JOURNAL OF BACTERIOLOGY, Aug. 1995, p. 4385–4391 Vol. 177, No. 15
0021-9193/95/$04.00+0

Differential DNA Secondary Structure-Mediated Deletion Mutation in the Leading and Lagging Strands

WILLIAM A. ROSCHE,[1,2] THUAN Q. TRINH,[1]† AND RICHARD R. SINDEN[1,2]*

Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati College of Medicine, Cincinnati, Ohio 45267-0524,[1] and Department of Biochemistry and Biophysics, Institute of Biosciences and Technology, Texas A&M University, Houston, Texas 77030-3303[2]

Received 30 January 1995/Accepted 30 May 1995

The frequencies of deletion of short sequences (mutation inserts) inserted into the chloramphenicol acetyltransferase (CAT) gene were measured for pBR325 and pBR523, in which the orientation of the CAT gene was reversed, in *Escherichia coli*. Reversal of the CAT gene changes the relationship between the transcribed strand and the leading and lagging strands of the DNA replication fork in pBR325-based plasmids. Deletion of these mutation inserts may be mediated by slipped misalignment during DNA replication. Symmetrical sequences, in which the same potential DNA structural misalignment can form in both the leading and lagging strands, exhibited an approximately twofold difference in the deletion frequencies upon reversal of the CAT gene. Sequences that contained an inverted repeat that was asymmetric with respect to flanking direct repeats were designed. With asymmetric mutation inserts, different misaligned structural intermediates could form in the leading and lagging strands, depending on the orientation of the insert and/or of the CAT gene. When slippage could be stabilized by a hairpin in the lagging strand, thereby forming a three-way junction, deletion occurred by up to 50-fold more frequently than when this structure formed in the leading strand. These results support the model that slipped misalignment involving DNA secondary structure occurs preferentially in the lagging strand during DNA replication.

The cellular replication machinery has evolved to accurately replicate double-stranded B-form DNA while allowing mutation to occur at a low rate (9). Defined ordered sequence DNA (dosDNA), such as inverted repeats, direct repeats, mirror repeats, and homopurine homopyrimidine runs, may present an enormous challenge to the replication process (10, 13, 32, 33, 37, 49). Simple repetitive sequences in DNA, such as runs of mono-, di- or trinucleotide repeats, are dosDNA elements which provide multiple opportunities for slipped misalignment that may lead to duplication or deletion mutations (40). Alternative DNA secondary structures including cruciforms, hairpins, slipped mispaired structures, intramolecular triplex structures, and Z-DNA can form in dosDNA and may create unusual non B-DNA or nonduplex structures as replication templates (12, 38). Specific misalignments leading to base substitution, frameshift, deletion, or duplication mutations can be caused by alternative DNA secondary structures, which can form in dosDNA during replication (18, 19, 29, 33, 38, 44, 45, 50). Interstrand misalignment can also occur during replication of inverted repeats, leading to mutation (26, 32). Aberrant replication at dosDNA may be responsible for mutations that occur in many human diseases (6, 7, 17).

The concurrent model for DNA replication suggests that a dimeric DNA polymerase III holoenzyme complex simultaneously synthesizes both the leading and the lagging strands (5, 8, 24, 39). While the leading-strand polymerase can be very processive, the lagging-strand polymerase must synthesize DNA discontinuously and temporally later than synthesis of the complementary sequences in the leading strand. In support of this model, two functionally asymmetric polymerase III holoenzyme complexes can be purified from *Escherichia coli* (23, 52). Asymmetric polymerases with different binding affinities and processivities may have intrinsic differences in the frequency and/or spectra of spontaneous mutations. Moreover, during replication of the leading strand, a single-strand loop is generated in the lagging strand. The regions of single-stranded DNA, equal in length to an Okazaki fragment, provide an opportunity for the formation of intrastrand hairpins or other structures. This secondary-structure formation may be in competition with the binding of single-strand binding proteins that might prevent the formation of hairpins or other alternative structures. The free ends of the Okazaki fragments and the regions of single-stranded DNA provide a greater opportunity for misalignments stabilized by DNA secondary structure in the lagging strand than in the leading strand (43).

