Available online at www.sciencedirect.com
SCIENCE DIRECT®
Archives of Biochemistry and Biophysics xxx (2004) xxx–xxx
ABB
www.elsevier.com/locate/yabbi

A gene cluster for biosynthesis of kanamycin from *Streptomyces kanamyceticus*: comparison with gentamicin biosynthetic gene cluster

Madan K. Kharel,[a] Bimala Subba,[a] Devi B. Basnet,[a] Jin Suk Woo,[b] Hei Chan Lee,[a] Kwangkyoung Liou,[a] and Jae Kyung Sohng[a,*]

[a] Institute of Biomolecule Reconstruction, SunMoon University, Asan, Chung-nam, 336-708, South Korea
[b] GeneChem Inc. 461-17, Jeonmin-Dong, Yusong-Gu, Daejon, 305-390, South Korea

Received 30 April 2004, and in revised form 7 June 2004

Abstract

Gene clusters for the biosynthesis of kanamycin (Km) and gentamicin (Gm) were isolated from the genomic libraries of *Streptomyces kanamyceticus* and *Micromonospora echinospora*, respectively. The sequencing of the 47 kb-region of *S. kanamyceticus* genomic DNA revealed 40 putative open reading frames (ORFs) encoding Km biosynthetic proteins, regulatory proteins, and resistance and transport proteins. Similarly, the sequencing of 32.6 kb genomic DNA of *M. echinospora* revealed a Gm biosynthetic gene cluster flanked by resistant genes. Biosynthetic pathways for the formation of Km were proposed by the comparative study of biosynthetic genes. Out of 12 putative Km biosynthetic genes, *kanA* was expressed in *Escherichia coli* and determined its function as a 2-deoxy-*scyllo*-inosose synthase. Furthermore, the acetylations of aminoglycoside–aminocyclitols (AmAcs) by Km acetyltransferase (KanM) were also demonstrated. The acetylated derivatives completely lost their antibacterial activities against *Bacillus subtilis*. The comparative genetic studies of Gm, Km, tobramycin (Tm), and butirosin (Bn) reveal their similar biosynthetic routes and provide a framework for the further biosynthetic studies.

Keywords: Kanamycin; Gentamicin; 2-Deoxystreptamine; *Streptomyces kanamyceticus*; *Micromonospora echinospora*

Aminoglycoside–aminocyclitols (AmAcs)[1] constitute a large family of broad-spectrum antibacterial compounds that have been proven to be efficient for the treatment of severe infections caused mostly by *staphylococci* and *enterococci*. They were among the earliest to reach clinical usage [1]. The specific interaction of AmAcs to 16S RNA of 30S bacterial ribosomal subunit interferes with the protein synthesis process, which eventually leads to cell death [2,3]. Despite the emergence of several pathogenic bacteria resistant to these compounds, their potent bactericidal activity and their ability to interact synergistically with other antimicrobial agents have sustained their clinical utility. Combinations of such AmAcs with β-lactams are still preferred for the treatment of nosocomial infections caused by the opportunistic pathogens, such as *Pseudomonas* spp. [4–6]. Alternatively, many efforts have been carried out to develop the semisynthetic analogues of naturally existing Ams antibiotics with better antibacterial properties [7–10]. Dibekacin (Dbk), amikacin (Amk), nitilmicin (Ntl), and isepamicin (Ipm) are the typical examples of such semisynthetic analogues.

Despite 2-deoxystreptamine (DOS)-containing AmAcs constituting the largest group of clinically important antibiotics, e.g., Gm, km, tobramycin (Tm), neomycin (Nm), ribostamycin (Rm), etc., genetic and biochemical studies of these compounds are limited and still provide an incomplete view of these processes.

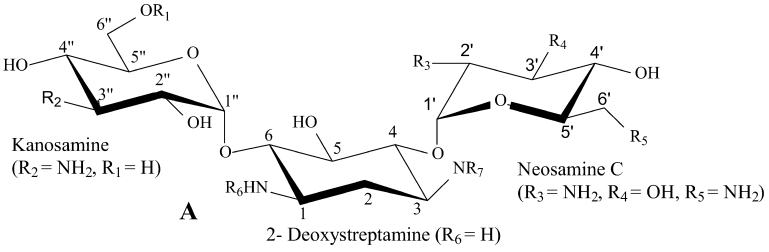
* Corresponding author. Fax: +82-41-544-2919.
E-mail address: sohng@sunmoon.ac.kr (J.K. Sohng).

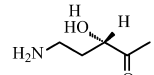
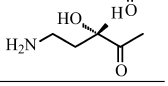
[1] *Abbreviations used:* AmAc, aminoglycoside–aminocyclitol; Km, kanamycin; Spm, spectinomycin; Gm, gentamicin; Tm, tobramycin; Bn, butirosin; Amk, amikacin; Dbk, dibekacin; Ntl, nitilmicin; Nm, neomycin; Rm, ribostamycin; Ipm, isepamicin; DOS, 2-deoxystreptamine; ORF, open reading frame; G-6-P, glucose-6-phosphate; DOI, 2-deoxy-*scyllo*-inosose; DHQ, 3-dehydroquinate; LBBS, LB medium containing 2.5 mM betaine and 1 M sorbitol; IPTG, isopropyl-β-D-thiogalactopyranoside; PLP, pyridoxal phosphate.

0003-9861/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.abb.2004.06.009

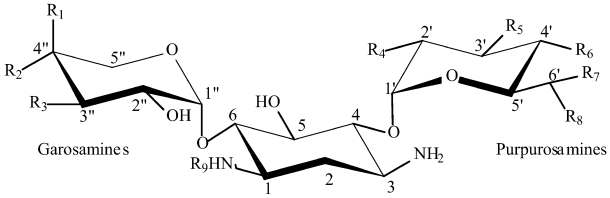
Recent isolation of a gene cluster for butirosin (Bn) biosynthesis from *Bacillus circulans* and a partial gene cluster for Tm biosynthesis from *Streptomyces tenebrarius* represent the sole genetic studies on these classical antibiotics reported so far [11,12]. However, there are several reports of genetic and biochemical analyses on streptamine-containing AmAcs, especially on streptomycin [13–18]. Gms (Gm A, B, and C series) and Kms (Km A, B, and C), the synthetic precursor of several semisynthetic analogues (Ipm, Amk, Dbk, etc.) are isolated from the culture broths of *M. echinospora* and *S. kanamyceticus*, respectively. Such diversities in the natural products of these organisms spurred us to investigate their biosynthesis and resistance.

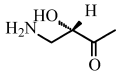
The DOS subunit of Km substitutes at positions 4 and 6 for neosamine and kanosamine subunits whereas Gms contain purpurosamines and garosamines (Fig. 1). Biosynthetic studies on Bn and Tm have demonstrated the beginning of DOS biosynthesis with the formation of carbocycle through the activity of 2-deoxy-*scyllo*-inosose (DOI) synthase utilizing glucose-6-phosphate (G-6-P) [19]. The L-glutamine-dependent aminotrans-



Tobramycin and kanamycin derivatives	Functional groups					
	R_1	R_2	R_3	R_4	R_5	R_6
Nebramycin factor 4'	$CONH_2$	NH_2	NH_2	OH	NH_2	H
Nebramycin factor 5'	$CONH_2$	NH_2	NH_2	H	NH_2	H
Nebramycin factor 6' (Tobramycin)	H	NH_2	NH_2	H	NH_2	H
Kanamycin A	H	NH_2	OH	OH	NH_2	H
Kanamycin B	H	NH_2	NH_2	OH	NH_2	H
Kanamycin C	H	NH_2	NH_2	OH	OH	H
[a]Amikacin	H	NH_2	OH	OH	NH_2	
[a]Arbekacin	H	NH_2	NH_2	OH	NH_2	

[a]Semisynthetic antibiotics derived from kanamycin.



