# kasT Gene of Streptomyces kasugaensis M338-M1 Encodes a DNA-binding Protein

# which Binds to Intergenic Region of kasU-kasJ in

# the Kasugamycin Biosynthesis Gene Cluster

Souichi Ikeno\*, Daisuke Aoki, Koji Sato, Masa Hamada<sup>†</sup>, Makoto Hori<sup>†</sup> and Kayoko S. Tsuchiya

Showa Pharmaceutical University, 3-3165 Higashi Tamagawagakuen, Machida-shi, Tokyo 194-8543, Japan <sup>†</sup> Institute of Microbial Chemistry, 3-14-23 Kamiosaki, Shinagawa-ku, Tokyo 141-0021, Japan

(Received for publication June 21, 2002)

We previously reported that a 4.2 kb Sacl-EcoRI DNA region from Streptomyces kasugaensis M338-M1, a kasugamycin (KSM) producer, included KSM transporter genes (kasKLM). As an extension of that study, a 3.7 kb Pstl-SacI DNA region, located at  $1.5 \sim 5.2$  kb upstream of kasK, was cloned and sequenced, revealing three complete open reading frames, designated kasT, kasU and kasJ. The kasJ gene encodes a protein (KasJ) with a conserved dinucleotide (FAD)-binding motif. Homology search for KasJ showed its similarity to NADH: N-amidino-scyllo-inosamine oxidoreductase (StsB) which is involved in biosynthesis of the streptidine moiety of streptomycin (SM) in S. griseus. The kasT gene encodes a DNA-binding protein (KasT), including a helix-turn-helix motif near the center of the sequence. This protein is similar in structure to a pathway-specific activator protein (StrR) that plays a role in regulating the SM biosynthesis gene cluster of S. griseus. A fusion protein (Trx-KasT) clearly showed DNA binding activity with the intergenic region of kasU-kasJ, suggesting that KasT is a pathway-specific regulator of the KSM biosynthesis gene cluster.

The genus Streptomyces is a group of Gram-positive bacteria that produce a wide variety of secondary metabolites including antibiotics and other pharmacologically active agents. Mechanisms and regulation of biosynthesis of these bioactive compounds are poorly understood. Elucidation of biosynthesis pathways and regulatory mechanisms of production may help to increase antibiotic productivity and/or to design new antibiotics with hybrid structures with enhanced properties.

Kasugamycin (KSM) is an aminoglycoside antibiotic produced by *Streptomyces kasugaensis* M338-M1<sup>1)</sup>, effective against *Piricularia oryzae* and widely used in agriculture in Japan to prevent the rice blast. From *S. kasugaensis* MB273-C4, another KSM producing strain, a gene encoding the enzyme that acetylates the 2'-NH<sub>2</sub> of KSM and thereby inactivates the antibiotic was cloned and named *kac* by HIRASAWA *et al.* (JP. A-05-23187, 1993). Genes required for antibiotic biosynthesis are usually clustered in streptomycetes and linked to genes for selfprotection against the endogenous antibiotic. Genes for regulation of antibiotic biosynthesis are also often found in the clusters. We described in the previous papers that some KSM biosynthetic genes<sup>2)</sup> and KSM transporter genes<sup>3)</sup> were found in the upstream region of KSM acetyltransferase gene ( $kac^{338}$ )<sup>4)</sup>, a self-protection gene, from *S. kasugaensis* M338-M1.

In the present paper, we report the cloning, sequencing and characterization of a putative transcriptional regulator gene for the KSM biosynthetic pathway.

<sup>\*</sup> Corresponding author: ikeno@ac.shoyaku.ac.jp

## **Materials and Methods**

## Strains and Plasmids

Strains and plasmids used in this work are summarized in Table 1. *S. kasugaensis* M338-M1 has been maintained at the Institute of Microbial Chemistry. The following strains and plasmids were of commercial origins: *E. coli* TH2<sup>5</sup> (TaKaRa), *E. coli* DH5 $\alpha^{6}$  (TOYOBO), *E. coli* BL21(DE3) (STRATAGENE), pKF 3<sup>5,7</sup> (TaKaRa), pUC118<sup>8</sup> (TaKaRa) and pET-32a<sub>(+)</sub><sup>9</sup> (Novagen). Other plasmids were produced in the present study.

### Growth Conditions

S. kasugaensis M338-M1 was grown in MR medium (KSM producing medium)<sup>3)</sup> under shaking at 27°C for 72 hours. E. coli TH2 transformants were grown at 37°C in L-broth containing  $12 \,\mu$ g/ml chloramphenicol and  $50 \,\mu$ g/ml streptomycin. E. coli DH5 $\alpha$  and BL21(DE3) transformants were grown at 37°C in Luria-Bertani (LB) medium containing 50  $\mu$ g/ml ampicillin.

# Cloning and Sequencing

Isolation of genomic DNA from *S. kasugaensis* M338-M1 was conducted as described previously<sup>4)</sup>.

The *E. coli* TH2/pKF 3 cloning system<sup>5,7)</sup> was used to clone pSKE 6 (Fig. 1). The cloned DNA region (1.1 kb *Eco*RI-*Sac*I fragment) was digested with appropriate restriction endonucleases, and the fragments were

subcloned into pUC118. Nucleotide sequence was determined for both strands using an ALF*red*<sup>TM</sup> DNA sequencer (Amersham Biosciences). Sequencing reactions were carried out with Cy5<sup>TM</sup> AutoCycle<sup>TM</sup> Sequencing Kit (Amersham Biosciences) according to the supplier's instructions.

