JOURNAL OF BACTERIOLOGY, Oct. 1987, p. 4589–4596
0021-9193/87/104589-08$02.00/0

Vol. 169, No. 10

Nucleotide Sequence Analysis of Tn*4551*: Use of *ermFS* Operon Fusions To Detect Promoter Activity in *Bacteroides fragilis*

C. JEFFREY SMITH†

Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases, Frederick Cancer Research Facility, Frederick, Maryland 21701

Received 4 February 1987/Accepted 13 July 1987

The *Bacteroides* pBI136 clindamycin resistance (Cc^r) determinant from the composite transposon Tn*4551* was cloned onto the shuttle plasmid pFD160, and the regions necessary for expression in *Bacteroides fragilis* were determined. These results suggested that transcriptional regulatory signals required for Cc^r were located in the Tn*4551* direct repeat sequence (DRS) adjacent to the resistance determinant. Analysis of the nucleotide sequence of this region revealed that the Cc^r structural gene, 798 base pairs (bp), was located 17 bp from the terminus of the DRS and that this gene (*ermFS*) differed from *ermF* (pBF4) by one amino acid. The DRS element was found to be 1,155 bp and appeared to contain the *ermFS* transcription start signals. The DRS structure was typical of insertion sequence elements isolated from other bacterial species, and its termini were characterized by 25-bp regions of imperfect dyad symmetry. The DRS was dominated by a 978-bp open reading frame, which terminated in the left inverted repeat 27 bp from the *ermFS* start codon, and weak amino acid sequence homology was observed with the putative transposase of IS*3*. Promoter activity of the DRS in *B. fragilis* was demonstrated by in vitro construction of operon fusions with a promoterless *ermFS* gene followed by transformation of the recombinant plasmids with selection for resistance to clindamycin. The location of one DRS promoter was identified by using the *ermFS* fusions and then verified by in vitro mutagenesis of the site with single-stranded linkers. Northern blot (RNA blot) analysis of total RNA from *B. fragilis* strains containing pBI136 or *ermFS* recombinant plasmids confirmed the location of this promoter and indicated that it was used in vivo by Tn*4551*. A second DRS promoter, which activated *ermFS* transcription by readthrough of the large DRS open reading frame, was also identified by the Northern blot analysis. This bicistronic *ermFS* message was not observed in strains containing a complete copy of Tn*4551*, and the possibility of transcriptional regulation is discussed.

Our understanding of basic genetic exchange mechanisms and the dissemination of genetic information among intestinal *Bacteroides* species has deepened during the last few years through the characterization of several *Bacteroides* R plasmids (19, 34, 40, 44). These plasmids are unrelated, except that they each encode resistance to clindamycin (Cc^r) and they share DNA sequence homology in the Cc^r region (12, 30, 38). Restriction analyses and DNA-DNA hybridization studies have demonstrated that this interplasmid homology is due entirely to the presence of three related transposons (30, 37). These Cc^r transposons are classic composite structures bounded by 1.2-kilobase-pair (kb) direct repeat sequences (DRSs), and the DRSs probably function as individual insertion sequence (IS) elements as shown for Tn*4400* in *Escherichia coli* (25). Two of the transposons, Tn*4351* (pBF4 [32, 33]) and Tn*4400* (pBFTM10 [25]), are very closely related, sharing more than 90% DNA sequence homology (37). On the other hand, Tn*4551* (pBI136 [38a]) is nearly 3 kb larger and is completely dissimilar, except for the Cc^r determinant and the DRSs. This dissimilar structural arrangement of Tn*4551* suggests that it has either diverged significantly from or arisen independently of the other two transposons.

IS elements and transposons can dramatically affect the expression of nearby genes, and this effect may take the form of transcriptional activation of a gene adjacent to the site of IS integration (16, 17, 46). The DRSs of the *Bacteroides* transposons may play a similar role, and results from two laboratories have implicated these repeats in the control of Cc^r expression (23, 35, 37). These observations and the importance of the *Bacteroides* DRSs in R-plasmid organization have prompted an analysis of their structure at the nucleotide level. In this paper the primary DNA sequence of the Tn*4551* right DRS (designated here as DRS4551-R) and the adjacent Cc^r determinant is reported. In addition, operon fusions with the Cc^r structural gene, *ermFS*, were used to demonstrate promoter activity of DRS4551-R. These results, together with Northern blot (RNA blot) analyses, indicated that there were at least two promoters associated with DRS4551-R which directed transcription toward *ermFS*.

MATERIALS AND METHODS

Bacterial strains and culture conditions. *Bacteroides fragilis* 638 (21) and its plasmid-bearing derivatives were used exclusively for analyses requiring *Bacteroides* species. *Bacteroides* cultures were grown in supplemented brain heart infusion broth as described previously (34). When required, clindamycin (The Upjohn Co., Kalamazoo, Mich.) was added to a final concentration of 5 μg/ml. Transformation of *B. fragilis* with plasmid DNA was performed by the polyethylene glycol-facilitated method described by Smith (36).

E. coli TB1, a restriction-negative (*hsdR17*) derivative of JM83 (45), was used for all plasmid cloning, and *E. coli* JM107 (45) was the host for M13 bacteriophage work. Cultures were grown in a tryptone-yeast extract broth (35), with ampicillin and 5-bromo-4-chloro-3-indolyl-β-D-galac-

† Present address: Department of Microbiology and Immunology, School of Medicine, East Carolina University, Greenville, NC 27858.