Gentamicins	Functional groups								
	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9
Gentamicin C series	OH	CH_3	$NHCH_3$	NH_2	H	H	CH_3, H	$NHCH_3$, NH_2	H
Gentamicin B series	OH	CH_3	$NHCH_3$	OH	OH	OH	H	NH_2	H
Gentamicin A series	OH, H	OH, H	$NHCH_3$,OH, $N(CHO)CH_3$	NH_2. OH	OH	OH	H	NH_2, OH	H
[a]Isepamicin	OH	CH_3	$NHCH_3$	OH	OH	OH	H	NH_2	

[a]Semisynthetic derivative of gentamicin B

Fig. 1. Structures of 4,6-disubstituted aminoglycoside–aminocyclitol antibiotics. A and B represent kanamycin and gentamicin derivatives, respectively.

ferase catalyzes the transamination of DOI to form *scyllo*-inosamine [20]. The rest of the biosynthetic enzymes are yet to be investigated.

Here, we report the isolation of Km and Gm biosynthetic gene clusters from *S. kanamyceticus* and *M. echinospora*, respectively. A pathway for the biosynthesis of Km is also proposed using a comparative sequence analysis with the Gm, Tm, and Bn biosynthetic genes. The functions of KanA and KanM have been determined as DOI synthase and Km acetyltransferase, respectively, by in vitro assay of the overexpressed proteins.

Experimental procedures

Microorganisms, culture conditions, and vectors

The Km and Gm producing-strains *S. kanamyceticus* ATCC12853 and *M. echinospora* ATCC15835 were grown in the ISP2 (Difco, USA) and N-Z amine A (glucose, 10 g; soluble starch 20 g; yeast extract 5 g; N-Z-amine 5 g; and calcium carbonate 5 g in 1 L) media or on respective agar plates at 28 °C. *B. circulans* IFO 13157 was cultured in the TSB medium at 28 °C as described previously [21]. *E. coli* XL1 blue MRF (Stratagene, USA) was used as host for constructing cosmid libraries and manipulating recombinant plasmids. His-tagged proteins were expressed in *E. coli* BL21(DE3)pLysS (Stratagene, USA). pOJ446 was used for constructing genomic libraries, whereas pET-32a(+) (Novagen, USA) without the thioredoxin fusion protein and pRSETC (Invitrogen, USA) were used for expressing genes. *E. coli* strains were grown in the Luria–Bertani (LB) or LB containing 2.5 mM betaine and 1 M sorbitol (LBBS) [22] supplemented with ampicillin (100 μg/ml) and chloramphenicol (25 μg/ml) for the selection or maintenance of plasmids.

DNA manipulations

General genetic procedures such as genomic and plasmid DNA isolations, restriction endonuclease digestions, alkaline phosphatase treatments, DNA ligations, and other DNA manipulations were performed according to the standard protocols [23]. For the construction of the genomic library, the genomic DNA of *M. echinospora* and *S. kanamyceticus* was partially digested with diluted *Sau*3AI over various times (0.5–5 min), and the aliquots were analyzed by agarose gel electrophoresis. The cleaved sample containing 35–45 kb fragments was pooled, treated with alkaline phosphatase, and then ligated to the pOJ446 digested with *Bam*HI and *Hpa*I. In vitro packaging was carried out using Gigapack III XL packaging extract (Stratagene, USA).

Screening of the cosmid library

To isolate gene clusters for Km and Gm biosynthesis, cosmid libraries were constructed and then screened with the two different probes. The partial sequences of L-glutamine: DOI aminotransferase (260 bp) obtained from the genomic DNA of *S. kanamyceticus* and *M. echinospora* by the homology-based PCR primers were used as the first probe as described previously [11]. For generating the second probe, internal nucleotide sequences of Km resistance ribosomal RNA methyltransferase (GenBank Accession No. Y15838) were obtained from genomic DNA of *S. kanamyceticus* with the primers KAN-c (5′-ATGTCGCAGTCCGCGTCC GACGAGGAC-3′) and KAN-d (5′-TCAGCCCTTCG TGACCACGTAGACGAG-3′). Similarly, the internal nucleotide sequence of Gm resistance ribosomal RNA methyl transferase was also obtained. Each probe was labeled with [^{32}P]dCTP (Perkin–Elmer Life Sciences, USA) using the random primer labeling kit (Stratagene, USA), purified by gel filtration, and used to screen the cosmid libraries of *M. echinospora* and *S. kanamyceticus*. Hybridization was carried out for 6 h with each probe at 65 °C in 10 ml of 2× SSC [23]. Sequencing was carried out in an automatic sequencer by the dideoxy chain termination method and the data were assembled using the DNA Star program package (DNASTAR, Madison, WI). Potential open reading frames were identified using FramePlot [24] and searched for homologies using the BLAST server [25]. Phylogenetic analyses of amino acid sequences were carried out using *MEGA* version 2.1 [26].

Construction of the plasmids expressing KanA, BtrC, and KanM

Plasmids were constructed to produce DOI synthase (KanA) fused with histidine (His_6) at the amino-terminus. The *kanA* (1.17 kb) obtained from the pSKC2 with primers DOI-a (5′-CACGCCTTCAGGAATTCGGAC AGCAT-3′) and DOI-b (5′-CCGACGCGAAGCTTG GCGCTCA-3′) was cloned into *Eco*RI- and *Hin*dIII-digested pRSET C resulting in the formation of pDOI-1. The construction of recombinant pDOI-2 for the expression of *btrC* (1.1 kb) was described previously [11]. The thioredoxin (Trx) tag was removed from pET-32a(+) deleting the 0.32 kb-*Nde*I fragment to construct pET-32a(ΔTrx). The *kanM* (0.44 kb), amplified from the cosmid with primers KMR-a (5′-GAGGAATTCCTG ATGCGGCCCGTGCTCGA-3′) and KMR-b (5′-AG GAAGCTTGGGTCGGACGGGCCGACCCG-3′), was cloned into *Eco*RI- and *Hin*dIII-digested pET-32a(ΔTrx) to form pKMR1. PCR was performed in a thermocycler (Takara, Japan) under the following conditions: 25 cycles of 30 s at 95 °C, 1 min at 55 °C, and 1 min at 72 °C. All of the PCR products were cloned into

pGEM-T vector (Promega, USA) and sequenced prior to cloning into the expression vectors to verify that no mutation had been introduced during the PCR amplification. All of these expression constructs were introduced into *E. coli* BL21(DE3)pLysS by the heat pulse transformation, and ampicillin- and chloramphenicol-resistant transformants were selected.

Gene expression

For the expression of KanA, *E. coli* BL21(DE3)-pLysS/pDOI-1 was grown in 10 ml LB medium containing ampicillin and chloramphenicol at 37 °C and 250 rpm for 8 h. The culture was then transferred to 600 ml LBBS medium and incubated at 37 °C and 250 rpm. At an OD_{600} of 0.6, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.4 mM and the incubation was continued at 30 °C for 20 h. The cells were harvested by centrifugation at 6000*g* for 10 min and stored at −20 °C for 6 h. Instead of LBBS medium, KanM was expressed in LB medium following the induction with IPTG (0.4 mM) at 20 °C for 12 h.