To clone the DNA region of pSKE 7 (Fig. 1), the genomic DNA was digested with SacI, and the restriction fragments were ligated themselves with T4 ligase. After self-ligation, an unknown DNA region (1.2 kb, SacI-EcoRI) was amplified by PCR using oligonucleotides 5'-pKS26b (5'-CGGCTCACACAGGCCCGCACCGAAA-3') and 3'-(5'-GCACAGTTCAGCCTCATTTCCTTGT-3'). pKS26a For cloning of pSKE 8, the genomic DNA was digested with PstI. After self-ligation, an unknown DNA region (1.5 kb, PstI-SacI) was amplified with 5'-SKE72 (5'-CCAAGATCGCACATGCGGCAGGGGT-3') 3'and SKE73 (5'-TCCGCTCCGTCGAAGAACCGCACCT-3'). To avoid misreading of the nucleotide sequences of the amplified fragments for pSKE 7 and pSKE 8, the PCR products were used as templates for the dideoxy chaintermination method. As sequencing primers, we used M13-40 universal primer, M13-reverse primer and some synthesized oligonucleotide primers (labeled with Cy5, purchased from Amersham Biosciences). After direct sequencing, these PCR products were cloned into pUC118 to make pSKE 7 and pSKE 8 (Table 1).

Fig. 1. Restriction map of the cloned DNA region from *S. kasugaensis* M338-M1, including kasugamycin transporter genes (*kasKLM*).



On the restriction map, the striped region is dealt with the present paper. The solid region (right) has been reported<sup>3)</sup>. The open arrows indicate the deduced ORFs and direction of transcription. The inserts of plasmids pSKE 6, pSKE 7 and pSKE 8 are indicated above the map. Abbreviations: E, *Eco*RI; K, *Kpn*I; P, *Pst*I; Sa, *Sac*I; Sm, *Sma*I; Sp, *Sph*I.

# Table 1. Strains and plasmids.

Strains and plasmids	Genotype and genetic construct	Source or reference
Strains		
S. kasugaensis		1)
M338-M1	Kasugamycin producing strain	
E. coli		
DH 5α	$\phi$ 80/acZ $\Delta$ M15 $\Delta$ (lacZYA-argF) U169 deoR recA1 endA1 hidR17(r_{\kappa}^{-}, m_{\kappa}^{+})	TOYOBO
	phoA supE44 λ thi-1 gyrA96 relA1	
TH 2	supE44 hsdS20(r <sub>B</sub> , m <sub>B</sub> <sup>+</sup> ) recA13 ara-14 proA2 lacY1 galK2 rpsL20	TaKaRa
	xyl-5 mtl-1 thi-1 trpR624	
BL21(DE3)	F' <i>ompT</i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> )	Novagen
Plasmids		
pUC118	Cloning vector. Amp'. 3.1 kb.	TaKaRa
pKF 3	Cloning vector. Sm <sup>s</sup> , Cm <sup>r</sup> . 2.2 kb.	TaKaRa
pET-32a(+)	Amp', expression vector derived from pBR322, containing trxA [1-109]	Novagen
	and His-tag expressed from T7 promoter for construction of Trx hybrid proteins.	
pSKE 6	pKF 3 derivative containing 1.0 kb EcoRI-SacI fragment. 3.2 kb	This work
pSKE 7	pUC118 derivative containing 1.2 kb SacI-EcoRI fragment. 4.3 kb.	This work
pSKE 8	pUC118 derivative containing 1.5 kb Pstl-Sacl fragment. 4.7 kb.	This work
pET-KasT	pET-32a(+) derivative containing 1.0 kb <i>Ncol-Bam</i> HI fragment.	This work
	Trx-KasT expression plasmid.	

### Computer Analysis of DNA and Protein Sequences

DNA and protein sequences were analyzed with the DNASIS-Mac version 3.7 (Hitachi Software Engineering). FramePlot  $2.3.1^{10}$  was used to search for open reading frames (ORF). Amino acid sequences of potential gene products were compared with those in the databases (SWISS-PROT and PIR) by means of FASTA<sup>11</sup> and BLAST<sup>12</sup>.

#### Nucleotide Sequence Accession Number

The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases with the accession number AB076838.

# Rapid Amplification of cDNA Ends (RACE)

The transcription start site of *kasJ* was determined by the 5' RACE method with 5'-Full RACE Core Set (TaKaRa) according to the manufacturer's instructions<sup>13)</sup>. Isolation of total RNA from *S. kasugaensis* M338-M1 was conducted as described previously<sup>3)</sup>. RT reaction was performed in 15  $\mu$ l containing 200 pmol of 5'-phosphorylated primer 3'-PKasJ (5'-CCAGCACGATGTACTCGCGGCCCAGTCGTT-3'), 5 U of AMV reverse transcriptase XL and 5  $\mu$ g of the total

RNA. The primer was annealed stepwise (80°C for 2 minutes and 45°C for 10 minutes) before the reverse transcriptase was added. The RT reaction mixture was incubated at 50°C for 1 hour. A negative control lacked reverse transcriptase. The template mRNA was hydrolyzed with RNase H (60 U) at 30°C for 1 hour, and the resulting cDNA was ligated with T4 RNA ligase (40 U) at 15°C for 18 hours. The first strand cDNA was circularized and/or connected into a concatemeric form. The first PCR mixture contained 20 pmol each of 5'-RACE26 (5'-GGCTCGAGTGGATACGGCGGTAATCGGGTC-3') and 3'-RACE26 (5'-GGGCATGCGCACAGTTCAGCCTCAT-TTCCT-3'). The reaction was carried out with 20 cycles under the following conditions: 97°C for 30 seconds (denaturation), 66°C for 1 minute (annealing), and 72°C for 1 minute (extension). The resulting PCR product  $(1 \mu l)$  was used as a template for the second nested PCR amplification with primers 5'-RACE26b (5'-GGGCATGCCCGAACTG-GAACGACT-3') and 3'-RACE26b (5'-CGCTCGAGTC-CGGCGTGCCGCTGT-3') under the same conditions as described above. The 140 bp product was digested with SphI and XhoI (recognition sites are underlined in the primers) and cloned into a plasmid vector pUC118. The

DNA sequence of the product was determined by the dideoxy chain-termination method.