Differences in the frequency of mutations in the leading and lagging strands have been previously observed (14, 34, 35, 44, 47). The deletion of direct repeats stabilized by DNA secondary structure occurred preferentially in the lagging strand for one mutation insert containing an inverted repeat flanked asymmetrically by direct repeats within the chloramphenicol acetyltransferase (CAT) gene of pBR325 (44). The presence of inverted repeats can increase the frequency of deletion between flanking direct repeats. This is consistent with a model in which the formation of a hairpin, creating a three-way junction, stabilizes slippage between direct repeats and promotes replication of the misaligned progeny strand, leading to deletion (45). Deletion between direct repeats involving DNA misalignment intermediates that have an unpaired loop between the direct repeats can occur at a lower frequency than misalignment involving secondary structures (45). In an asymmetric mutation construct (called p1), the misalignment intermediates that can lead to deletion during DNA replication are different in the leading and lagging strands (44). In the lagging strand, the intermediate giving rise to deletion of the p1 insert

* Corresponding author. Mailing address: Institute of Biosciences and Technology, Texas A&M University, 2121 W. Holcombe Blvd., Houston, TX 77030-3303. Phone: (713) 677-7664. Fax: (713) 677-7689. Electronic mail address: RSinden@IBT.TAMU.edu.

† Present address: Molecular Biology Research and Development, Life Technologies, Inc., Gaithersburg, MD 20877.

involves stabilization by a hairpin forming a three-way junction, whereas an unpaired loop is involved in the misaligned replication intermediate in the leading strand. In the leading strand, the direction of replication and the position of the direct repeats with respect to the inverted repeat preclude the formation of a three-way junction. Reversal of the CAT gene within plasmid pBR325 reverses the polarity by which these sequences are replicated, such that in this reversed orientation, the hairpin-stabilized three-way junction occurs in the leading strand and the unpaired loop occurs in the lagging strand (44). When the hairpin-stabilized three-way junction occurred in the lagging strand, a 20-fold increase in the reversion frequency was observed (44).

This phenomenon of preferential deletion involving DNA secondary structure has previously been reported for only one construct (44). The present report analyzes additional DNA secondary-structure mutational inserts and presents further evidence for the role of DNA secondary structure in preferential mutagenesis in the lagging strand. We have designed asymmetric palindromic inserts in which either a hairpin-stabilized three-way junction or an unpaired loop structure can form in opposite strands. If a three-way junction can form in the lagging strand, the frequency of deletion is higher than if an unpaired loop is formed. In one construct, we observed an approximately 50-fold increase in the deletion frequency associated with the three-way junction compared with that for an unpaired loop in the lagging strand.

MATERIALS AND METHODS

Bacterial strains and media. The bacterial strains used were the following *E. coli* K-12 derivatives: HB101 [F$^-$ *hsdS20 mcrB mrr recA13 ara-14 proA2 galK2 rpsL20* (Sm^r) *xyl-5 mtl-1 supE44*], RR1 [F$^-$ *hsdS20 mcrB mrr ara-14 proA2 galK2 rpsL20* (Sm^r) *xyl-5 mtl-1 supE44*], and DH5 [F$^-$ *recA1 supE44 endA1 hsdR17* ($r_K^-m_K^+$) *thi-1 gyrA96 relA*]. Reversion assays were performed in either Luria broth (10.0 g of Bacto Tryptone, 5.0 g of yeast extract, 10.0 g of NaCl, 1.5 ml of 1 N NaOH per liter of H_2O) or K medium (M9 buffer [1.0 g of NH_4Cl, 11.0 g of Na_2HPO_4 $7H_2O$, 3 g of KH_2PO_4 per liter of H_2O], to which was added 20 ml of 10% Casamino Acids, 10 ml of 20% glucose, 0.4 ml of 0.1 M $CaCl_2$, 0.2 ml of 1 M $MgSO_4$, and 0.2 ml of 0.01% thiamine per 200 ml of M9 buffer).

Oligo-deoxyribonucleotide synthesis, fragment preparation, and cloning. The mutation inserts to be cloned (Fig. 1) were chemically synthesized by using β-phosphoramidite chemistry. The oligonucleotides were purified either on a 20% denaturing polyacrylamide gel as previously described (45) or on an Oligonucleotide Purification Cartridge (Applied Biosystems) according to the manufacturer's instructions. The F14C inverted repeat has been described previously (53).