Enzyme purification and assay

For the purification of KanM, cell pellets were treated with 12 ml sodium phosphate buffer (50 mM, pH 7.5), thoroughly mixed and sonicated for 12 s and then debris was removed by centrifugation at 12,000*g* for 20 min. The protein was purified by Ni^{2+}-affinity chromatography (Invitrogen, USA) and the concentration was determined following the Bradford assay procedure [27]. The cell-free extract for the crude KanA was prepared by the sonication of thawed cell pellets and then centrifugation at rpm of 12,000*g* followed by the dialysis of soluble fractions for 8 h at 4 °C. The concentration of $CoCl_2$ was maintained at 0.1 mM in the suspension buffer throughout the KanA manipulations.

The *btrC* from *B. circulans* was expressed in *E. coli* and the crude extract was used to synthesize DOI in vitro according to the method reported by Kakinuma and co-workers [19] with the slight modifications. The reaction was carried out in 20 ml containing 50 mM NaH_2PO_4 (pH 7.5), 2.5 mM NAD^+, 2 mM $CoCl_2$, and 1 ml crude BtrC. The mixture was incubated at 40 °C for 30 min and heating at 80 °C for 5 min quenched the reaction. The supernatant obtained by centrifugation was treated with 320 μl *O*-(4-nitrobenzyl)hydroxylamine hydrochloride solution in pyridine (250 mg/ ml) and further incubated at 72 °C for 2 h. The mixture was dried under reduced pressure and separated by the silica gel column chromatography. Pure DOI-oxime derivative was eluted with a mixture of methanol and chloroform (1:16). The assay for KanA was carried out under identical conditions where the crude BtrC was replaced by KanM. The formation of oxime derivative was detected by TLC (MercK, Germany) and HPLC (SHIMADZU, Japan) at 362 nm. For the HPLC analysis, isocratic elution was carried out with methanol and water (3:7) using C-18 column (MIGHTYSIL –RP-18, Japan) at the flow rate of 1 ml/min at 30 °C. The activity of KanM was studied in 100 μl of 50 mM Tris–HCl buffer (pH 7.5) containing 1 mM acetyl coenzyme A (acetyl-CoA), 0.5 mM AmAc antibiotic, and 31.2 μg KanM. The acetylation was monitored on TLC. The antibacterial susceptibilities of AmAcs and their acetylated derivatives against *B. subtilis* were assayed on the soft agar.

Results

Screening of the cosmid libraries and isolation of biosynthetic gene clusters

The ^{32}P-labeled core sequences of L-glutamine: DOI aminotransferase were used for the screening of the cosmid library of *S. kanamyceticus* and revealed 11 positive clones. Considering the general fact that the clustered organizations of antibiotic biosynthetic genes are often associated with resistance genes, the partial sequence of Km resistance ribosomal RNA methyltransferase from *S. kanamyceticus* was used for further screening. Three colonies were hybridized with the second probe for *S. kanamyceticus* whereas a single colony harboring a cosmid (designated as pGEN01) was isolated for *M. echinospora* following the similar procedures. The restriction analysis revealed over 36 kb insertion in each cosmid. Out of three positive cosmids, pSKC1 and pSKC2 were sequenced. Similarly, the cosmid, pGEN01, was also sequenced. The nucleotide sequences discussed in this paper are deposited in the EMBL nucleotide sequence database with Accession Nos. AJ582817 and AJ575934 for Km and Gm clusters, respectively.

Nucleotide sequence analysis of kanamycin biosynthetic gene cluster

A total of 46.9 kb of DNA sequence was determined in the cosmids pSKC1 and pSKC2 with a significant overlapping region (16.7 kb, Fig. 2). The sequence was analyzed for the putative ORFs using the FramePlot and ORF Finder programs and the deduced amino acid sequences were compared to the proteins in various databases using the BLAST program. The analysis revealed 40 ORFs organized in a cluster including putative kanamycin biosynthetic genes, resistance and regulatory genes (Fig. 2). Most of the Km biosynthetic genes were well oriented in the same direction and were flanked by two resistance genes (*kanM* and *kmr*). A similar cluster

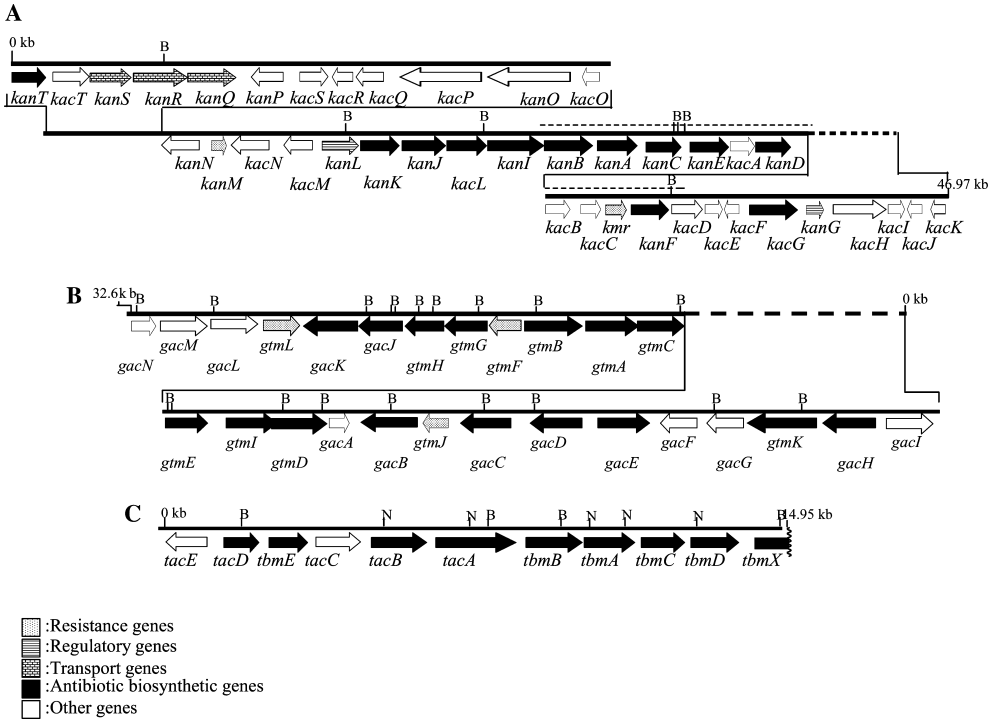


Fig. 2. The biosynthetic gene clusters for 4,6-disubstituted aminoglycoside–aminocyclitols: (A) kanamycin biosynthesis, (B) gentamicin biosynthesis, and (C) tobramycin biosynthesis. The dotted-thin line in the Km cluster indicates the overlapped region in the pSKC1 and pSKC2.

was found for Gm biosynthesis but some of the biosynthetic genes were transcribed in the opposite direction (Fig. 2). All of the ORFs detected showed typical biased codon usage (mean G + C content of 70.9 and 70.2% for the Km and Gm clusters, respectively) of actinomycetes genes [28].