TABLE 1. Relevant characteristics of some important plasmids used in this study

Plasmid	Description	Source or reference
pBI136	80.6 kb, Cc^r, containing Tn*4551*	34
pBI136Δ1	73.4 kb, Cc^s, deletion containing a single DRS from Tn*4551*	34
pFD160	5.4 kb, Ap^r *lac*$^+$, pUC19::pBI143 chimeric shuttle vector	35
pFD167	6.3 kb, Ap^r *lac*$^+$ Cc^r, 0.85-kb *Eco*RI-*Hae*II pBI136 fragment ligated to the pFD160 *Cla*I site	35
pFD179	7.1 kb, Ap^r *lac*, Cc^s, 0.85-kb *Eco*RI-*Hae*II pBI136 fragment ligated to pFD160 *Sma*I site	This report
pFD214	6.3 kb, Ap^r *lac* Cc^s, 0.85-kb *Eco*RI pFD179 fragment ligated to *Eco*RI site of pFD160	This report
pFD224	8.6 kb, Ap^r *lac* Cc^r, 3.2-kb *Xba*I-*Hae*II pBI136 fragment ligated to *Sma*I site of pFD160	This report
pFD225	6.8 kb, Ap^r *lac* Cc^r, 1.4-kb *Ava*I-*Hae*II pBI136 fragment ligated to *Sma*I site of pFD160	This report
pFD237	6.55 kb, Ap^r *lac* Cc^r, 0.25-kb *Ban*I pBI136 fragment ligated to MCS of pFD214	This report
pFD240	6.65 kb, Ap^r *lac* Cc^r, 0.35-kb *Dde*I-*Eco*RI pBI136 fragment ligated to MCS of pFD214	This report
pFD241	7.9 kb, Ap^r *lac* Cc^r, 1.6-kb *Pst*I-*Eco*RI pBI136 fragment ligated to MCS of pFD214	This report
pUC18, pUC19	2.7 kb, Ap^r *lac*$^+$	20

topyranoside (X-Gal) added to a final concentration of 50 μg/ml when indicated. *E. coli* strains were transformed with plasmid DNA or M13 replicative-form DNA by the method of Hanahan for frozen competent cells (14).

Preparation and analysis of nucleic acids. Plasmid DNA and M13 replicative-form DNA were purified by density gradient centrifugation (22) of *E. coli* (10) and *B. fragilis* (8, 34) cell lysates. The rapid alkaline lysis method (5) was used for routine screening of bacterial transformants. Restriction endonuclease digestion and analysis of DNA were performed by standard methods (18, 35). Total cellular RNA for Northern blot analysis was extracted from 10-ml cultures (28). RNA samples were electrophoresed in 1% agarose gels containing 2.2 M formaldehyde (18) and transferred to nitrocellulose filters by the method of Southern (39). Filters were hybridized overnight with [^{32}P]DNA at 67°C under the conditions described by Thayer (41). Labeled probe DNA was prepared by the in vitro nick translation reaction (24).

Cloning strategies and plasmids. The relevant characteristics of the plasmids used in this study are presented in Table 1. One of these, pFD179, was used extensively throughout this work and was constructed as follows. A 0.85-kb *Eco*RI-*Hae*II fragment bearing the pBI136 Cc^r determinant (Tn*4551 ermFS*) was purified, blunt ended, and ligated to the *Sma*I site at the multiple cloning sequence (MCS) of pFD160. Plasmid pFD179 was among the transformants obtained from this ligation reaction; it contained two copies of the Cc^r fragment aligned in a head-to-tail fashion ligated in the *Sma*I site. The *Eco*RI-to-*Hae*II junction which resulted from this ligation caused the regeneration of the *Eco*RI recognition site, and this allowed for the convenient excision of the 0.85-kb Cc^r fragment with *Eco*RI. The promoter probe fusion vector, pFD214, was constructed by ligating the 0.85-kb *Eco*RI fragment from pFD179 to the *Eco*RI site of pFD160. The pFD238 fusion vector resulted from the ligation of a promoterless *ermFS* gene into the blunt-ended *Eco*RI site of pFD160. This vector was constructed by reconstituting the *ermFS* gene from two DNA restriction fragments, and the construction followed standard cloning strategies. A 0.45-kb pFD179 *Eco*RI-*Hin*dIII fragment bearing the C-terminal half of *ermFS* and a *Dra*I-*Hin*dIII fragment (0.5 kb from pFD224) bearing the N-terminal half of *ermFS* were the starting materials for this vector.

In vitro mutagenesis of the pFD237 *Ban*I fragment with single-stranded linkers was done basically as described by Barany (2, 3). The linker used for this study, AATTCG (Pharmacia, Inc., Piscataway, N.J.), converts a *Hha*I site to an *Eco*RI site, and it was inserted into the appropriate *Hha*I site of pFD237 by standard procedures. Transformants thought to contain the mutagenized plasmid were screened for the addition of a new *Eco*RI site, and isolates thought to contain the inserted linker were subjected to detailed restriction analyses to ensure the identity of the altered *Ban*I fragment.

DNA sequence analysis. The method of Sanger et al. (27) was used for sequencing of clones prepared in the M13mp18 and M13mp19 vectors (20). Reagents and template primer were obtained from Bethesda Research Laboratories and used with the modifications for [^{35}S]dATP suggested by the supplier. Clones to be used for sequencing were obtained either by subcloning small fragments from pFD214 or by exonuclease III treatment (15) of a 1.6-kb *Pst*I-*Eco*RI fragment from pFD224 (see Fig. 1A). In this way the *ermFS* structural gene and the DRS4551-R region were sequenced in their entirety. Sequence data were analyzed by using computer software described by Conrad and Mount (7) and Beckman Instruments, Inc., Palo Alto, Calif.

RESULTS

Cloning and expression of the clindamycin resistance gene. The *Bacteroides* Cc^r determinants do not express in *E. coli* (13, 35). Therefore, the basic experimental approach used in this study relied on the use of *E. coli* for plasmid construction followed by transformation of *B. fragilis* to determine expression of the Cc^r determinant. Results presented here, together with those from a previous report (35), showed that the Tn*4551* Cc^r determinant, *ermFS*, was located on a 0.85-kb *Eco*RI-*Hae*II fragment (coordinates 6.7 to 7.6; Fig. 1A) but that its expression depended on position or orientation within the vector. For example, plasmid pFD167 contained the *Eco*RI-*Hae*II fragment cloned into the *Bacteroides* component of shuttle vector pFD160 (a pUC19::pBI143 chimera; Table 1), and this expressed resistance to clindamycin. However, the same fragment cloned into the MCS (i.e., the pUC19 component) did not transform *B. fragilis* to Cc^r (pFD178; Fig. 1A). These results suggested that specific control regions were missing from the cloned fragment. This possibility was tested by subcloning larger regions of Tn*4551* into the MCS of pFD160 to identify regions capable of orientation-independent expression of Cc^r. Tn*4551* sequences to the left of the *Hae*II site did not restore expression of the resistance determinant. This was made apparent by the lack of Cc^r transformants obtained with pFD171 which encompassed coordinates 0.4 to 7.6 (Fig. 1A). Thus attention focused on the region of the DRS located adjacent to *ermFS*. Two fragments originating at the *Xba*I site (coordinate 10.0; Fig. 1A) outside of Tn*4551* were