Transcription start site of kasU was also determined using 5'-PKasU (5'-CGACTTCCGGACCGCGTCGCAC-AGT-3') for the first strand synthesis. Pairs of primers [5'-RACE7a (5'-TTGGTTGAACACAGGTGGAG-3')/3'-RACE7a (5'-GGAATTTCCTGGGTTGAACA-3') and 5'-RACE7b (5'-GGGCATGCGGTGCGCACGTACGGGGA-C-3')/3'-RACE7b (5'-GCGTCGACCCGCCCGTTGCAG-CCA-3')] were used for the first and second PCR, respectively. The 5'-RACE for *kasU* mRNA was performed as described above for the 5'-RACE for *kasJ* mRNA.

## Overexpression and Purification of Trx-KasT

The coding region of kasT was amplified by PCR using 5'-THIS6 (5'-GGCCATGGTTCATCATCATCATCATCATCA CGTGGCTGAAACA-3') and 3'-TBgl1 (5'-CCAGATC-TTCACGCACTTGCCCGTCCACGC-3'), designed to include NcoI and BglII recognition sites, respectively. The amplified fragment was digested with NcoI and BglII. The 1073 bp NcoI-BglII fragment was ligated into NcoI-BamHI digested pET-32 $a_{(+)}$ . The construct (pET-KasT, Table 1) was transformed into E. coli BL21(DE3). In pET-KasT, kasT was fused in frame to the 3' end of trxA (thioredoxin gene) under the control of the T7 promoter. In this host strain, the T7 polymerase gene was controlled by the tac promoter that could be induced by isopropyl- $\beta$ -Dthiogalactopyranoside (IPTG). The fusion protein (Trx-KasT) contained two His(6x)-Tag (Fig. 4A). The transformants were grown at 37°C to 0.5 OD<sub>660</sub>, then induced with 1 mM IPTG and cultured further at 30°C for 20 hours. The cells were collected by centrifugation and disrupted by sonication. Insoluble materials were separated by centrifugation for 30 minutes at 10,000 g. Trx-KasT in the supernatant was bound to Ni-nitrilotriacetic acid (Ni-NTA) resin and eluted with imidazole, as described by the supplier (Qiagen). The eluate (Trx-KasT) was inspected by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE, Fig. 4B). The Trx-KasT was stored at  $-70^{\circ}$ C in the presence of 10% glycerol.

# Gel Retardation Assay

The intergenic region of *kasU-kasJ* was separated into two regions; 26A (nt 2713~2917, Fig. 2) and 26B (nt 2918~3060, Fig. 2), each amplified by PCR using pairs of primers 5'-GS26a (5'-GGGAATTCCGGGGGTCGGCGGC-ACAC-3') and 3'-GS26a (5'-CACACACCTGAGTCCGG-GCAAGG-3') for 26A and 5'-pKS26a (5'-CTCGACACT-TCTCGACGCTCTCCCT-3') and 3'-pKS26a (5'-GCACA-GTTCAGCCTCATTTCCTTGT-3') for 26B. The upstream region of kasT (nt 705~1313) was amplified in three separate segments (8A, 8B and 8C, Figs. 2, 5). Pairs of PCR primers were 5'-GS8a (5'-GGCGATGCGGTCGTCGGA-AC-3') and 3'-GS8a (5'-GTTTCGACTGGTCACGCGTT-3') for 8A (nt 705~902, Fig. 2), 5'-GS8b (5'-GCCTAC-CGAACCGTGGGAGATG-3') and 3'-GS8b (5'-AGATCG-GCCATTTTCCCCGAGG-3') for 8B (nt 903~1131, Fig. 2), and 5'-GS8c (5'-CGTGCTATGGCCGAACTTTTGC-3') and 3'-GS8c (5'-GAGCCTCTCGGTCGATGATGTT-3') for 8C (nt 1132~1313, Fig. 2). These PCR fragments were <sup>32</sup>P labeled at the 5' ends with  $[\gamma^{-32}P]$ ATP and T4 polynucleotide kinase. For the gel retardation assay, 0.3 to 0.5 ng of the <sup>32</sup>P-labeled DNA fragments (10,000 to 15,000 cpm) were incubated with 3 to  $5 \mu g$  of Trx-KasT at 30°C for 1 hour in a buffer containing 10 mM Tris-HCl (pH 8.0), 50 mM KCl, 5 mM DTT, 10% (v/v) glycerol, 12.5 mM spermidine and  $1 \mu g$  of poly(dI-dC) poly(dI-dC) in a total volume of 20 µl. Protein-DNA complexes were separated from free DNA on 4% polyacrylamide gels by electrophoresis at 150 V for about 1 hour (Figs. 5, 6).

## Results

# Cloning and DNA Sequence Analysis

The newly cloned 3.7 kb *PstI-SacI* region was located upstream of *kasKLM*, the KSM transporter genes. As shown in Fig. 1, the region was separated into three short fragments and cloned separately. The inserted fragment of pSKE 6 was shotgun-cloned as described previously<sup>3)</sup>. These inserted fragments of pSKE 7 and pSKE 8 were cloned by a method using PCR technique (see Methods).

The 3755 bp *PstI-SacI* region from pSKE 6~8 was sequenced (Fig. 2). The GC content of the region was calculated to be 67.9% from the sequence. ORFs were searched for based on the codon usage and third codon position bias which are characteristic of streptomyces genes<sup>14</sup>). Within this region two complete ORFs, designated *kasT* (nt 1318~2358) and *kasU* (nt 2737~2426, complementary), and the 5' region of already reported ORF J<sup>3</sup> were recognized. We designated ORF J as *kasJ*. No ORFs were found upstream (nt 1~1317) of *kasT*. There was only a relatively low GC content region (nt 870~1230, 57.6% GC) for streptomyces genomes.