2-Deoxystreptamine biosynthetic genes

Out of several Km biosynthetic genes, *kanA*, *B*, *C*, and *D* probably participate in the DOS biosynthesis. The *kanA* product (M_r 45.83 kDa and p*I* 5.73) would code for 2-deoxy-*scyllo*-inosose synthase that catalyzes a key reaction of the DOS biosynthesis. This reaction involves the formation of DOI, **2**, from G-6-P, **1**, in the presence of NAD^+ and Co^{2+} cofactors (Fig. 3). The amino acid sequence of KanA showed significant similarity to that of TbmA, a recently characterized DOI synthase from *S. tenebrarius* [11]. Fewer identities (<40%) of amino acids were reported to the 3-dehydroquinate (DHQ) synthase and amino DHQ synthase involved in the biosynthesis of shikimic acid [29] and aminohydroxybenzoic acid [30], respectively. The sequence ^{103}GGGVTGNIAGL^{113} agrees with the consensus NAD^+-binding motif GXXGXXXG found in 3-dehydroquinate (DHQ) synthase. The formation of DOI follows the transamination reaction (catalyzed by BtrS) to form *scyllo*-inosamine, **3**, in the butirosin biosynthetic pathway [20]. A BtrS-homologous protein, KanB (M_r. 45.83 kDa, p*I* 5.73), was similar to the TbmB (74%) from the tobramycin biosynthetic gene cluster. The StsC (GenBank Accession No. CAA70012) and SpcS2 (GenBank Accession No. AAD28492) are the other homologous proteins (*scyllo*-inosose aminotransferases) involved in streptamine-containing AmAcs biosynthesis [31,32]. All of these proteins utilize L-glutamine as amino source and pyridoxal phosphate (PLP) as a co-factor. Here, it is important to note that none of the further DOS biosynthetic genes have been cloned and characterized so far.

KanC encodes a polypeptide (38.16 kDa, p*I* 5.42), which is 51% identical to the TbmC from the Tm biosynthetic gene cluster (Table 1). The other homologues include putative *myo*-inositol dehydrogenase (MDH) from various sources including *S. griseus* (GenBank Accession No. PO9400) and *Rhizobium* sp. NGR234 (GenBank Accession No. AAQ87377). *Myo*-inositol dehydrogenase catalyzes the reversible oxidation–reduction of *myo*-inositol and *scyllo*-inosose in the presence of NAD^+. The BtrS homologues KanB and KanD were 36% identical to each other pointing to their similar functions. The existence of GtmD and GtmB is another such example in Gm biosynthesis, instead of the single aminotransferase reported in butirosin biosynthetic gene cluster so far. Because the DOS subunit contains two amino functional groups, dehydrogenation of **3** by KanC and its subsequent transamination by KanD are predictable (Fig. 3). Similarly, GtmA, B, C,

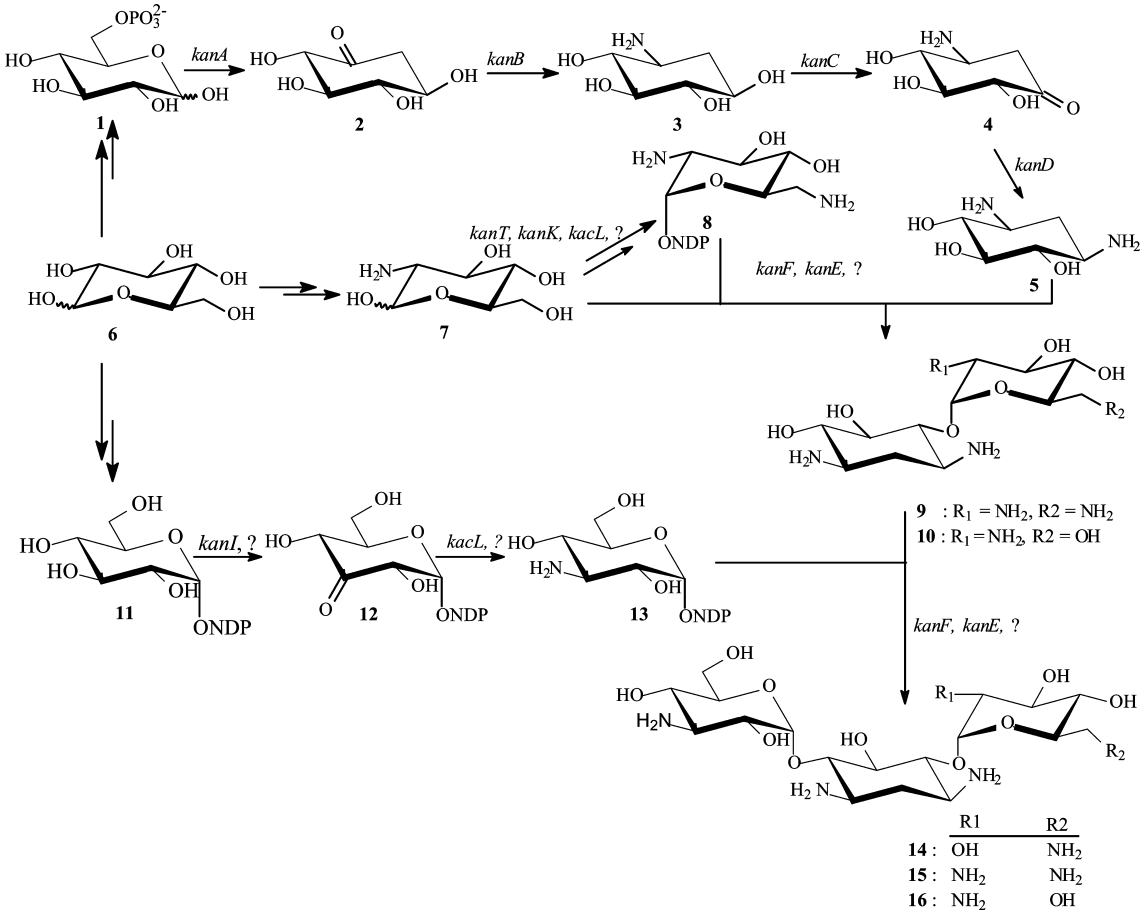


Fig. 3. Proposed pathway for kanamycin biosynthesis. (**1**) Glucose-6-phosphate, (**2**) 2-deoxy-*scyllo*-inosose, (**3**) 2-deoxy-*scyllo*-inosamine, (**4**) 2-deoxy-3-amino-*scyllo*-inosose, (**5**) 2-deoxystreptamine, (**6**) glucose, (**7**) 2-amino-2-deoxyglucose, (**8**) NDP-2,6-diamino-2,6-dideoxyglucose, (**9**) neamine, (**10**) paromamine, (**11**) UDP-glucose, (**12**) UDP-3-keto-glucose, (**13**) UDP-3-amino-3-deoxyglucose (UDP-kanosamine), (**14**) kanamycin A, (**15**) kanamycin B, and (**16**) kanamycin C.

and D might be involved in DOS biosynthesis in *M. echinospora*.