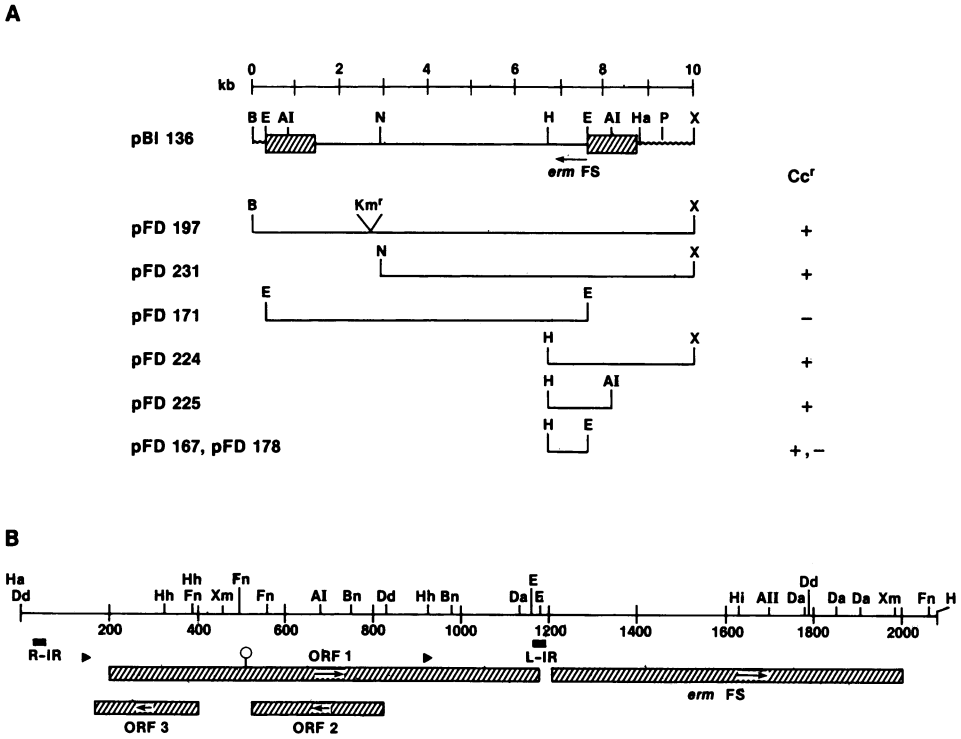

FIG. 1. Cloning and physical map of the pBI136 (Tn*4551*) *ermFS* region. (A) Various regions of pBI136 were cloned by using *E. coli* and then tested for Cc^r by transformation into *B. fragilis*. pFD197 was constructed by cloning the designated region (plus a kanamycin resistance determinant, Km^r) into the shuttle vector pFD165 (38a). pFD167 was formed by ligation to the *Cla*I site of pFD160, and all other recombinant plasmids were formed by ligations to the MCS of pFD160 as described in the text. The Cc^r region from pBI136 used for the cloning experiments is presented on the top line. ▨, DRSs. (B) The pBI136 region extending from coordinates 6.7 to 8.8 was sequenced, and the map corresponding to the nucleotide sequence of this region is presented. The map orientation is opposite that in Fig. 1A. Below the map, the four open reading frames and the direction of translation are shown (▨). The DRS right inverted repeat (R-IR) and left inverted repeat (L-IR) (■), two putative promoter sequences (▶), and a second Met codon with a ribosome-binding site within ORF-1 (○) are also shown. Restriction sites for panels A and B are as follows: AI, *Ava*I; AII, *Ava*II; B, *Bam*HI; Bn, *Ban*I; Da, *Dra*I; Dd, *Dde*I; E, *Eco*RI; Fn, *Fnu*4HI; H, *Hae*II; Ha, *Hae*III; Hh, *Hha*I; Hi, *Hin*dIII; N, *Nde*I; P, *Pst*I; X, *Xba*I; Xm, *Xmn*I.

cloned, and both of these, pFD231 and pFD224, showed resistance to clindamycin. An *Ava*I-*Hae*II fragment extending from coordinates 6.7 to 8.2 (pFD225 in Fig. 1A) and wholly contained within Tn*4551* was also found to confer Cc^r. Together, these results indicated that orientation-independent expression of Cc^r required the presence of the DRS sequences; therefore this region was subjected to nucleotide sequence analysis to search for transcription regulatory signals.

Nucleotide sequence of the Cc^r region. The primary DNA sequence was determined for the Cc^r region extending from coordinates 6.7 to 8.8 (Fig. 1A), and the results are summarized in Fig. 1B. Owing to the orientation of the major open reading frames in this region, these results and the sequence (Fig. 2) are presented in the opposite orientation to the results of cloning experiments shown in Fig. 1A. As expected from the cloning work, the entire *ermFS* structural gene was located within the *Eco*RI-*Hae*II fragment. This gene occupied an open reading frame (ORF) of 798 base pair (bp) and could encode a 266-amino-acid polypeptide (Fig. 1B and 2). This sequence differed from the *ermF* gene of plasmid pBF4 (23) by a single base and thus was designated *ermFS*. The difference between the two genes was an A-to-G transition at bp 1811, which resulted in a tyrosine-to-cysteine change in the predicted polypeptide of *ermFS*. The *ermFS* initiation codon was just 17 bp from the 3′ end of the DRS and 26 bp from the *Eco*RI recognition site. Consistent with the cloning experiments, there were no obvious transcription start signals downstream of the *Eco*RI site. There was, however, a ribosome-binding site-like sequence (31) that was complementary to the terminus of the 16S rRNA from *B. fragilis* (Fig. 2) (43). The *ermFS* stop codon, TAG, was located at bp 2001, and examination of the sequences downstream from this site did not reveal any stem-loop structures typical of factor-independent transcription termination (1, 26).

