JOURNAL OF BACTERIOLOGY, Dec. 1992, p. 7762–7769
0021-9193/92/237762-08$02.00/0

Vol. 174, No. 23

Discovery and Characterization of a New Transposable Element, Tn*4811*, in *Streptomyces lividans* 66

CARTON W. CHEN,[1]* TIN-WEIN YU,[1]† HUI-MING CHUNG,[1]‡ AND CHUN-FEN CHOU[2]

Institute of Genetics[1] and Institute of Microbiology and Immunology,[2] National Yang-Ming Medical College, Shih-Pai, Taipei, Taiwan, Republic of China

Received 29 July 1992/Accepted 28 September 1992

Transposition of a new 5.4-kb transposon, Tn*4811*, of *Streptomyces lividans* to the *melC* operon of *Streptomyces antibioticus* on plasmid pIJ702 was discovered. The nucleotide sequence of this copy of Tn*4811*, which contained an imperfect (9 of 11 bp) terminal inverted repeat, five putative *Streptomyces* coding sequences for an oxidoreductase and its transcription regulator, and three transposition-related proteins, was determined. SLP$^-$ strains of *S. lividans* contained one copy (*A*) of Tn*4811*, while SLP2$^+$ strains contained an additional copy (*B*) on the SLP2 plasmid. The nucleotide sequences at three insertion junctions of Tn*4811* were determined. Copy *B* lacked 41 bp from the left end. At the other five junctions the duplication of a putative 3-bp target sequence (TGA) was observed. A sequence of less than 3 kb homologous to Tn*4811* was present in *S. antibioticus*. DNA homologous to Tn*4811* was not detected in 14 other *Streptomyces* species.

Structural instability of genomic DNA is a widespread feature of gram-positive, filamentous, soil bacteria of the genus *Streptomyces* (32). Many genetic traits undergo spontaneous loss in laboratory cultures at frequencies of 10^{-4} to 10^{-2}. Exposure to certain stresses such as UV irradiation, DNA intercalating agents, cold temperature, or protoplasting and regeneration increases the frequency of these instabilities. The mutations involved in the instabilities are attributed to large deletions of chromosomal DNA, which are frequently accompanied by tandem amplifications of particular sequences nearby (32). The genetic principle(s) underlying the structural instability is not clear. Possible mechanisms involved are homologous recombination, site-specific recombination, and transposition. Homologous recombination systems in *Streptomyces* species and their contribution to genetic instability have received little study. Tsai and Chen (41) isolated a *rec* mutant defective in intraplasmid recombination but not in chromosomal recombination (25). Chou and Chen (11), in their investigation of an unstable *arg* gene, found no effect of this *rec* mutation on its instability. Site-specific recombination and transposition, although representing forms of fluidity of genomic DNA, have not been implicated in the instability of other DNA sequences (8).

About one-third of *Streptomyces* species can produce melanin pigment (44). The melanin (*melC*) operon of *Streptomyces antibioticus* has been cloned (24), sequenced (3), and widely used in many recombinant vectors. The *melC* operon in the *Streptomyces* species examined is genetically unstable, undergoing spontaneous deletions at relatively high frequencies—about 10^{-3} (32). Similar to those observed in other unstable genes in *Streptomyces* species, the deletions of *melC* were frequently accompanied by extensive tandem amplifications of specific sequences (18, 32).

The structure of the *melC* sequence on recombinant plasmids is also relatively unstable, being frequently observed in many laboratories. In this study we characterized the structural instabilities of *melC* on a recombinant plasmid, pIJ702, and discovered a new 5.4-kb transposon, designated Tn*4811* (independently discovered by Tanaka et al. [38]). The nucleotide sequence of Tn*4811*, which exhibited an imperfect terminal inverted repeat and several potential coding sequences with putative transpositional and accessory functions, was determined. Two copies of Tn*4811* were found in wild-type *Streptomyces lividans*. Putative duplications of target sequences were detected at the insertion junctions. Screening of 15 other *Streptomyces* species for sequences homologous to Tn*4811* by hybridization revealed only a less than 3-kb homolog in *S. antibioticus*.

MATERIALS AND METHODS

Bacterial cultures, plasmids, and phages. The bacterial strains and plasmids used in this study are in Tables 1 and 2, respectively. *S. lividans* TK64 *argX* was an Arg$^-$ mutant of TK64 that required arginine or ornithine for growth (11). *Streptomyces albus* B3381 and *Streptomyces griseorubens* IAM0060 were a gift from Ying-Chieh Tsai (National Yang-Ming Medical College, Taipei, Taiwan). *S. antibioticus* strains were grown on S agar or in S liquid medium (9) at 30°C. *S. lividans* was grown on R2YE agar or in TSB or YEME liquid medium (22) at 30°C. *Escherichia coli* cultures were grown on L agar or in L broth (27) at 37°C. VCSM13 helper phage was from Stratagene (La Jolla, California).

Enzymes and reagents. Restriction enzymes and T4 DNA ligase were from Boehringer Mannheim (Mannheim, Germany), Bethesda Research Laboratories (Gaithersburg, Md.), and New England Biolabs (Beverly, Mass.) and were used according to the specifications supplied. Thiostrepton was a gift from S. J. Lucania of Squibb Institute for Medical Research (Princeton, N.J.). Sequenase and *Taq*I polymerase sequencing kits were from United States Biochemical (Cleveland, Ohio) and Promega (Madison, Wis.), respec-

* Corresponding author.
† Present address: John Innes Institute, John Innes Center, Norwich NR4 7UH, United Kingdom.
‡ Present address: Department of Genetics and Development, Columbia University, New York, NY 10032.

