamycin. *Chem. Biol.* **15,** 790–798 (2008).
11 Li, L. *et al.* The mildiomycin biosynthesis: initial steps for sequential generation of 5-hydroxymethylcytidine 5′-monophosphate and 5-hydroxymethylcytosine in *Streptoverticillium rimofaciens* ZJU5119. *Chembiochem* **9,** 1286–1294 (2008).
12 Ohnuki, T., Muramatsu, Y., Miyakoshi, S., Takatsu, T. & Inukai, M. Studies on novel bacterial translocase I inhibitors, A-500359s. IV. Biosynthesis of A-500359s. *J. Antibiot.* **56,** 268–279 (2003).
13 Sambrook, J. & Russell, D. W. *Molecular Cloning: A Laboratory Manual* 3rd edn, (Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2001).
14 Weber, J. M., Wierman, C. K. & Hutchinson, C. R. Genetic analysis of erythromycin production in *Streptomyces erythreus*. *J. Bacteriol.* **164,** 425–433 (1985).
15 Shen, B & Hutchinson, CR Deciphering the mechanism for the assembly of aromatic polyketides by a bacterial polyketide synthase. *Proc. Natl Acad. Sci. USA* **93,** 6600–6604 (1996).
16 Muramatsu, Y. *et al.* Studies on novel bacterial translocase I inhibitors, A-500359s. V. Enhanced production of capuramycin and A-500359A in *Streptomyces griseus* SANK60196. *J. Antibiot.* **59,** 601–606 (2006).
17 Decker, H. *et al.* A general approach for cloning and characterizing dNDP-glucose dehydratase gene from actinomycetes. *FEMS Microbiol. Lett.* **141,** 195–201 (1996).
18 Link, A. J., Phillips, D. & Church, G. M. Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J. Bacteriol.* **179,** 6228–6237 (1997).
19 Kieser, T., Bibb, M., Buttner, M., Chater, K. F. & Hopwood, D. A. *Practical Streptomyces Genetics* (The John Innes Foundation: Norwich, UK, 2000).
20 Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F. & Miller, N. E. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res.* **29,** 1097–1106 (2001).
21 Flatt, P. M. & Mahmud, T. Biosynthesis of aminocyclitol–aminoglycoside antibiotics and related compounds. *Nat. Prod. Rep.* **24,** 358–392 (2007).
22 Koronakis, V. TolC-the bacterial exit duct for proteins and drugs. *FEBS Lett.* **555,** 66–71 (2003).
23 Huang, F. *et al.* The neomycin biosynthetic gene cluster of *Streptomyces fradiae* NCIMB 8233: characterisation of an aminotransferase involved in the formation of 2-deoxystreptamine. *Org. Biomol. Chem.* **3,** 1410–1418 (2005).
24 Subba, B. *et al.* The ribostamycin biosynthetic gene cluster in *Streptomyces ribosidificus*: comparison with butirosin biosynthesis. *Mol. Cells* **20,** 90–96 (2005).
25 Distler, J. *et al.* Gene cluster for streptomycin biosynthesis in *Streptomyces griseus*: nucleotide sequence of three genes and analysis of transcriptional activity. *Nucleic Acids Res.* **15,** 8041–8056 (1987).
26 Palaniappan, N., Ayers, S., Gupta, S., Habib, e. I.-S. & Reynolds, K. A. Production of Hygromycin A analogs in *Streptomyces hygroscopicus* NRRL 2388 through identification and manipulation of the biosynthetic gene cluster. *Chem. Biol.* **13,** 753–764 (2006).
27 Yin, X., O'Hare, T., Gould, S. J. & Zabriskie, T. M. Identification and cloning of genes encoding viomycin biosynthesis from *Streptomyces vinaceus* and evidence for involvement of a rare oxygenase. *Gene* **312,** 215–224 (2003).
28 Felnagle, E. A., Rondon, M. R., Berti, A. D., Crosby, H. A. & Thomas, M. G. Identification of the biosynthetic gene cluster and an additional gene for resistance to the antituberculosis drug capreomycin. *Appl. Environ. Microbiol.* **73,** 4162–4170 (2007).