JOURNAL OF BACTERIOLOGY, Jan. 1992, p. 638–642
0021-9193/92/020638-05$02.00/0

Vol. 174, No. 2

Structural Organization of pBC1, a Cryptic Plasmid from *Bacillus coagulans*†

EDDA DE ROSSI,[1] ANNA MILANO,[1] PATRIZIA BRIGIDI,[2] FRANCESCA BINI,[1] AND GIOVANNA RICCARDI[1]*

Department of Genetics and Microbiology "A. Buzzati Traverso," University of Pavia, Via Abbiategrasso 207, 27100 Pavia,[1] and Division of Fermentation Chemistry and Industrial Microbiology, Department of Pharmaceutical Sciences, University of Bologna, Via Belmeloro 6, 40126 Bologna,[2] Italy

Received 6 June 1991/Accepted 31 October 1991

The complete nucleotide sequence of the *Bacillus coagulans* plasmid pBC1 was determined. The sequence revealed an open reading frame encoding a polypeptide of 259 amino acids. This open reading frame shows sequence similarity to genes coding for replication-associated proteins in a group of gram-positive bacterial plasmids known to replicate via single-stranded intermediates. A region required for replication in *cis*, when the intact replicon is supplied in *trans*, was identified as well.

The 1.6-kbp cryptic plasmid pBC1 was originally isolated from a thermophilic strain of *Bacillus coagulans* (5). A genetically marked derivative, pLM6 (6), was shown to replicate in a wide variety of gram-positive bacteria (such as bacilli, lactococci, and staphylococci). The elucidation of the plasmid replication mechanism is fundamental to further vector development and to use of pLM6 as a cloning vector.

In this report, we describe the nucleotide sequence and genetic organization of plasmid pBC1. A 777-bp open reading frame (ORF), a promoter for the ORF, and a possible replication origin upstream of the promoter were identified.

Cultures of *Escherichia coli* HB101 (20) and *Bacillus subtilis* PB1424 (*hisH2 metD4 trpC2*; a gift of A. Galizzi) were grown in Luria-Bertani broth (20) at 37°C, while *B. coagulans* Zu196I (from our laboratory collection) culture was grown in Trypticase soy broth (Difco Laboratories) at 60°C. Luria-Bertani agar (*E. coli*) or NB agar (*B. subtilis*) (Difco Laboratories) was supplemented with antibiotics as indicated. Erythromycin (5 μg/ml for *B. subtilis*) and chloramphenicol (12.5 and 5 μg/ml for *E. coli* and *B. subtilis*, respectively) were used. Transformation of *B. subtilis* or *E. coli* competent cells with plasmid DNA and all manipulations of DNA were performed as previously described (3, 20). pBC1 plasmid was extracted from *B. coagulans* Zu196I as previously described (5). The DNA sequence was determined by the dideoxy-chain termination method (26) after different DNA restriction fragments were subcloned in the pGEM4Z plasmid (Promega Biotec, Madison, Wis.) by using the strategy shown in Fig. 1. A T7 DNA polymerase sequencing kit (Pharmacia, Uppsala, Sweden) and SP6 and T7 promoters (Promega) were used for sequencing and as primers, respectively.