The DRS (designated DRS4551-R) was located upstream from and directly adjacent to *ermFS*. The DNA sequence of this region revealed that the entire DRS structure was 1,155 bp in length and was characterized by 25-bp regions of imperfect (20 of 25 bp) dyad symmetry at its termini (Fig. 1B and 2). Identification of these termini was possible by comparison of the DRS4551-R nucleotide sequence with that for the ends of the left Tn*4551* DRS (data not shown) and with the N-terminal region of the pBF4 *ermF* gene (23). The structural features of DRS4551-R (summarized in Fig. 1B) showed that the element was dominated by a large ORF (ORF-1, 978 bp) located on the same DNA strand and in the same reading frame as *ermFS*. ORF-1 started 68 bp inside the DRS and terminated in the left inverted repeat. This sequence could encode a basic polypeptide of 326 amino acids with a predicted molecular weight of 37,909. A ribosome-binding site with a 5-of-6-bp match to the putative *ermFS* ribosome-binding site preceded ORF-1 by the proper

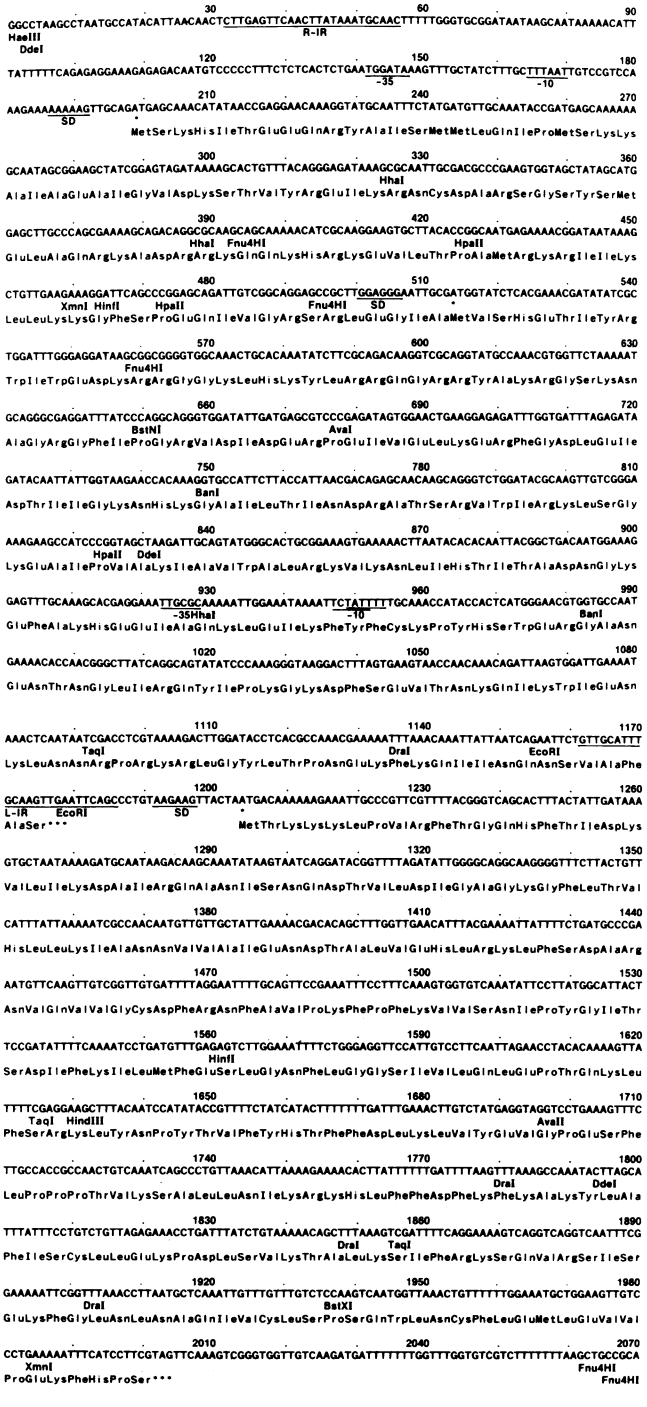

FIG. 2. Nucleotide sequence of DRS4551-R and the Cc^r determinant *ermFS*. The sequence is shown 5′ to 3′ with respect to the *ermFS* reading frame and corresponds to the map shown in Fig. 1B. Base no. 1 is the first base of the *Hae*III recognition site (Fig. 1A, coordinate 8.8), and the last base shown, no. 2080, is the first G of the *Hae*II recognition site (Fig. 1A, coordinate 6.7). Tn*4551* sequences start at base 32, and the right (R-IR) and left (L-IR) inverted repeats are underlined. The predicted amino acid sequences for ORF-1 and *ermFS* are shown below the corresponding DNA sequence. Putative ribosome-binding sites are underscored and labeled SD; potential promoter sites for *ermFS* and ORF-1 are indicated by underscoring of the sequences which best approximate the *E. coli* consensus −10 and −35 regions (26).

number of bases (31). In addition, a promoterlike sequence which closely approximated the *E. coli* consensus (26) was located 30 bp upstream from the ORF-1 initiation codon (Fig. 2). Within ORF-1 at bp 514 was a second Met codon preceded by a ribosome-binding site-like sequence. This would encode a polypeptide of only 221 amino acids, and its significance, if any, is not known. Further computer analysis of the DRS4551-R sequences revealed two additional ORFs on the opposite strand from that of the larger ORF-1 (Fig. 1B). ORF-3 had a coding potential for a 79-amino-acid polypeptide and was in codon register with ORF-1, but extended 30 bp beyond the ORF-1 N terminal. ORF-2 could encode a 99-amino-acid polypeptide and was wholly contained within ORF-1. Neither of the smaller ORFs were preceded by ribosome-binding sites with strong homology to the *Bacteroides* 16S rRNA, and their importance is not known.

The nucleotide sequence of DRS4551-R was compared with the sequences of 10 IS elements available from the GenBank database, but no significant homologies were found. In addition, the predicted polypeptide of ORF-1 was compared with the largest open reading frames from IS*1*, IS*2*, IS*3*, IS*4*, IS*5*, IS*26*, IS*50*, IS*102*, IS*903*, and Tn*3* at the amino acid level. Of these, only a weak relationship with the putative transposase of IS*3* could be found.

