JOURNAL OF BACTERIOLOGY, Apr. 1989, p. 1846–1853
0021-9193/89/041846-08$02.00/0

Vol. 171, No. 4

Deletion Hot Spots in Chimeric *Escherichia coli* Plasmids

BÉNÉDICTE MICHEL,† EMMANUELLE D'ALENÇON, AND S. DUSKO EHRLICH*

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

Received 18 August 1988/Accepted 29 December 1988

Deletions form frequently in chimeric plasmids composed of M13mp2, pBR322, and pC194 (B. Michel and S. D. Ehrlich, Proc. Natl. Acad. Sci. USA 83:3386–3390, 1986). They are generated by joining of the nucleotide neighboring the nick site in the M13 replication origin to a nonadjacent nucleotide. This nucleotide is most often located within particular short plasmid regions, named deletion hot spots. Three natural hot spots were present in the chimeric plasmids. Two were active only when the DNA replication initiated at the M13 origin was allowed to progress; the third was active only in the presence of wild-type amounts of DNA ligase. Three artificial hot spots were generated by creating palindromic sequences in the plasmids.

Recombination between sequences that share little or no homology, termed illegitimate by Franklin (17), may be due to errors of enzymes that replicate or break and join DNA (2, 17). Recombination between short direct repeats has often been attributed to errors of the replication enzymes (1, 13, 14, 16, 28), which were shown to occur during excision of a transposon related to Tn*10* (6). In contrast, recombination involving DNA topoisomerases (7, 25, 26) provides an example of the errors of break and join enzymes.

The involvement of DNA breaking and joining enzymes in illegitimate recombination is also illustrated by the formation of deletions at nicked replication origins, such as plus origins of single stranded DNA phages (21, 35, 41) and plasmids (36) or the conjugal transfer origin (*oriT*) of plasmid F (22). A nick is introduced in the origin by the genome-specific replication protein, and the deletions are formed by joining the nucleotide adjacent to the nick (either at the 5′ or the 3′ side) to a nucleotide present elsewhere in the molecule. Previous work has indicated that nicking and joining may be carried out by the same protein (36). To further analyse this process, the endpoints of deletions initiated at the M13 replication origin were examined. They were found to cluster in short regions, named hot spots. Three naturally occurring hot spots were characterized, and three artificial hot spots were generated by creating palindromic sequences in the genomes analyzed.

MATERIALS AND METHODS

Bacterial strains and plasmids. The bacterial strains and plasmids are listed in Table 1. Strains N2603 and N2604 were rendered *hsdR* by P1 cotransduction of *hsdR* with *zjj*::Tn*10* marker (the donor strain was GC2885; R. D'Ari, personal communication). All plasmids were derived from pBR322, the *Bacillus subtilis* plasmid pC194, and M13mp2. Nucleotide coordinates used for pBR322 are from Sutcliffe (42) as modified by Backman and Boyer (3), those used for M13 are from van Wezenbeek et al. (43), and those used for pC194 are from Horinouchi and Weisblum (20), modified and numbered from the *Hin*dIII site (11). Restriction enzymes, ligase, and polymerase were from commercial sources and were used according to suppliers' recommendations.

DNA preparation and analysis. For the construction of hybrid genomes, plasmid DNA was extracted from *B. subtilis* and *Escherichia coli* by the cleared-lysate method (8, 37) and purified by chromatography on hydroxyapatite (9). For analytical purposes and sequencing, plasmid DNA was extracted by a rapid procedure (4). The dideoxynucleotide sequencing technique was used (40) on double-stranded templates with the polymerization reaction performed for 5 min at 42°C with the Klenow fragment of DNA polymerase I (P. Lecomte, personal communication).

Competence induction and transformation. Transformation of *B. subtilis* and *E. coli* cells that are not thermosensitive has been described previously (10, 37). Thermosensitive *E. coli* cells [*lig*(Ts7)] and the isogenic thermoresistant control cells [*lig*(Tr)] were grown overnight at 30°C in rich medium, diluted 200 times, grown at 37°C until the A_{650} reached 0.1, and incubated at 42°C for 30 min. Competence was induced by a standard $CaCl_2$ procedure; the cells were then incubated with DNA for 10 min at 0°C followed by 5 min at 42°C. They were then inoculated in rich liquid medium prewarmed to 42°C, incubated at that temperature for 1 h, and plated on selective plates prewarmed to 42°C. The plates were incubated for 90 min at 42°C and then overnight at 37°C. The viability of the *lig*(Ts7) and *lig*$^+$ cells was not affected by this procedure. Cell competence and the level of ligase were routinely controlled by transformation with intact and *Hin*dIII-linearized pBR322 DNA. Intact DNA yielded regularly 10^5 to 10^6 transformants per μg with both *lig*$^+$ and *lig*(Ts7) cells, whereas linear DNA gave 10^2 to 10^3 and 0 transformants per μg of DNA with *lig*$^+$ and *lig*(Ts7) cells, respectively.

Plasmid nomenclature. The plasmids constructed for this study which were structurally unstable in *E. coli* were labeled by the letter U. Three parental plasmids used were named pU1, pU2, and pU3 (Fig. 1). Two progeny plasmids were made by deleting the hot spot A from pU1 and pU3. They are identified by the letter D (e.g., pU1D). The insertion of a natural hot spot in the parental or progeny plasmids is indicated by the name of the hot spot (e.g., pU1D-B), and the insertion of a synthetic hot spot is signaled by the letter S (e.g., pU1-BS). Inversion of the orientation of a plasmid segment is indicated by a prime (e.g., A and A′). pB and pC stand for pBR322 and pC194, respectively, and P1 and P2 refer to artificial palindromes created at *Pvu*II and *Nru*I sites of pBR322, respectively (Table 1).

Plasmid constructions. Plasmids comprised of pBR322, pC194, and M13mp2 sequences were all unstable in *E. coli*. To construct them we followed a two-step procedure. First,

* Corresponding author.
† Present address: Rockfeller University, New York, NY 10021.