TABLE 1. Bacterial strains used in this study

Bacterial strain	Genotype or characteristics	Source[a] or reference
E. coli XL-1 Blue	*recA1 lac endA1 gyrA96 thi hsdR17 supE44 relA1* (F' *proAB lacI*^{q} *lacZ*ΔM15 Tn*10*)	Stratagene
Streptomycetes		
S. lividans		
1326	Wild type (SLP2$^+$ SLP3$^+$)	22
TK19	SLP2$^-$ SLP3$^+$	22
TK20	SLP2$^+$ SLP3$^-$	22
TK21	SLP2$^-$ SLP3$^-$	22
TK23	SLP2$^-$ SLP3$^-$ *spc-1*	22
TK24	SLP2$^-$ SLP3$^-$ *str-6*	22
TK64	SLP2$^-$ SLP3$^-$ *pro-2 str-6*	22
TK64 *argX*	SLP2$^-$ SLP3$^-$ *pro-2 str-6 argX*	11
JT46	SLP2$^-$ SLP3$^-$ *pro-2 str-6 rec-46*	41
S. alboniger ATCC 12461	Wild type	CCRC
S. albus B3381	Wild type	Y.-C. Tsai
S. antibioticus IMRU 3720	Wild type	24
S. argenteolus ATCC 11009	Wild type	CCRC
S. cattleya IFO 14057	Wild type	CCRC
S. coelicolor M130	*uraA1 argA1 strA1* SCP1$^-$ SCP2$^-$	22
S. glaucescens ATCC 23622	Wild type	CCRC
S. griseorubens IAM0060	Wild type	Y.-C. Tsai
S. griseus ATCC 10137	Wild type	CCRC
S. hydrogenans ATCC 19631	Wild type	CCRC
S. kanamyceticus ATCC 12853	Wild type	CCRC
S. paryulus 2283 (from ATCC 12434)	*str-1*	22
S. moderatus ATCC 23443	Wild type	CCRC
S. nogalater ATCC 27451	Wild type	CCRC
S. sparsogenes ATCC 25498	Wild type	CCRC

[a] CCRC, Culture Collection and Research Center, Food Industry Research and Development Institute, Hsinchu, Taiwan.

tively. KS and SK sequencing primers were from Stratagene. [α-^{32}P]d-eoxynucleoside triphosphate and [α-^{35}S]dATP were from Amersham (Buckinghamshire, England).

DNA manipulations and genetic transformations. Standard recombinant DNA techniques were the methods of Sambrook et al. (29). Preparation, transformation, and regeneration of *Streptomyces* protoplasts were by the methods of Hopwood et al. (22).

DNA sequencing. pLUS729 was digested with *Bam*HI into three fragments, subcloned into pBluescript II KS(−) and propagated in *E. coli* XL1-B (6). These subclones were subsequently dissected with other restriction enzymes and further subcloned into 25 recombinant plasmids (based on the pBluescript II vectors) for sequencing. Single-stranded plasmid DNA was rescued by using VSCM13 helper phage and sequenced with the dideoxy method of Sanger et al. (30) by using the Sequenase or the *Taq*I Polymerase Sequencing Kit. Primers were either the SK or KS primer or synthetic oligonucleotides based on the sequences determined. Oligonucleotides were synthesized with an oligonucleotide synthesizer (model 380B; Applied Biosystems). pLUS796, pLUS797, pLUS798, and pLUS799 DNA containing the various insertion junctions on pBluescript II KS(+)/KS(−) (Table 2) were sequenced directly without subcloning.

Computer analyses of DNA and protein sequences. DNA and protein sequences were analyzed by the University of Wisconsin Genetic Computing Group (UWGCG) package (version 7.0), PC/Gene (IntelliGenetics, release 6.60), or GeneWorks (IntelliGenetics, release 2.01). The protein and DNA data bases used for homology search were SwissProt

TABLE 2. Plasmids used in this study

Plasmid	Description	Source or reference
SLP2	50-kb linear conjugative plasmid of *S. lividans*	10
pBluescript II KS(+)/KS(−)	*E. coli* phagemid cloning vectors, *bla lacZ*α	Stratagene
pIJ702	*Streptomyces* plasmid pIJ350 containing *melC* of *S. antibioticus*	24
pLUS729	pIJ702 *melC2*::Tn*4811*	This study
pLUS791	pBluescript II KS(−) carrying the 5.0-kb *Kpn*I-*Sst*I fragment of Tn*4811* from pLUS729	This study
pLUS796	pBluescript II KS(−) carrying the 0.45-kb *Kpn*I-*Eco*RI fragment encompassing the left terminus of copy *A* of Tn*4811*	This study
pLUS797	pBluescript II KS(+) carrying the 0.40-kb *Bam*HI-*Pst*I fragment containing the right terminus of copy *A* of Tn*4811*	This study
pLUS798	pBluescript II KS(+) carrying the 4.0-kb *Eco*RI-*Cla*I fragments of SLP2 containing the left terminus of copy *B* of Tn*4811*	This study
pLUS799	pBluescript II KS(+) carrying the 4.7-kb *Eco*RI-*Cla*I fragments of SLP2 containing the right terminus of copy *B* of Tn*4811*	This study

(version 18.0) and EMBL (version 27.0). Prosite (version 4.00 in PC/Gene) was used for searching for signals and motifs in proteins.

RESULTS AND DISCUSSION

***melC* on pIJ702 was structurally unstable.** Recombinant plasmid pIJ702 (24) contains the cloned *melC* sequence of *S. antibioticus*, which confers melanin production to Mel$^-$ hosts, *S. lividans*, and others. Spontaneous Mel$^-$ variants were detected at high frequencies in pIJ702 transformants of *S. lividans* TK64. The frequencies were in the order of 10^{-3} in fresh cultures from single Mel$^+$ reisolates or transformants but appeared to rise as the cultures were progressively propagated. In some cases the Mel$^-$ phenotype was deduced to be due to an unidentified mutation in the *S. lividans* host, because plasmid DNA isolated from these cultures still conferred melanin production to secondary transformants. This genetic determinant also appeared to be unstable; mutations in it were readily detected and accumulated in cultures propagated without reisolation (up to a frequency of 30% in very old cultures) (40). These mutants were not explored further. In other cases, however, the mutations were evidently on the plasmid (see below), because secondary transformants by these plasmids were Mel$^-$.

Analysis of plasmids isolated from these Mel$^-$ transformants showed that rearrangements had occurred in many of them. Of 11 Mel$^-$ plasmids analyzed, 3 had 1.0- to 1.8-kb deletions including part of *melC*, 2 showed no apparent changes in size, and 6 contained a 5.4-kb insert.

