A 32 kb nucleotide sequence from the region of the lincomycin-resistance gene (22°–25°) of the *Bacillus subtilis* chromosome and identification of the site of the *lin-2* mutation

Miyuki Kumano, Atsuo Tamakoshi and Kunio Yamane

Author for correspondence: Kunio Yamane. Tel/Fax: +81 298 53 6680.
e-mail: kyamane@sakura.cc.tsukuba.ac.jp

Institute of Biological
Sciences, University of
Tsukuba, Tsukuba-shi,
Ibaraki 305, Japan

A 32 kb nucleotide sequence in the region of the lincomycin-resistance gene, located from 22° to 25° on the *Bacillus subtilis* chromosome, was determined. Among 32 putative ORFs identified, four [*lipA* for lipase, *natA*, *natB* and *yzsE* (renamed *yccK*)] have already been reported, although the functions of *NatA*, *NatB* and *YccK* remain to be characterized. Six putative products were found to exhibit significant similarity to known proteins in the databases, namely L-asparaginase precursor, protein aspartate phosphatase, α -glucosidase, two tellurite-resistance proteins and a hypothetical protein from *B. subtilis*. The region of the tellurite-resistance gene, consisting of seven ORFs, seems to correspond to an operon. The products of 14 ORFs exhibited considerable or limited similarity to known proteins. The sequenced region seems to be rich in membrane proteins, since at least 16 gene products appeared to contain membrane-spanning domains. The site of the *lin-2* mutation (two nucleotide replacements) was mapped and identified by sequencing. This site is located between a putative promoter and the SD sequence of *lmrA* (*yccB*) [a putative repressor of the *lmr* operon, which consists of *lmrA* and *lmrB* (*yccA*)]. *LmrB* is a homologue of proteins involved in drug-export systems and seems likely to be the protein responsible for resistance to lincomycin.

Keywords: genome sequencing, *Bacillus subtilis*, 22°–25° region, lincomycin-resistance gene, *lin-2* mutation

INTRODUCTION

A lincomycin-resistance mutation (*lin-2*) has been used as a fundamental genetic marker in the determination of map position of genes around the 25° region of the *Bacillus subtilis* chromosome (Goldthwaite *et al.*, 1970; Harford & Sueoka, 1970; Nomura *et al.*, 1978). Two different determinants of lincomycin-resistance have been reported in *Streptomyces lincolnensis* (Zhang *et al.*, 1992; Peschke *et al.*, 1995). One involved in the S-adenosyl-methionine-dependent methyltransferase which methylates 23S rRNA and mediates so-called MLS (macrolide/lincosamide/streptogramin B) resistance (Skinner *et al.*, 1983), whereas the other is involved in drug export (Neyfakh *et al.*, 1991; Rouch *et al.*, 1990).

As part of the *B. subtilis* genome cooperative project between European and Japanese laboratories to determine the entire sequence of the *B. subtilis* genome, we analysed a region of about 200 kb, from 20° to 36°, in which genes for α -amylase and the production of surfactin are located. The region that includes the lincomycin-resistance gene, comprising 35 kb, is located between the 21.5 kb of the pNEXT9 region (20°–22°; Ogawa *et al.*, 1995; accession no. D30808) and the 146 kb of the *amyE-srfA*-pNEXT59 region (25°–36°; Yamane *et al.*, 1996; accession nos D38181 and D50543). The sequence of a 32 kb portion of this 35 kb region was determined and analysed with respect to the amino acid sequences encoded by each ORF and their homologies to known proteins in the databases. The remaining two 1.5 kb sequences at the 5' and 3' ends of the 35 kb region were reported previously (Ogawa *et al.*, 1995; Yamane *et al.*, 1996). Thirty two genes were identified in the sequenced region. We then determined the site of the

The GenBank/EMBL/DBJ accession number for the sequence reported in this paper is AB000617.

lin-2 mutation in the chromosome of the lincomycin-resistant strain *B. subtilis* 1A221 and discuss a possible mechanism for resistance to lincomycin.

METHODS

Bacterial strains and a plasmid. *B. subtilis* 168 (*trpC2*) from our laboratory was used as the standard strain for sequencing. A lincomycin-resistant strain, *B. subtilis* 1A221 (*lin-2*), was obtained from the *Bacillus* Genetic Stock Center (Columbus, OH, USA). A lincomycin-resistant transformant of *B. subtilis* 168 [*B. subtilis* 168L (*lin-2*, *trpC2*)] was prepared in this study. Chromosomal DNA was extracted by the phenol/pH 9 method of Saito & Miura (1963). Plasmid pUC18 and *E. coli* JM109 were used for the preparation of randomly overlapping libraries for shotgun sequencing.

