Plus-Origin Mapping of Single-Stranded DNA Plasmid pE194 and Nick Site Homologies with Other Plasmids

S. SOZHAMANNAN,† PATRICK DABERT, VERONIQUE MORETTO, S. DUSKO EHRICH, AND ALEXANDRA GRUSS*

Laboratoire de Génétique Microbienne, Institut de Biotechnologie, Institut National de la Recherche Agronomique—Domaine de Vilvert, 78350 Jouy en Josas, France

Received 28 November 1989/Accepted 3 May 1990

*Present address: National Cancer Institute, Bethesda, MD 20892.
†Corresponding author.

Staphylococcus aureus plasmid pE194 manifests a natural thermosensitivity for replication and can be established in several species, both gram positive and gram negative, thus making it attractive for use as a delivery vector. Like most characterized plasmids of gram-positive bacteria, pE194 generates single-stranded DNA. The direction of pE194 replication is clockwise, as determined by the strandedness of free single-stranded DNA. Significant homology exists between a 50-base-pair sequence in the origin of pE194 and sequences present in plasmids pMV158 (Streptococcus agalactiae), pADB201 (Mycoplasma mycoides), and pSH71 (Lactococcus lactis). We used an initiation-termination reaction, in which pE194 initiates replication at its own origin and is induced to terminate at the related pMV158 sequence, to demonstrate that pE194 replicates by a rolling-circle mechanism; the initiation nick site was localized to an 8-base-pair sequence.

Numerous small multicopy plasmids of gram-positive bacteria have been analyzed and employed extensively for cloning and manipulation. Many of the plasmids currently used were isolated from Staphylococcus aureus (17, 21) and used for cloning in other hosts (8, 10, 14); examples are pUB110 (24), pT181 (19), pE194 (15), and pE194 (16). Plasmid pE194 is of particular use; its thermosensitive character (33) allows its application as a delivery vector or for inducible replication. A super-thermosensitive mutant (38, 41; A. Gruss, unpublished data) and copy mutants (38, 39) of pE194 have been generated and widely used. The plasmid is readily established in Bacillus subtilis (14) and has also been reported to replicate in Escherichia coli (H. te Riele, personal communication). Numerous characteristics of pE194 are known; it has been fully sequenced (16), it encodes its own protein required for replication, RepF (14, 38), and it is erythromycin resistant (16, 17). Interestingly, pE194 shows homology with other plasmids isolated from gram-positive bacteria; e.g., the pre protein of pE194 (Fig. 1) is also present on plasmids pT181, pUB110, and pMV158 (9, 37; see reference 12 for a review).

There are indications that pE194 replicates via a rolling-circle mechanism (RCR): (i) it generates single-stranded DNA (ssDNA) (36); and (ii) it has a minus origin, palA (active only in S. aureus), which is homologous to the minus origins of other plasmids replicating by RCR and in the absence of which ssDNA is accumulated (13). However, the data supporting these implications and more detailed analyses of the origin of replication are inconclusive, even controversial (6, 33, 38). Since pE194 is a prototype representing an important family of plasmids in gram-positive bacteria (the plus origins of only two types of plasmids have been analyzed), we characterized the pE194 plus origin. A 248-base-pair (bp) sequence was shown unequivocally to contain origin activity. A subsequence of the plus origin was found to have imperfect homology with the origin of Streptococcus agalactiae plasmid pMV158. Previous studies showed homology in the replication proteins of these two plasmids (26; pMV158 is the parent of plasmid pLS1). It has been observed that homology among plus origins is most conserved around the initiation nick site (3, 11, 12). We make use of these homologies to prove that pE194 replicates by RCR, which initiates by a specific nick within the origin.

MATERIALS AND METHODS

Plasmids and strains. Plasmids were established in E. coli JM101 (40) and HVC45 (laboratory collection), or B. subtilis SB202 and recE4 (laboratory collection). Standard DNA transformation and cloning techniques were used in all cases. Plasmids pE194 (16), pC194 (3, 15), and pMV158 (22; pLS1 is a deletion derivative of pMV158 missing 1.1 kilobases [28]) have all been sequenced; the published nucleotide numbering is maintained in this work. To facilitate plasmid preparations, all pE194 constructs described in this work utilize pE194cop6, a high-copy-number derivative of pE194 (16). A plasmid construct containing all pMV158 sequences was the kind gift of V. Hirschfeld. Plasmid pHV1060 is a marker plasmid composed of pUC18 plus an erythromycin resistance gene that can be expressed in B. subtilis (M.-F. Gros, unpublished data). pRepF-2 is composed of the replicative segment of pE194 on a TaqI B fragment (bp 387 through 1939) and a segment of pUB110 encoding kanamycin resistance on a TaqI A fragment of pUB110 (bp 4377 through 3446). pRepF-1 contains the SphI-Clai (bp 1202 through 1939) segment of pE194 encoding RepF cloned into the SphI-EcoRI fragment (bp 1018 through 4544) of pUB110. Both pRepF-2 and pRepF-1 express the pE194 repF gene. Other plasmids used in this work are represented in Fig. 1 and 4.