By examination of possible ORFs, we found only one large frame (Fig. 2), designated ORF E, which extended from nucleotides 610 to 1386 over 777 bp. The 5′ upstream region contained a 5′-GAGGT-3′ sequence (indicated as SD in Fig. 2) complementary to the 3′ end of 16S rRNA (3′-UCUUUCCUCCAGUAG-5′) of *B. subtilis* (23), at position −9 to −13 from the ATG initiation codon. Potential promoter regions following consensus sequences pointed out by Graves and Rabinowitz (10) were found. The most plausible area for transcriptional initiation seems to be the 5′-TATATC-3′ sequence (at positions −33 to −38 from the ATG codon) for the −10 region and the 5′-GTGTTT-3′ sequence (at positions −64 to −69 from the ATG codon) for the −35 region. Translation of ORF E resulted in a polypeptide of 259 amino acids with a predicted molecular mass of 30.3 kDa. The amino acid sequence of ORF E was compared with those of the replication proteins of gram-positive bacterial plasmids such as pFTB14 from *Bacillus amyloliquefaciens* (24), pC194 (16) and pUB110 from *Staphylococcus aureus* (21), pLP1 from *Lactobacillus plantarum* (2), and pBAA1 from *B. subtilis* (7) and also with the phage φX174 replication protein by using the CLUSTAL homology search system (14). As shown in Fig. 3, significant homologies were observed between the ORF E product and the proteins involved in gram-positive bacterial plasmid replication. This result strongly suggests that the ORF E-encoded protein is implicated in pBC1 replication; thus, this protein was named RepE. As in other published replication protein sequences, a tyrosine residue is present at position 171. This amino acid was described as the linkage site of the φX174 protein A and the DNA when nicking occurs at the plus-strand origin (29). At position 161 began a sequence (5′-TTTCTTATCTTGATA-3′) (Fig. 2) which was identical to the known plus-strand origin of replication sequences of pC194, pUB110 (11), pBAA1 (7), and pLP1 (2) and which was highly similar to that of the *E. coli* phage φX174 (5′-TCCCCCAACTTGATA-3′ [19]). Within this 15-bp conserved sequence were two additional features. (i) The sequence CTTGATA was the sequence at which nicking of the plus strand occurs in the initiation of replication of φX174. (ii) An 18-bp sequence (including these conserved 15 bp) has been shown by Gros et al. (11) to contain a signal sufficient to terminate replication of plus-strand synthesis. These data strongly suggest that this 15-bp conserved region was a part of the pBC1 plus-strand origin of replication. To identify the minimal region of pBC1

* Corresponding author.

† This paper is no. 205, of Special Project RAISA, subproject no. 4, of the National Research Council of Italy.

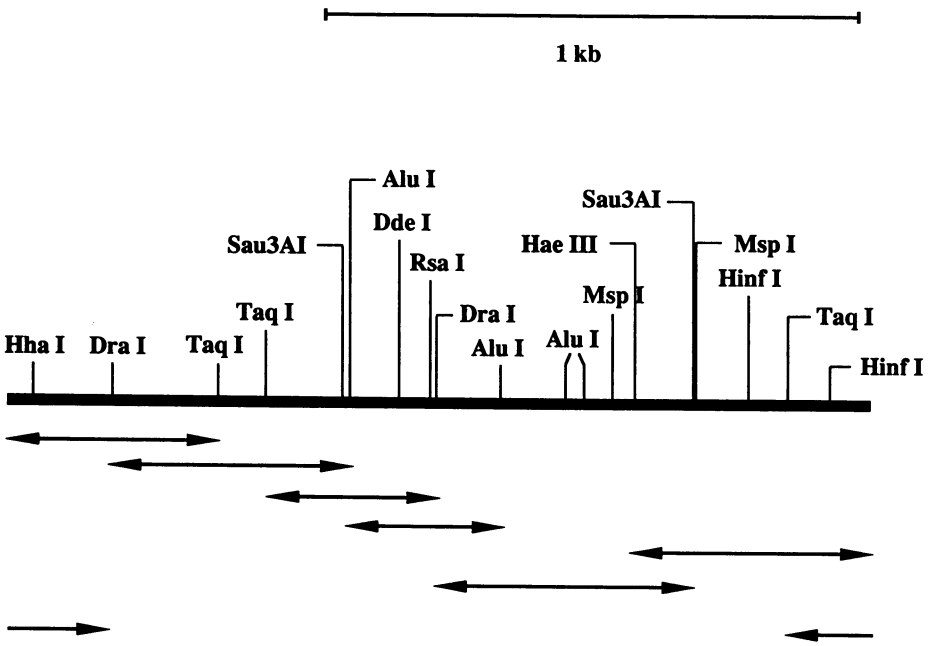

FIG. 1. Physical map of pBC1. The sequencing strategy is shown by the arrows below the map.