Transformation of *B. subtilis*. The method used for DNA-mediated transformation with competent cultures of *B. subtilis* 168 was a slightly modified version of that reported by Wilson & Bott (1968). Lincomycin (Sigma) was used at 75 µg ml⁻¹ for assaying resistance for the drug.

Sequencing, handling of data, and computer analysis. We constructed overlapping libraries of short fragments obtained by sonic oscillation (30 W, 30 s) as described by Ogasawara *et al.* (1994) and Ogawa *et al.* (1995). Sequences of fragments from the libraries were determined with an automated DNA sequencer (model 373, Applied Biosystems) with sequencing kits (Applied Biosystems). DNA sequences were compiled using the ATSQ program from Software Development, and putative ORFs were located using the GENETYX-MAC (Software Development) and GENEWORKS (IntelliGenetics) programs. The amino acid sequences encoded by ORFs were searched for similarity to sequences reported in non-redundant sequence

databases using the BLAST e-mail servers at the Human Genome Center, Institute of Medical Sciences, University of Tokyo, Tokyo, Japan.

RESULTS AND DISCUSSION

Preparation of fragments of the region from 22° to 25° of the *B. subtilis* chromosome for sequencing

The strategy for the cloning of DNA fragments from the region 22°–25° of the *B. subtilis* chromosome is summarized in Fig. 1. Since the 21.5 and 146 kb sequences in the pNEXT9 (20°–22°) and *amyE-srfA*-pNEXT59 (25°–36°) regions, respectively, were known (Ogawa *et al.*, 1995; Yamane *et al.*, 1996), four fragments of DNA for sequencing were prepared by inverse and standard long-range PCR (Triglia *et al.*, 1988; Saiki *et al.*, 1985). F1 (5.35 kb) and F4 (5.30 kb) fragments were prepared by inverse PCR with *NspV*- or *Clal*-digested and re-ligated preparations as template, respectively. After sequencing of the F1 and F4 fragments, the chromosomal DNA was digested by *SphI* and the F3 fragment (6.8 kb) was obtained by inverse PCR. Finally, the F2 fragment (15 kb) was obtained by standard long-range PCR.

Assignment of ORFs

A search was made for putative ORFs in the region from 22° to 25° region in all six possible translation frames. Based on the length of a sequence starting from ATG, GTG or TTG and preceded by a putative Shine–

Fig. 1. Cloning of DNA fragments for sequencing and localizing of ORFs in the lincomycin-resistance region (22°–25°) of the *B. subtilis* chromosome. The upper part shows the location of the DNA fragments prepared by inverse PCR (striped boxes) and standard long-range PCR (solid black box); white boxes indicate previously sequenced regions (Ogawa *et al.*, 1995; Yamane *et al.*, 1996). In the lower part, genes shown above the line are those whose transcription and translation occur in the same direction as the replication fork, and those shown below the line are transcribed in the opposite direction. Solid black boxes, function of the gene product known or reported; chequered boxes, the product shows strong similarity to known protein(s) in the databases (BLAST score ≥ 400); hatched boxes, considerable similarity (BLAST score 200–399); striped boxes, very low similarity (BLAST score ≤ 199); white boxes, gene product with no similarity to any known proteins.

Table 1. Putative ORFs in the 35 kb sequence around the 22°–25° region of the chromosome