Biosyntheses of kanosamine and neosamine

It has previously been verified by labeling experiments that all the subunits of Km A are derived from glucose [33]. Frost and co-worker [34] have demonstrated the enzymatic synthesis of kanosamine from the UDP-glucose in the 3-amino-5-hydroxy benzoic acid biosynthetic pathway. In contrast, no genes encoding such proteins were found in the Km and Tm clusters sequenced so far but glucose 3-dehydrogenase (GenBank Accession No. JC7628) homologues were identified in the both clusters. The glycine-rich nucleotide-binding βαβ motifs, the characteristic of FAD-dependent dehydrogenase, were conserved at the N-termini of KanI (^{19}VGSGAS GAIT^{28}) and TacB (^{20}GSGASGSVA^{28}). Thus, the involvement of KanI in the dehydrogenation of glucose and its subsequent transamination to yield kanosamine could be expected. The neosamine and its analogous subunits have amino groups at either C-6′ or at C-2′ and C-6′ positions and are conserved in most of the Gms, Kms, and Bns. Though these amino sugars are believed to derive from glucose via glucosamine, none of the related biosynthetic genes were identified in either of the clusters but there were alcohol dehydrogenase homologues (*kanT*, *kanK*, and *gacH*) in the both clusters. Furthermore, a separate group of aminotransferases (KacL, GacE, and BtrB) was found while analyzing the phylogenetic relationships of aminotransferases from Gm, Km, and Bn biosynthetic gene clusters (Fig. 4). They might be involved in transamination following the oxidation of glucosamine to form neosamine or its analogous subunit. The genes, *gacJ* and *gtmH*, encode the other dehydrogenase homologous proteins in the Gm cluster. The presence of three *btrB* homologous genes (*gacK*, *gacC*, and *gacB*) in the Gm cluster is another feature. Interestingly, a very high identity (89%) of nucleotide sequences was found for *gacB* and gac*C*, indicating the possibility of a gene duplication event. These proteins could be involved in the biosynthesis of

Table 1
Summary of proteins identified from the Km cluster

Km[d] ORFs	Related clusters identities (%)			Best match		Proposed function
	Gm[d]	Tm[b]	Bn[b]	Source	Entry[c]	
kanA[a]	GtmA (56)	TbmA (61)	BtrC (35)	*S. tenebrarius*[e]	CAE22471.1	DOI synthase
kanB	GtmB (59)	TbmB (74)	BtrS (40)	*S. tenebrarius*[e]	CAE06512.1	DOI aminotransferase
kanC	GtmC (44)	TbmC (51)	?	*S. tenebrarius*[e]	CAE22470.1	Dehydrogenase
kanD	GtmD (58)	TbmB (36)	?	*M. echinospora*[f]	CAE06507.1	Aminotransferase
kanE	GtmE (58)	TbmD (74)	BtrM (26),—	*S. tenebrarius*[e]	CAE22469.1	Glycosyltransferase
kanF	GtmG (50)	TbmD (33),?	BrtM (33), ?	*M. echinospora*[f]	CAE06514.2	Glycosyltransferase
kanG	—	?	?	*S. cinnamonensis*[e]	AAK19882.1	Regulatory protein
kmr	GtmJ (47)	?	?	*S. kanamyceticus*[e]	CAA75800.1	rRNA methylase
kanI	GacJ (39)	TacB (58)	—	*S. tenebrarius*[e]	CAE22474.1	Dehydrogenase
kanJ	?	TbmE (42)	?	*S. tenebrarius*[e]	CAE22476.1	Transport protein
kanK	GacH (26)	TacD (69)	?	*S. tenebrarius*[e]	CAE22477.1	Dehydrogenase
kanL	GacH	—	—	*S. griseus*[e]	CAA68515.1	Regulatory protein
kanM[a]	—	—	—	*S. albus*[e]	BAD10948.1	Km acetyltransferase
kanN[R]	—	—	—	*S. avermitilis*[e]	NP_828653.1	Efflux protein
kanO[R]	—	—	—	*S. coelicolor*[e]	CAA22731.1	Efflux protein
kanP[R]	—	—	—	*S. griseus*[e]	BAA32132.1	Methyltransferase
kanQ	—	—	—	*S. coelicorol*[e]	NP_626652.1	Transport protein
kanR	—	—	—	*S. avermitilis*[e]	NP_826945.1	ATP-binding protein
kanS	—	—	—	*S. coelicolor*[e]	NP_626650.1	Receptor protein
kanT	GacF (26)	?	?	*S. coelicolor*[e]	NP_626648.1	Dehydrogenase
kacA	—	—	—	*A. mediterranei*[g]	CAC48365.1	Hypothetical protein
kacB	—	—	—	*M. grisea*[h]	EAA53554.1	Hypothetical protein
kacC	—	—	—	—	—	Unknown
kacD	—	—	—	*M. avium*[i]	AAS05884.1	Hypothetical protein
kacE	—	—	—	*S. coelicolor*[e]	NP_630561.1	Hypothetical protein
kacF[R]	—	—	—	*S. avermitilis*[e]	NP_823084.1	Hypothetical protein
kacG	—	—	—	*S. avermitilis*[e]	NP_823085.1	Dehydrogenase
kacH	—	—	—	*S. avermitilis*[e]	NP_823087.1	Oxidoreductase
kacI	—	—	—	—	—	Unknown
kacJ[R]	—	—	—	—	—	Unknown
kacK[R]	—	—	—	*S. coelicolor*[e]	NP_624733.1	Hypothetical protein
kacL	GacE (35)	TacC (71)	BtrB (30)	*S. tenebrarius*[e]	CAE22475.1	Aminotransferase
kacM[R]	—	—	—	*S. cinnamoneus*[e]	CAD60534.1	Hypothetical protein
kacN[R]	—	—	—	*S. cinnamoneus*[e]	CAD60536.1	Hypothetical protein
kacO[R]	—	—	—	*S. coelicolor*[e]	NP_630565.1	Hypothetical protein
kacP[R]	—	—	—	*S. avermitilis*[e]	NP_823078.1	Hypothetical protein
kacQ[R]	—	—	—	*S. avermitilis*[e]	NP_823077.1	Hypothetical protein
kacR[R]	—	—	—	*R. solanacesrum*[j]	NP_522161.1	Hypothetical protein
kacS	—	—	—	*S. coelicolor*[e]	CAB58266.1	NAD synthatase
kacT	—	—	—	*S. coelicolor*[e]	CAB62735.1	Hypothetical protein

— Clearly missing in cluster.
? Perhaps present (may not have been sequenced or published yet).
[R] Gene on opposite strand.
[a] Functions verified.
[b] DOS-containing AmAc biosynthetic gene clusters.
[c] Protein identifier in the database.
[d] ORFs determined in this study.
[e] *Streptomyces* sp.
[f] *Micromonospora* sp.
[g] *Amicolatopsis* sp.
[h] *Meganopothe* sp.
[i] *Micobacterium* sp.
[j] *Ralstonia* sp.

garosamine or purpurosamine. The deduced products of *kanE* and *kanF*, and GtmE and GtmG were similar to the previously characterized glycosyltransferases (Table 1) suggesting a two-step glycosylation during Gm and Km biosyntheses.