containing the plus-strand origin, pBC1 derivatives were constructed in vitro and analyzed for the ability to replicate in the gram-positive *B. subtilis* host in the presence of a helper plasmid supplying the pBC1 RepE initiator protein. Gel-purified pBC1 restriction fragments were inserted into the *E. coli* vector pJH101, which is devoid of a gram-positive replicon but which contains the chloramphenicol resistance determinant of pC194 (9). Plasmids pORI1, pORI2, and pORI3 contained, in the *Eco*RV site of pJH101, the *Hae*III-*Dra*I (bp 1 through 198), the *Hae*III-*Hinc*II (bp 1 through 288) and the *Hae*III-*Alu*I (bp 1 through 640) fragments of pBC1, respectively. Plasmid pBC2 was obtained by inserting into the *Hha*I site of pBC1 the 2.6-kb *Hha*I fragment of pE194, encoding erythromycin resistance (15). The plasmid constructions were initially used to transform *E. coli* HB101 (20) to chloramphenicol resistance, and the predicted structures of pORI1, pORI2, and pORI3 were verified by restriction mapping (data not shown). The resultant plasmids were then tested for origin activity by their ability to transform *B. subtilis* PB1424 in which RepE was provided in *trans* by plasmid pBC2. Only pORI2 and pORI3 replicated efficiently in *B. subtilis* cells harboring plasmid pBC2; they failed to replicate in cells devoid of the helper plasmid providing the RepE protein. The smallest cloned segment containing origin activity was 286 bp, an interval located between position 1 (*Hae*III recognition site) and the *Hinc*II recognition site at position 287, although the actual site of the origin, presumably the site of interaction with the RepE product, is possibly much smaller. Of the transformant clones whose plasmid DNAs were examined on an agarose gel, all harbored two different plasmid molecules with the same molecular sizes as pBC2 and pORI2, but no recombinant plasmids were generated between them (data not shown).

A body of evidence supports the hypothesis that pBC1 replicates by a rolling-circle-type mechanism. Key features of this mode of replication include (i) an origin of plus-strand synthesis containing a site at which the DNA is nicked; (ii) a replication protein which nicks the plus strand, forms a covalent link with the DNA, and renicks and ligates the newly synthesized plus strand after one round of replication; and (iii) an origin sequence for initiation of minus-strand synthesis. The replication functions of pBC1 have several features homologous to the replication functions of *E. coli* phage φX174 and *S. aureus* plasmids pC194 and pUB110, which replicate by a rolling-circle-type mechanism (1, 11, 18). The similarities between φX174, pC194, pUB110, and pBC1 include the features i and ii, which are essential to this mode of replication. Indeed, there are conservation of sequence at the plus-strand origin and conservation of amino acids of the active site of the replication protein in all four cases.

Numerous plasmids from gram-positive bacteria which replicate by this mechanism accumulate single-stranded DNA molecules as replication intermediates (27, 28). If pBC1 replicates by a rolling-circle-type mechanism, we would predict that pBC1 would accumulate strand-specific single-stranded DNA in *B. subtilis*. Cells containing pLM6 plasmid (pBC1 plasmid marked with the chloramphenicol acetyltransferase gene [6]) were tested for the presence of single-stranded plasmid DNA. Whole-cell lysates were prepared and electrophoresed by the method of te Riele et al. (27), and DNA was transferred to Hybond-N (Amersham Corp.), without or after denaturation, according to the instructions of the manufacturer. Southern hybridization was performed with 50% formamide at 37°C (20) by using as a probe pLM6 labeled with [α-^{32}P]dCTP by the multiprime random-labeling system (8). The filters were washed in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) (pH 7.0)–0.1% sodium dodecyl sulfate at 60°C with two wash changes and fluorographed at −70°C by using Kodak X-Omat AR films and DuPont Lightning Plus intensifying screens. As shown in Fig. 4, in lysates prepared from *B. subtilis* harboring pLM6, a sequence homologous to pLM6