The putative initiation codon for kasT is nt 1318 GTG, preceded by a potential ribosome binding site (RBS) (nt 1307~1310, GAGG). The stop codon for kasT is nt 2356 TGA. The kasU gene runs in the opposite direction to kasTand kasJ. The putative start and stop codons for kasU are nt 2737 GTG and nt 2428 TAG (complementary, Fig. 2), PstI

1 CTGCAGGGCTCCAGCCACATGCCTGGATGTCAAGGAGTGTGCACCGGAGTGGTGTCGTAG

61 GAGTGCAGACGAGTTCCGTCCGAGCCGTACAGGTTGAGCAGGTAGCCCTGATGGAGGCGG

121 CCTCCTTCCCGCGTCAGGAAGGAGGCCATACCCGTGGCGACGTCGTCGAAATTCTTGTAG

181 TAGGTCTTCTGGTGAGTCCACACGGTCAGCACGAGTACGCCCTCGCGGTCCTTCGCGCTG

241 CGGATGCCGTATACGTCGGCGGCTCCTCCGGCCTCGCGGTAGCGGTCGGCCAGTTCCTTG

361 CCGCCCGCGCCACATCCCGTCAGCACAACGGCGGCCGCGAGAAAGAGAACCGGCAGGATT 421 ATTCGTCTCATCGCGCCGGGGTTCCCCGGCCCTACGGCGCCAGGACGGCAGCGAAGGGCCG

481 GCAACCGCCCCAGCACCTTGTCCGGAGGTCTCGTGCCAGGTGCGGATACTGCGCACCATT 541 TCGTCTGCCACGGGGGGGGGGGGGCGGCTTCTCCCTTCGATAGACGATCTTCCCCGCACCATCCGAGCT

601 GACGACACCGAACAACTCCTCGGCCGGGGTGCTCACGCGTACACGTGACCGGACATGTCC

661 GATCGTCCGCCACCGCCTTGACGGTCGATGAACGTCCCGTTCCCGGCGATGCGGTCGTCG

721 GAACCGCGCGCGAAGCTCCCGGCCCGCCCTCGCCATCGCGCCCAGGAAAGATCATCGGGC 781 GCCCACCGACGCCCTTGCCGCCGAAGGCGGATGGGGGCAGTTTTCCGCCACCGCGCCACAC 841 CCATCCCGCCCCGGCATTCCCACGCCAGAAACCAGCCCGAACAACGCGTGACCAGTCGAA

901 ACGCCTACCGAACCGTGGGAGATGGGTGAAACCTGTGATCGGCACGACTATTCTCAAATT

961 ATCCCCACAATGCCTCGGACCACCCGATACGGCAGCGTTGTCAAGCTCGTGCACGCCCTT

1021 GTCGAAAAATCACCGTCGCCCCAGCTCACACCCGACGCGCACCCCCAACCGGCAGCCGTGG

1081 CGGGACATCATCGGGCCGTTCCCGGGTTCCCCTCGGGGAAAATGGCCGATCTCGTGCTATG 1141 GCCGAACTTTTGCCTTGTTATTCGGTGTTGTTGACGAAACGATTGTTGCTCAAAGTAAGC

1201 TCTGTGAATCCGATAGCCGCTTGAGCCGAGAGTGGCTCGGGTCAGGTCGAGAGTGCTTCT 1261 GTTCCGCTTCGTGACGGGTCCGGCAAGTGACAACATCATCGACCGA GAGGCTCCCCGGTG