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Description	Origin or reference
B. subtilis HVS49	*trpC2 hisA aroB2 tyrA*	14
E. coli		
HVC748	*thrA1 leu-6 thi-1 proA2 his-4 argE3 lacY1 galK2 ara-14 xyl-15 mtl-1 tsx-33 rpl-31 supE44 hsdR* (F' *lacI*^{q}Z Δ M15)	36
HVC749	As HVC747 but Rep3 *pro*$^+$	36
N2603	*strA ptsI105 gal*$^+$	18
N2604	As N2603 but *lig*(Ts7)	18
HVC750	As N2603 but Tc^r *hsdR*	This work
HVC751	As N2604 but Tc^r *hsdR*	This work
Plasmids		
pC194	Natural isolate	27
pE194	Natural isolate	27
pBR322	Cloning vector	5
pHV33	pBR322 linked to pC194 by *Hin*dIII sites	38
pHV677	pBR322 linked to pC194 by *Pvu*II sites	This work
pHV727	Replication origin of pC194 (bp 1430 through 1485) and Em^r gene of pE194 linked to pBR322 *Cla*I site	This work
pU1	M13mp2 linked to pHV33 by *Eco*RI sites; previously called pHV672	35
pU1-pC'	Inversion of the pU1 *Hin*dIII segment corresponding to pC194	This work
pU1-pB'	Inversion of the *Aat*II-*Pvu*II (bp 4286 through 2066) segment of pBR322 in pU1	This work
pU1-pB'pC'	Inversion of the *Aat*II-*Pvu*II (bp 4286 through 2066) segment of pBR322 in pU1-pC'	This work
pU1-A'	Inversion of the *Pvu*II segment (pBR322 bp 2066 through pC194 bp 316) of pU1	This work
pU1D	Deletion of the *Pvu*II segment (pBR322 bp 2066 through pC194 bp 316) carrying hot spot A in pU1	This work
pU1D-A	Insertion of the pBR322 *Dde*I segment (bp 1581 through 1743) in the *Pvu*II site of pU1D, same orientation as in pU1	This work
pU1D-A'	Same construction as pU1D-A, opposite orientation of the *Dde*I segment	This work
pU2	M13mp2 linked to pHV677 by *Eco*RI sites	This work
pU3	M13mp2 linked to pHV727 by *Eco*RI sites	This work
pU3D	Deletion of the *Pvu*II-*Ava*I (bp 2066 through 1425) segment of pBR322 in pU3	This work
pU3D-BS	Insertion of a synthetic oligonucleotide (bp 1353 through 1391) of pC194 in the *Nru*I site (pBR322 bp 974) of pU3D, same orientation as in pU1	This work
pU3D-BS'	Same construction as pU3D-BS, opposite orientation of the synthetic oligonucleotide	This work
pU3D-CS	Insertion of a synthetic oligonucleotide (bp 224 through 256) of pC194 in the *Nru*I site (pBR322 bp 974) of pU3D; same orientation as in pU1	This work
pU3D-CS'	Same construction as pU3D-CS, opposite orientation of the synthetic oligonucleotide	This work
pU3-P1	Insertion of a synthetic oligonucleotide (bp 2067 through 2096) of pBR322 in the *Pvu*II site (pBR322 bp 2068) of pU3	This work
pU3D-P2	Insertion of a synthetic oligonucleotide (bp 944 through 970) of pBR322 in the *Nru*I site (pBR322 bp 974) of pU3D	This work

suitable derivatives lacking the M13mp2 sequences, which were stable in *E. coli*, were constructed. Then, the *Eco*RI-cleaved M13mp2 DNA was joined to them (invariably at the *Eco*RI site of pBR322), and the hybrids were introduced into *B. subtilis*, where they were stable. More details about the construction and structure of the hybrids carrying M13mp2 are given in Table 1 and Fig. 1.

Deletion frequencies. *E. coli* transformants obtained with each of the unstable genomes used in this work contained at least one and sometimes several species of smaller derivative plasmids. The deletion frequencies correspond to the number of times a given derivative plasmid was observed among all plasmids deriving from the same unstable genome. The amount of each derivative plasmid, which varied between 10 and 50% of the total plasmid content when several derivatives were present in the same transformant, was not used to correct the deletion frequencies, since it could reflect the rate of plasmid segregation during cell growth rather than the frequency of plasmid formation.

RESULTS

Plasmid pU1, composed of pBR322, pC194, and M13mp2 (Fig. 1), is stable in *B. subtilis* but not in *E. coli*, where it undergoes deletions. Of 20 analyzed deletions 18 had one end at the nick in the M13 replication origin, and 9 of those had the other end either 3.15 kilobases (kb) (between base pairs [bp] 1610 and 1700 of pBR322) or 6.3 kb (between bp 1360 and 1400 of pC194) from the nick (35). These regions were named hot spots A and B, respectively (Fig. 1).

Hot spot A. Over 20% of the deletions arising in pU1 (10 out of 44 studied) ended in hot spot A. To test whether this hot spot corresponds to a motif (a sequence or a secondary structure) particularly prone to recombine with the origin, we first deleted from pU1 a 2.3-kb *Pvu*II segment carrying the hot spot (the plasmid was named pU1D; Fig. 1) and then inserted into the *Pvu*II site of pU1D a 160-bp segment of pBR322 expected to contain the hot spot (bp 1582 through 1743) either in the orientation corresponding to that of pU1 or in the opposite orientation (the plasmids were named pU1D-A and pU1D-A', respectively). Of deletions arising in pU1D-A and pU1D-A', 27% (12/44) and 35% (15/43), respectively, had an end at the hot spot, which was located 2.8 kb from the M13 replication origin. In contrast, no deletions arising in the parental plasmid devoid of hot spot A (pU1D) had an end at this distance from the origin. This suggests that a motif particularly prone to recombine with the M13 repli-