Fragment purification was effected by using a Gene Clean kit (Bio101, La Jolla, Calif.). Techniques used in cloning (23), Southern blot hybridization (34), and sequencing (32) were performed by published procedures. Agarose gel electrophoresis was performed with 0.7% agarose in Tris-borate-EDTA buffer containing 40 μg of ethidium bromide per ml (23). Densitometric analyses to measure plasmid copy number per chromosome equivalent were performed on gel photo negatives with a Shimazdu scanner; whole-cell lysates...
of strains containing plasmids were run in three dilutions on the gels.

RESULTS

Direction of replication of pE194. Plasmid pE194 generates ssDNA in *B. subtilis* (36), which should correspond to the displaced or plus-strand replication intermediate (35). To determine which strand is displaced, we constructed two derivatives of pE194. Each derivative was composed of pE194 and pBR322 joined at their unique CfoI sites in one of the two possible orientations. Depending on the orientation, a different strand of the pBR322 sequences will be released as ssDNA during plasmid replication. We used strand-specific oligonucleotide probes hybridizing to each strand of pBR322 to determine which single strand was accumulated for each construct. Total DNAs of *B. subtilis* strains containing the two plasmids were separated by agarose gel electrophoresis, and the gel was then analyzed by Southern blot hybridization. Each ssDNA hybridized exclusively to one probe (data not shown). The strand detected as free ssDNA indicated that the direction of pE194 replication is clockwise with respect to the published map (Fig. 1).

This direct demonstration of the direction of replication of pE194 resolves previous uncertainties (13, 33); our determination is in agreement with that deduced from the known orientation of the minus origin (13).

Localization of the origin of replication of pE194. The minimal replicon of plasmid pE194 had been previously localized between bp 762 and 1926 on the published map (Fig. 1A) (14). This segment contains the RepF protein, which is required for replication (14, 38), and the plus origin. More precise measurements of the plus origin, previously reported, were either in conflict or not entirely conclusive (6, 33, 38). To identify unambiguously a minimal region of pE194 containing the plus origin, segments of the minimal replicon (Fig. 1B) were subcloned in *E. coli* onto marker plasmid pHV1060, which does not replicate in *B. subtilis*. The resulting plasmids were tested for origin activity by their ability to transform *B. subtilis* strains in which RepF was provided in trans. RepF was provided by two different plasmid constructs, pRepF-2 and pRepF-1 (see Materials and Methods). In initial experiments, homology existed between the complementing plasmid, pRepF-2, and the test clones. We therefore tested transforming activity in both wild-type (rec') and recE4 (defective for homologous recombination) backgrounds. The plasmid DNA content of the transformants was examined by gel hybridization. Depending on the DNA insert, two types of transformants were obtained; either they appeared in the rec' strain only and contained plasmids larger than the input ones, or they appeared in both rec' and recE4 strains and had a plasmid of the monomer size (in this case they also generated ssDNA...
protein of pMV158 (parent of pLS1) (isolated from S. agalactiae [26]), and an open reading frame of pADB201 (isolated from Mycoplasma mycoides [2]). These similarities in replication function induced us to look for sequence homologies in a region corresponding to plasmid origins. Within the identified pE194 origin sequence described above, we found a highly conserved subsequence in a 50-bp DNA stretch in these plasmids (Fig. 3A). In addition, the direction of replication determined for pLS1 (31) with respect to the homologous origin sequences is the same as that found above for pE194, further supporting their relatedness. A fourth plasmid, pHV271, originating in Lactococcus lactis (7) was also reported to share this sequence homology (Fig. 3; W. M. de Vos, personal communication).

We speculated that the highly conserved sequence and secondary structure contains the initiation site recognized by the RepF protein of pE194, and we applied this idea to search for the pE194 initiation nick site.

Using the initiation-preface termination reaction to map the nick site of pE194. Previously, it was shown that the insertion of a second plus origin into a plasmid replicating by RCR provokes premature termination at this second sequence. For example, the pC194 origin nick site was analyzed by insertion of a second sequence having homology with the origin (11). In the presence of the pC194-encoded replication protein, deletion derivatives formed as a result of initiation at one origin sequence and premature termination at the other, proving that pC194 replicates by RCR and allowing localization of the nick site in the origin (11). A similar approach was used subsequently to map the pUB110 origin nick site (3) and to analyze relatedness of pT181 and pC221 origins (18).