ORF	Endpoints (nt)	Size of product (aa/kDa)	SD consensus sequence (upper core) and initiation codon (bold)
<i>pcp</i>	— < 513	—	AAAGGAGatgggagt ATG
<i>ycbU</i>	592 > 1701	370/40.8	AAAGGAGtgaatcata ATG
<i>yccA</i>	1749 < 3185	479/51.7	AGaGAGGgaa ATG
<i>yccB</i>	3228 < 3791	188/20.7	AAGGtGGGAatgtt ATG
<i>yccC</i>	4008 > 5132	375/40.1	AAAGGAGagaaatag ATG
<i>lipA</i>	5298 > 5933	212/22.8	AAAGGAGGatatt ATG
<i>yccE</i>	6073 > 6315	81/9.4	GGAGaacac ATG
<i>yccF</i>	6592 > 7665	358/40.0	GGAGGgtttatt GTG
<i>yccG</i>	7711 < 8664	318/35.0	GAGGAGtttagtac GTG
<i>yccH</i>	8680 < 9378	233/27.0	AGGGGAatgata ATG
<i>natA</i>	9522 > 10259	246/27.9	AGGGAGattgtgac ATG
<i>natB</i>	10778 > 11416	213/23.3	GAGGAGAAAGGGActtccgctatt ATG
<i>yccK</i>	11555 > 12484	310/34.8	AAAGGAGaaggaaca ATG
<i>ycdA</i>	12530 < 13591	354/39.2	AGGAGGaaaaatga ATG
<i>ycdB</i>	13919 > 15334	472/53.9	GGAGGgaacttaagaagg ATG
<i>ycdC</i>	15524 > 16886	453/52.2	AAAGGAGtgacttca TTG
<i>ycdD</i>	16896 < 17396	167/19.1	AAAGtGAGaccgat ATG
<i>ycdE</i>	17519 > 18637	373/44.4	GAGGAAGGAGGcagaccg ATG
<i>ycdF</i>	18747 > 19520	258/27.8	GAGGAGGatatcaagt ATG
<i>ycdG</i>	19548 > 21230	561/65.8	AGGgGGAAttgcggg ATG
<i>ycdH</i>	21421 > 22377	319/35.7	GAGGGGAtatacgat ATG
<i>ycdI</i>	22436 > 23128	231/26.3	AAAGGAagattaac ATG
<i>yceA</i>	23089 > 23928	280/30.4	GAGGAGaaaaagcgcg ATG
<i>yceB</i>	23972 < 24910	313/35.0	AAAGGAGaaagccct GTG
<i>yceC</i>	25248 > 25844	199/21.8	AGGAGtgaatcgt GTG
<i>yceD</i>	25869 > 26447	193/20.7	AAAGGAGtggtgaaac ATG
<i>yceE</i>	26485 > 27060	192/20.9	AAAGGAGaGAGGcagt ATG
<i>yceF</i>	27114 > 27944	277/31.4	AGGAGGaaaaaac TTG
<i>yceG</i>	27972 > 29582	537/62.8	AAAGGAGGatatg ATG
<i>yceH</i>	29601 > 30689	363/41.7	AAGGAGagaagaaaagc ATG
<i>yceI</i>	30814 > 32013	400/43.7	GAGGTgttggca ATG
<i>yceJ</i>	32272 < 33441	390/41.9	AGGAGtgataaaat TTG
<i>yceK</i>	33513 < 33812	100/11.5	ATG
<i>opuAA</i>	34102 > —	—	GGAGGgaaaat ATG

Dalgarno (SD) sequence, we identified 34 putative ORFs. Of these, the *pcp* and *opuAA* genes and part of the *ycbU* gene were reported previously (Fig. 1; Ogawa *et al.*, 1995; Yamane *et al.*, 1996). The initiation codon for 27 ORFs was ATG, for four ORFs it was GTG, and for three it was TTG. No suitable SD sequence was found for *yceK*. The direction of transcription and translation of 24 ORFs was the same as that of movement of the replication fork. In the other ten cases, this direction was the opposite. The organization of the 34 ORFs is summarized in Fig. 1 and Table 1.

Identification of previously reported genes

In addition to *pcp* (Awade *et al.*, 1992) and *opuAA* (Kempf & Bremer, 1995), four genes had previously

been reported (Table 2): the *lipA* gene for lipase (Dartois *et al.*, 1992); and the *natA*, *natB* and *yzae* genes (sp: NATA_BACSU, NATB_BACSU, and YZAE_BACSU). The designation of *yzae* has been changed to *yccK* in this report. The nucleotide sequences of the four genes that we determined were identical to those reported previously. The amino acid sequence of NatA showed considerable similarity to that of an ABC transporter ATP-binding protein of *Synechocystis* sp. PCC 6803 (BLAST score, 283; gp: D90905_77) and to that of a hypothetical ATP-binding transport protein of *Bacillus firmus* (sp: YATP-BACFI; BLAST score, 259). Nevertheless, the deduced amino acid sequences of NatB and YccK did not show any similarity to any proteins in the databases. Therefore the function of the three gene products remains to be characterized.

Table 2. Similarity of the predicted ORF products to other proteins in the databases and identified genes