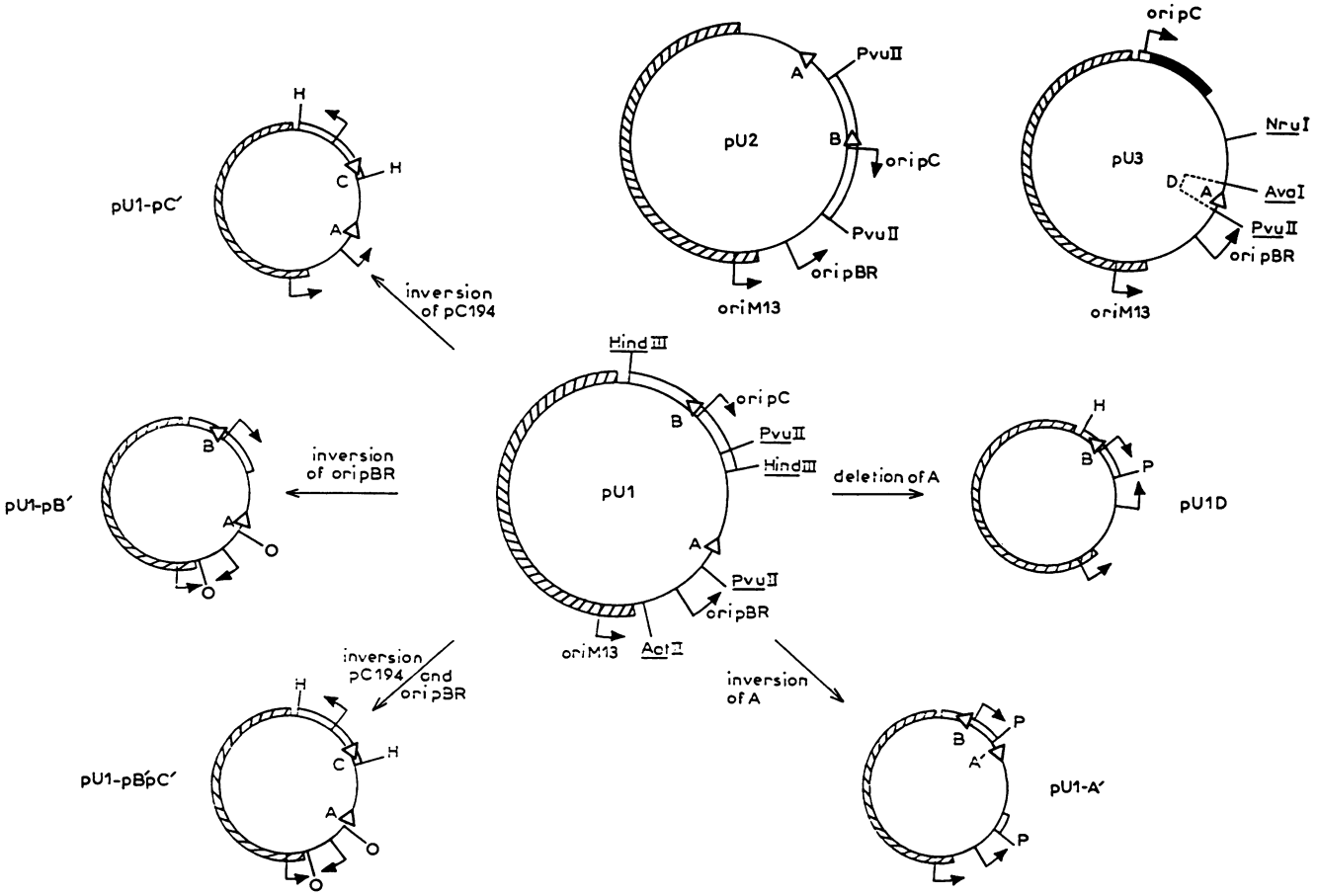

FIG. 1. Plasmids used. Lines: (——) pBR322 sequences, (▭) pC194 sequences, (▨) M13mp2 sequences, (■) Em^r gene of plasmid pE194. oriM13, oripBR, and oripC denote replication origin of M13, pBR322, and pC194, respectively. The arrows show the direction of replication of each replicon. Triangles indicate the position and the (arbitrary) orientation of the active hot spots. Only the restriction sites used to construct pU1 and pU3 derivatives are shown. H and P, *Hin*dIII and *Pvu*II sites, respectively; O, a sequence generated by joining *Aat*II and *Pvu*II sites. The deletion of hot spot A in pU3, which generates pU3D, is indicated by the interrupted line and the letter D. The pU1 derivatives are not drawn to scale.

cation origin is present at hot spot A and that it is active in both orientations.

Restriction analysis of a number of independent plasmids resulting from deletion at A indicated that A is composed of two subregions, A1 and A2, corresponding to bp 1680 through 1690 and bp 1620 through 1645, respectively. When A was carried in the original orientation (as in pU1) A1 was dominant (35 of 40 deletions ending within were at A1, 5 of 40 ended at A2); when it was carried in the opposite orientation (in pU1D-A') A2 was dominant (7 of 7 deletions ended at A2). Sequence analysis of 10 deletions ending at A1 and 7 deletions ending at A2 (Fig. 2) confirmed the localization of the subregions A1 and A2.

Hot spot B. About 30% of the deletions (16 of 52) arising in pU1 ended in hot spot B (6.3 kb from the M13 origin, bp 1360 through 1400 of pC194). Several arguments suggested that this hot spot also corresponds to a sequence (or a structural element) prone to recombine with the M13 replication origin. (i) Deletions arising in plasmids pU1D and pU2 (Fig. 1) ended frequently (in 36 of 84 cases analyzed) 3.9 kb from the

FIG. 2. Deletions in hot spot A. The nucleotide sequence of pBR322 between nucleotides 1604 and 1701 is shown. The asterisks indicate the position of deletion endpoints found with the original (top) or inverted (bottom) hot spot orientation. The number of asterisks corresponds to the number of independent deletions analyzed. Palindromic sequences are indicated with inverted arrows. A1 and A2 refer to the subregions of hot spot A.