We used this strategy to map the pE194 origin. We constructed plasmid pHV2616 (Fig. 4A), which contains the following: (i) the pE194 plus origin; (ii) the plus origin of pMV158, which has homology (39 out of 48 bp) with pE194 (Fig. 3); and (iii) a pC194 replicon, which serves as the carrier vector. The pE194 and pMV158 origin segments are oriented such that replication from either origin proceeds in the same direction. pHV2616 was maintained without detectable deletion in cells lacking the RepF protein. Plasmid DNA was prepared from pHV2616 and used to transform recipient recE4 strains containing either pUB110, repF-2, or pRepF-1. Only the latter two strains provide RepF and

[Fig. 2]. Plasmids giving the latter result were considered to have origin activity. These experiments placed the plus-origin sequence upstream of the repF gene. We subsequently constructed a RepF donor plasmid, pRepF-1, that had no homology with the origin (see Materials and Methods) and performed further studies of complementation in the rec- background.

A summary of the origin fragments tested and their activities is given in Fig. 1. The smallest cloned segment containing origin activity is 287 bp, whereas the smallest sequence common to all the positive clones is 248 bp. These results are in agreement with those reported by Dempsey and Dubnau during the course of our work (6). This region is separated by about 170 bp from the RepF open reading frame (16).

The pE194 origin has homology with origin sequences of plasmids from other species. The pE194 RepF protein was previously found to share amino acid homology with the Rep

FIG. 2. Origin activity of cloned segments of pE194 in wild-type and recE4 backgrounds. Plasmids pHV2005 and pHV2004 (Fig. 1) were tested in a strain providing RepF in trans from pRepF-2, which has homology with the test plasmids. pHV2005 could be established in a recE4 background, whereas pHV2004 could not. ds, Double-stranded monomeric plasmid DNA; ss, single-stranded monomeric plasmid DNA; hybrid, plasmid that has homology with the pHV2004 but does not migrate at the position of the monomeric plasmid.

[Fig. 3]. Homologous sequences (left to right, 5' to 3') within the plus origins of pE194-like plasmids. Sequences are lined up to optimize the homologies. Sequence numbering is given according to published maps. Nucleotides in capital letters are homologous; those in lowercase letters are nonhomologous. Letters in boldface type indicate the localization of the nick site by the pE194-pMV158 initiation-termination reaction; sequence under the hyphenated line localizes the nick site by the premature termination reaction (see text). Arrows above indicate palindromic structure. Restriction sites AvrII and Rsal are present in pE194 (as shown) but not in pMV158 and were used to localize the origin in the initiation-termination reaction (see text). &. There is a 23-bp intervening sequence AAAGGAAGCGAATTTTGCTTCCG, not present in the other origins. (B) Mutations at the AvrII restriction site affect plasmid copy number. A plasmid pE194::pBR322 cointegrate, joined at the ClaI site (the pE194 cops derivative is used), was linearized with AvrII, and the site was filled in with Klenow enzyme. Mutated sequences of two plasmids are shown below the wild-type (w.t.) sequence around the relevant region (additional bases are in boldface type). Plus-origin activity, as estimated by plasmid copy number in B. subtilis, is given at right.
should activate the pE194 origin. Whole-cell DNA was extracted from transformant colonies and screened by agarose gel electrophoresis (Fig. 4B). pHV2616 was established in strains containing pUB110. In strains containing pRepF-2 or pRepF-1 (i.e., the RepF donor) and pHV2616, the donor plasmid was present in normal copy number, but pHV2616 was present at low copy number or absent, and an extra band was detected corresponding to a plasmid smaller than pHV2616. The smaller plasmid was isolated from eight independent transformants, purified by transformation in *B. subtilis*, and analyzed by DNA sequencing. In every case, pHV2616 suffered the same deletion event. The junctions mapped to within a 10-bp overlap sequence, TACTACGA CC, present in pE194 and pMV158 (Fig. 3A). Since these deletions were detected only in the presence of RepF protein, we conclude that they result from specific nicking by RepF at the pE194 plus origin and precocious termination at the pMV158 origin.

In similar experiments, we cloned the PstI fragment containing the pMV158 origin into the PstI site of pE194 and transformed *B. subtilis* competent cells. Two types of plasmid transformants were obtained; in one type the pMV158 origin segment was cloned in orientation opposite to that of the pE194 origin, and in the other type a smaller plasmid was present (data not shown). Sequence analysis of four such smaller plasmids mapped the junction to the same site as identified in pHV2616 deletion derivatives. This would be expected if clones containing the pMV158 origin in the same orientation as the pE194 origin underwent a deletion by initiation at the pE194 origin and termination at the pMV158 origin. Interestingly, the pE194-pMV158 hybrid origin formed by deletion formation was active with the RepF protein, despite sequence differences in the two origins (Fig. 3A).