Regulatory and resistance proteins

The genes, *kanL* and *kanG*, encode polypeptides homologous to regulatory proteins from various actinomycetes and are located at either end of the Km

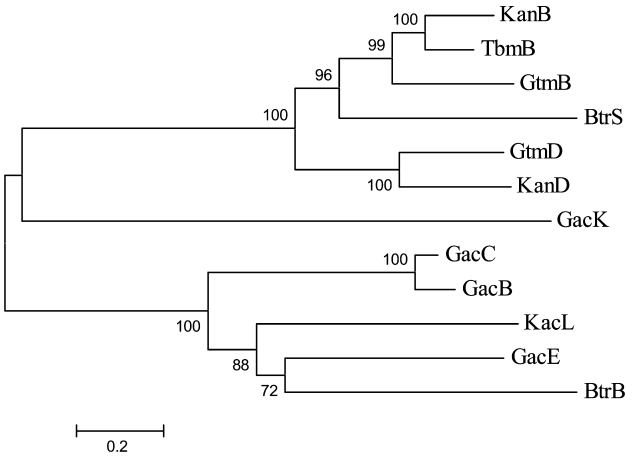


Fig. 4. Phylogenetic analysis of different aminotransferases found in 2-deoxystreptamine-containing aminoglycoside–aminocyclitol biosynthetic gene clusters. The bar in the lower left corner indicates 0.2 amino acid substitutions per amino acid for the branch length.

biosynthetic gene cluster (Fig. 2) pointing to their regulations on Km production in *S. kanamyceticus*. Instead, no such regulatory gene could be found in the Gm gene cluster sequenced so far. While preparing this manuscript, nucleotide sequences deposited by Unwin and coworkers for gentamicin biosynthetic gene cluster were also available (GenBank Accession No. AY524043). Besides the 27 ORFs determined in this study, the authors have determined additional two ORFs in the downstream region of *gacI*, and the putative methyltransferase (*gtmI*) has been annotated as putative oxidoreductase. Indeed, experimental evidences are required to validate such functional assignments. A putative regulatory gene (*gntR*) has been reported in this submission at the downstream region of *gacI* but the protein encoded by this gene does not share any significant homology to Km regulatory proteins. Overall, the genetic organizations of the Km and Gm biosynthetic gene clusters are in fair agreement with the corresponding structures.

The similarities of amino acid sequences encoded by *kmr* and *kanM* to the various databases showed their plausible resistance functions against Km in *S. kanamyceticus*. The *kmr* gene has been reported previously from *S. kanamycetiucs* (GenBank Accession No. CAA75800). The products of this gene confer resistance to Km by the methylation of ribosomal RNA. In contrast, the gene product of *kanM* was similar to Ams-acetylating enzymes. Homology searches revealed the significant amino acid identity (68%) of KanM to putative Km-6′-*N*-acetyltransferase (GenBank Accession No. BAD10948) from *S. albus* whereas poor identities to several AAC(6′Ib) from the various pathogens were observed: 32% to *Pseudomonos aeruginosa* (GenBank Accession No. AAK49459) and *Klebsiella pneumoniae* (GenBank Accession No. AAK51924), 31% to *Salmonella typhimurium* (GenBank Accession No. AAN41403), etc. These enzymes acetylate the C-6′ NH_2 group in the presence of acetyl-CoA and diminish the activities of antibiotics. Several genes encoding hypothetical proteins separated the Km biosynthetic and transport genes (*KanS*, *KanR*, and *KanQ*) (Fig. 2). The putative translational start codon of *kanR* was fused with the stop codon of *kanS* indicating the coupled translation of their respective proteins. A few non-coding nucleotides (4–8 bp) separated the *kacT* from *kanS*, and *kanR* from *kanQ*. The proteins encoded by those genes showed significant similarities (57–64% identities) to the previously reported sugar transport proteins indicating their involvement in pumping out of Km from the cell. Similarly, *gtmL*, *gtmJ*, and *gtmF* represent the resistance genes in the Gm gene cluster. The presence of putative Km biosynthetic genes between hypothetical ORFs outlines the Km biosynthetic gene cluster but further sequence data are required to outline the Gm cluster. The brief description of ORFs determined in the Km cluster is summarized in Table 1.

Functional characterization of KanA

Prior to the heterologous expression of KanA, several models were generated using DHQ synthase as a template and compared the active site topologies to that of TbmA using the different programs described previously [11]. All the amino acid residues reported around the active site of TbmA were conserved for KanA too. The overall folds and topologies were in good agreement with DHQS reported previously [35]. To verify that *kanA* encodes DOI synthase, the protein was expressed in *E. coli* as His_6-fusion protein. The molecular weight observed in the SDS–PAGE (~45 kDa) was similar to that calculated from amino acid sequences of KanA (data are not shown). In vitro enzyme assay was carried out and the product was derivatized to spectrophotometrically visible oxime-derivative as mentioned in Experimental procedures, and analyzed by HPLC and TLC. For identifying the product formation, standard DOI-oxime derivative was used as a reference. The activity of KanA was detected on TLC by comparing the retention factor of product with that of pure compound, but no conversion was noticed in either control sample without substrate or without KanA. The crude product was fractionated using silica-gel column chromatography. A peak was detected at the same retention time (3.70 min) as that of standard DOI-oxime derivative (Fig. 5). The enzyme was rapidly denatured at above 55 °C, and no conversion was noticed above this temperature. Optimal reaction conditions for the formation of DOI were determined at 42 °C and 30 min incubation. These

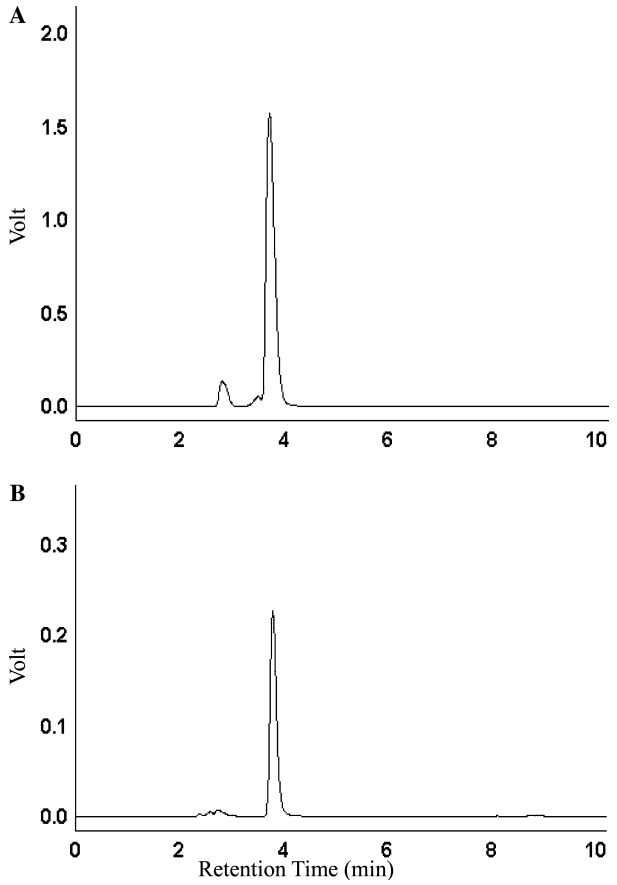


Fig. 5. HPLC analysis of oxime derivative of DOI. (A) An oxime derivative of KanA reaction product eluted in a fraction from silica gel column chromatography; (B) standard DOI-oxime derivative.

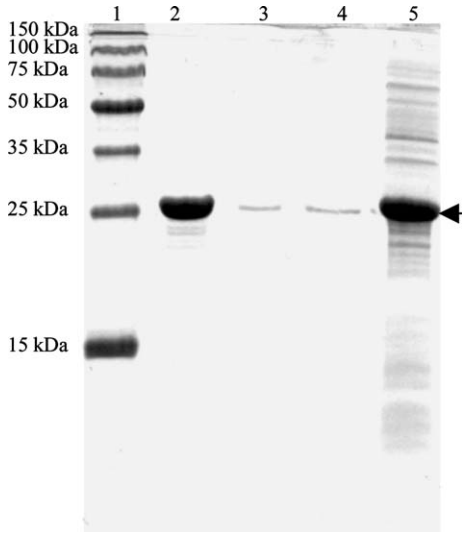


Fig. 6. SDS–PAGE analysis of KanM. Lane 1, Protein marker; lanes 2–4, fractions containing KanM from Ni-affinity chromatography; and lanes 5, cell free extract containing KanM. The arrow indicates a band referring to KanM.