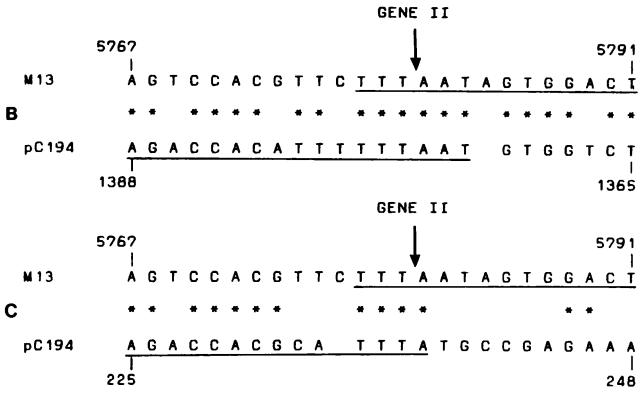

FIG. 3. Deletions in hot spots B (top) and C (bottom). Sequences are shown in the 5'-to-3' direction. The underlined sequence is that of the novel junction. The asterisks indicate homologous nucleotides. The arrow indicates the position at which the gene II product introduces a nick in the M13 replication origin.

M13 replication origin, which corresponds to the location of pC194 sequence 1360 through 1400. (ii) Plasmid pU3 (Fig. 2), which is composed of pBR322, M13mp2, the Em^r gene of pE194, and the replication origin of pC194 (bp 1430 through 1485) (19), lacks the pC194 sequences 1360 through 1400. It is stably maintained in *B. subtilis* (provided that the pC194 replication protein is synthesized in the host cell; unpublished data) but undergoes deletions in *E. coli*. No deletions arising in that plasmid ended either 6.3 or 3.9 kb from the M13 replication origin.

Hot spot B was active only in the orientation present in pU1. In plasmid pU1-pC' (Fig. 1), where it is localized 6.1 kb from the M13 replication origin but in the opposite orientation, no deletions out of 48 analyzed ended at that distance.

Restriction analysis of 31 independent plasmids indicated that they all arose from identical events. This was confirmed by sequence analysis of four plasmids. Recombination at B corresponds to joining of the M13 origin nick site to a partially homologous sequence in pC194 (Fig. 3). One such event was described previously (35).

To determine whether a shorter sequence would have the properties of the hot spot B, we made use of a plasmid devoid of hot spots A and B, named pU3D (Fig. 1; it was obtained by deleting in vitro from pU3 the segment *Pvu*II-*Ava*I, corresponding to pBR322 sequence 2066 through 1425). A 32-bp oligonucleotide, designated BS and corresponding to bp 1363 through 1391 of pC194, was inserted into the *Nru*I site of pU3D (localized 3.2 kb from the M13 replication origin) either in the orientation corresponding to that of pU1 or in the opposite orientation (the resulting plasmids were named pU3D-BS and pU3D-BS', respectively). The inserts are represented in Fig. 4. Of 56 independent deletions arising in pU3D-BS, 4 ended at BS, as shown by restriction analysis, whereas none of the 154 deletions arising in pU3D ended 3.2 kb from the M13 replication origin. Sequence analysis confirmed that the deletion endpoint at BS was identical to that observed at B (Fig. 4A). This shows that a 32-bp sequence can act as hot spot B. A longer sequence may, however, be needed for a full hot spot activity.

Unexpectedly, 2 of 25 deletions ended in the BS region of plasmid pU3D-BS', where its orientation should have rendered it inactive. Sequencing showed, however, that the joining took place not at BS but at the foot of the hairpin structure that could form upon insertion of BS in pU3D (Fig. 4B). This observation prompted us to test whether palindromes can recombine with the M13 replication origin (see below, palindromic hot spots).

Hot spot C. Plasmid pU1-pC', which contains pC194 in the orientation opposite to that found in pU1, carried another hot spot, named C (Fig. 1). Restriction analysis has shown that 20% (8 of 41) of the deletions arising in pU1-pC' end 5 kb from the M13 replication origin (between bp 220 and 250 of pC194), whereas no deletion out of 44 analyzed ended at that distance from the M13 replication origin in the parental plasmid, pU1. The endpoints of four such deletions were analyzed by sequencing and found to be identical. The deletions were generated by joining the M13 replication origin to a partially homologous pC194 sequence (Fig. 3).

To test whether a shorter sequence has hot spot C activity, a 33-bp synthetic oligonucleotide (designated CS and corresponding to bp 224 through 256 of pC194) was inserted in the *Nru*I site of pU3D, either in the original or in the inverted orientation (the plasmids were designated pU3D-CS and pU3D-CS', respectively). Restriction analysis showed that 0 of 154, 2 of 19, and 0 of 28 deletions occurred in the *Nru*I region of pU3, pU3D-CS, and pU3D-CS', respectively. Sequencing has demonstrated that the two deletions arising in pU3D-CS are analogous to those identified at the hot spot C, which shows that a 33-bp sequence can act as that hot spot.

Palindromic hot spots. Two artificial palindromes, named P1 and P2, were tested for hot spot activity. The first, 26 bp long (Fig. 4C), was created at the *Pvu*II site of pU3; the second, 27 bp long (Fig. 4D), was created at the *Nru*I site of pU3D (the plasmids were designated pU3-P1 and pU3D-P2, respectively). Restriction analysis has indicated that none of the 86 deletions arising in pU3 ended close to the *Pvu*II site. Similarly, none of the 154 deletions arising in pU3D ended close to the *Nru*I site of pU3D. However, when the palindromes were present, 9 of 57 and 5 of 58 deletions ended close to P1 in pU3-P1 and P2 in pU3D-P2, respectively. This indicates that palindromes can recombine with the M13 replication origin.