1321 GCTGAAACAGTTCGGGCCGATTCACCTCTGAAGAGCAGTTACAGAAACGTTCCTGCTGCT A E T V R A D S P L K S S Y R N V P A A

1381 GAGGTCCAGGGGGGCGGCCTATCGGTCGGGCAACGGACCACTCGTATAGCGATCAGCTCG E V Q G S G L S V G Q R T T R I A I S S

1441 CTGCTGGCGGCCCGACTCGCCTGGCAGCGGGGAGAGAGCGCCGAGCACATCCGCTTGCTG L L A A D S P R S A G E N A E H I R L L <u>Saci</u>

1501 GCCGATTCCGGAGCTCGGCTGCCACCCATCGTCGTACAGCGCTCGACCATGCGCGTCATC

1561 GACGGGATGCACCGGCTGCGGGGCCGCCGCGCGCGAGACCGAGATCGAGGTGCGG

1621 TTCTTCGACGGAGCGGAGGAGGACTCCTTCCTCCTCGCGGTCAGGTCCAATATCGCCCAC

1681 GGACTGCCGCTGTCCCAGGAGGAGCGCGCCGCCGCCGCCCCAACGCATCATCCGTTCCCAT

G L P L S Q E E R A A A A Q R I I R S

1741 GCCCAGTGGTCCAACCAGGCGATCGGCGAGGTCACCGGGCTGGACGCGAAGACGATTGCG

A Q W S N Q A I G E V T G L D A K T I A 1801 GCACTGCGACGAGACGCCAAGGACGTACCGCAGCTGGACGCCCGGATCGGCAGGGACGGA

A L R R D A K D V P Q L D A R I G R D G Sall 1861 CGGGTACGGCCCGTCGACGGGCCCAAGGGCGCAAGGCTGGCCGGCGAATTGATGGCCGAG

2041 GCCGACCCGCAGCCACCGGCCCGGTACGCCGCGTCGGAGGACCGAAGCGGTACGACCGCT A D P Q P P A R Y A A S E D R S G T T A

2101 CCGCGGACCGCGGAGCAGAACCGGCGCATCCTGCTGCAGAAACTGCGCAAGGACCCCTCA R T A E Q N R R I L L Q K L R K D P

2221 GGCGAGGACTGGGAGCGGCTGCTGGACAGCGTGCCGATGCACTGTGCGGCAACCATCGTG G E D W E R L L D S V P M H C A A T

2281 GAACT GGCGCGGGGCT GTT CCGGGGT CT GGCAGGATTT CGCCGCGCAACT GGAACGGCGT

E L A R G C S G V W Q D F A A Q L E R

2341 GGACGGGCAAGTGCGTGACACGCGCGAAGATCAGATCCCCGCGGTGACGTAGTACGAGAT

2401 CCCAGGCCAAGCTTCCGGGGCGCGCCTACGGAAAGCCGGCCCCGTACGTCCCCATGAGGC

2461 CGGCGGAGAGGCGCCGGCGGCGGCGAGACGGACTCCTTGCCCGAGGCGGCCAGGGCATATGG G A S L R R R G L R V G Q G L R G P C I

PFGAGYTGML

GRASA\* <u>HindII</u>I

L R C N E A G R A L L R W L E V Q A V E

V R P V D G A Q G R R L A G E L M A E

A D S G A R L P P I V V Q R S T M R V I

DGMHRLRAAALRGETEIEVR

F F D G A E E D S F L L A V R S N I A H

kasT →

2521 GCCGGCTCCGGAACAGCGACTTCCGGACCGCGCCGCACAGTAGGTGGCTGCCAACGGGCGG P R S R F L S K R V A D C L L H S C R A A T N L G P F E K T S C L H L P R T R H 2641 CACGTACGGGGACTTGTACGTGCGGGGGGAACGAACGGGGATTGGCTGCCGGGTGACAAGG A R V P V Q V H P P V F P S Q S G P S L <u>Ecori</u> 2701 GGAGGTTTCCGCGGGAATTCCGGGGTCGGCGGCACACCGATGTCGGGGGTGCCGAAAGGG PLNGRSNRPRCV + kasu 2761 ACGCACCTGGGTCGGTAATCCACGTAACACCCCCATGCTG, CTGCTTAGCGGGAAACACA kasU mRNA←1 -10 -35 2821 CGACGCCGCGAGACGGAGAAAGACCTGACGGTGTCGTACTGAGAAGCCGGCGTCGGTGTC 2881 TTCCTGACCCCGACOTTGCCCGGGACTCAGGTGTGTGCTCGACACTTCTCGACGCTCTCC

l →*kasj* mRNA -35 -10 <u>BamHI</u> 2941 CTCGTCACACCGGGATCCACCGGTGAAGACACTGGAATGAGAAAGCCCTCCACAAGTGCA

3001 CCCGCGCCACAGCGGCACGCCGGAAGTGCCTTGGGACAAGGAAAT GAGGCTGAACTGTGC V I kas J

3061 TGGATACGGCGGTAATCGGGTCGGGTATGGCGGGGGCGGCGGCGGCGCCACCGAACTGGAAC T A V I G S G M A G T A V A T E L E R Dinucleotide (FAD)-binding motif

3121 GACTGGGCCGCGAGTACATCGTGCTGGAGGCAGGCCCGCACCAGGGGCGTGCGCACATCG 
L
G
R
Y
I
V
L
A
G
P
H
Q
G
R
A
H
I
A
BanHI
Ba ADEGSAHLADPAQDPAFTP <u>Smai</u><u>Kpni</u> <u>Smai</u> 3241 TCCTCGCGCAGCCCGGGACGCCGTACGGCCCGCTGTCGGGGTACCGGCGACGCCGTCGGGACGCCGTGTCGGCGACGCGTCGGCGACGCGTCGGCG L A Q P G T P Y G P L S G Y R R V G G 

A E V R R A L L G T D G E P G L Y A R 3421 GAACCGAGGCGGCCCTGAGGGCGTGGTCCGGCGGAACCCGGCTCACACAGGCCCGCACCG T E A A L R A W S G G T R L T Q A R T E 3481 AAACAGAGGAGGCGATGGCGAAGGAGCTGGCCGGGCTCGGATACCCGGCCGTTCCGACCC ΤΕΕΑΜΑΚΕΙΑGΙGΥΡΑΥΡΤΡ 3541 CGCGGGCCATACGGACGCTGGGCCAGGGCCGTTGGGAGGCGTACTCCCCGGTCACCGGCC R A I R T L G Q G R W E A Y S P V T G L 3601 TGCCGTCCGAGCGCATCCGCACGGGGCACCGCGTCAGCGCGGTGCGGAAGCGGGCCGAAG P S E R I R T G H R V S A V R K R A E G 3661 GCGGCTACTCGCTGATCCTCGACGACGGGGACGGGCCCGCACTCGATGCCCGGCACATCG G Y S L I L D D G D G P A L D A R H I V <u>Saci</u><u>Sali</u> 3721 TCCTGTGCGCGGGCACCGTCGGCAATATCGAGCTCGTCGACGGACTGGTGCACGGGCAGC LCAGTVGNIELVDGLVHGQH <u>Sali</u><u>Sphi</u> 3781 ACCGGGAGCCGCGGCGGTACTCCCTCGTCGACCACGTGGCCTGCGGCATGCTCGTCGTGG R E P R R Y S L V D H V A C G M L V V E

3841 AGGACGGTCGGCACGGGGCGATGGAGTCCTCGGTGCACGCGGGGTTCGACGCGCCGTCCC DGRHGAMESSVHAGFDAPSR 3901 GCTCCAATCTGATAGTCGAGCGCCAACGGCACGGCAGCGGTGTCCTCATGGACGCCTGGG S N L I V E R Q R H G S G V L M D A W A 3961 CCATGGGGGAGCAGCCGCCGGAGGCGGCGTCCACCGTCCTGGTCGGTGACCGGAGCAGCA M G E Q P P E A A S T V L V G D R T S I