Insertional mutations by a new transposable element, Tn*4811*. Digestion of the six plasmids with a 5.4-kb insert with *Bcl*I produced the same seven fragments (not shown). Four of the restriction sites were on pIJ702 (24), and the other three must be on the insert. It thus appeared that either the insert had integrated at approximately the same position on these six plasmids or these plasmids were of the same origin. One of them, pLUS729 (Fig. 1), was chosen for further characterization. Restriction mapping of pLUS729 showed that the insertion was between the *Sst*I and *Sal*I sites within the *melC2* (tyrosinase) gene.

The size and restriction map of the inserted sequence were identical to a mobile DNA element in a preliminary report by Tanaka et al. (36, 38). This element was originally designated IS*456*, but because of the complexity of its genetic content, we (Tanaka's and our laboratories) now regard it a type II transposon (see below) and have renamed it Tn*4811*.

Determination of nucleotide sequence of Tn*4811*. The copy of Tn*4811* on pLUS729 was subcloned into 25 phagemids and sequenced in both directions by using synthetic oligonucleotide primers. The Tn*4811* sequence was determined by comparison of the compiled sequence with the *melC* sequence. Between the putative duplicated target sequences, TGA (see below), the total length of Tn*4811* is 5,396 bp (Fig. 2). The average G+C content is 68%. There is an imperfect inverted terminal repeat (a match of 9 of 11 bp).

Codon preference analysis (16; data not shown), based on the preference for G/C in the third position of codons by typical *Streptomyces* genes (5), revealed five putative coding sequences, open reading frame 1 (ORF1) to ORF5. Prediction with the TESTCODE algorithm (14; data not shown) also identified the same regions as coding sequences. The sizes of the proteins encoded by ORF1 to ORF5 were 336, 138, 320, 297, and 304 amino acids, respectively.

ORF4 and ORF5 as a putative accessory unit. The protein data bases were searched for sequences similar to the

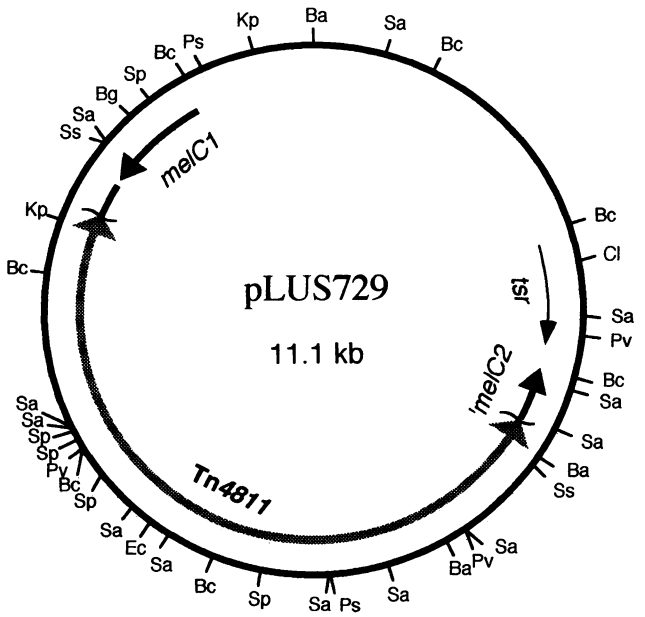

FIG. 1. Restriction map of pLUS729. The *melC1* and *melC2* sequences are indicated by the thick solid arrows, and the *tsr* gene is indicated by the thin solid arrow. Tn*4811* is indicated by the gray double-headed arrow. The two sequenced restriction fragments that contain the insertion junctions are indicated by the two outermost arcs. Abbreviations for restriction enzymes: Ba, *Bam*HI; Sa, *Sal*I; Bc, *Bcl*I; Cl, *Cla*I; Pv, *Pvu*II; Ss, *Sst*I; Ps, *Pst*I; Sp, *Sph*I; Ec, *Eco*RI; Kp, *Kpn*I; Bg, *Bgl*II.

potential protein products encoded by the five ORFs. Only the ORF4 and ORF5 products found significantly similar protein sequences. The ORF4 product is similar to a group of oxidoreductases (Fig. 3), including the protochlorophyllide reductase (Pcr) of barley (33), the glucose dehydrogenase of *Bacillus megaterium* (19), the ribitol dehydrogenase of *Klebsiella aerogenes* (28), and the NodG protein of *Rhizobium meliloti* (which is homologous to β-ketoacyl reductase [34]). The similarity among these oxidoreductases was stronger at the N termini. The ORF4 product was most similar to Pcr, with an identity of 35% and a similarity of 48% (Fig. 3). Although the putative ORF4 product is likely also an oxidoreductase, its function is not clear.

The ORF5 product is similar to a group of bacterial transcription regulators of the AraC (arabinose operon regulator) family. Figure 4 shows an optimized alignment of the ORF5 product with five members of the AraC family. AraC of *Erwinia carotovora* exhibited the highest homology with the ORF5 product (26% identity and 50% similarity). Along the polypeptide sequences the strongest similarity lies in the C termini that contain potential helix-turn-helix DNA-binding domains (17), also seen in the ORF5 product.

The divergently oriented ORF4 and ORF5 are separated by 188 bp of relatively low G+C content. This region is likely to contain the promoter and regulatory sequences involved in the divergent transcription of the two putative genes. The distances between the translation start sites in the known divergent promoters of prokaryotes range from 103 to 611 bp (2). Beck and Warren (2) reported that in the 45 compiled regions of divergent transcription in prokaryotes, 18 were of the regulatory-structural type. With one exception (*micF-ompC* of *E. coli*), in which the regulator RNA