Endpoints were determined by sequencing in 9 of the above 14 deletions (4 arising in pU3-P1, 5 arising in pU3D-P2). They were localized close to the foot or at the side of the potential hairpin structure (asterisks, Fig. 4C and D), and, relative to the center of the palindrome, they were all proximal to the M13 replication origin. A similar conclusion was obtained from the fine restriction mapping of the remaining five deletion endpoints, since all but one were located at the foot of the palindrome proximal to the M13 origin, the remaining one being 30 bp further on the other side of the palindrome.

Effect of DNA replication on hot spot activity. Previous studies of deletions arising at the nick in the replication origin of plasmid pC194 have indicated that the progression of the replication fork can influence deletion endpoints (36). Three different replication origins known to be active in *E. coli* are present in pU1 (pC194, pBR322, and M13; Fig. 1). We tested the effect of the progression of the replication fork initiated at each of them on the activity of hot spot A.

The replication fork initiated at the pC194 replication origin enters the hot spot A present in pU1 from the clockwise direction, whereas that initiated at the pBR322 replication origin reaches it from the counterclockwise direction (Fig. 1). We constructed plasmids in which the orientation of the two replication origins was such that pC194 (in pU1-pC'), pBR322 (in pU1-pB'), or both (in pU1-pB'-pC') forks reached hot spot A from the direction

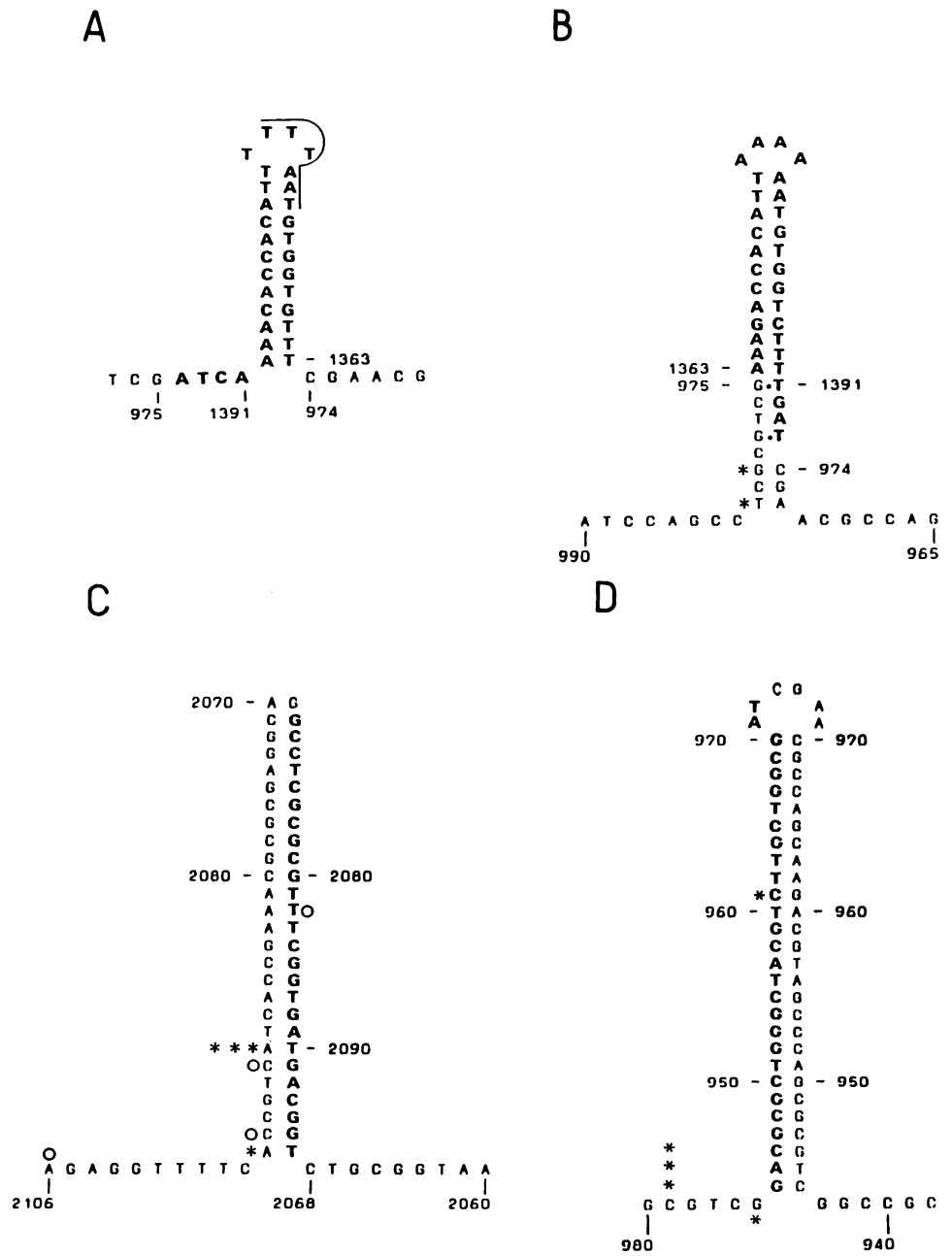

FIG. 4. Hot spots generated by insertion of synthetic oligonucleotides. Sequences are shown in the 5′-to-3′ direction; the direction of M13 replication is in all cases from left to right. Bold letters denote synthetic sequences. (A and B) Synthetic nucleotide BS, corresponding to pC194 sequence 1363 through 1391 and three additional nucleotides creating a *Sau*3A site (GATC) needed for cloning, was inserted between pBR322 nucleotides 974 and 975. (A) Original orientation of BS. The line indicates a sequence homologous to the M13 replication origin where recombination takes place. (B) Inverted orientation of BS. The points signal G · T base pairs. (C) Artificial palindrome P1. The synthetic oligonucleotide corresponding to pBR322 sequence 2071 through 2096 was inserted between pBR322 nucleotides 2068 and 2069. (D) Artificial palindrome P2. The synthetic oligonucleotide corresponding to pBR322 sequence 944 through 970, to which two nucleotides were added to generate a *Sau*3A site needed for cloning, was inserted between nucleotides 974 and 975 of pBR322. The nucleotides indicated by a star or a circle were linked to M13 replication origin in the Rep^+ and Rep^- strains, respectively; the number of stars at a given position indicates the number of independent deletions analyzed.

opposite to that of pU1. A similar proportion of deletions ended at A in pU1 and in the other three plasmids (10 of 44 in pU1, 15 of 48 in pU1-pB′, 14 of 41 in pU1-pC′, 13 of 51 in pU1-pB′-pC′). This indicates that the activity of hot spot A is independent of the progression of the pC194 and pBR322 replication forks, or that the forks synthesizing pC194 and pBR322 leading and lagging strands activate A in a similar way. An argument in favor of the first hypothesis is that in plasmid pU3 (Fig. 1), where the pC194 replication origin is inactive because of the absence of the cognate replication protein, a high proportion of deletions (38 of 80) ended at A.