4081 TGCTCAGGGAATTCGCCAGGCGCACAGGACTCAAGGCGCGCTTCCGTGAGGAGTACGTCT L R E F A R R T G L K A R F R E E Y V S 4141 CGTTCGACGAAGCCCTGGCCCAGGCCGAGGCGCACCGGCCACGCCGTGGTCTACGACG F D E A L A Q A E A Q P G T A V V Y D A 4201 CGACGATCGGGGAACTCGACCACGAGAGCGGGGGGTTGCCGCTGGACGGGGAGCACGTAG TIGELDHESGGLPLDGEHVD

4261 ACGTCGCAGGACACCTTCGGGAGGCCCCTGGTGTGGCCGTCGCCGGCCCGTGCCTCATCC VAGHLREAPGVAVAGPCLIP

4321 CCCGGGCGGGAGCGGCCAACCCGAGCCTCACCACGCTGGCTCTGGCGCGCGGTACGTCGTCC RAGAANPSLTTLALARYVV 4381 AGCACATGAGCTGAGCGCCCCCCGGTGCCG

HMS

Deduced amino acid sequences of KasT, KasU and KasJ were shown in single letter code below the nucleotide sequence. Thick arrows indicate the putative transcriptional terminators for kasT and kasU. Boxed nucleotides indicate -35 and -10 hexamer of the potential promoters for kasU and kasJ. The helix-turn-helix motif (KasT) and the dinucleotide (FAD)-binding motif (KasJ) are boxed.



Fig. 3. Homology plot.

Deduced amino acid sequences of KasT and KasJ were compared with  $StrR^{15}$  and  $StsB^{19}$  from *S. griseus*, respectively. These data were compared using a window of 30 and a stringency of 8.





(A) The putative primary structure of Trx-KasT. (B) Purification of Trx-KasT with Ni-NTA. Lane 1, Size marker; Lane 2, Cell lysate; Lane 3, Flow-through; Lane  $4\sim 5$ , Wash (20 mM imidazole); Lane  $6\sim 9$ , Elution (250 mM imidazole).



Fig. 5. Gel retardation assays demonstrating the binding of Trx-KasT.

Thick line and open arrows indicate the cloned DNA and the putative structural genes, respectively. The location of the DNA regions used as probe is shown under the map.

Fig. 6. Gel retardation assays demonstrating the specific binding of Trx-KasT.



An open triangle indicates shifted band. A 259 bp *SphI-Sal*I fragment from pKF 3 was used as noncompetitor. Lane 1, free probe (26A); Lane 2, Trx-KasT; Lane 3, Trx-KasT and competitor (200 ng); Lane 4, Trx-KasT and noncompetitor (400 ng).

respectively. There are two imperfect palindromic sequences (nt  $2363 \sim 2383$ , nt  $2416 \sim 2444$ ) in the intergenic region of *kasT-kasU*. These sequences might be transcriptional terminators for each gene. The putative start

codon for *kasJ* is nt 3057 GTG, preceded by a potential RBS (nt 3046~3049, GAGG). The stop codon for *kasJ* is nt 4392 TGA. The intergenic region of *kasU-kasJ* (nt 2738~3056) includes two possible promoters for each direction.

# Characterization of Putative KasT, KasU and KasJ Proteins

The deduced product (KasT) of *kasT* was a protein containing 346 amino acid residues, with a molecular mass of 37,551 Da and a pI of 10.3. Homology search using the FASTA and BLAST programs suggested that KasT was a DNA-binding protein with a conserved helix-turn-helix (HTH) motif (Fig. 2). The deduced amino acid sequence of KasT showed 50% similarity with StrR<sup>15</sup>), a pathway-specific activator of SM biosynthesis from *S. griseus* (Fig. 3). KasT was also similar to various other pathway-specific transcriptional regulators controlling the expression of biosynthetic gene clusters for antibiotics in *Streptomyces*, such as SpcR<sup>16</sup> (44%) from *S. flavopersicus*, NovG<sup>17</sup> (45%) from *S. spheroides* and ComG<sup>18</sup> (44%) from *S. lavendulae*.

The deduced KasU protein containing 103 amino acid residues showed no significant similarities with any proteins in the databases.

1060

The *kasJ* gene encoded protein (KasJ) containing 445 amino acid residues, with a molecular mass of 47,662 Da and a pI of 5.68. The deduced amino acid sequence contained a dinucleotide (FAD)-binding motif (Fig. 2) and showed 36% similarity with NADH:N-amidino-*scyllo*-inosamine oxidoreductase (StsB)<sup>19</sup> for SM biosynthesis in *S. griseus* (Fig. 3).

# Determination of Transcription Start Sites of *kasJ* and *kasU* by Rapid Amplification of cDNA Ends (RACE)

Transcription start sites of *kasJ* and *kasU* were determined by the RACE method. The results suggested that the transcription start sites of *kasU* and *kasJ* were located at G (nt 2788, complementary), 51 nucleotides upstream of the putative start codon of *kasU* and at G (nt 2934), 123 nucleotides upstream of the predicted start codon of *kasJ*, respectively (Fig. 2).

Most *Escherichia coli* promoters, recognized by RNA polymerase holoenzyme containing  $\sigma^{70}$ -factor ( $E\sigma^{70}$ ), have two conserved hexamer sequences at -35 and -10. Some streptomycete promoters have similar structures, thus called streptomycete *E. coli*-like promoter (SEP)<sup>20)</sup>. The consensus sequences of SEP are TTGAC[Pu] (-35 hexamer) and TAG[Pu][Pu]T (-10 hexamer)<sup>21)</sup>, 16 to 18 nucleotides apart.