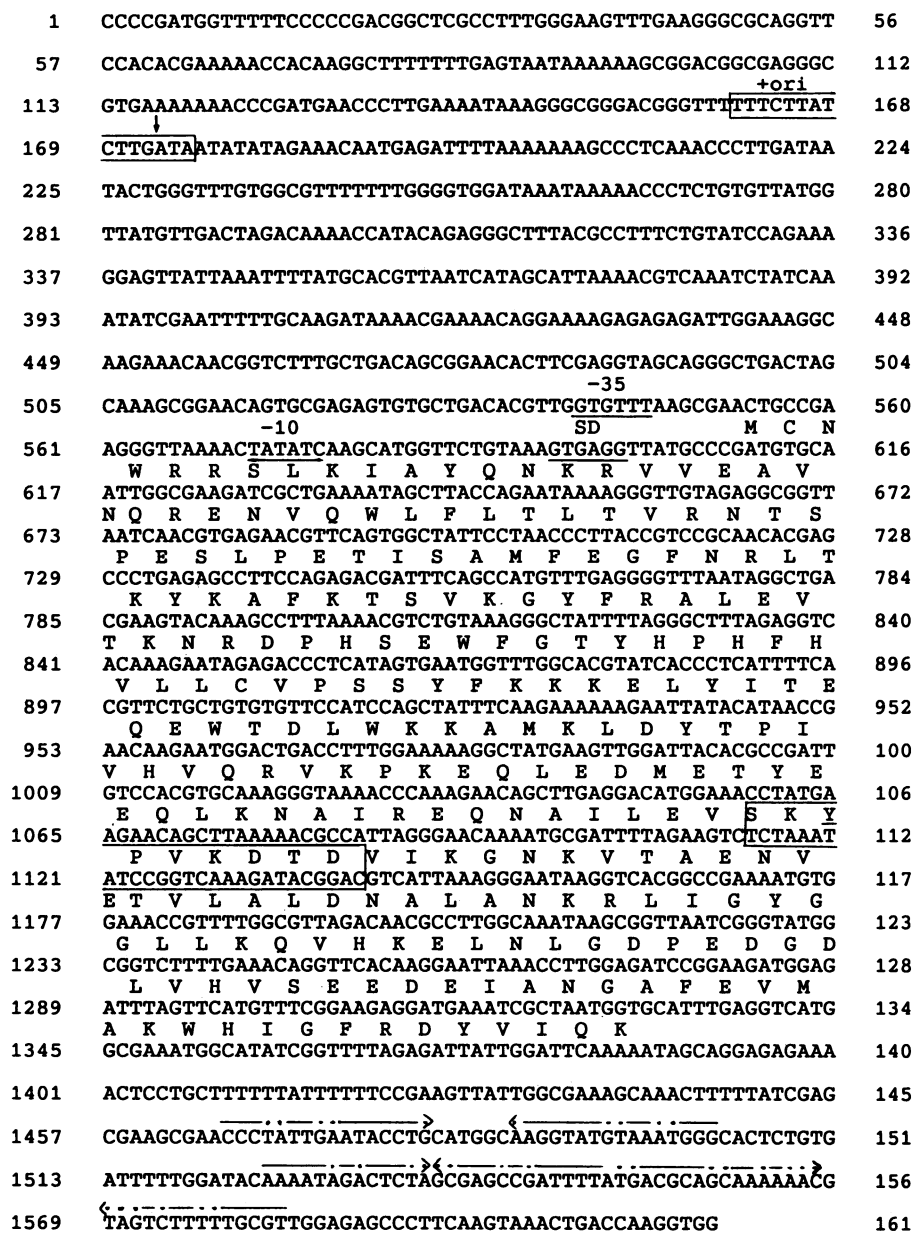

FIG. 2. pBC1 plasmid nucleotide sequence and predicted RepE protein amino acid sequence. The putative promoter sequences (−10, −35) and ribosome binding site (SD) are underlined. The possible plus-strand origin of replication (+ori), conserved in plasmids pC194, pUB110, pBAA1, and pLP1, is shown; the arrow within this sequence indicates the site at which nicking of the plus strand occurs in phage ϕX174. The putative active site of the RepE protein is boxed, and the conserved tyrosine residue is underlined. Palindromes are indicated by arrows above the DNA sequence; dots show positions of nonpalindromic bases.

that migrated faster than supercoiled monomeric plasmid DNA (Fig. 4A, lane a) and bound to a nylon filter in the absence of denaturation (Fig. 4B, lane c) was present. This fast-migrating band was preferentially degraded by S1 nuclease under conditions which cause no detectable degradation of the double-stranded forms (Fig. 4, lanes b and d). These data indicate that this band is single-stranded plasmid DNA.