The replication fork initiated at the M13 replication origin does not progress in *E. coli rep* mutants, which are deficient in a DNA helicase (33). Deletions formed at the M13 origin

TABLE 2. Effect of the *rep* and *lig* mutations on different hot spots[a]

Hot spot	No. of deletions/total (%) with the following:		
	rep$^+$ *lig*$^+$	*rep lig*$^+$	*rep*$^+$ *lig*
A	90/295 (31)	4/186 (2)	7/28 (25)
A′	15/43 (35)	0/30 (<3)	ND[b]
B	67/167 (40)	8/37 (21)	0/91 (<1.1)
C	8/41 (20)	0/79 (<1.5)	ND
P1	9/57 (16)	4/50 (8)	3/14 (21)

[a] The numbers of instances deletions ended at a given hot spot among all instances analyzed are indicated. Results with different parental plasmids are pooled; percentage values are shown within parentheses. The results with each individual parental plasmid were not significantly different from the indicated mean. *rep*$^+$ *lig*$^+$ refers to HVC748 or HVC750 (the proportion of deletions at different hot spots was not significantly different in the two strains); *rep lig*$^+$ refers to HVC749, isogenic with HVC747; *rep*$^+$ *lig* refers to HVC751, isogenic with HVC750.

[b] ND, Not determined.

with similar overall efficiency in the *rep*$^+$ and the mutant *rep* isogenic strains (35). The proportion of deletions ending in hot spot A was, however, drastically reduced in the *rep* strain (entries A and A′; Table 2). This indicates that the activity of hot spot A depends on the progression of the M13 replication fork.

The activity of hot spot C, but not that of hot spots B and P1, also depended on the progression of the M13 replication fork (Table 2). Sequence analysis of deletions arising in a *rep* host at the last two hot spots confirmed this conclusion. One deletion ending at B was analyzed and was found to be identical to that observed in the *rep*$^+$ hosts (Fig. 3). Out of four deletions at P1, three ended in the same region as the deletions formed in the *rep*$^+$ host (proximal to the M13 replication origin relative to the center of the palindrome); the fourth was at the other side of the palindrome (Fig. 4C).

The activity of hot spots B and C was independent of the direction of replication of pBR322 (in pU1-pB′ and pU1-pB′-pC′; data not shown). The effect of the replication of pC194 on hot spot B or of pBR322 and pC194 on P1 was not tested.

Effect of ligase on hot spot activity. One of the possibilities considered to account for deletions at nicked replication origins was that a ligase may join the 5′-phosphate nucleotide adjacent to the nick to a distant 3′-hydroxyl nucleotide (35, 36). We therefore tested the effect of ligase on the activity of several hot spots. For that purpose a *lig*(Ts7) *E. coli* strain, HVC751, was kept temporarily at a restrictive temperature and used as a recipient for transformation with different plasmids derived from pU1 (see Materials and Methods). Under these conditions deletions ended frequently at hot spots A and P1 but were not found at hot spot B (Table 2).

DISCUSSION

The nick in the replication origin of phage M13 is a deletion endpoint independent of various host recombination functions (35). In the present work we characterized short DNA sequences (hot spots) frequently found at the other deletion endpoint. In addition, we analyzed the effect of DNA replication and DNA ligase on the activity of various hot spots.

We detected several naturally occurring deletion hot spots and generated several others artificially. The natural hot spots have little in common besides the short but noticeable

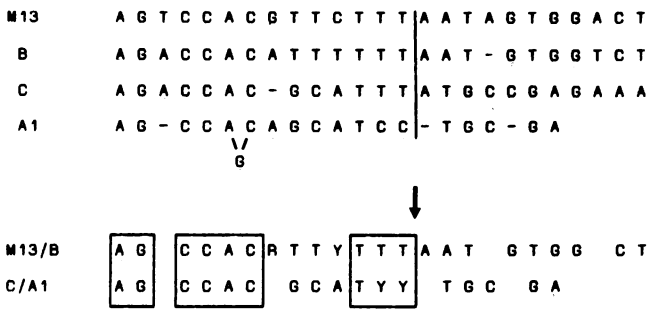

FIG. 5. Homologies between natural hot spots and the M13 replication origin. The sequence alignment is shown at the top. The junction site between a hot spot and the origin is indicated by the vertical line (for the hot spot A1, this corresponds to the site where six identical deletions were found; Fig. 2). Gaps (indicated by a dash) and the inserted base were introduced to optimize the alignments. The homologies between M13 and B and between C and A1 are highlighted at the bottom. The four common sequences are boxed. The arrow indicates the junction site, which corresponds to the M13 origin nick site.

sequence homology (Fig. 5). A is complex, spans almost a 100 bp, and is composed of two subregions, A1 and A2. B and C are simpler; a core sequence of about 30 bp is sufficient for their activity, but the flanking sequences possibly increases the efficiency of the core. Hot spot B and one subregion of hot spot A (A2) are palindromic, whereas hot spot C and the other subregion of A (A1) are not. The orientation of A relative to the M13 origin did not affect its overall activity, although it changed the relative importance of the two subregions (A1 was dominant in the original, A2 was dominant in the opposite orientation). B and C were active in one orientation only. The three artificial hot spots (P1, P2, BS′) share no sequence homology, but all have a palindromic structure. Different hot spots responded differently to DNA replication and the presence of ligase. Hot spots A and C required the progression of the M13 replication fork for their activity, whereas the other hot spots did not. Hot spot B, on the other hand, was dependent on DNA ligase.