Product	Similar protein in database	BLAST score	Reference
Pcp	Pyrrolidone-carboxylate peptidase	947	Awade <i>et al.</i> (1992)
YcbU	NifS-like protein of <i>Bacillus subtilis</i>	148	pir:D64064
YccA(lmrB)	Lincomycin-resistance protein	—	This study
YccB(lmrA)	Probable repressor of the lincomycin-resistance operon (<i>lmr</i>)	—	This study
YccC	L-Asparaginase precursor of <i>Erwinia chrysanthemi</i>	630	Minton <i>et al.</i> (1986)
LipA	Lipase	1134	Dartois <i>et al.</i> (1992)
YccE	Unknown	—	
YccF	Unknown	—	
YccG	Unknown	—	
YccH	Sensor-response regulator protein of <i>Staphylococcus aureus</i>	100	Brunskill & Bayles (1996)
NatA	ATP-binding transport protein (function unknown) of <i>Bacillus subtilis</i>	934	sp:NATA_BACSU
NatB	NatB protein (function unknown) of <i>Bacillus subtilis</i>	825	sp:NATB_BACSU
YccK	Hypothetical YzaE protein (function unknown) in the <i>natB</i> region of <i>Bacillus subtilis</i>	1152	sp:YZAE_BACSU
YcdA	Unknown	—	
YcdB	Unknown	—	
YcdC	Unknown	—	
YcdD	<i>ply</i> gene product (phage lysin) of bacteriophage A500	169	Loessner <i>et al.</i> (1995)
YcdE(RapJ)	Protein aspartate phosphatase (<i>gsiA</i>) of <i>Bacillus subtilis</i>	847	Mueller <i>et al.</i> (1992)
YcdF	Glucose 1-dehydrogenase of <i>Bacillus subtilis</i>	255	Lampel <i>et al.</i> (1986)
YcdG(GlcA)	Oligo-1,6-glucosidase of <i>Bacillus thermoglucosidasius</i>	1419	Watanabe <i>et al.</i> (1991)
	α -Glucosidase of <i>Bacillus</i> sp.	1324	Nakao <i>et al.</i> (1994)
YcdH	Adhesion protein precursor of <i>Streptococcus pneumoniae</i>	141	Sampson <i>et al.</i> (1994)
YcdI	<i>Streptococcus parasanguis</i> adhesin (<i>fimA</i>)	191	gp:STRSTRA_1
YceA	Hypothetical 32.8 kDa protein in YCF23-APCF intergenic region of <i>Cyanophora paradoxa</i>	317	gp:CPU30821-133
YceB	Hypothetical 36.6 kDa protein in <i>qoxD-vpr</i> region of <i>Bacillus subtilis</i>	475	sp:YWCH_BACSU
YceC	Tellurite-resistance protein (TerZ) of plasmid R478 from <i>Serratia marcescens</i> (tellurite-resistance-related protein)	334	Whelan <i>et al.</i> (1995)
YceD	Tellurite-resistance protein (TerD) of plasmid R478 from <i>Serratia marcescens</i> (tellurite-resistance-related protein)	527	Whelan <i>et al.</i> (1995)
YceE	Tellurite-resistance protein (TerE) of plasmid R478 from <i>Serratia marcescens</i> (tellurite-resistance-related protein)	494	Whelan <i>et al.</i> (1995)
YceF	Tellurite-resistance protein (TerC) of plasmid R478 from <i>Serratia marcescens</i> (tellurite-resistance-related protein)	95	Whelan <i>et al.</i> (1995)
YceG	Unknown (tellurite-resistance-related protein)	—	

Table 2. (cont.)

Product	Similar protein in database	BLAST score	Reference
YceG	Unknown (tellurite-resistance-related protein)	—	
YceH	Hypothetical 43·8 kDa protein in <i>xpaC-abrB</i> intergenic region of <i>Bacillus subtilis</i> (tellurite-resistance related protein)	278	sp:YAAN_BACSU
YceI	Transmembrane transporter SV2 of <i>Discopyge ommata</i> (tellurite-resistance-related protein)	232	Bindra <i>et al.</i> (1993)
YceJ	Multidrug-resistance protein 2 of <i>Bacillus subtilis</i>	158	Ahmed <i>et al.</i> (1995)
YceK	Unknown	—	
OpuAA	Glycine betaine/L-proline transport ATPase	2078	Kempf & Bremer (1995)

* pir, PIR protein sequence database; sp, SWISS-PROT protein sequence database; gp, translated protein sequence from GenBank.

Similarity of the gene products to known proteins

The amino acid sequences of the putative products of each ORF were compared with those in non-redundant protein databases and the results are summarized in Table 2. Among the 28 putative ORFs (the four previously reported ORFs were not included), the predicted amino acid sequences of six showed strong similarity (BLAST scores above 400) to known proteins in the databases, eight showed considerable similarity (BLAST scores 399–200) and six showed very limited similarity (BLAST scores 199–100). The remaining eight showed no similarity to known proteins. Among the 20 gene products that exhibited some similarity to known proteins in the databases, YceA was similar to a hypothetical protein in *Cyanophora paradoxa* (gp:CPU30821_133), and YceB and YceH were similar to proteins of unknown function in *B. subtilis* (sp:YWCH_BACSU; and sp:YAAN_BACSU). Thus the functions of 14 gene products, including NatA, NatB and YceK, remain unknown.