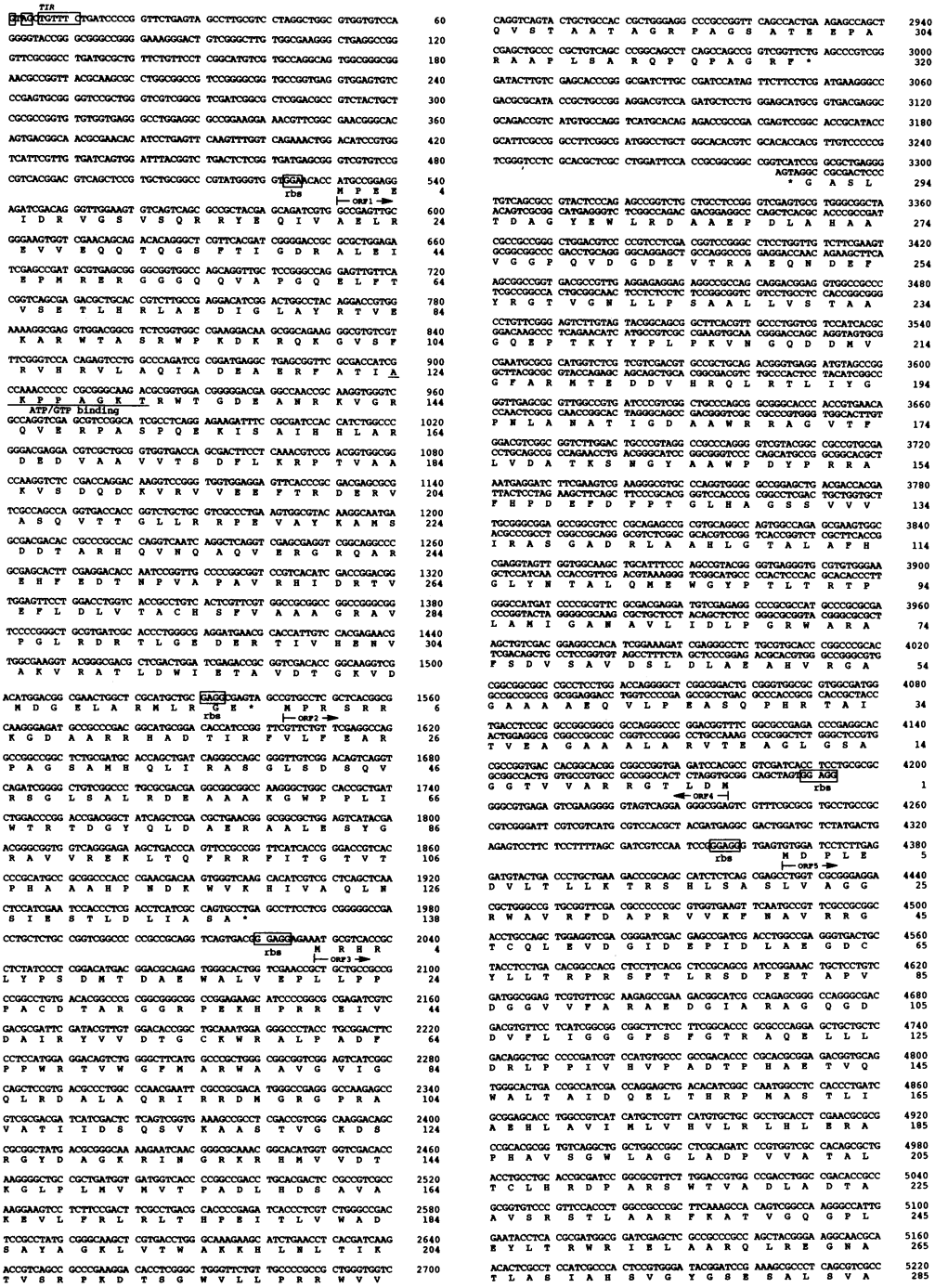

FIG. 2. Nucleotide sequence and potential protein products of Tn*4811* isolated from pLUS729. The nucleotide sequence between the putative duplicated target sequences (TGA) is shown. The imperfect terminal inverted repeats (*TIR*) and the putative ribosome-binding sequences (rbs) are boxed. The five putative coding sequences (ORF1 to ORF5) and their translation products are indicated. The asterisks denote translation termination codons. The putative motif A of an ATP and GTP binding sequence in the ORF1 product is underlined. For ORF4, which runs in the reverse orientation, both strands are shown.

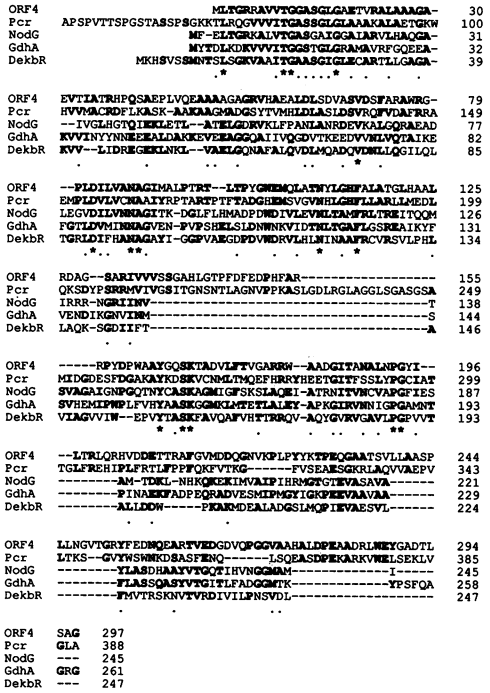

FIG. 3. Similarity of the ORF4 product to oxidoreductases. The protein product of ORF4 was aligned with the amino acid sequences of Pcr of barley (33), nodulation protein G (NodG) of *R. meliloti* (13), glucose dehydrogenase (GdhA) of *B. megaterium* (19), and ribitol dehydrogenase (DekbR) of *K. aerogenes* (28) by using the CLUSTAL program of PC/GENE. A K-tuple value of 1, a gap penalty of 5, a window of 15, a filtering level of 1.5, an open gap cost of 10, and a unit gap cost of 10 were used. The bold characters denote homologous amino acids. Among the aligned sequences, perfectly conserved residues are indicated by an asterisk below, and well conserved residues are indicated by a dot. Pcr is most similar to the ORF4 product, with an identity score of 35% and a similarity of 48% by using the GAP program of the UWGCG package (gap weight of 3.0, length weight of 0.1, average match of 0.54, and average mismatch of 0.396). The first 74 amino acids of Pcr is supposedly a transit peptide (33). Pcr is shown starting from amino acid 51, whereas the others start at amino acid 1.