Differences in the structure of the hot spots and in the functions necessary for their activity suggest that there may be a variety of mechanisms by which deletions can form at the nick in the M13 replication origin. One of the possible mechanisms is aberrant termination of DNA replication by the M13 gene II protein, like that observed with the replication protein of another single-stranded DNA replicon, pC194 (36). Homology between the M13 replication origin and several natural hot spots (A1, B, and C; Fig. 5) as well as the dependence of the activity of two hot spots (A, C) on the progression of the M13 replication fork (Table 2), which could generate single-stranded DNA substrate for the M13 gene II protein, supports this interpretation. However, the ligase requirement at the hot spot B and the absence of any homology between the M13 replication origin and the palindromic hot spots cannot be easily explained by the involvement of the gene II protein in deletion formation at these hot spots.

Deletions could possibly also be generated by joining of the 5′-phosphate nucleotide at the nick to a nonadjacent 3′-hydroxyl nucleotide by DNA ligase (35, 36). The deletion hot spots might correspond to replication pause sites, since 3′-hydroxylnucleotides at such sites would be available for joining for longer time than at other sites. The support for this hypothesis derives from the analysis of artificial palin-

dromic hot spots P1, P2, and BS′ (Fig. 4). In these, 90% of deletions (18 of 20) ended at the side (or at the base) of the palindrome proximal to the M13 replication origin. It is known from in vitro studies that *E. coli* polymerase III (29, 32), T4 DNA polymerase (23, 24), and the eucaryote polymerase α (29, 31, 39, 44) are arrested at palindromic sequences. Alternatively, DNA cleavage at palindromes, rather than the arrest of replication, could generate a 3′-hydroxyl nucleotide. Bacterial enzymes that can resolve recombination intermediates, such as *recBCD* nuclease (34), T4 endonuclease VII (30), and T7 gene 3 protein (12), could mediate such cleavage.

The putative 3′-hydroxyl nucleotide arising at different hot spots would have to be joined to the 5′-phosphate nucleotide adjacent to the nick in the M13 replication origin by DNA ligase. A ligase requirement for deletion formation at hot spot B was observed (Table 2). However, this requirement was not found at other hot spots (Table 2). Among possible explanations is that the ligase mutation may be leaky and that its residual activity was insufficient for joining at hot spot B but sufficient for joining at the other hot spots. Alternatively, upon introduction of plasmids in competent cells, deletions might be generated at hot spots other than B during cell growth at the permissive temperature, necessary for colony formation, when the ligase is active. Further work is needed to test different facets of the proposed models.

ACKNOWLEDGMENTS

We thank L. Jannière for numerous useful discussions. This work was supported by grants from Ministère de la Recherche et de l'Enseignement Supérieur (grant 510070), Fondation Médicale Française, and the Commission of the European Communities (grant BAP-0141-F).