Among the gene products with strong similarity to other proteins, YcdG exhibited very strong similarity to oligo-1,6-glucosidase of *Bacillus thermoglucosidasius* (49% identity; Watanabe *et al.*, 1991) and to α -glucosidase of *Bacillus* sp. (51·6% identity; Nakao *et al.*, 1994). Therefore, YcdG might be the α -glucosidase (GlcA) of *B. subtilis*. YccC was very similar to L-asparaginase of *Erwinia chrysanthemi* (Minton *et al.*, 1986). The amino acid sequence of YcdE was very similar to that of protein aspartate phosphatase (GsiA) (44% identity) of *B. subtilis* (Mueller *et al.*, 1992). Perego *et al.* (1996) reported that six genes which were homologues of *gsiA* encode the family of aspartate-phosphate phosphatase (RapA–RapF). The sequence of YcdE was not identical to any of these gene products. No gene encoding a short peptide that resembled PhrA (phosphatase regulator) was found in the sequence corresponding to the C-

terminal region or downstream of the gene as in the cases of *rapB* and *rapD*. YcdE seems to be a new member of the protein aspartate phosphatase family and the corresponding gene was designated *rapJ*. *ycdH* and *ycdI* seem to encode homologues of adhesion proteins.

The IncHI2 plasmid R478 from *Serratia marcescens* encodes the tellurite-resistance operon, which consists of seven ORFs (*terZ*, -A, -B, -C, -D, -E and -F; Whelan *et al.*, 1995). The amino acid sequences of YceC, YceD, YceE and YceF showed strong and considerable similarity to TerZ (37·6% identity), TerD (57·0%), TerE (53·6%) and TerC (22·6%), respectively. YceD and YcdE were also quite similar to TerE (51·0%) and TerD (52·9%), respectively. The seven genes from *yceC* to *ycel* seemed to form an operon and the number of genes in the operon is equal to that in the tellurite-resistance operon of plasmid R478. The first four genes in the putative operon in *B. subtilis* were similar to those in plasmid R478, as described above, whereas the latter three were not.

Determination of the site of the *lin-2* mutation

To isolate it from other possible mutations in *B. subtilis* 1A221, the *lin-2* mutation was transferred to *B. subtilis* 168 by DNA-mediated transformation. A lincomycin-resistant transformant (*B. subtilis* 168L) was selected, and its chromosomal DNA was used as a template for preparing four DNA fragments by PCR (F5–F8), which covered the 32 kb sequenced region (Fig. 2). These DNA fragments were introduced into competent *B. subtilis* 168 cells. Lincomycin-resistant transformants were obtained only with the F5 fragment, which included *yccA*, the gene for a homologue of the multidrug-resistance protein (EmrB) of *E. coli*. In a computer search for promoters and terminators, we found that

Fig. 2. Location of DNA fragments for mapping of the site of the *lin-2* mutation. Nine fragments (F5, 7.7 kb; F6, 15.8 kb; F7, 5.9 kb; F8, 5.0 kb; F9, 2.0 kb; F10, 2.3 kb; F11, 4.0 kb; F12, 3.7 kb; F13, 1.5 kb) were prepared by standard PCR using the chromosomal DNA of *B. subtilis* 168L as the template, and their ability to confer resistance to lincomycin was determined by DNA-mediated transformation with *B. subtilis* 168 as the recipient strain. ++, More than 10^3 transformants ($\mu\text{g DNA}$)⁻¹ appeared; this is similar to the number of transformants obtained with *B. subtilis* 1A221 chromosomal DNA. -, No transformants.

Fig. 3. Nucleotide replacements at the site of the *lin-2* mutation in the genome of *B. subtilis* 1A221. Black squares indicate replaced nucleotides. The initiation codon and possible -35 and -10 regions of the putative promoter are boxed. The putative ribosome-binding sites (SD) are underlined.

covered the 32 kb sequenced region (Fig. 2). These DNA fragments were introduced into competent *B. subtilis* 168 cells. Lincomycin-resistant transformants were obtained only with the F5 fragment, which included *yccA*, the gene for a homologue of the multidrug-resistance protein (EmrB) of *E. coli*. In a computer search for promoters and terminators, we found that *yccA* formed an operon with *yccB*, and we designated this operon the *lmr* operon. We next prepared four other DNA fragments (F9-F12) by PCR. F9 included *pcp-yccU*, F10 included *yccA-yccB*, F11 included *yccB-yccF*, and F12 included *yccC-yccF*. The F10 and F11 fragments gave lincomycin-resistant transformants whereas F9 and F12 did not. These results indicated that the site of the *lin-2* mutation was located at the overlapping region of fragments F10 and F11, which includes a putative promoter region for the *lmr* operon. No lincomycin-resistant transformants were obtained

by F13, although it encodes another homologue (YceJ) of the multi-drug resistance protein, Bmr.