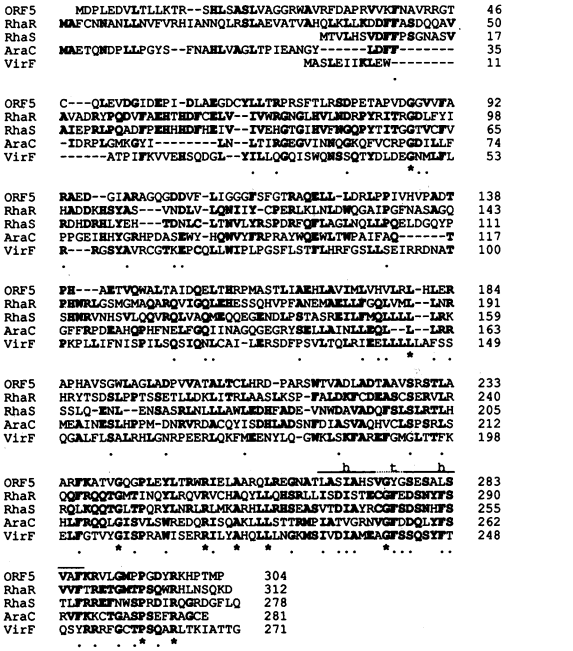

FIG. 4. Similarity of the ORF5 product to the AraC family of transcription regulators. The CLUSTAL program of PC/GENE was used to align the ORF5 product with four proteins with the highest similarity: the rhamnose operon transcription activators, RhaR and RhaS, of *E. coli* (39), the arabinose operon regulator AraC of *E. carotovora* (26), and the virulence regulon transcriptional activator VirF of *Yersinia* species (12). The parameters and notations used are as in Figure 3. The predicated helix-turn-helix (h t h) DNA-binding domains are indicated by the overline. AraC is most similar to the ORF5 product, with an identity score of 26% and a similarity of 50% by using the GAP program of the UWGCG package (gap weight of 3.0, length weight of 0.1, average match of 0.54, and average mismatch of 0.396).

directly regulates a gene located elsewhere, all the regulatory polypeptides act within the intergenic region. It is most likely that the ORF5 product also regulates the divergent transcription of ORF4 and ORF5. The notion that ORF4 and ORF5 form a functional unit is supported by our finding of a 2.2-kb sequence in *S. antibioticus* containing highly conserved sequences of ORF4 and ORF5 in the same arrangement (45) (see below).

ORF1 to ORF3 as possible transposition proteins. Although no significant similarity to protein sequences in the data bases was observed for ORF1 to ORF3, a low level of similarity of the putative ORF1 and ORF2 products to some DNA-binding proteins and retroviral DNA polymerases was detected. ORF1 contained a consensus ATP/GTP-binding motif A (43) or P-loop (31), A/G-X_4-G-K-S/T, at residues 124 to 131 (Fig. 2). Proteins possessing this motif include several bacterial helicases (such as DnaB, RecB, Rho, and UvrD), viral proteins essential for replication (UL5 of herpes simplex virus, BBLF4 of Epstein-Barr virus, PS3 of human cytomegalovirus, and the product of gene 55 of varicella-zoster virus), and the yeast PIF protein involved in mitochondrial DNA recombination (21). All these proteins are known or presumed to be involved in nucleic acid replication and/or recombination. The ORF1 product may take part in the replication and/or recombination process during transposition.

Furthermore, the predicted pIs of the ORF2 and ORF3 products were very high (10.9 and 11.2, respectively). This is consistent with the fact that many characterized transposase proteins are very basic (15). Thus, we tentatively suppose that these are genes involved in transposition and its regulation. The lack of homology between these ORFs and other transposon-encoded products is not uncommon. The putative transposition-related proteins encoded by *Streptomyces* transposable elements, IS*110* (7), IS*117* (20), and IS*493* (35), also showed no homology to other known transposases.

Putative transcriptional and translational signals. The basic features of Tn*4811*, including the distributions of restriction sites, potential coding sequences, and G+C content, are summarized in Fig. 5. Two distinct regions of relatively low G+C contents were discerned: one between ORF4 and ORF5 (53% G+C) and the other at 170 to 60 bp upstream from ORF1 (50% G+C). The low G+C is typical for putative promoter-containing regions in *Streptomyces* species. We propose that these regions contain the promoter sequences.

ORF3, ORF4, and ORF5 had a putative 5-bp ribosome-binding sequence (1, 4), GGAGG, located 4, 11, and 6 bp, respectively, upstream from the initiation codon. ORF1 had

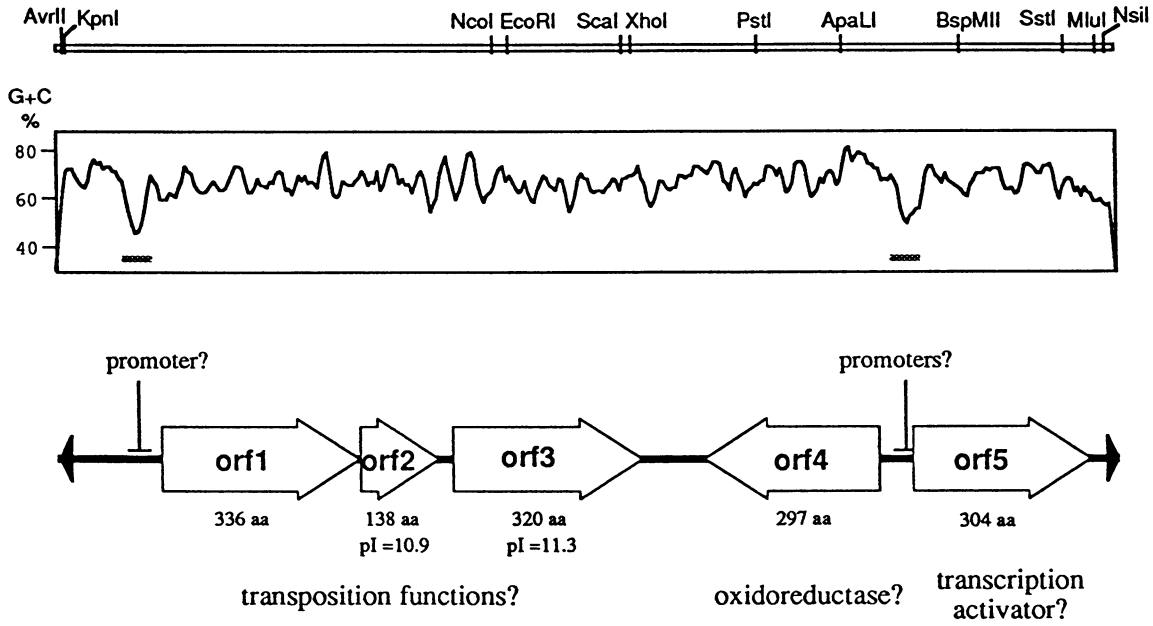

FIG. 5. Overview of the organization of Tn*4811*. The top panel shows the unique cleavage sites for several restriction enzymes. The middle panel shows the G+C content averaged over 50 nucleotides. The two particularly low G+C regions proposed to be promoter regions are indicated by the gray horizontal bars. In the lower panel, solid arrows at the termini indicate the imperfect inverted repeats. The five potential coding sequences are shown by the open arrows together with their sizes (in amino acids) and putative functions. The high pI values of ORF2 and ORF3 are indicated.