***ermFS* operon fusions.** Promoter activity of DRS4551-R in *B. fragilis* was demonstrated by in vitro construction of fusions between a promoterless *ermFS* structural gene and various DRS4551-R fragments. The work dealt primarily with the vector pFD214, which contained the *ermFS* gene plus 25 bp upstream from the N-terminal amino acid and included the ribosome-binding site (Fig. 2 and 3). This particular plasmid vector had translation stop signals in two reading frames prior to *ermFS* but not in the *ermFS* reading frame itself. The source of DRS4551-R sequences to be tested was a 1.6-kb *Pst*I-*Eco*RI fragment from pFD224, which was blunt ended and ligated to the *Hin*cII site of pUC19. Fragments from this plasmid were isolated by digestion with *Xba*I plus a second enzyme (as indicated in Fig. 3) and then subcloned into the *Sma*I site of pUC19 by blunt-end ligation. Once the appropriate fragments were identified by restriction mapping, they were excised with *Pst*I-*Sst*I and ligated to pFD214 digested with the same enzymes. This cloning strategy was particularly useful because the *Hin*cII-*Eco*RI junction formed in the original pUC19::pFD224 clone introduced a translation stop signal, TGA, in the *ermFS* reading frame.

Transformations with the vector pFD214 did not result in Cc^r *B. fragilis* transformants. The system was tested further by transformations with pFD241 which contained nearly the entire DRS element cloned in the same orientation, relative to *ermFS*, as found in Tn*4551* (Fig. 3). As expected, this clone conferred resistance to clindamycin. Since the cloning results with pFD225 (Fig. 1A) indicated that sequences downstream from the DRS4551-R *Ava*I site were important for transcription of *ermFS*, further analysis centered on this region. Promoter activity was found on a 234-bp *Ban*I fragment (pFD237) and was probably located on sequences bounded by the *Dde*I site on the 5′ side and the *Ban*I site on the 3′ side (Fig. 3). Careful inspection of the nucleotide sequence in this region revealed a promoterlike sequence between bp 923 and 954 (Fig. 2). This putative promoter spanned a *Hha*I recognition site (bp 925), consistent with the fact that pFD239 did not transform *B. fragilis* to clindamycin resistance. Subsequently, the pFD237 *Ban*I fragment was subcloned further by digestion with *Hha*I and ligation of

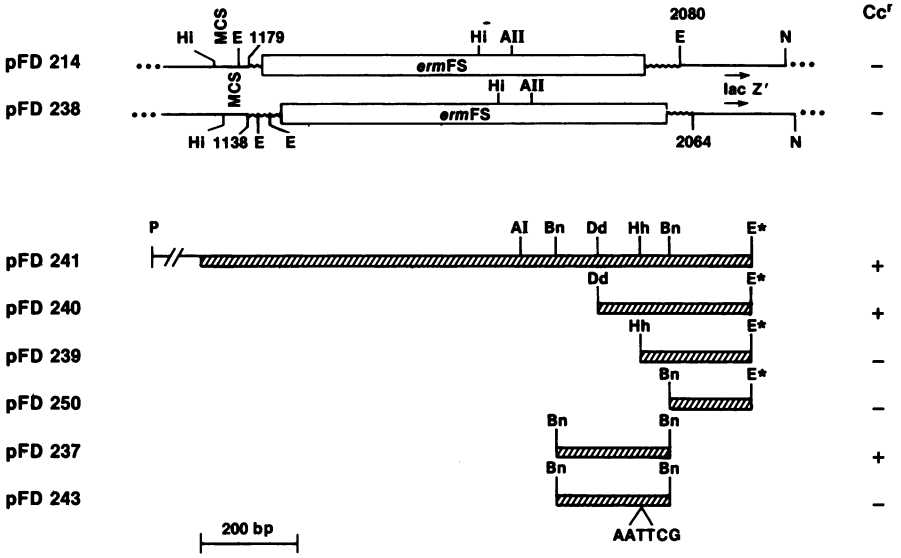

FIG. 3. Structure of the *ermFS* operon fusion vectors and their use for detecting promoter activity in DRS4551-R. The vectors pFD214 and pFD238 were constructed as described in Materials and Methods. The *ermFS* structural gene (▭) and other Tn*4551* sequences (〰) are indicated. For each of the vectors, numbers corresponding to the nucleotide sequence show the exact limits of the Tn*4551* sequences present. The MCS and the remainder of *lacZ'* are shown, and N indicates the location of the *Nde*I-*Hae*II junction between the pUC19 and pBI143 components of the shuttle vectors (35). Fragments from DRS4551-R (▨) were cloned into the MCS of pFD214 and transformed into *B. fragilis* to determine the expression of clindamycin resistance. The symbol E* indicates the location of the *Eco*RI-*Hin*cII junction formed during the construction of the DRS4551-R recombinants and is shown as a point of reference (see text). Insertion of the 6-bp sequence AATTCG into the *Hha*I site of pFD243 is indicated by the arrowhead. Restriction site designations are the same as in Fig. 1.

each half into pFD214. Neither of these two plasmids yielded Cc^r *B. fragilis* transformants (data not shown). One additional test of this promoter site centered on in vitro mutagenesis of the *Hha*I site with the single-stranded linker AATTCG (2, 3). The 6-bp linker was inserted into the *Hha*I site of pFD237 and changed the sequence from TTGCGCAA to TTGCGAATTCGCAA. The resulting plasmid, pFD243, did not transform *B. fragilis*.

A second fusion vector, pFD238, was used as a control in several experiments. This plasmid was similar to pFD214, except that it included additional sequences upstream from the *ermFS* initiation codon. These sequences started at the Tn*4551* *Dra*I site, bp 1138 (Fig. 2), and contained translation stop codons in all three reading frames prior to *ermFS*. pFD238 did not transform *B. fragilis* to Cc^r, indicating that there was no promoter activity within the left inverted repeat of DRS4551-R (Fig. 3). This was an important observation, because cloning strategies used with the pFD214 fusions excluded this region from the analysis. In addition, the promoter activity of the pFD237 *Ban*I fragment was verified by cloning this fragment and its mutated derivative (from pFD243) into pFD238. In both cases the original observations were confirmed (data not shown).

Transcriptional analysis of *ermFS*. The role of DRS4551-R in *ermFS* transcription was examined further by Northern blot analyses of total RNA from several of the Cc^r clones and the *ermFS* fusions. It is important to note that the plasmids used for these analyses (except pBI136) were all comparable because they contained *ermFS* sequences with the same 3' terminus, cloned into the MCS of pFD160, and in the same orientation as shown for pFD214 in Fig. 3. The ^{32}P-labeled probe used for these studies was a 0.45-kb *Eco*RI-*Hin*dIII fragment bearing the N-terminal portion of *ermFS* (bp 1179 to 1634; Fig. 2). The specificity of this probe for *ermFS* transcripts was demonstrated by its lack of hybridization to RNA isolated from pBI136Δ1, a pBI136 deletion derivative bearing a single copy of the DRS but no *ermFS* gene (Fig. 4A, lane 1). On the other hand, results in Fig. 4A, lane 2, showed that a single fragment from a pBI136-containing strain did hybridize to the probe.