We sequenced the F10 fragment to determine the exact site of the *lin-2* mutation. The nucleotide sequences of *yccA* and *yccB* of *B. subtilis* 168L were identical to those of *B. subtilis* 168. However, two nucleotides, A and G residues at positions 3834 and 3819 in the chromosomal DNA of *B. subtilis* 1A221, were replaced by C and T, respectively. The mutations were located between the putative promoter and the SD sequence (Fig. 3). The organization of genes in the *lmr* operon, which consists of *yccA* and *yccB*, was identical to those of the *emrA/emrB* and *emrK/emrY* operons of *E. coli*. YccA consists of 479 amino acids (molecular mass, 51.7 kDa) and it contains at least 15 transmembrane domains. Furthermore, its amino acid sequence showed considerable similarity to the lincomycin-resistance protein

Table 3. Consensus amino acid sequences in antibiotic-transport proteins involved in efflux systems

Gene product*	Consensus amino acid sequences for		Reported phenotypic character	Reference
	Translocase	Drug-consensus		
BS LmrB	66 GILVPVSGLLQ	153 APAIGPTFSG	Lincomycin resistance	This study
BS Bmr	62 GRWVDRFGRKI	142 GFI IGPGGIG	Multidrug resistance	Neyfakh <i>et al.</i> (1991)
SL LmrA	87 GRIADLFGRRR	166 GLVLGFLLSG	Lincomycin resistance	Zhang <i>et al.</i> (1992)
EC EmrB	70 GWLAKRVGEUK	150 APICGPILGG	Multidrug resistance	Lomovskaya & Lewis (1992)
HI EmrB	74 GWLAKRFGEVR	164 APIFGPILGG	Multidrug resistance	Fleischmann <i>et al.</i> (1995)
EC EmrY	74 GRLAQRIGELK	154 APICGPILGG	Multidrug resistance	sp:EMRY_ECOLI
SG TcmA	87 GRIADLFGRRR	166 GLVLGFLLSG	Tetracenomycin resistance	Guilfoile & Hutchinson (1992a)
SA Pur8	83 GRVADAFGRRR	164 GAAVGLLAGG	Puromycin resistance	Tercero <i>et al.</i> (1993)
Tn10 TetA	62 GKMSDRFGRRP	142 GFI IGPGIGG	Tetracycline resistance	Postle <i>et al.</i> (1984)

* BS, *Bacillus subtilis*; EC, *Escherichia coli*; HI, *Haemophilus influenzae*; SA, *Streptomyces alboniger*; SG, *Streptomyces glaucescens*; SL, *Streptomyces lincolnensis*.

(LmrA; 23.2% identity) of *S. lincolnensis* 78-11, a producer of lincomycin (Zhang *et al.*, 1992). The sequence was also similar to those of multidrug-resistance proteins EmrB (25.0%; Lomovskaya & Lewis, 1992) and EmrY (25.2%; sp:EMRY_ECOLI) of *E. coli*, the QacA protein (27.2%) of *Staphylococcus aureus* (Rouch *et al.*, 1990), and others (Table 3), which consist of drug-efflux systems. Therefore, YccA is likely to be the lincomycin-resistance protein of *B. subtilis* and *yccA* was renamed *lmrB*. The amino acid sequence of LmrB from *B. subtilis* was very different from that of LmrB from *S. lincolnensis*, a homologue of rRNA methyltransferase, and from those of other methyltransferases. Furthermore, disruptants of the *lmrB* gene of *B. subtilis* 1A221 were sensitive to lincomycin. Therefore, it is possible that LmrB endows resistance to lincomycin through a drug-efflux system. Table 3 summarizes the consensus amino acid sequences of translocases and the drug-consensus sequences of various drug-resistance proteins, which are involved in efflux systems. The deduced molecular mass of YccB was similar to that of EmrA and EmrK of *E. coli*, whereas the amino acid sequence of YccB was different from that of those proteins. The sequence showed 21.8% identity to that of TcmR of *Streptomyces glaucescens*, which is a repressor of the *tcm* operon (tetracenomycin resistance) that consists of 11 genes. In TcmR, a DNA-binding motif (a helix–turn–helix structure) has been identified (Guilfoile & Hutchinson, 1992a, b). Using the rules developed by Chou & Fasman (1978), we found that YccB contains a putative helix–turn–helix structure from Ala70 to Leu90. This suggests that YccB might be a regulator of the *lmr* operon. Therefore, *yccB* was renamed *lmrA* of *B. subtilis*. Our results suggest that the *lin-2* mutation resulted in overexpression of *lmrB* by disruption of a binding site for a regulatory protein, which might be encoded by *lmrA*. To examine the mechanism of this overexpression of the *lmr* operon, we are now studying the disruption and expression of *lmrA* and *lmrB*.

The 35 kb region that we have sequenced is rich in genes for membrane proteins. At least 16 gene products (YcbU, lmrB, YccC, YccG, YccH, NatA, NatB, YccK, YcdA, YcdH, YcdI, YceA, YceF, YceG, YceI and YceJ) seem to be proteins with transmembrane domains. It is also of interest that this region is rich in genes that confer resistance to drugs and a metal. The products of ten genes (LmrA, LmrB, YccC–YceI and YceJ) appear to be associated with these properties.