Oligonucleotides were phosphorylated with [γ-^{32}P]ATP (NEN-Dupont) as described previously (46). Complementary strands were annealed by slowly cooling in TEN buffer as detailed elsewhere (45) or by incubation at 88°C for 2 min, at 65°C for 10 min, at 37°C for 10 min, and then at room temperature (22 to 24°C) for 5 min in annealing buffer (67 mM Tris-HCl [pH 7.6], 13 mM $MgCl_2$, 6.7 mM dithiothreitol, 13 mM spermidine, 1.3 mM EDTA [2a]). Duplex fragments were purified from acrylamide gels and were cloned into the *Eco*RI site of pBR325 (45).

Plasmid pBR325 contains three genes conferring antibiotic resistance to tetracycline, ampicillin, and chloramphenicol (4, 30). The synthetic oligonucleotides were designed such that cloning them into the unique *Eco*RI site in the CAT gene disrupted the reading frame. Bacteria harboring the plasmid containing an intact insert were chloramphenicol sensitive (Cm^s). In-frame deletions of the insert restored chloramphenicol resistance (Cm^r). Because of the unidirectional ColE1 origin of replication in pBR325 (16, 21, 41, 42), assignment of the leading and lagging strands is possible.

pBR523 derivatives, in which the orientation of the CAT gene was reversed, were created by digesting the plasmid with *Csp*45, which cuts on either side of the CAT gene, and by ligating the fragments (44). The orientation of the CAT gene was screened by analysis of the different electrophoretic migrations of *Pvu*II digestion products in 0.8% agarose.

Reversion frequencies and analysis of revertants. Cm^r reversion frequencies were determined as described elsewhere (44, 45) from cultures growing in exponential phase in K medium. Overnight cultures were started from a single colony grown on Luria broth plates, and this culture was used to grow an exponentially dividing culture for reversion analysis. For constructs with very high reversion frequencies, cultures were typically inoculated from a frozen culture. The frequency of reversion is the mean of no fewer than six independent experiments. Although variations in reversion frequencies were observed in these mutational analyses as expected (22), the individual reversion frequencies used for the calculation of the Cm^r reversion frequencies typically varied by less than a factor of 2 to 4. "Jackpot" frequencies that were much higher than the other values were occasionally observed but were not included in the calculation of the mean reversion frequency. Following purification of DNA from cells transformed with DNA from the original Cm^r revertant, the Cm^r reversion events were analyzed by analysis of *Alu*I restriction digests and DNA sequence analysis (45).

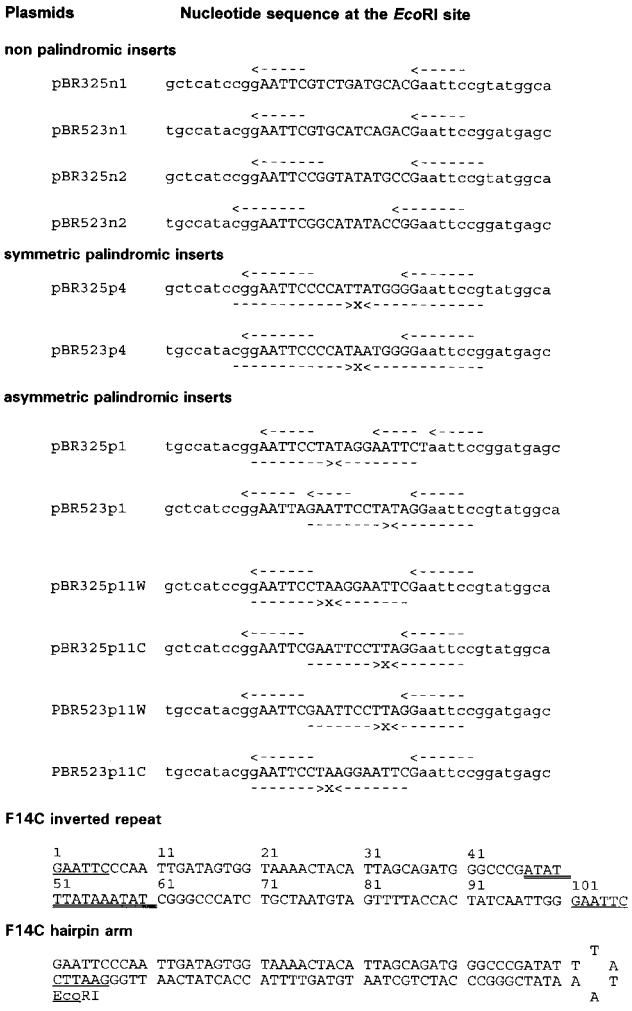

FIG. 1. Nucleotide sequence around the *Eco*RI site in the CAT gene of pBR325 and pBR523 derivatives containing the deletion inserts. The sequences of the mutation inserts are shown in uppercase letters, while the surrounding CAT gene sequences are in lowercase letters. Arrows above the sequences indicate the direct repeats, while inverted arrows below the sequences indicate the positions of the inverted repeats. x's at the center of the inverted repeat arrows in the p11 and p4 sequences indicate the central bases of these 17-bp inverted repeats. The sequence of the F14C inverted repeat and the structure of the hairpin arm are shown. The *Eco*RI sites, which are used for cloning into pBR325, are underlined. The 14-bp A+T-rich center is double underlined.