In the rolling-circle model of plasmid replication, synthesis of the second strand initiates at a minus-strand origin (MO) sequence, generating a double-stranded plasmid DNA molecule. The MOs of several single-stranded DNA plasmids have been analyzed (1, 4, 7, 13) and located within regions (200 to 300 bp) containing imperfect palindromic structures. Similar structures were identified in a region downstream from ORF E of the pBC1 plasmid (bp 1455 through 1585) (Fig. 2). However, this region showed no extensive sequence similarity with any known MOs from other rolling-circle-replicating plasmids. Further experiments will be needed to clarify whether pBC1 is a plasmid with a new type of MO.

Nucleotide sequence accession number. The sequence

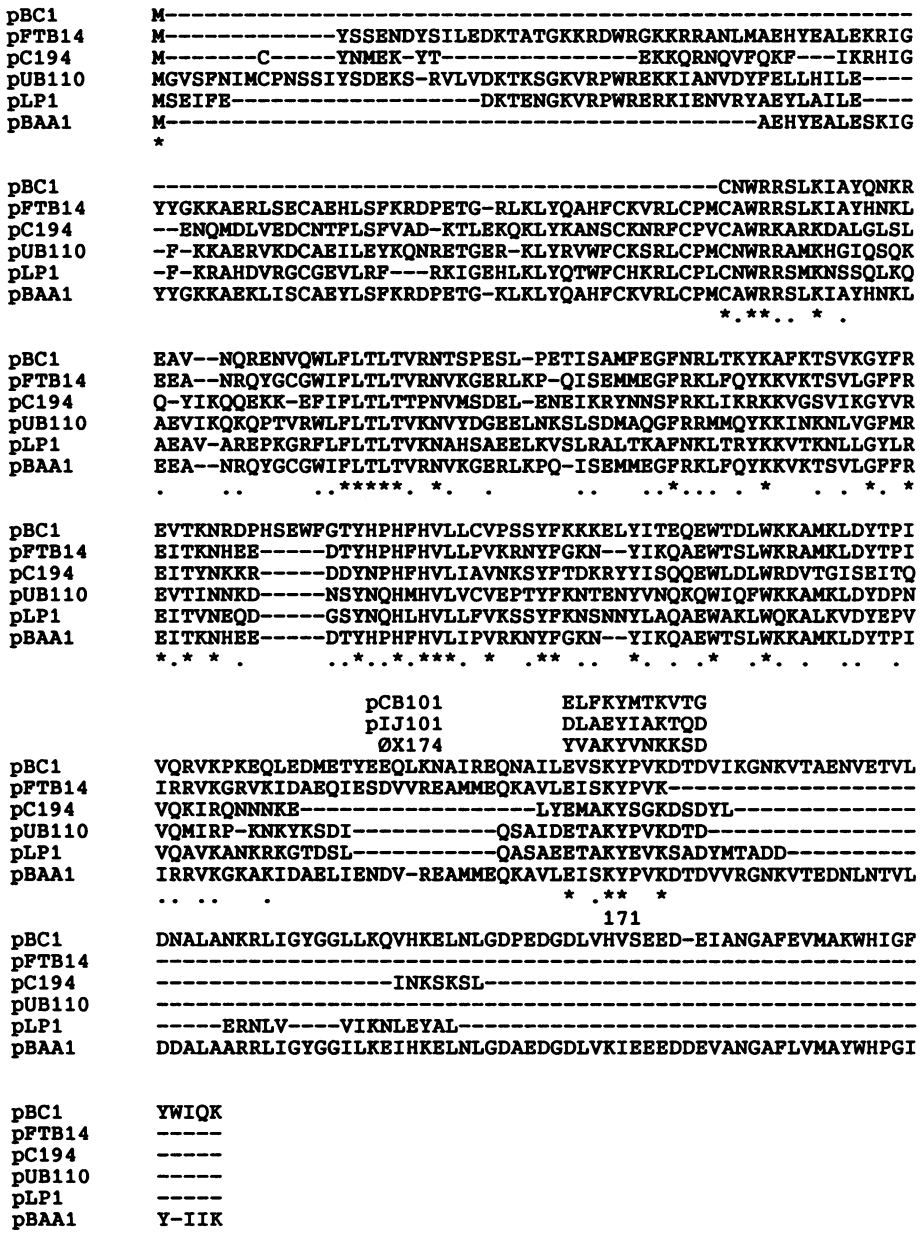

FIG. 3. Comparison of the amino acid sequences of the replication proteins of pBC1, pFTB14, pC194, pUB110, pLP1, and pBAA1. Asterisks indicate amino acids common to all replication proteins; dots indicate conservative amino