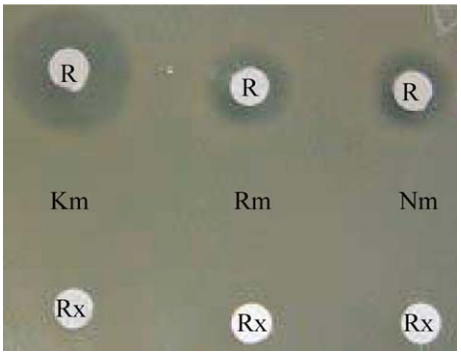


Fig. 7. Antibacterial assay of KanM reaction products against *B. subtilis*. R and Rx stand for inhibition zone in the reference (without acetyl-CoA) and reaction samples correspondingly. Km, Rm, and Nm represent kanamycin, ribostamycin, and neomycin used in the reaction mixture, respectively.

results suggest that *kanA* encodes DOI synthase and catalyzes the first step of the DOS biosynthetic reactions.

Acetylation of aminoglycoside–aminocyclitols by KanM

Murakami et al. [36] reported acetyl transferase activity in Km-resistant *S. lividans* harboring cloned *S. kanamyceticus* genomic fragments. Goo and co-worker [37] recently reported AAC(6′) activity in *S. lividans* by transforming the 1.8 kb-region of *S. kanamyceticus* genomic DNA, but the nucleotide sequences and their location were mysterious. The gene encoding AAC(6′) homologous protein was located at the one end of the Km biosynthetic gene cluster. The protein was expressed in *E. coli* as His_6-fusion protein and purified by Ni-affinity chromatography (Fig. 6). Acetylation of Km, Nm, and Rm was detected on TLC when the pure enzyme was incubated with AmAcs in the presence of acetyl-CoA. The acetylated antibiotics were no longer active against *B. subtilis* as evident from the antibacterial assay of the product (Fig. 7). No acetylation was observed when DOS was subjected to the enzyme assay indicating the specificity of KanM to the C-6′ NH_2 of Ams.

Discussion

In the work presented here, we report the sequence identification, analysis, and comparison of Km and Gm biosynthetic gene clusters. Similar organizations of genes in the cluster for Gm and Km biosynthesis demonstrate their structural similarities as well as similar biosynthetic routes. The biochemical characterization of KanA and KanM with the prior characterization of BtrC and TbmA supports the involvement of this gene cluster in Km biosynthesis.

Glucose has been discovered as the biosynthetic origin of all of the Km subunits by feeding experiments in *S. kanamyciticus* [33]. The DOI synthase triggers the conversion of **1** to **2** and paves the way towards DOS biosynthesis from the pool of primary metabolites in the presence of NAD^+ [11,19]. The transamination of DOI by KanB or GtmB in DOS biosynthesis can be compared to their streptidine biosynthetic counterparts; StsC from streptomycin (Sm) or SpcS2 from Spm biosynthetic pathways. The latter proteins transaminate *scyllo*-inosose to form inosamine [31,38]. Such conversion by pure DOI aminotransferase isolated from the cell-free extract of *M. echinospora* further demonstrates their substrate flexibilities [39]. Also the transamination of **3** to **4** by DOI aminotransferase has been explained by carrying out the mimic reaction using 2-ketoglutaramate as an amino acceptor and DOS as a donor [39]. In contrast, the existence of two DOI aminotransferase-homologous genes in both clusters (*kanA* and *kanD*, and *gtmB* and *gtmD* in the Km and Gm clusters, respectively) allows a possibility of their involvement at different stages of DOS biosynthesis. It is still not clear whether these two genes are prerequisite for DOS biosynthesis or the product of a gene can catalyze both transamination reactions. Unlike streptidine biosynthesis where the involvement of phosphorylases and phosphatases is prerequisite [40], no such genes were found in the Gm and Km clusters, thus indicating the transformation of DOI to DOS without phosphorylation of cyclitol at any stage. Feeding experiments in the Nm producer *S. fradiae* have revealed glucose as the biosynthetic origin of neosamine without any labilization of the C–H bonds, and glucosamine as an intermediate compound [41]. Loss of a C-6′ proton of glucose during transformation also elaborates its oxidation prior to the transamination. The absence of any glucosamine biosynthetic genes in the Km or Gm cluster indicates their other locations on the chromosome or glucosamine are utilized from the primary metabolite pool but the involvement of dehydrogenases (KanT, KanK or GacH) and conserved aminotransferases (KacL, GacE, and BtrB) on neosamine or its analogues' biosynthesis cannot be ignored. Also the absence of any previously reported UDP-kanosamine biosynthetic genes in either cluster allows the possibilities of their location at different chromosomal regions, but dehydrogenation at C-3′ by KanI and its subsequent transamination are also predictable. Because of the resistance of *S. kanamyceticus* to a high concentration of commonly using antibiotics (150 μg/ml neomycin and thiostrepton), our efforts to inactivate *kanF* were not successful. However, the existence of two glycosyltransferases in the both clusters (KanE, KanF, and GtmE and GtmG in the Km and Gm clusters, respectively) is the first report in the AmAcs genetics, and supports a stepwise glycosylation process, but the nucleotide-activation of aminosugars prior to glycosidic linkage is still unclear.

In most cases, antibiotic resistance genes are clustered with the biosynthetic genes: *tlrA*, *B*, *C*, and *D* in tylosin biosynthetic gene cluster; *pikR1*, *pikR2*, and *desR* in pikromycin biosynthetic gene cluster and *drrC* in doxorubicin biosynthetic gene cluster, etc., [42–45]. It has been well established that the antibiotic producer organisms confer resistance to the antibiotics they produce to avoid suicide by three different methods: (1) by pumping out the antibiotics, (2) by destroying the antibiotic's warheads, and (3) by reprogramming the target structure [46]. All these three are found in *S. kanamyceticus*, as well as in *M. echinospora*. Demydchuk et al. [47] have demonstrated the resistance conferred by Kmr in *S. lividans* 66. The characterization of *gtmF* by Kelemen et al. [48] represents its counterpart in Gm cluster. These enzymes utilize the third mechanism by methylating the 16S rRNA. Acetylation of many AmAcs by KanM demonstrates its resistance function and broad substrate specificity by the second mechanism. In contrast to the previous suggestion for the role of AAC(6′) in Km production [49], the loss of activities of acetylated AmAcs by KanM supports its resistance function, but clearly, the discrepancies raised from the expression of AAC(6′) in *S. kanamyceticus* prior to the kanamycin production [50] need to be resolved. No acetylation of Spm and DOS under identical experimental conditions reflects its limited specificity for C-6′ NH_2 group. The aminoglycoside 3′-phosphotransferase (Aph3′) homologous gene (*gtmJ*) in Gm cluster represents the KanM counterpart of the Km cluster. The similarity of KanS, KanR, and KanQ to the transport proteins and their locations in the cluster illustrates their efflux functions by the first mechanism.

To sum up, the gene clusters described in this study are likely to be involved in Km and Gm biosynthesis, their resistance, and regulation. Recent feeding experiments on *B. circulans* to investigate (2*R*)-4-amino-2-hydroxybutyric acid biosynthesis [51], and further genetic and biochemical investigation of its amide linkage to DOS in butirosin biosynthesis in conjunction with this study can be valuable for developing the new derivatives of classical DOS-containing AmAcs or for the replacement of complicated synthetic modifications of Gms or Kms by genetic engineering.