a 3-bp GGA sequence 5 bp upstream, and ORF2 had a 4-bp GAGG sequence 9 bp upstream, as a potential ribosome-binding sequence. The initiation codon of ORF2 was separated by only 2 bp from the termination codon of ORF1 (Fig. 2).

Occurrence of Tn*4811*. The 5.0-kb *Kpn*I-*Sst*I fragment that spanned more than 90% of Tn*4811* was subcloned on pBluescript II KS(+) to yield pLUS791 and used as a hybridization probe in subsequent experiments. Southern hybridization (Fig. 6) revealed a single copy of Tn*4811* in the genome of TK64, which had been genetically characterized, but not physically proven, to be plasmidless (23). Thus, the location of copy *A* cannot be certain. On the other hand, an additional copy of Tn*4811* was present in wild-type *S. lividans* 1326 and its SLP2$^+$ derivatives (see legend to Fig. 6). We designated the copy of Tn*4811* in TK64 copy *A* and designated the other copy *B*. Copy *B* appeared to cosegregate with SLP2. We have recently isolated SLP2 DNA and confirmed the presence of Tn*4811* on it (10).

The restriction and hybridization patterns of the two copies of Tn*4811* were indistinguishable. The restriction maps of the region spanning the two copies were determined by a series of Southern hybridization analyses (Fig. 6). Unexpectedly, the homology between the two copies extended further from the right end of Tn*4811*, whereas the sequences flanking the left ends were different. The homology to the right extended for about 10 kb and approximately coincides with the right end of SLP2 DNA (10).

Tn*4811* in *S. lividans* appeared quite stable. During this and other studies, we have never detected its movement on the chromosome (by using the not-very-sensitive Southern hybridization assay) or to a plasmid other than the ones reported here. Tanaka et al., however, reported the transpositions of Tn*4811* to plasmids (38) and the rearrangement of Tn*4811*-containing sequences in ethidium bromide-induced mutants of *S. lividans* 1326 (37).

Fifteen other *Streptomyces* species (Table 1) were probed for the presence of sequences homologous to Tn*4811*. Only *S. antibioticus* contained a sequence (less than 3.0 kb) of significant homology (data not shown). This *S. antibioticus* sequence, designated V1, is homologous to the ORF4 and ORF5 sequences of Tn*4811* (45).

The junction sequences of three Tn*4811* insertions. A cosmid clone containing the copy *B* sequence was isolated from a library of SLP2 DNA (10) by hybridization with a Tn*4811* probe. A cosmid clone containing the genomic copy *A* sequence was isolated from a library of TK64 (SLP2$^-$). Junction fragments from these two insertions were subcloned in pBluescript II KS(+) (Table 2) and their nucleotide sequences were determined. These junction sequences together with those for copy *A* on pLUS729 are aligned and shown in Fig. 7.

Copy *A* of Tn*4811* from the genomic DNA of TK64 and the copy from pLUS729 are identical in their terminal sequences, which is consistent with the fact that the copy on pLUS729 originated from the TK64 genome. Interestingly, the putative (duplicated) target sequence (TGA) of the insertion on TK64 genome is identical to that on pLUS729. Copy *B* isolated from SLP2 DNA is 41 bp shorter from the left terminus than copy *A*. Again, the right junction is flanked by the TGA triplet.

The insertion of Tn*4811* into a TGA (*opal*) termination codon is not likely to affect the expression of the gene. This potentially allows Tn*4811* to transpose with a reduced probability of damaging the host and being detected by genetic screening. In the insertional inactivation of *melC* on pLUS729, the putative target was an out-of-frame TGA triplet. If the insertion were at the (in-frame) termination