As shown in Fig. 2, the -35 and -10 hexamer sequences of the putative kasU promoter (kasU-p) were TCGTGT (nt 2823~2818, complementary) and CAGCAT (nt 2800~2795, complementary), respectively, 17 nucleotides apart. The -35 and -10 hexamer sequences of putative kasJ promoter (kasJ-p) were TTGCCC (nt 2897~2902) and GACACT (nt 2921~2926), respectively, 18 nucleotides apart. These results suggested that kasU-p and kasJ-p belonged to SEP. The transcription start site of kasT remains to be determined.

### Overexpression and Purification of Trx-KasT

To demonstrate the function of KasT, we overexpressed in *E. coli* BL21(DE3) a fusion protein in which Trx-KasT was preceded by a thioredoxin peptide and six histidine residues, and purified the product by affinity chromatography with a Ni-NTA column. Purity was confirmed by SDS-polyacrylamide gel electrophoresis (Fig. 4B). The molecular mass of 56 kDa agreed well with that (Mr=55,819) deduced from the nucleotide sequence.

# Trx-KasT Binds Specifically to the Intergenic Region of kasU-kasJ

DNA binding ability of the purified fusion protein was demonstrated by gel retardation assay. As shown in Figs. 5 and 6, the protein bound to the intergenic region of *kasU*-*kasJ* (nt 2713~2917, Fig. 2) but not to the upstream region of *kasT* (nt 705~1313, Fig. 2).

It was confirmed that the thioredoxin peptide part alone did not bind to DNA (data not shown).

## Discussion

KSM, an aminoglycoside antibiotic, is composed of carboxyformidoyl, kasugamine and D-inositol moieties (Fig. 7). We have found three genes responsible for biosynthesis of the kasugamine moiety (kasA, kasC, and kasD) located upstream of  $kac^{338}$ . In this region of the S. kasugaensis genome, we also found the KSM transporter genes (kasKLM) but did not find any biosynthetic genes for the carboxyformidoyl or the D-inositol moiety. In the present study, we report that KasJ, containing a dinucleotide- (FAD-) binding motif in the N-terminal region, is similar to StsB which is a putative oxidase involved in keto group formation during the streptidine biosynthesis in S. griseus. The substrate of StsB should be an inositol derivative (N-amidino-scyllo-inosamine). It was inferred, therefore, that KasJ might be responsible for biosynthesis of the D-inositol moiety (Fig. 7). FUKAGAWA et al.<sup>22)</sup> reported that a precursor of the D-inositol moiety was not D-inositol, but myo-inositol. It remains to be established whether KasJ catalyzes the epimerization (from mvoinositol to D-inositol).

KasT showed high degrees of similarity with StrR, a pathway-specific activator protein of SM biosynthesis. The HTH motif of KasT shared 75% identity with that of StrR. As shown in Fig. 5, KasT did not bind to the upstream region of its own gene, *kasT*, but did bind to an intergenic region of *kasU-kasJ*. This result suggests that KasT might be a pathway-specific regulator of the KSM biosynthesis gene cluster, but not autoregulator.

The binding sites of StrR have a partial palindromic consensus sequence  $[GTTCGActG(N)_{11}CagTcGAAc]^{23}$  and are located upstream of the *str/sts* promoters in SM producer strains. Although we found a partial palindromic sequence GAAGCCGGCGTCGGTGTCTTC (nt 2863~2883, Fig. 2) in the intergenic region, the sequence had no similarity with that of StrR-binding site.

Use of a rare codon UUA is a characteristic of some

Fig. 7. A tentative pathway of kasugamycin biosynthesis.



Abbreviations: DG-6-P, D-glucose 6-phosphate; myo-Ino, myo-Inositol; Gly, Glycine; UDP-GlcNAc, UDP-N-acetylglucosamine.

actinomycete pathway-specific regulators<sup>24,25)</sup> including *strR*. This codon is not used in *kasT*, however, indicating that *kasT* is not regulated by *bldA*<sup>26)</sup>.

In the biosynthesis of SM, expression of *str* is regulated by a complicated regulatory cascade involving pleiotropic and pathway-specific factors. The start point of the cascade is controlled by A-factor<sup>27)</sup>, an autoinducer produced by SM producing strains. Future studies of *kasT* will be necessary to show whether the transcription is regulated by A-factor or related  $\gamma$ -butyrolactone(s)<sup>28)</sup>.

### Acknowledgements

We thank HUMI KAMEYAMA, NAOKO KAWASHIMA and IKUKO KUMANO for excellent technical assistance.

#### Reference

- UMEZAWA, H.; M. HAMADA, Y. SUHARA, T. HASHIMOTO & T. IKEKAWA: Kasugamycin, a new antibiotic. Antimicrobial Agents Chemother. 5: 753~757, 1965
- 2) IKENO, S.; T. TSUJI, K. HIGASHIDE, N. KINOSHITA, M. HAMADA & M. HORI: A 7.6kb DNA region from *Streptomyces kasugaensis* M338-M1 includes some genes responsible for kasugamycin biosynthesis. J. Antibiotics 51(3): 341~352, 1998
- 3) IKENO, S.; Y. YAMANE, Y. OHISHI, N. KINOSHITA, M. HAMADA, K. S. TSUCHIYA & M. HORI: ABC transporter genes, *kasKLM*, responsible for self-resistance of a kasugamycin producer strain. J Antibiotics 53(4): 373~384, 2000
- 4) IKENO, S.; K. HIGASHIDE, N. KINOSHITA, M. HAMADA & M. HORI: Correlation between the presence of *kac*,

kasugamycin acetyltransferase gene, and the productivity of kasugamycin in *Strepromyces*. Actinomycetologica 10(2):  $73 \sim 79$ , 1996