RESULTS

Symmetrical and asymmetrical mutation inserts for detecting differences in mutagenesis between leading and lagging strands. Symmetrical and asymmetrical 17- and 18-bp mutation inserts have previously been synthesized and cloned into the *Eco*RI site within the CAT gene of pBR325, and their spontaneous frequencies of reversion to Cm^r have been measured (45). Cm^r reversion can occur by deletion of the insert

FIG. 2. Potential structural intermediates involved in deletion of the p11W insert in pBR325. (A) Potential events leading to deletion of the p11 insert in the leading strand of pBR325p11W. 1, DNA sequence surrounding the *Eco*RI site with direct repeats and inverted repeats shown as described in the legend to Fig. 1. The arrows over the direct repeats indicate the directions of replication of the leading strand. 2a, synthesis occurring through the first copy of the direct repeat up to the inverted repeat in its linear form; 2b, potential structure if the inverted repeat forms a hairpin in the template strand; 3, misalignment of the first direct repeat in the progeny strand with the second copy of the direct repeat in the template strand would necessitate a disruption of the hairpin shown in 2b. This would result in formation of a loop consisting of the sequences between the direct repeats in the template strand. (B) Potential events leading deletion of the p11 insert in the lagging strand of pBR325p11W. 1, sequence of the mutation insert, with arrows indicating the direct repeats and the direction of replication; 2, replication occurring through the first copy of the direct repeat in the lagging template strand; 3, misalignment of the first copy of the direct repeat in the lagging strand with the second copy of the direct repeat in the template strand could be stabilized by the formation of a perfect three-way junction formed by hairpin formation within the inverted repeat in the lagging template strand. Continued replication of this intermediate or structure no. 3 in panel A would lead to precise deletion of the mutation insert in the nascent strand.

mediated by slipped misalignment between flanking direct repeats during DNA replication or by other frameshift mutations that restore the reading frame of the CAT gene. Symmetrical mutation inserts, which include n1, n2, and p4, are those in which the structural intermediate that could lead to misalignment-mediated mutagenesis is the same in both the leading and the lagging strands. In the n1 and n2 nonpalindromic sequences, the direct repeats provide the only symmetry element involved in misalignment, while in the p4 palindromic insert, the direct repeats constitute part of an inverted repeat. (Although the DNA sequence between the direct repeats in n1 and n2 is asymmetric, it forms an unpaired loop during misalignment, and, although the sequences in the loop of the leading or lagging strand are complementary and thus different, the sequence itself does not contribute to the misalignment). In the asymmetric p1 and p11 inserts, an inverted repeat is positioned asymmetrically between two flanking direct repeats. Thus, during misalignment of the direct repeats in one strand, a stable hairpin can form creating a three-way junction, while misalignment in the other strand involves formation of an unpaired loop (Fig. 2; see Fig. 3 in reference 44).

The Cm^r reversion frequencies measured for the n1, n2, p1, and p4 mutation inserts are listed in Table 1. For the symmetrical n1, n2, and p4 inserts, an approximately twofold difference in reversion frequencies was observed upon reversal of the CAT gene (Table 1). Analysis of the nature of the Cm^r reversion event showed that the frequencies of deletion compared with other mutational events for the n1, n2, and p4 inserts were similar for either orientation (Table 1). For example, for the n1 insert, 58 and 50% of the total revertants for the pBR325 and pBR523 constructs were precise deletions, respectively. The percentages of deletions for the n2 and p4 inserts were also similar when the CAT gene was reversed. As reported previously, when the p1 insert was reversed, the Cm^r reversion frequency decreased about 20-fold (44). The molecular events responsible for reversion of the p1 insert were dependent on the orientation of the CAT gene. For the p1 insert in pBR325p1, 100% (12 of 12) of the Cm^r revertants