acid substitutions. Only the replication protein region around the tyrosine (Y) residue, related to the enzymatic active site of ϕX174 protein A (25), is shown for plasmid pCB101 from *Clostridium butyricum* (22) and plasmid pIJ101 from *Streptomyces lividans* (17) (these short regions of sequence similarity have already been described by Gruss and Ehrlich [12]). Dashes represent gaps inserted to optimize the protein alignment.

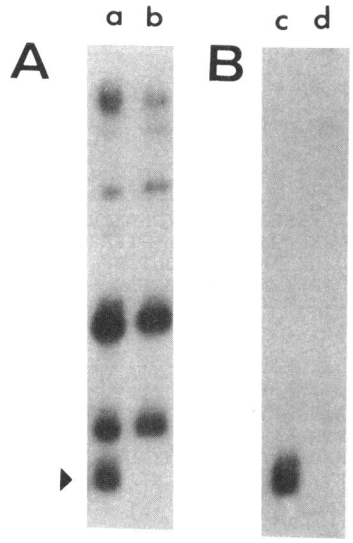

FIG. 4. Detection of single-stranded pLM6 DNA. Whole-cell lysates of *B. subtilis* containing pLM6 were treated (lanes b and d) or not treated (lanes a and c) with S1 nuclease. The DNA was denaturated (A) or not (B) prior to transfer to a nylon filter and was then hybridized to [α-^{32}P]dCTP-labelled pLM6. The band of single-stranded plasmid DNA (▶) is shown.

shown in Fig. 2 has been submitted to GenBank and assigned the accession number M64604.

This work was supported by National Research Council of Italy Special Project RAISA.