In contrast to the pBI136 results, two major *ermFS* transcripts were observed in strains containing pFD224 and pFD241 (Fig. 4A, lanes 3 and 5). These transcripts were approximately 1.6 and 2.3 kb in size, and the 2.3-kb species was seen only with plasmids having one copy of the entire DRS. Strains with plasmids composed of less than half of the DRS, such as pFD225 and pFD240, possessed only the smaller, 1.6-kb, transcript (Fig. 4B). pFD237 was the only plasmid tested which did not give rise to the 1.6-kb *ermFS* mRNA, and in this case the transcript appeared to be 100 to 150 bp smaller (Fig. 4B, lane 8). This result was consistent with the proposed location for a promoter at the *Hha*I site (bp 925), because pFD237 was the only plasmid in which a significant amount of DNA was deleted between the putative promoter and the start of *ermFS*.

DISCUSSION

The genetic and nucleotide sequence analyses of Tn*4551* presented in this communication have provided an opportunity to explore the relationship between the *Bacteroides* Cc^r transposons and the biological mechanisms involved in their evolution. The genetic basis of the Cc^r resistance phenotype, which also confers resistance to other lincosamide-, macrolide-, and streptogramin B-type antibiotics, was shown to be the product of a single gene, *ermFS*. This determinant differed from the Tn*4351* *ermF* by a single amino acid (Fig. 2) (23). These data have confirmed and extended previous studies showing the close relationship between Cc^r determinants on pBF4 and pBI136 (37). How-

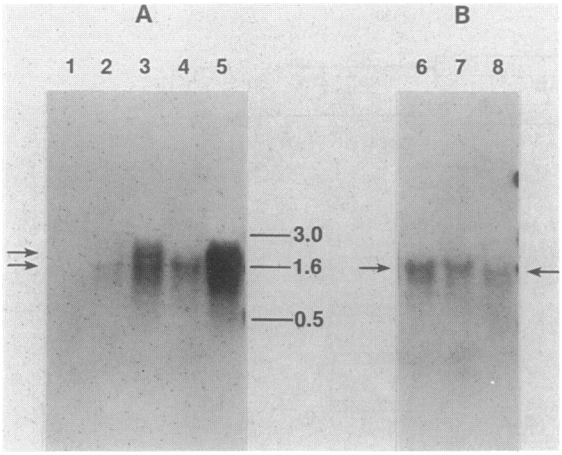

FIG. 4. Northern blot analysis of total RNA from plasmid-containing strains of *B. fragilis* 638. Total cellular RNA (5 to 10 μg) was electrophoresed, transferred to nitrocellulose filters, and hybridized to a ^{32}P-labeled *ermFS* probe as described in the text. The resulting autoradiographs are shown in panels A and B. Lane designations are as follows. (A) 1, pBI136Δ1; 2, pBI136; 3, pFD241; 4, pFD225; 5, pFD224; (B) 6, pFD240; 7, pFD225; 8, pFD237. The numbers to the right of panel A indicate the location of molecular size markers, in kilobases, from the 1-kb ladder (Bethesda Research Laboratories, Inc., Gaithersburg, Md.). Arrows to the left of the autoradiographs indicate the locations of the 2.3- and/or 1.6-kb *ermFS* transcript. The arrow to the left of panel B shows the location of the 1.5-kb transcript from pFD237.

ever, a comparison of sequences upstream from the *erm* start codons revealed some important differences between the two (Fig. 5). The sequences were identical for 17 bp, including the ribosome-binding site, but the pBF4 sequence contained an additional 9 bp between the start codon and the terminus of the Tn*4351* DRS. Furthermore, the DRS terminus which abutted the pBI136 *ermFS* gene was the opposite one from that found adjacent to *ermF* in pBF4. These differences strongly suggest that Tn*4351* and Tn*4551* arose independently through the action of the DRSs on the same progenitor *erm* sequences.

The nucleotide sequence of DRS4551-R revealed structural features typical of known IS elements (Fig. 1B and 2). This is consistent with the role proposed for DRS4551 and with the observations of Robillard et al. (25), who showed that the DRSs of Tn*4400* are IS elements capable of independent transposition in *E. coli*. The DRSs of Tn*4551* also may be capable of independent integration into the *Bacteroides* chromosome (38a). The 25-bp imperfect inverted repeats at the ends of DRS4551-R defined its limits and set its size at 1,155 bp. This size is close to the 1.2-kb estimate obtained previously by heteroduplex analysis of pBI136 (34). The DRS termini were identified by comparison of DNA sequences from the ends of both Tn*4551* DRS elements. These data also revealed a possible 3-bp duplication of the pBI136 sequence ACT in direct orientation at the outside ends of Tn*4551* (Fig. 2, bp 29 to 31, and data not shown). Although no other Tn*4551* insertion sites have been sequenced for comparison, Macrina et al. have found that the related Tn*4351* DRS generates a 3-bp duplication of the target site when it integrates in *E. coli* (F. L. Macrina, J. L. Rasmussen, and D. A. Odelson, personal communication).

DRS4551-R contained a major open reading frame, ORF-1, which traversed nearly the whole length of the element, terminating in the left inverted repeat. It is likely that ORF-1 encodes a transposase, and, typical of many proteins that interact with DNA, the product of ORF-1 would be a basic protein. The overall size of ORF-1 is similar to the size of major ORFs from other IS elements (4, 9, 17, 42). A computer-assisted comparison of the deduced amino acid sequences from many of these ORFs suggested a distant relationship between IS*3* and DRS4551-R. This observation is noteworthy primarily because IS*3* is the only other example of an element that generates a 3-bp duplication of its target site upon insertion, and it has been the premier example of IS elements that are mobile promoters (6, 46).