ACKNOWLEDGEMENTS

The authors thank R. Ohki and A. Yamaguchi for discussion on multidrug-resistance genes. This study was supported by a Grant-in-Aid for Creative Research (Human Genome Project) from the Ministry of Education, Science and Culture of Japan.

REFERENCES

- Ahmed, M., Lyass, L., Markham, P. N., Taylor, S. S., Vazquez-Laslop, N. & Neyfakh, A. A. (1995). Two highly similar multidrug transporters of *Bacillus subtilis* whose expression is differentially regulated. *J Bacteriol* 177, 3904–3910.
- Awade, A., Cleuziat, P., Gonzales, T. & Robert-Baudouy, J. (1992). Characterization of the *pcp* gene encoding the pyrrolidone carboxy-peptidase of *Bacillus subtilis*. *FEBS Lett* 305, 67–73.
- Bindra, P. S., Knowles, R. & Buckley, K. M. (1993). Conservation of the amino acid sequence of SV2, a transmembrane transporter in synaptic vesicles and endocrine cells. *Gene* 137, 299–302.
- Brunskill, E. W. & Bayles, K. W. (1996). Identification and molecular characterization of a putative regulatory locus that affects autolysis in *Staphylococcus aureus*. *J Bacteriol* 178, 611–618.
- Chou, P. Y. & Fasman, G. D. (1978). Prediction of the secondary structure of proteins from their amino acid sequences. *Adv Enzymol* 47, 45–148.
- Dartois, V., Baulard, A., Schanck, K. & Colson, C. (1992). Cloning, nucleotide sequence and expression in *Escherichia coli* of a lipase gene from *Bacillus subtilis* 168. *Biochim Biophys Acta* 1131, 253–260.
- Fleischmann, R. D., Adams, M. D., White, O. & 37 other authors