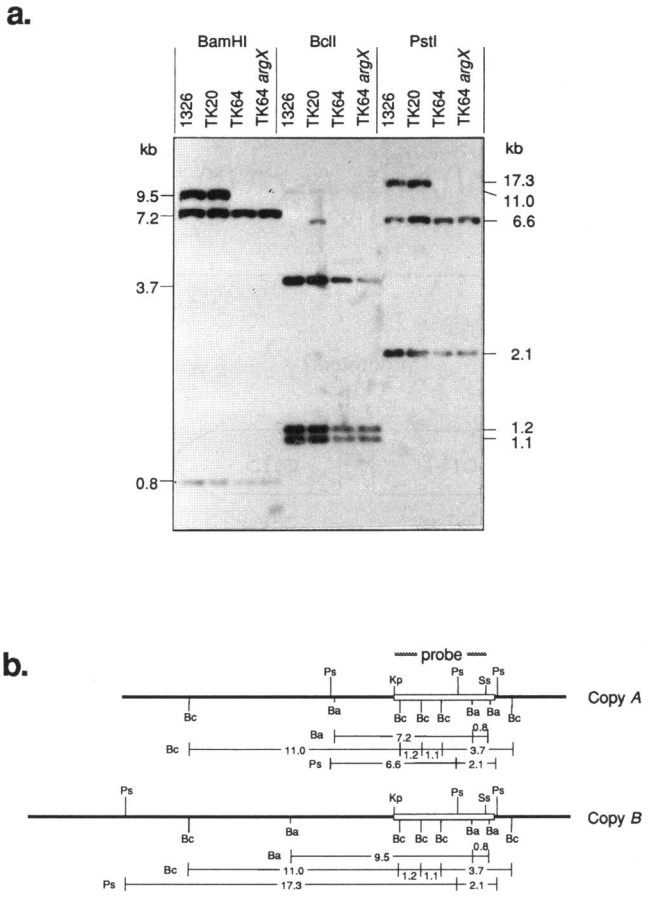

FIG. 6. Two copies of Tn*4811*. (a) Genomic DNA was digested with the restriction enzymes indicated and hybridized with the radioactively labeled 5.0-kb *Kpn*I-*Sst*I fragment from pLUS791. The sizes of the hybridizing fragments are indicated. All the *S. lividans* strains in Table 1 were tested, but only selected results are shown here. All the SLP2$^+$ strains gave the same hybridization pattern as 1326 (with both copies of Tn*4811*), whereas all the SLP2$^-$ strains, like TK64, had copy *A* only. Hybridization to the 11.0-kb *Bcl*I fragments was very weak but real. The hybridizing 6.6-kb *Bcl*I fragment in TK20 DNA was due to incomplete digestion. (b) The restriction maps of regions spanning the two copies of Tn*4811* based on the Southern hybridization data. Tn*4811* is represented by the open bars, and the neighboring sequences are represented by the solid lines. The probe and the fragments that hybridized (sizes in kilobases) are indicated. Abbreviations for restriction sites are as in the legend to Fig. 1.

codon TGA, its presence might not have been detected. In this connection, TGA is by far the most frequently used termination codon in *Streptomyces* species (42). Analyses of more insertion sites will be necessary to determine the precise target specificity of Tn*4811*.

ACKNOWLEDGMENTS

This work was supported by research grants from National Science Council to C.W.C. (NSC78-0412-B010-08 and NSC79-0412-B010-08). C.W.C. was a recipient of a research award from the Medical Research and Advancement Foundation in memory of Chi-Shuen Tsou.

We thank David A. Hopwood and Stanley N. Cohen for helpful comments and suggestions for improvement of the manuscript, M. Tanaka for exchange of information and discussion, and Helen M. Kieser and Ying-Chieh Tsai for providing the *Streptomyces* strains.

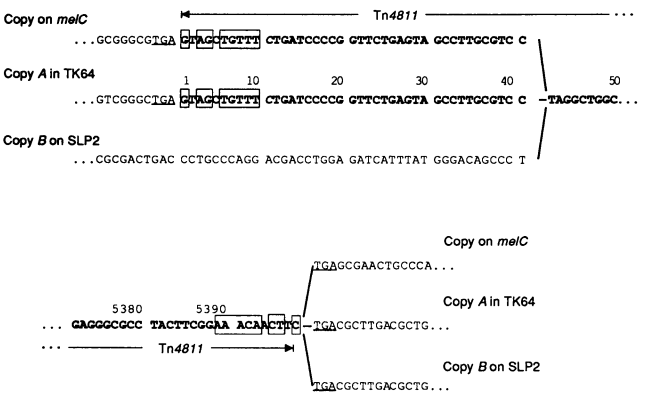

FIG. 7. Junction sequences at the three insertion sites of Tn*4811*. The junction sequences of copy *A* in *melC* on pLUS729, copy *A* on the TK64 genome (cloned in pLUS796 and pLUS797), and copy *B* on SLP2 (cloned in pLUS798 and pLUS797) were determined. The relevant nucleotide sequences at these junctions are aligned. The Tn*4811* sequences are shown in boldface type and numbered, and the flanking sequences are in plain type. The sequence shared by all three copies was shown only in one (middle) line. The imperfect terminal repeats of Tn*4811* are boxed. The putative target sequences (TGA) are underlined.

REFERENCES

1. **Baylis, H. A., and M. J. Bibb.** 1988. Transcriptional analysis of the 16S rRNA gene of the *rrnD* gene set of *Streptomyces coelicolor* A3(2). Mol. Microbiol. **2**:569–579.
2. **Beck, C. F., and R. A. J. Warren.** 1988. Divergent promoters, a common form of gene organization. Microbiol. Rev. **52**:318–326.
3. **Bernan, V., D. Filpula, W. Herber, M. Bibb, and E. Katz.** 1985. The nucleotide sequence of the tyrosinase gene from *Streptomyces antibioticus* and characterization of the gene product. Gene **37**:101–110.
4. **Bibb, M. J., and S. N. Cohen.** 1982. Gene expression in *Streptomyces*: construction and application of promoter-probe plasmid vectors in *Streptomyces lividans*. Mol. Gen. Genet. **187**:265–277.
5. **Bibb, M. J., P. R. Findlay, and M. W. Johnson.** 1984. 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**:157–166.
6. **Bullock, W. O., J. M. Fernandez, and J. M. Short.** 1987. XL1-B: a high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. BioTechniques **5**:376–379.
7. **Burton, C. J., and K. F. Chater.** 1987. Nucleotide sequence of IS*110*, an insertion sequence of *Streptomyces coelicolor* A3(2). Nucleic Acids Res. **15**:7053–7065.
8. **Chater, K. F., D. J. Henderson, M. J. Bibb, and D. A. Hopwood.** 1988. Genetic flux in *Streptomyces coelicolor* and other streptomycetes and its possible relevance to the evolution of mobile antibiotic resistance determinants, p. 8–42. *In* A. J. Kingston, S. M. Kingsman, and K. F. Chater (ed.), Transposition. Cambridge University Press, Cambridge.
9. **Chater, K. F., D. A. Hopwood, T. Kieser, and C. J. Thompson.** 1982. Gene cloning in *Streptomyces*. Curr. Top. Microbiol. Immunol. **96**:69–95.
10. **Chen, C. W., T.-W. Yu, Y.-S. Lin, H. Kieser, and D. A. Hopwood.** Submitted for publication.
11. **Chou, C.-F., and C. W. Chen.** Submitted for publication.
12. **Cornelis, G., C. Sluiters, C. L. de Rouvroit, and T. Michiels.** 1989. Homology between VirF, the transcriptional activator of