Genetic results from the initial cloning experiments showed that in Tn*4551*, expression of Cc^r was dependent on the adjacent DRS (Fig. 1A). The observation that there were no obvious transcriptional start signals between the DRS terminus and the start of *ermFS* was consistent with this (Fig. 5). To demonstrate this promoter activity experimentally, plasmid vectors with the *ermFS* structural gene and ribosome-binding site were constructed so that fusions with DRS sequences could be generated in vitro and then tested in *B. fragilis* for activity. This particular approach was chosen over the many promoter detection systems available for use in *E. coli*, because it is not known whether *E. coli* RNA polymerase recognizes the same sequences that are used by the *Bacteroides* enzyme. The *Bacteroides* spp. are phylogenetically distant from *E. coli* (43), and there are presently no examples of an *E. coli* gene which functions in these organisms (11, 13, 35; unpublished observations). The results summarized in Fig. 3 were useful for the identification of one DRS4551-R promoter which directed transcription inward toward *ermFS*. This promoter activity was localized to a region between bp 830 (*Dde*I) and bp 982 (*Ban*I), but the only promoterlike sequence observed in this area was sufficiently different from established consensus sequences that it could not be confidently identified as the promoter (26). Fortunately, the putative −35 region spanned a *Hha*I site, and by disrupting the sequence as in the case of pFD239 or pFD243, expression of *ermFS* was abolished.

IS elements can activate transcription of adjacent genes by several mechanisms. With IS*2* for example, this activation depends upon the actual insertion event, whereby an efficientpromoter is generated at the site of insertion (16). This does not seem to be the case with Tn*4551*, because pFD238, which contained the DRS-*ermFS* junction plus 49 bp upstream of this site, did not express Cc^r. Elements such as IS*3* and IS*21*, on the other hand, activate transcription from promoters within the elements themselves (6, 29, 46). The *ermFS* fusion work described here suggests that DRS4551-R resembles these two elements, since it contains the *ermFS* promoter within its own sequences.

A second method used to detect promoter activity was Northern blot analysis. These results were difficult to interpret because the site of transcription termination was not

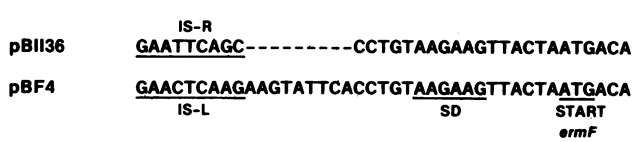

FIG. 5. Alignment of sequences upstream from the *erm* genes of pBF4 (23) and pBI136. The ribosome-binding site (SD) and Met initiation codon (START) for *ermF* are underscored and labeled. Also shown and underscored are sequences at the termini of the right DRS (IS-R) and the left DRS (IS-L) which abut to the *erm* genes in pBI136 and pBF4, respectively.

known. However, if the results with the recombinant plasmids are considered first and it is assumed that transcription terminated at the same point in these related plasmids, then certain predictions can be made; that is, if the promoter site was located where proposed (bp 923 to 954), then *ermFS* transcripts should be of equal size in all of the recombinant plasmids tested except pFD237. The pFD237 mRNA would be smaller by some 173 bp owing to the deletion of DNA sequences between the promoter and the start of *ermFS* (see Fig. 3). This prediction was fulfilled in part, since a common 1.6-kb *ermFS* mRNA was observed for pFD224, pFD225, pFD240, and pFD241, but in pFD237 a single transcript of about 1.5 kb was seen (Fig. 4). In addition, these studies revealed the presence of a second promoter in DRS4551-R which directed transcription toward *ermFS*. This was identified by the generation of a second *ermFS* transcript in plasmids harboring the entire DRS element (e.g., pFD224 and pFD241; Fig. 4A). It is possible that this larger, 2.3-kb, transcript represents a bicistronic mRNA generated by readthrough of ORF-1 into *ermFS*. The DNA sequence is consistent with this interpretation, since there are no obvious transcription termination signals in the 27 bp between the end of ORF-1 and the start of *ermFS* (Fig. 2). Furthermore, the size of this larger transcript is close to what would be required for such a bicistronic mRNA if the start of ORF-1 transcription was located at the proposed site (Fig. 2).

Also noteworthy was the presence of only one major *ermFS* mRNA from the parental plasmid, pBI136, which contains the Tn*4551* element complete with both flanking DRSs. As mentioned above, recombinant plasmids with just one DRS copy generated two transcripts, the larger of which might include the putative transposase encoded by ORF-1. The single pBI136 mRNA was too small (1.6 kb) to be the bicistronic mRNA proposed above, and thus it is likely that *ermFS* was transcribed from the *Hha*I (bp 923) promoter. The lack of this second mRNA in pBI136-containing strains suggests the possibility of ORF-1 regulation at the level of transcription. If ORF-1 encodes the transposase, it would be an attractive and testable hypothesis that its transcription is regulated by a second copy of the DRS as is present in the complete Tn*4551* element.

ACKNOWLEDGMENTS

The helpful advice of and critical discussions with D. J. LeBlanc and L. N. Lee are gratefully acknowledged. I also thank F. L. Macrina, D. A. Odelson, and J. L. Rasmussen for sharing their data prior to publication.