Acknowledgments

The work was supported by medium-term strategic technology development (to JKS), Republic of Korea. We thank Prof. Michael J. Huntington for helpful comments on the manuscript.

References

[1] A. Schatz, E. Bugie, S.A. Waksman, Proc. Soc. Exp. Biol. 55 (1944) 66–69.
[2] J.E. Davies, in: V. Lorian (Ed.), Antibiotics in Laboratory Medicine, Williams and Wilkins, Baltimore, 1991, pp. 691–713.
[3] D. Moazed, H.F. Noller, Nature 327 (1987) 389–394.
[4] R.J. Dworkin, Drug Resist. Updates 2 (1999) 173–179.
[5] D.M. Livermore, Clin. Infect. Dis. 34 (2002) 634–640.
[6] T.R. Zembower, G.A. Noskin, M.J. Postelnick, C. Nguyen, L.R. Perterson, Int. J. Antimicrob. Agents 10 (1998) 95–105.
[7] H. Umezawa, S. Umezawa, T. Tsuchiya, Y. Okazaki, J. Antibiot. 24 (1971) 485–487.
[8] H. Kawaguchi, T. Naito, S. Nakagawa, K. Fujisawa, J. Antibiot. 25 (1972) 695–708.
[9] J.J. Wright, J. Am. Chem. Soc. Chem. Commun. (1976) 206–208.
[10] T.L. Nagabhusan, A.B. Copper, H. Tsai, P.J.L. Daniels, G.H. Miller, J. Antibiot. 31 (1978) 681–687.
[11] M.K. Kharel, D.B. Basnet, H.C. Lee, K. Liou, J.S. Woo, B.-G. Kim, J.K. Sohng, FEMS Microbiol. Lett. 230 (2004) 185–190.
[12] Y. Ota, H. Tamegai, F. Kudo, H. Kuriki, A. Koike-Takeshita, T. Eguchi, K. Kakinuma, J. Antibiot. 53 (2000) 1158–1167.
[13] K. Mansouri, W. Piepersberg, Mol. Gen. Genet. 228 (1991) 459–469.
[14] J. Distler, C. Braum, A. Ebert, W. Piepersberg, Mol. Gen. Genet. 208 (1987) 204–210.
[15] K. Pissowotzki, K. Mansouri, W. Piepersberg, Mol. Gen. Genet. 231 (1991) 113–123.
[16] J. Distler, K. Mansouri, G. Mayer, M. Stockmann, W. Piepersberg, Gene 115 (1992) 105–111.
[17] J. Distler, A. Ebert, K. Mansouri, K. Pissowotzki, M. Stockmann, W. Piepersberg, Nucleic Acid Res. 15 (1987) 8041–8056.
[18] S. Beyer, J. Distler, W. Piepersberg 250 (1996) 775–784.
[19] F. Kudo, H. Tamegai, T. Fujiwara, U. Tagami, K. Hirayama, K. Kakinuma, J. Antibiot. 52 (1999) 559–571.
[20] H. Tamegai, E. Nango, M. Kuwahara, H. Yamamoto, Y. Ota, H. Kuriki, T. Eguchi, K. Kakinuma, J. Antibiot. 55 (2002) 707–714.
[21] F. Kudo, N. Yamauchi, R. Suzuki, K. Kakinuma, J. Antibiot. 50 (1997) 424–428.
[22] C.-G. Kim, T.-W. Yu, C.B. Fryhle, S. Handa, H.G. Floss, J. Biol. Chem. 273 (1998) 6030–6040.
[23] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, second ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989.
[24] J. Ishikawa, K. Hotta, FEMS Microbiol. Lett. 174 (1999) 251–253.
[25] S.F. Altschul, W. Gish, W. Miller, E.W. Myess, D.J. Lipman, J. Mol. Biol. 215 (1990) 403–410.
[26] S. Kumar, K. Tamura, I.B. Jakobsen, M. Nei, Bioinformatics 17 (2001) 1244–1245.
[27] M.M. Bradford, Anal. Biochem. 72 (1976) 248–254.
[28] M.J. Bibb, P.R. Findlay, M.W. Johnson, Gene 30 (1985) 157–166.
[29] J.W. Frost, J.L. Bender, J.T. Kadonaga, J.R. Knowles, Biochemistry 23 (1984) 4470–4475.
[30] P.R. August, L. Tang, Y.J. Yoon, S. Ning, R. Mueller, T.W. Yu, M. Taylor, D. Hoffmann, C.G. Kim, X. Zhang, C.R. Hutchinson, H.G. Floss, Chem. Biol. 5 (1998) 69–79.
[31] J. Ahlert, J. Distler, K. Mansauri, W. Pipersberg, Arch. Microbiol. 168 (1997) 102–113.
[32] C.G. Hyun, S.S. Kim, J.K. Sohng, J.J. Hahn, J.W. Kim, J.W. Suh, FEMS Microbiol. Lett. 183 (2000) 183–189.
[33] M. Koijama, Y. Yamada, H. Umezawa, Agric. Biol. Chem. 32 (1968) 467–473.
[34] J. Guo, J.W. Frost, J. Am. Chem. Soc. 124 (2002) 10642–10643.
[35] E.P. Carpenter, A.R. Hawkins, J.W. Frost, K.A. Brown, Nature 394 (1998) 299–302.
[36] T. Murakami, C. Nojiri, H. Toyama, E. Hayashi, K. Katumata, H. Anzai, Y. Matsubhashi, Y. Yamda, K. Nagaoka, J. Antibiot. 36 (1983) 1305–1311.
[37] Y.A. Joe, Y.M. Goo, Arch. Pharm. Res. 21 (1998) 470–474.
[38] Y.Y. Jo, S.H. Kim, Y.Y. Yang, C.M. Kang, J.K. Sohng, J.W. Suh, J. Microbiol. Biotechnol. 13 (2003) 906–911.
[39] L.A. Lucher, Y.-C. Chen, J.B. Walker, Antimicrob. Agents. Chemother. 33 (1989) 452–459.
[40] J.B. Walker, Methods Enzymol. 43 (1975) 429–470.
[41] S.K. Goda, A.-F. Walid, M. Akhtar, J. Chem. Soc. PERKIN TRANS I (1986) 1383–1389.
[42] M. Liu, S. Douthwaite, Biochemistry 12 (2002) 14658–14663.
[43] R. Fouces, E. Mellado, B. Diez, J.L. Baffedo, Microbiology 145 (1999) 855–868.
[44] Y. Xue, D.H. Sherman, Metab. Eng. 3 (2001) 15–26.
[45] K. Furuya, C.R. Hutchinson, FEMS Microbiol. Lett. 15 (1998) 243–249.
[46] C.T. Walsh, Nature 406 (2000) 775–781.
[47] J. Daemydchyuk, Z. Oliynyk, V. Fedorenko, J. Basic. Microbiol. 38 (1998) 231–239.
[48] G.H. Kelemen, E. Cundliffe, T. Financsek, Gene 98 (1991) 53–60.
[49] A. Satoh, H. Ogawa, Y. Satomura, Agric. Biol. Chem. 39 (1975) 1593–1598.
[50] A. Satoh, H. Ogawa, Y. Satomura, Agric. Biol. Chem. 39 (1975) 2331–2336.
[51] T. Yukita, H. Nishida, T. Eguchi, K. Kakinuma, J. Antibiot. 56 (2003) 497–500.