- (1995). Whole-genome random sequencing of *Haemophilus influenzae* Rd. *Science* **269**, 496–512.
- Goldthwaite, C., Dubnau, D. & Smith, I. (1970). Genetic mapping of antibiotic resistance in markers *Bacillus subtilis*. *Proc Natl Acad Sci USA* **65**, 96–103.
- Guilfoile, P. G. & Hutchinson, C. R. (1992a). Sequence and transcriptional analysis of the *Streptomyces glaucescens* tcmAR tetracenomycin C resistance and repressor gene loci. *J Bacteriol* **174**, 3651–3658.
- Guilfoile, P. G. & Hutchinson, C. R. (1992b). The *Streptomyces glaucescens* TcmR protein represses transcription of the divergently oriented tcmR and tcmA genes by binding to an intergenic operator region. *J Bacteriol* **174**, 3659–3666.
- Harford, N. & Sueoka, N. (1970). Chromosomal location of antibiotic resistance markers in *Bacillus subtilis*. *J Mol Biol* **51**, 267–286.
- Kempf, B. & Bremer, E. (1995). OpuA, an osmotically regulated binding protein-dependent transport system for the osmo-protectant glycine betaine in *Bacillus subtilis*. *J Biol Chem* **270**, 16701–16713.
- Lampel, K. A., Uratani, B., Chaudhry, G. R., Ramaley, R. F. & Rudikoff, S. (1986). Characterization of the developmentally regulated *Bacillus subtilis* glucose dehydrogenase gene. *J Bacteriol* **166**, 238–243.
- Loessner, M. J., Wendlinger, G. & Scherer, S. (1995). Heterogeneous endolysins in *Listeria monocytogenes* bacteriophages. A new class of enzyme and evidence for conserved holin genes within the siphoviral lysis cassettes. *Mol Microbiol* **16**, 1231–1241.
- Lomovskaya, O. & Lewis, K. (1992). emr, an *Escherichia coli* locus for multidrug resistance. *Proc Natl Acad Sci USA* **89**, 8938–8942.
- Minton, N. P., Bullman, H. M. S., Scawen, M. D., Atkinson, T. & Gilbert, H. J. (1986). Nucleotide sequence of the *Erwinia chrysanthemi* NCPPB1066 L-asparaginase gene. *Gene* **46**, 25–35.
- Mueller, J. P., Bukusoglu, G. & Sonenshein, A. L. (1992). Transcriptional regulation of *Bacillus subtilis* glucose starvation-inducible genes: control of *gsiA* by the ComP-ComA signal transduction system. *J Bacteriol* **174**, 4361–4373.
- Nakao, M., Nakayama, T., Kakudo, A., Inohara, M., Harada, M., Omura, F. & Shibano, Y. (1994). Structure and expression of a gene coding thermostable α -glucosidase with a broad substrate specificity from *Bacillus* sp. SAM1606. *Eur J Biochem* **220**, 293–300.
- Neyfakh, A. A., Bidnenko, V. E. & Chen, L. B. (1991). Efflux-mediated multidrug resistance in *Bacillus subtilis*: similarities and dissimilarities with the mammalian system. *Proc Natl Acad Sci USA* **88**, 4781–4785.
- Nomura, S., Yamane, K., Sasaki, T., Yamasaki, M., Tamura, G. & Maruo, B. (1978). Tunicamycin-resistant mutants and chromosomal locations of mutational sites in *Bacillus subtilis*. *J Bacteriol* **136**, 818–821.
- Ogasawara, N., Nakai, S. & Yoshikawa, H. (1994). Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* chromosome containing the replication origin. *DNA Res* **1**, 1–14.
- Ogawa, K., Akagawa, E., Nakamura, K. & Yamane, K. (1995). Determination of a 21 548 bp nucleotide sequence around the 24° region of the *Bacillus subtilis* chromosome. *Microbiology* **141**, 269–275.
- Perego, M., Glaser, P. & Hoch, J. A. (1996). Aspartyl-phosphate phosphatases deactivate the response regulator components of the sporulation signal transduction system in *Bacillus subtilis*. *Mol Microbiol* **19**, 1151–1157.
- Peschke, U., Schmidt, H., Zhang, H.-Z. & Piepersberg, W. (1995). Molecular characterization of the lincomycin-production gene cluster of *Streptomyces lincolnensis* 78-11. *Mol Microbiol* **16**, 1137–1156.
- Postle, K., Nguyen, T. T. & Bertrand, K. P. (1984). Nucleotide sequence of the repressor gene of the TN10 tetracycline resistance determinant. *Nucleic Acids Res* **12**, 4849–4863.
- Rouch, D. A., Cram, D. S., Diberardino, D., Littlejohn, T. G. & Skurray, R. A. (1990). Efflux-mediated antiseptic resistance gene *qacA* from *Staphylococcus aureus*: common ancestry with tetracycline- and sugar- transport proteins. *Mol Microbiol* **4**, 2051–2062.
- Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985). Enzymic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350–1354.
- Saito, H. & Miura, K. (1963). Preparation of transforming deoxyribonucleic acid by phenol treatment. *Biochim Biophys Acta* **72**, 619–629.
- Sampson, J. S., O'Connor, S. P., Stinson, A. R., Tharpe, J. A. & Russell, H. (1994). Cloning and nucleotide sequence analysis of *psaA*, the *Streptococcus pneumoniae* gene encoding a 37-kilodalton protein homologous to previously reported *Streptococcus* sp. adhesins. *Infect Immun* **62**, 319–324.
- Skinner, R., Cundliffe, E. & Schmidt, F. J. (1983). Site of action of a ribosomal RNA methylase responsible for resistance to erythromycin and other antibiotics. *J Biol Chem* **258**, 12702–12706.
- Tercero, J. A., Lacalle, R. A. & Jimenez, A. (1993). The *pur8* gene from the *pur* cluster of *Streptomyces alboniger* encodes a highly hydrophobic polypeptide which confers resistance to puromycin. *Eur J Biochem* **218**, 963–971.
- Triglia, T., Peterson, M. G. & Kemp, D. J. (1988). A procedure for *in vitro* amplification of DNA segments that lie outside the boundaries of known sequences. *Nucleic Acids Res* **16**, 8186.
- Watanabe, K., Chishiro, K., Kitamura, K. & Suzuki, Y. (1991). Proline residues responsible for thermostability occur with high frequency in the loop regions of an extremely thermostable oligo-1,6-glucosidase from *Bacillus thermoglucosidasius* KP1006. *J Biol Chem* **266**, 24287–24294.
- Whelan, K. F., Collieran, E. & Taylor, D. E. (1995). Phage inhibition, colicin resistance, and tellurite resistance are encoded by a single cluster of genes on the IncHI2 plasmid R478. *J Bacteriol* **177**, 5016–5027.
- Wilson, G. A. & Bott, K. F. (1968). Nutritional factors influencing the development of competence in the *Bacillus subtilis* transformation system. *J Bacteriol* **95**, 1439–1449.
- Yamane, K., Kumano, M. & Kurita, K. (1996). The 25°–36° region of the *Bacillus subtilis* chromosome: determination of the sequence of a 146 kb segment and identification of 113 genes. *Microbiology* **142**, 3047–3056.
- Zhang, H.-Z., Schmidt, H. & Piepersberg, W. (1992). Molecular cloning and characterization of two lincomycin-resistance genes, *lmrA* and *lmrB*, from *Streptomyces lincolnensis* 78-11. *Mol Microbiol* **6**, 2147–2157.

Received 5 February 1997; revised 15 April 1997; accepted 21 April 1997.