Premature termination (described above with termination in the pMV158 origin) can also occur at secondary, less accurate termination sites; for example, hybrid plasmids composed of pC194, pBR322, and phage f1 sequences frequently underwent deletions in *E. coli*, of which one endpoint coincided with the pC194 replication nick site (25). It was proposed that pC194 initiates replication at its normal origin but terminates prematurely at false termination sites; subsequent observations (1) were consistent with this proposal. We substituted pE194 for pC194 in analogous constructions (pE194 is replicative in *E. coli*, as shown by the presence of ssDNA in a pE194::pBR322 hybrid [H. te Riele, personal communication]) and obtained deleted plasmid derivatives (data not shown). Three independent deletions with endpoints in the pE194 origin region were found. The proportion of such deletions was much lower than those described for pC194 (about 6% compared to about 50%), possibly because pE194 is less active than pC194 in *E. coli*. The deletion endpoints determined by DNA sequencing were the same in the three plasmids; the sequence containing the pE194 origin nick site, as mapped above, is joined to an M13 sequence with 8 bp of sequence homology, TAC TACGA, with the plus origin. We interpret these deletions as arising from proper initiation at the pE194 origin and premature termination at an M13 sequence resembling the normal termination signal. These results confine the pE194 nick site to an 8-bp sequence of the plus origin, TAC TACGA, present in the loop of a potential stem-loop structure (Fig. 3A).

A sequence adjacent to the palindromic structure of the origin affects plasmid copy number. Plasmid pE194 contains a unique AvrII site at the border of the palindromic region.
flanking the nick site (Fig. 3B). To see whether the sequence recognized by AvrII is necessary for origin activity, DNA of plasmid pE194::pBR322, joined at the Clai site (the pE194 replicon is the cop6 derivative [16]) was linearized by AvrII, and overhanging ends were filled in with DNA polymerase I (Klenow fragment). Plasmids were established in E. coli by virtue of the pBR322 replicon, and the modified region was sequenced, thereby identifying plasmids with two different mutations (Fig. 3B). These were transferred to B. subtilis to determine their replicative activity, as measured by plasmid copy number. Compared with the parental pE194::pBR322 construct (copy number of about 25 per chromosome equivalent), the plasmids mutated at AvrII had markedly decreased copy numbers (plasmid AvrII-6 copy number was about 15, and AvrII-9 could only be detected by Southern blot hybridization and grew very slowly under selection [data not shown]). It is likely that the sequence mutated in these plasmids is required for origin recognition. We cannot rule out, however, an effect of this region on RepF expression, as there is an open reading frame overlapping the origin that crosses the AvrII site (16) (Fig. 1).

DISCUSSION

Four types of replicons isolated from S. aureus are the known prototypes for all ssDNA replicons found thus far in gram-positive bacteria (13): the pT181 (30), pC194 (11), pSN2 (29), and pE194 types. Each type is composed of a series of plasmids with amino acid homologies in the replication proteins and good sequence homologies in the plus origins (the pSN2 plus origin has not yet been identified). The best example of related plasmids originating from different species can be viewed in the pE194-type replicon described here; related plasmids were isolated from M. mycoides (pADB201) (2), S. agalactiae (pMV158) (4, 22), and L. lactis (pSH71) (7). Despite their similarities and the fact that pE194, pMV158, and pSH71 can all replicate in B. subtilis, pE194 cannot be established autonomously in L. lactis (4a). However, there is sufficient homology between origin sequences of pE194 and pMV158 that the latter is recognized as a termination signal by the pE194 RepF protein. Premature termination provoked by the presence of a sequence homologous to the origin is strictly a property of RCR. Using this reaction, we showed that pE194 replicates by RCR and that the nick site is contained within an 8-bp sequence, TACTACGA. In addition, knowledge of the nick site of pE194, by inference, allows one to predict the plus origin nick sites of pMV158, pADB201, and pSH71 with confidence.

The pE194 plus origin sequence has potential secondary structure, forming a strong stem-and-loop structure; the nick site is within the loop (Fig. 3A). A sequence alteration at the AvrII site (Fig. 3B), i.e., at the base of this structure, decreases the replicative activity.

Thanks to its thermosensitive character (33), pE194 has been successfully used for inducible deletion (27) and amplification of adjacent sequences when present in the chromosome (M.-A. Petit et al., submitted for publication) and also as a delivery vector (41). We observe that the related replicon pSH71 (L. lactis) is also thermosensitive for replication when assayed in B. subtilis at 51°C. Like pE194, derivatives of the pSH71 replicon (or cognate replicon pWVO1 [20]) are readily established in high copy number in a variety of hosts (7, 20); the common feature of thermosensitivity of these plasmids may be a useful tool in building a delivery vector with a more widespread use.

ACKNOWLEDGMENTS

We are grateful to V. Hirschfield for providing us with the pMV158 replicon and to W. M. de Vos for communicating the pSH71 replicon sequence before publication. We thank V. Akueson for the artwork.

LITERATURE CITED