the *Yersinia* virulence regulon, and AraC, the *Escherichia coli* arabinose operon regulator. J. Bacteriol. **171:**254–262.
13. **Debelle, F., and S. R. Sharma.** 1986. Nucleotide sequence of *Rhizobium meliloti* RCR2011 genes involved in host specificity of nodulation. Nucleic Acids Res. **14:**7453–7472.
14. **Fickett, J. W.** 1982. Recognition of protein coding regions in DNA sequences. Nucleic Acids Res. **10:**5303–5318.
15. **Galas, D. J., and M. Chandler.** 1989. Bacterial insertion sequences, p. 109–162. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
16. **Gribskov, M., J. Devereux, and R. R. Burgess.** 1984. The codon preference plot: graphic analysis of protein coding sequences and prediction of gene expression. Nucleic Acids Res. **12:**539–549.
17. **Harrison, S. C., and A. K. Aggarwal.** 1990. DNA recognition by proteins with the helix-turn-helix motif. Annu. Rev. Biochem. **59:**933–969.
18. **Haüsler, A., A. Birch, W. Krek, J. Piret, and R. Hütter.** 1989. Heterogeneous genomic amplification in *Streptomyces glaucescens*: structure, location and junction sequence analysis. Mol. Gen. Genet. **217:**437–446.
19. **Heilman, H. J., H. J. Maegert, and H. G. Gassen.** 1988. Identification and isolation of glucose dehydrogenase genes of *Bacillus megaterium* M1286 and their expression in *Escherichia coli*. J. Biochem. **174:**485–490.
20. **Henderson, D. J., D. J. Lydiate, and D. A. Hopwood.** 1989. Structural and functional analysis of the mini-circle, a transposable element of *Streptomyces coelicolor* A3(2). Mol. Microbiol. **3:**1307–1318.
21. **Hodgman, T. C.** 1988. A new superfamily of replicative proteins. Nature (London) **333:**23. (Erratum, **333:**578.)
22. **Hopwood, D. A., M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. Schrempf.** 1985. Genetic manipulation of *Streptomyces*: a laboratory manual. John Innes Foundation, Norwich, England.
23. **Hopwood, D. A., T. Kieser, H. M. Wright, and M. J. Bibb.** 1983. Plasmids, recombination, and chromosomal mapping in *Streptomyces lividans* 66. J. Gen. Microbiol. **129:**2257–2269.
24. **Katz, E., C. J. Thompson, and D. A. Hopwood.** 1983. Cloning and expression of the tyrosinase gene from *Streptomyces antibioticus* in *Streptomyces lividans*. J. Gen. Microbiol. **129:**2703–2714.
25. **Kieser, H. M., D. J. Henderson, C. W. Chen, and D. A. Hopwood.** 1989. A mutation of *Streptomyces lividans* which prevents intraplasmid recombination has no effect on chromosomal recombination. Mol. Gen. Genet. **220:**60–64.
26. **Lei, S. P., H. C. Lin, L. Heffernan, and G. Wilcox.** 1985. *araB* gene and nucleotide sequence of the *araC* gene of *Erwinia carotovora*. J. Bacteriol. **164:**717–722.
27. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Morris, R. M., D. H. Williams, G. G. Midwinter, and B. S. Hartley.** 1974. A mass-spectrometric sequence study of the enzyme ribitol dehydrogenase from *Klebsiella aerogenes*. Biochem. J. **141:**701–713.
29. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
31. **Saraste, M., P. R. Sibbald, and A. Wittinghofer.** 1990. The P-loop—a common motif in ATP- and GTP-binding proteins. Trends Biochem. Sci. **15:**430–434.
32. **Schrempf, H., A. Kessler, V. Brönneke, W. Dittrich, and M. Betzler.** 1989. Genetic instability in streptomycetes, p. 133–140. *In* C. L. Hershberger, S. W. Queeners, and G. Hegeman (ed.), Genetics and molecular biology of industrial microorganisms. American Society for Microbiology, Washington, D.C.
33. **Schulz, R., K. Steinmueller, M. Klass, C. Forreiter, S. Rasmussen, C. Hiller, and K. Apel.** 1989. Nucleotide sequence of a cDNA coding for the NADPH-protochlorophyllide oxidoreductase (PCR) of barley (*Hordeum vulgare* L.) and its expression in *Escherichia coli*. Mol. Gen. Genet. **217:**355–361.
34. **Sheldon, P. S., R. G. O. Kekwick, C. Sidebottom, C. G. Smith, and A. R. Slabas.** 1990. 3-Oxyacyl-(acyl-carrier protein) reductase from avocado (*Persea americana*) fruit mesocarp. Biochem. J. **271:**713–720.
35. **Solenberg, P. J., and S. G. Burgett.** 1989. Method for selection of transposable DNA and characterization of a new insertion sequence, IS*493*, from *Streptomyces lividans*. J. Bacteriol. **171:**4807–4813.
36. **Tanaka, M. (Toho University, Funabashi, Chiba, Japan).** 1991. Personal communications.
37. **Tanaka, M., N. Takabe, F. Kato, and Y. Koyama.** 1991. Analysis of mutants and genetic rearrangements in the region containing IS*456*(Tn*4811*), abstr. P1-056. Abstr. 8th Int. Symp. Biol. Actinomycetes. University of Wisconsin, Madison.
38. **Tanaka, M., H. Tanaka, F. Kato, and Y. Koyama.** 1988. IS-like element from *Streptomyces lividans*, abstr. P5-20, p. 157. Abstr. 7th Int. Symp. Biol. Actinomycetes. Academic Societies Japan, Tokyo.
39. **Tobin, J. F., and R. F. Schleif.** 1987. The positive regulation of the *Escherichia coli* L-rhamnose operon is mediated by the products of tandemly repeated regulatory genes. J. Mol. Biol. **196:**789–799.
40. **Tsai, J. F.-Y., and C. W. Chen.** Unpublished data.
41. **Tsai, J. F.-Y., and C. W. Chen.** 1987. Isolation and characterization of *Streptomyces lividans* mutants deficient in intraplasmid recombination. Mol. Gen. Genet. **208:**211–218.
42. **Wada, K., S. Aota, R. Tsuchiya, F. Ishibashi, T. Gojobori, and T. Ikemura.** 1990. Codon usage tabulated from the GenBank genetic sequence data. Nucleic Acids Res. **18:**2367–2411.
43. **Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay.** 1982. Distantly related sequenced in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1:**945–951.
44. **Williams, S. T., M. Goodfellow, G. Alderson, E. M. H. Wellington, P. H. A. Sneath, and M. J. Sackin.** 1983. Numerical classification of *Streptomyces* and related genera. J. Gen. Microbiol. **129:**1743–1813.
45. **Yu, T.-W., and C. W. Chen.** Submitted for publication.