- HASHIMOTO, G. T.; A. TSUJIMURA, K. KURIYAMA & S. MATSUDA: Construction and characterization of new host-vector systems for the enforcement-cloning method. Gene 137: 211~216, 1993
- 6) HANAHAN, D.: Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166(4): 557~580, 1983
- 7) TOBA, M. M. & T. G. HASHIMOTO: Characterization of the spontaneous elimination of streptomycin sensitivity (Sm<sup>s</sup>) on high-copy-number plasmids: Sm<sup>s</sup>-enforcement cloning vectors with a synthetic *rpsL* gene. Gene 121: 25~33, 1992
- VIELRA, J. & J. MESSING: Production of single-stranded plasmid DNA. Methods in Enzymology 153: 3~11, 1987
- 9) LAVALLIE, E. R.; E. A. DIBLASIO, S. KOVACIC, K. L. GRANT, P. F. SCHENDEL & J. M. MCCOY: Bio/Technology 11: 187~193, 1993
- 10) ISHIKAWA, J. & K. HOTTA: FramePlot: a new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G+C content. FEMS Microbiol. Lett. 174(2): 251~253, 1999
- PEARSON, W. R. & D. J. LIPMAN: Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85(8): 2444~2448, 1988
- 12) ALTSCHUL, S. F.; T. L. MADDEN, A. A. SCHAFFER, J. ZHANG, Z. ZHANG, W. MILLER & D. J. LIPMAN: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25(17): 3389~3402, 1997
- 13) MARUYAMA, I. N.; T. L. RAKOW & I. MARUYAMA: cRACE: a simple method for identification of the 5' end of mRNAs. Nucleic Acids Res. 23(18): 3796~3797, 1995
- 14) BIBB, M. J.; P. R. FINDLAY & M. W. JOHNSON: The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. Gene 30:  $157 \sim 166, 1984$
- 15) DISTLER, J.; A. EBERT, K. MANSOURI, K. PISSOWOTZKI, M. STOCKMANN & W. PIEPERSBERG: Gene cluster for streptomycin biosynthesis in *Streptomyces griseus*: nucleotide sequence of three genes and analysis of transcriptional activity. Nucleic Acids Res. 15(19): 8041~8056, 1987
- 16) LYUTZKANOVA, D.; J. DISTLER & J. ALTENBUCHNER: A spectinomycin resistance determinant from the spectinomycin producer *Streptomyces flavopersicus*. Microbiology 143: 2135~2143, 1997
- 17) STEFFENSKY, M.; A. MUHLENWEG, Z. X. WANG, S. M. LI & L. HEIDE: Identification of the novobiocin biosynthetic

gene cluster of *Streptomyces spheroides* NCIB 11891. Antimicrob. Agents Chemother. 44(5): 1214~1222, 2000

- 18) CHIU, H. T.; B. K. HUBBARD, A. N. SHAH, J. EIDE, R. A. FREDENBURG, C. T. WALSH & C. KHOSLA: Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster. Proc. Natl. Acad. Sci. USA 98(15): 8548~8553, 2001
- 19) AHLERT, J.; J. DISTLER, K. MANSOURI & W. PIEPERSBERG: Identification of stsC, the gene encoding the Lglutamine:scyllo-inosose aminotransferase from streptomycin-producing Streptomycetes. Arch. Microbiol. 168: 102~113, 1997
- 20) HOPWOOD, D. A.; M. J. BIBB, K. F. CHATER, G. R. JANSSEN, F. MALPARTIDA & P. SMITH: In Regulation of gene expression. 25 Years On Symp. Soc. Gen. Microbiol. pp. 251~276, Ed., I. R. BOOTH & C. F. HIGGINS, Cambridge, University of Cambridge Press, 1986
- STROHL, W. R.: Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. Nucleic Acids Res. 20(5): 961~974, 1992
- 22) FUKAGAWA, Y.; T. SAWA, T. TAKEUCHI & H. UMEZAWA: Studies on biosynthesis of kasugamycin. III. Biosynthesis of the D-inositol moiety. J. Antibiotics 21(3): 185~188, 1968
- 23) RETZLAFF, L. & J. DISTLER: The regulator of streptomycin gene expression, StrR, of *Streptomyces griseus* is a DNA binding activator protein with multiple recognition sites. Mol. Microbiol. 18(1): 151~162, 1995
- 24) FERNANDEZ-MORENO, M. A.; J. L. CABALLERO, D. A. HOPWOOD & F. MALPARTIDA: The act gene cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA* tRNA gene of *Streptomyces coelicolor*. Cell 66: 769~780, 1991
- 25) WHITE, J. & M. BIBB: *bldA* dependence of undecylprodigiosin production in *Streptomyces coelicolor* A3(2) involves a pathway-specific regulatory cascade. J. Bacteriol. 179: 627~633, 1997
- 26) LAWLOR, E. J.; H. A. BAYLIS & K. F. CHATER: Pleiotropic morphological and antibiotic deficiences result from mutations in a gene encoding a tRNA-like product in *Streptomyces coelicolor* A3(2). Gene Dev. 1: 1305~1310, 1987
- 27) KHOKHLOV A. S.; I. I. TOVAROVA, L. N. BORISOVA, S. A. PLINER, L. A. SCHEVCHENKO, E. Y. KORNITSKAYA, N. S. IVKINA & I. A. RAPOPORT: A-factor, responsible for streptomycin biosynthesis by a mutant strains of *Actinomyces streptomycini*. Dokl. Akad. Nauk. SSSR 177: 232~235, 1967
- 28) HORINOUCHI, S. & T. BEPPU: A-factor as a microbial hormone that controls cellular differentiation and secondary metabolism in *Streptomyces griseus*. Mol. Microbiol. 12(6): 859~864, 1994