LITERATURE CITED

1. **Albertini, A. M., M. Hofer, M. P. Calos, and J. M. Miller.** 1982. On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. Cell **29**:319–328.
2. **Anderson, P.** 1987. Twenty years of illegitimate recombination. Genetics **115**:581–584.
3. **Backman, K., and H. W. Boyer.** 1983. Tetracycline resistance determined by pBR322 is mediated by one polypeptide. Gene **26**:197–203.
4. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
5. **Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, and H. W. Boyer.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**:95–113.
6. **Brunier, D., B. Michel, and S. D. Ehrlich.** 1988. Copy choice illegitimate recombination. Cell **52**:883–892.
7. **Bullock, P., J. J. Champoux, and M. Botchan.** 1985. Association of crossover points with toposiomerase I cleavage sites: a model for nonhomologous recombination. Science **230**:954–958.
8. **Clewell, D. B., and D. R. Helinski.** 1969. DNA-protein complex in *E. coli*: purification and induced conversion to an open circular DNA form. Proc. Natl. Acad. Sci. USA **62**:1444–1448.
9. **Colman, A., M. J. Byers, S. B. Primrose, and A. Lyons.** 1978. Rapid purification of plasmid DNA by hydroxyapatite chromatography. Eur. J. Biochem. **91**:303–310.
10. **Dagert, M., and S. D. Ehrlich.** 1979. Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. Gene **6**:23–28.
11. **Dagert, M., I. Jones, A. Goze, S. Romac, B. Niaudet, and S. D. Ehrlich.** 1984. Replication functions of pC194 are necessary for efficient plasmid transduction by M13 phage. EMBO J. **3**:81–86.
12. **de Massy, B., R. A. Weisberg, and F. W. Studier.** 1987. Gene 3 endonuclease of bacteriophage T7 resolves conformationally branched structures in double-stranded DNA. J. Mol. Biol. **193**:359–376.
13. **Efstratiadis, A., J. W. Posakony, T. Maniatis, R. M. Lawn, C. O'Connel, R. A. Spritz, J. K. DeRiel, B. G. Forget, S. M. Weissman, J. L. Slightom, A. E. Blechi, O. Smithies, F. E. Baralle, C. C. Shoulders, and N. J. Proudfoot.** 1980. The structure and evolution of the human α-globin gene family. Cell **21**:653–668.
14. **Ehrlich, S. D.** 1978. DNA cloning in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **75**:1433–1436.
15. **Ehrlich, S. D., P. Noirot, M. A. Petit, L. Jannière, B. Michel, and H. te Riele.** 1986. Structural instability of *Bacillus subtilis* plasmids, p. 71–83. *In* J. K. Setlow and A. Hollaender (ed.), Genetic engineering, vol. 8. Plenum Publishing Corp., New York.
16. **Farabaugh, P. J., U. Schmeissner, M. Hofer, and J. H. Miller.** 1978. Genetic studies of the *lac* repressor. VII. On the molecular nature of spontaneous hostspots in the *lacI* gene of *Escherichia coli*. J. Mol. Biol. **126**:847–863.
17. **Franklin, N. C.** 1971. Illegitimate recombination, p. 175–194. *In* A. D. Hershey (ed.), The bacteriophage lambda. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Gottesman, M. M., M. L. Hicks, and M. Gellert.** 1973. Genetic and function of DNA ligase in Escherichia coli. J. Mol. Biol. **77**:531–547.
19. **Gros, M. F., H. te Riele, and S. D. Ehrlich.** 1987. Rolling circle replication of single-stranded DNA plasmid pC194. EMBO J. **12**:3863–3869.
20. **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.
21. **Horiuchi, K., J. V. Ravetch, and N. D. Zinder.** 1978. DNA replication of bacteriophage f1 *in vivo*. Cold Spring Harbor Symp. Quant. Biol. **43**:389–399.
22. **Horowitz, B., and R. C. Deonier.** 1985. Formation of Delta-tra F′ plasmids: specific recombination at oriT. J. Mol. Biol. **186**: 267–274.
23. **Huang, C.-C., and J. E. Hearst.** 1980. Pauses at positions of secondary structure during *in vitro* replication of single-stranded fd bacteriophage DNA by T4 DNA polymerase. Anal. Biochem. **103**:127–139.
24. **Huang, C.-C., J. E. Hearst, and B. M. Alberts.** 1981. Two types of replication proteins increase the rate at which T4 DNA polymerase traverses the helical regions in a single-stranded DNA template. J. Biol. Chem. **256**:4087–4094.
25. **Ikeda, H.** 1986. Bacteriophage T4 DNA topoisomerase mediates illegitimate recombination *in vitro*. Proc. Natl. Acad. Sci. USA **83**:922–926.
26. **Ikeda, H., K. Moriya, and T. Matsumoto.** 1980. *In vitro* study of illegitimate recombination: involvement of DNA gyrase. Cold Spring Harbor Symp. Quant. Biol. **45**:399–408.
27. **Iordanescu, S.** 1975. Recombinant plasmid obtained from two different compatible staphylococcal plasmids. J. Bacteriol. **124**: 597–601.
28. **Jannière, L., and S. D. Ehrlich.** 1987. Recombination between short repeated sequences is more frequent in plasmids than in the chromosome of *Bacillus subtilis*. Mol. Gen. Genet. **210**: 116–121.
29. **Kaguni, L. S., and D. Clayton.** 1982. Template-directed pausing in *in vitro* DNA synthesis by DNA polymerase α from *Drosophila melanogaster* embryos. Proc. Natl. Acad. Sci. USA **79**: 983–987.
30. **Kemper, B., F. Jensch, M. V. Depka-Prondzynski, H.-J. Fritz, U. Borgmeyer, and K. Mizuuchi.** 1985. Resolution of Holliday structures by endonuclease VII as observed in interactions with cruciform DNA. Cold Spring Harbor Symp. Quant. Biol. **49**: 815–825.
31. **König, H., H. D. Riedel, and R. Knippers.** 1983. Reactions *in vitro* of the DNA polylerase-primase from *Xenopus laevis* eggs. A role for ATP in chain elongation. Eur. J. Biochem. **135**: 435–442.
32. **LaDuca, R. J., P. J. Fay, C. Chuang, C. McHenry, and R. Bambara.** 1983. Site-specific pausing of deoxyribonucleic acid

synthesis catalyzed by four forms of *Escherichia coli* DNA polymerase III. Biochemistry **22**:5177–5188.

33. **Lane, D. H. E., and D. T. Denhardt.** 1974. The *rep* mutation. III. Altered structure of the replicating *E. coli* chromosome. J. Bacteriol. **120**:805–814.
34. **Leach, D. R. F., and F. W. Stahl.** 1983. Viability of lambda phages carrying a perfect palindrome in the absence of recombination nucleases. Nature (London) **305**:448–451.
35. **Michel, B., and S. D. Ehrlich.** 1986. Illegitimate recombination at the replication origin of bacteriophage M13. Proc. Natl. Acad. Sci. USA **83**:3386–3390.
36. **Michel, B., and S. D. Ehrlich.** 1986. Illegitimate recombination occurs between the replication origin of the plasmid pC194 and a progressing replication fork. EMBO J. **5**:3691–3696.
37. **Niaudet, B., and S. D. Ehrlich.** 1979. In vitro genetic labeling of *Bacillus subtilis* cryptic plasmid pHV400. Plasmid **2**:48–58.
38. **Primrose, S. B., and S. D. Ehrlich.** 1981. Isolation of plasmid deletion mutants and study of their instability. Plasmid **6**: 193–201.
39. **Reckmann, B., F. Grosse, C. Urbanke, R. Frank, H. Blöcker, and G. Krauss.** 1985. Analysis of secondary structures in M13mp8 (+) single-stranded DNA by the pausing of DNA polymerase α. Eur. J. Biochem. **152**:633–643.
40. **Sanger, F., S. Nicklens, and A. R. Coulsen.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
41. **Schaller, H.** 1978. The intergenic region and the origins for filamentous phage DNA replication. Cold Spring Harbor Symp. Quant. Biol. **43**:401–408.
42. **Sutcliffe, J. G.** 1978. pBR322 restriction map derived from the DNA sequence: accurate DNA size markers up to 4361 nucleotide pairs long. Nucleic Acids Res. **5**:2721–2728.
43. **van Wezenbeek, P. M. G. F., T. J. M. Hulsebos, and J. G. G. Schoenmakers.** 1980. Nucleotide sequence of the filamentous bacteriophage M13 DNA genome: comparison with phage fd. Gene **11**:129–148.
44. **Weaver, D. T., and M. L. de Pamphilis.** 1984. The role of palindromic and non-palindromic sequences in arresting DNA synthesis *in vitro* and in vivo. J. Mol. Biol. **180**:961–986.