REFERENCES

1. **Boe, L., M.-F. Gros, H. te Riele, S. D. Ehrlich, and A. Gruss.** 1989. Replication origins of single-stranded-DNA plasmid pUB110. J. Bacteriol. **171**:3366–3372.
2. **Bouia, A., F. Bringel, L. Frey, B. Kammerer, A. Belarbi, A. Guyonvarch, and J.-C. Hubert.** 1989. Structural organization of pLP1, a cryptic plasmid from *Lactobacillus plantarum* CCM 1904. Plasmid **22**:185–192.
3. **Contente, S., and D. Dubnau.** 1979. Characterization of plasmid transformation in *Bacillus subtilis*: kinetic properties and the effect of DNA conformation. Mol. Gen. Genet. **167**:251–258.
4. **del Solar, G. H., A. Puyet, and M. Espinosa.** 1987. Initiation signals for the conversion of single stranded to double stranded DNA forms in the streptococcal plasmid pLS1. Nucleic Acids Res. **15**:5561–5580.
5. **De Rossi, E., P. Brigidi, G. Riccardi, and D. Matteuzzi.** 1989. Plasmid screening in thermophilic *Bacillus*: physical characterization and molecular cloning. Curr. Microbiol. **19**:13–19.
6. **De Rossi, E., P. Brigidi, M. Rossi, D. Matteuzzi, and G. Riccardi.** 1991. Characterization of Gram-positive broad host-range plasmids carrying a thermophilic replicon. Res. Microbiol. **142**:389–396.
7. **Devine, K. M., S. T. Hogan, D. G. Higgins, and D. J. McConnell.** 1989. Replication and segregational stability of *Bacillus* plasmid pBAA1. J. Bacteriol. **171**:1166–1172.
8. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
9. **Ferrari, F. A., A. Nguyen, D. Lang, and J. A. Hoch.** 1983. Construction and properties of an integrable plasmid for *Bacillus subtilis*. J. Bacteriol. **154**:1513–1515.
10. **Graves, M. C., and J. C. Rabinowitz.** 1986. *In vivo* and *in vitro* transcription of the *Clostridium pasteurianum* ferredoxin gene. Evidence for "extended" promoter elements in Gram-positive organisms. J. Biol. Chem. **261**:11409–11415.
11. **Gros, M. F., H. te Riele, and S. D. Ehrlich.** 1987. Rolling circle replication of single-stranded DNA plasmid pC194. EMBO J. **6**:3863–3869.
12. **Gruss, A., and S. D. Ehrlich.** 1989. The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. Microbiol. Rev. **53**:231–241.
13. **Gruss, A. D., H. F. Ross, and R. P. Novick.** 1987. Functional analysis of a palindromic sequence required for normal replication of several staphylococcal plasmids. Proc. Natl. Acad. Sci. USA **84**:2165–2169.
14. **Higgins, D. G., and P. M. Sharp.** 1989. Clustal: a package for performing multiple sequence alignments on a microcomputer. Gene **73**:237–244.
15. **Horinouchi, S., and B. Weisblum.** 1982. Nucleotide sequence and functional map of pE194, a plasmid that specifies inducible resistance to macrolide, lincosamide, and streptogramin type B antibiotics. J. Bacteriol. **150**:804–814.
16. **Horinouchi, S., and B. Weisblum.** 1982. Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. J. Bacteriol. **150**:815–825.
17. **Kendall, K. J., and S. N. Cohen.** 1988. Complete nucleotide sequence of the *Streptomyces lividans* plasmid pIJ101 and correlation of the sequence with genetic properties. J. Bacteriol. **170**:4634–4651.
18. **Kornberg, A.** 1980. DNA replication, p. 473–570. W. H. Freeman & Co., San Francisco.
19. **Langeveld, S. A., A. D. M. van Mansfeld, P. D. Baas, H. S. Jansz, G. A. van Arkel, and P. J. Weisbeek.** 1978. Nucleotide sequence of the origin of replication in bacteriophage ϕX174 RF DNA. Nature (London) **271**:417–420.
20. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **McKenzie, T., T. Hoshino, T. Tanaka, and N. Sueoka.** 1987. A revision of the nucleotide sequence and functional map of pUB110. Plasmid **17**:83–85.
22. **Minton, N. P., and J. D. Oultram.** 1988. Host:vector systems for gene cloning in *Clostridium*. Microbiol. Sci. **5**:310–315.
23. **Moran, C. P., Jr., N. Lang, S. F. J. Le Grice, G. Lee, M. Stephens, A. L. Sonenshein, J. Pero, and R. Losick.** 1982. Nucleotide sequences that signal the initiation of transcription and translation in *Bacillus subtilis*. Mol. Gen. Genet. **186**:339–346.
24. **Murai, M., H. Miyashita, H. Araki, T. Seki, and Y. Oshima.** 1987. Molecular structure of the replication origin of a *Bacillus amyloliquefaciens* plasmid pFTB14. Mol. Gen. Genet. **210**:92–100.
25. **Sanger, F., A. R. Coulson, T. Friedmann, G. M. Air, B. G. Barrell, N. L. Brown, J. C. Fiddes, C. A. Hutchison III, P. M. Slocombe, and M. Smith.** 1978. The nucleotide sequence of bacteriophage ϕX174. J. Mol. Biol. **125**:225–246.
26. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
27. **te Riele, H., B. Michel, and S. D. Ehrlich.** 1986. Single-stranded plasmid DNA in *Bacillus subtilis* and *Staphylococcus aureus*. Proc. Natl. Acad. Sci. USA **83**:2541–2545.
28. **te Riele, H., B. Michel, and S. D. Ehrlich.** 1986. Are single-stranded circles intermediates in plasmid DNA replication? EMBO J. **53**:631–637.
29. **van Mansfeld, A. D. M., H. A. A. M. van Teeffelen, P. D. Baas, and H. S. Jansz.** 1986. Two juxtaposed tyrosyl-OH groups participate in ϕX174 gene A protein catalysed cleavage and ligation of DNA. Nucleic Acids Res. **14**:4229–4238.