LITERATURE CITED

1. **Adhya, S., and M. Gottesman.** 1978. Control of transcription termination. Annu. Rev. Biochem. **47**:967–996.
2. **Barany, F.** 1985. Single-stranded hexameric linkers: a system for in-phase insertion mutagenesis and protein engineering. Gene **37**:111–123.
3. **Barany, F.** 1985. Two-codon insertion mutagenesis of plasmid genes by using single-stranded hexameric oligonucleotides. Proc. Natl. Acad. Sci. USA **82**:4202–4206.
4. **Bernardi, A., and F. Bernardi.** 1981. Complete sequence of an IS element present in pSC101. Nucleic Acids Res. **9**:2905–2911.
5. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
6. **Charlier, D., J. Piette, and N. Glansdorff.** 1982. IS3 can function as a mobile promoter in *E. coli*. Nucleic Acids Res. **10**:5935–5948.
7. **Conrad, B., and D. W. Mount.** 1982. Microcomputer programs for DNA sequence analysis. Nucleic Acids Res. **10**:31–38.
8. **Currier, T. C., and E. W. Nester.** 1976. Isolation of covalently closed circular DNA of high molecular weight from bacteria. Anal. Biochem. **76**:431–444.
9. **Ghosal, D., H. Sommer, and H. Saedler.** 1979. Nucleotide sequence of the transposable DNA-element IS2. Nucleic Acids Res. **6**:1111–1122.
10. **Guerry, P., D. J. LeBlanc, and S. Falkow.** 1973. General method for the isolation of plasmid deoxyribonucleic acid. J. Bacteriol. **116**:1064–1066.
11. **Guiney, D. G., P. Hasegawa, and C. E. Davis.** 1984. Expression in *Escherichia coli* of cryptic tetracycline resistance genes from Bacteroides R-plasmids. Plasmid **11**:248–252.
12. **Guiney, D. G., P. Hasegawa, and C. E. Davis.** 1984. Homology between clindamycin resistance plasmids in *Bacteroides*. Plasmid **11**:268–271.
13. **Guiney, D. G., P. Hasegawa, and C. E. Davis.** 1984. Plasmid transfer from *Escherichia coli* to *Bacteroides fragilis*: differential expression of antibiotic resistance phenotypes. Proc. Natl. Acad. Sci. USA **81**:7203–7206.
14. **Hanahan, D.** 1983. Studies on the transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557–580.
15. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**: 351–359.
16. **Jaurin, B., and S. Normark.** 1983. Insertion of IS2 creates a novel *ampC* promoter in *Escherichia coli*. Cell **32**:809–816.
17. **Kleckner, N.** 1981. Transposable elements in prokaryotes. Annu. Rev. Genet. **15**:341–404.
18. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Matthews, B. G., and D. G. Guiney.** 1986. Characterization and mapping of regions encoding clindamycin resistance, tetracycline resistance, and a replication function on the *Bacteroides* R plasmid pCP1. J. Bacteriol. **167**:517–521.
20. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26**:101–106.
21. **Privitera, G., A. Dublanchet, and M. Sebald.** 1979. Transfer of multiple antibiotic resistance between subspecies of *Bacteroides fragilis*. J. Infect. Dis. **139**:97–101.
22. **Radloff, R., W. Bauer, and J. Vinograd.** 1967. A dye-buoyant-density method for the detection and isolation of closed circular duplex DNA: the closed circular DNA of HeLa cells. Proc. Natl. Acad. Sci. USA **57**:1514–1521.
23. **Rasmussen, J. L., D. A. Odelson, and F. L. Macrina.** 1986. Complete nucleotide sequence and transcription of *ermF*, a macrolide-lincosamide-streptogramin B resistance determinant from *Bacteroides fragilis*. J. Bacteriol. **168**:523–533.
24. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**: 237–251.
25. **Robillard, N. J., F. P. Tally, and M. H. Malamy.** 1985. Tn*4400*, a compound transposon isolated from *Bacteroides fragilis*, functions in *Escherichia coli*. J. Bacteriol. **164**:1248–1255.
26. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13**:319–353.
27. **Sanger, F., S. Nicklen, and A. R. Coulsen.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
28. **Sarimientos, P., J. E. Sylvester, S. Contente, and M. Cashel.** 1983. Differential stringent control of the tandem *E. coli* ribosomal RNA promoters from the *rrn*A operon expressed *in vivo* in multicopy plasmids. Cell **32**:1337–1346.
29. **Schurter, W., and B. W. Holloway.** 1986. Genetic analysis of promoters on the insertion sequence IS21 of plasmid R68.45. Plasmid **15**:8–18.
30. **Shimell, M. J., C. J. Smith, F. P. Tally, F. L. Macrina, and M. H. Malamy.** 1982. Hybridization studies reveal homologies between pBF4 and pBFTM10, two clindamycin-erythromycin

resistance transfer plasmids of *Bacteroides fragilis*. J. Bacteriol. **152**:950–953.
31. **Shine, J., and L. Dalgarno.** 1974. The 3′ terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementary to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
32. **Shoemaker, N. B., C. Getty, J. F. Gardner, and A. A. Salyers.** 1986. Tn*4351* transposes in *Bacteroides* spp. and mediates the integration of plasmid R751 into the *Bacteroides* chromosome. J. Bacteriol. **165**:929–936.
33. **Shoemaker, N. B., E. P. Guthrie, A. A. Salyers, and J. F. Gardner.** 1985. Evidence that the clindamycin-erythromycin resistance gene of *Bacteroides* plasmid pBF4 is on a transposable element. J. Bacteriol. **162**:626–632.
34. **Smith, C. J.** 1985. Characterization of *Bacteroides ovatus* plasmid pBI136 and structure of its clindamycin resistance region. J. Bacteriol. **161**:1069–1073.
35. **Smith, C. J.** 1985. Development and use of cloning systems for *Bacteroides fragilis*: cloning of a plasmid-encoded clindamycin resistance determinant. J. Bacteriol. **164**:294–301.
36. **Smith, C. J.** 1985. Polyethylene glycol-facilitated transformation of *Bacteroides fragilis* with plasmid DNA. J. Bacteriol. **164**: 466–469.
37. **Smith, C. J., and M. A. Gonda.** 1985. Comparison of the transposon-like structures encoding clindamycin resistance in *Bacteroides* R-plasmids. Plasmid **13**:182–192.
38. **Smith, C. J., and F. L. Macrina.** 1984. Large transmissible clindamycin resistance plasmid in *Bacteroides fragilis*. J. Bacteriol. **158**:739–741.
38a. **Smith, C. J., and H. Spiegel.** 1987. Transposition of Tn*4551* in *Bacteroides fragilis*: identification and properties of a new transposon from *Bacteroides* spp. J. Bacteriol. **169**:3450–3457.
39. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
40. **Tally, F. P., D. R. Snydman, M. J. Shimell, and M. H. Malamy.** 1982. Characterization of pBFTM10, a clindamycin-erythromycin resistance transfer factor from *Bacteroides fragilis*. J.Bacteriol. **151**:686–691.
41. **Thayer, R. E.** 1979. An improved method for detecting foreign DNA in plasmids of *Escherichia coli*. Anal. Biochem. **98**:60–63.
42. **Timmerman, K. P., and C. D. Tu.** 1985. Complete sequence of IS3. Nucleic Acids Res. **13**:2127–2139.
43. **Weisburg, W. G., Y. Oyaizu, H. Oyaizu, and C. R. Woese.** 1985. Natural relationship between bacteroides and flavobacteria. J. Bacteriol. **164**:230–236.
44. **Welch, R. A., and F. L. Macrina.** 1981. Physical characterization of the *Bacteroides* R plasmid pBF4. J. Bacteriol. **145**:867–872.
45. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
46. **Zafarullah, M., D. Charlier, and N. Glansdorff.** 1981. Insertion of IS*3* can turn on a silent gene in *Escherichia coli*. J. Bacteriol. **146**:415–417.