TABLE 1. Effect of reversal of the CAT gene on the Cm^r reversion frequency in the n and p series of plasmids

Insert	Cm^r reversion frequency (10^{-9})[a]	No. of occurrences of the following reversion events:			Deletion frequency (10^{-9})
		Deletion	Duplication	Other	
pBR325					
n1	3.7 ± 0.9	14	0	10	2.2
n2	129 ± 23[b]	9	0	3	97
p1	44 ± 20	12	0	0	44
p4	36 ± 5	24	0	0	36
pBR523					
n1	2.7 ± 0.9	5	1	4	1.4
n2	79 ± 1.7[b]	9	0	2	65
p1	2.0 ± 0.9	7	0	11	0.8
p4	18 ± 4.0	10	0	2	15

[a] Values are the arithmetic means ($\bar{x}$) ± the standard deviations.

[b] In this series of experiments, Cm^r reversion frequency for the n2 insert was higher than that reported previously (45).

TABLE 2. Effect of reversal of the CAT gene on the Cm^r reversion frequency in the p11 series of plasmids

Insert	Cm^r reversion frequency (10^{-9})[a]	No. of occurrences of the following reversion events:		% Deletions	Deletion frequency (10^{-9})
		Deletion	Other		
pBR325					
P11W	240 ± 6	23	1	97.6	234
P11 C	45 ± 13	21	7	78.8	35
pBR523					
P11W	25 ± 7	38	3	91.4	23
P11C	120 ± 9	21	2	93.5	112

[a] Values are arithmetic means ($\bar{x}$) ± standard deviations.

TABLE 3. Effect of reversal of the CAT gene on the Cm^r reversion frequency in plasmids containing the F14C inverted repeat

Insert (strain)	Cm^r reversion frequency (10^{-9})[a]	No. of occurrences of the following reversion events:	
		Deletion	Other
pBR325			
F14C (HB101)	239 ± 3	20	0
F14C (RR1)	363 ± 93	10	0
pBR523			
F14C (HB101)	579 ± 300	20	0
F14C (RR1)	406 ± 91	10	0

[a] Values are arithmetic means ($\bar{x}$) ± standard deviations.

were precise deletions. However, when the CAT gene was reversed in pBR523p1, only 39% (7 of 18) of the Cm^r revertants were precise deletions, which is similar to the result for the n1 insert. (The inherent differences in the deletion frequencies of the n1, n2, p1, and p4 inserts in pBR325 may be due to the different misaligned intermediates that can form during replication, as discussed elsewhere [45]).

Analysis of deletion frequencies in the p11 series inserts. The p11 insert was designed to extend the observation that an asymmetric palindromic insert, which can form different misaligned structural intermediates in the leading or the lagging strand, can have a differential deletion frequency depending on the polarity of replication caused by the orientation of the CAT gene. The p11 insert is similar to p1 but contains several important differences. p11 is 17 rather than 18 bp and has 7- rather than 6-bp direct repeats (Fig. 1). In addition, upon slipped misalignment in the lagging strand of pBR325p11W, a perfect three-way junction can form (Fig. 2), compared with the p1 misalignment intermediate that contains a 1-bp mismatch at the base of the three-way junction (44). Moreover, the p11 insert can be cloned in both directions (W or C), placing the three-way junction in either the leading or the lagging strand in both pBR325 and pBR523. Therefore, if misalignment between direct repeats is stabilized by a three-way junction and occurs preferentially in the lagging strand, then the frequency of deletion should be higher in pBR325p11W than in pBR523p11W and higher in pBR523 p11C than in pBR325p11C. This expected pattern of Cm^r reversion frequencies was observed for the p11 series inserts (Table 2). When the three-way junction can form in the lagging strand of pBR325p11W, >97% of the Cm^r revertants analyzed were deletions. Greater than 90% of the pBR523p11W and pBR523p11C Cm^r revertants contained precise deletions, while only 79% of the pBR325p11C revertants contained precise deletions. The deletion frequency of pBR325p11W was 10.1-fold greater than that for pBR523p11W, and that for pBR523p11C was 3.2-fold greater than that for pBR325p11C.

Effect of reversing the CAT gene on the deletion frequency of a 106-bp inverted repeat. F14C is a 106-bp perfect inverted repeat designed with a 14-bp A+T-rich center (53) (Fig. 1). This inverted repeat is the most active cruciform-forming sequence that we have made, which forms cruciforms very rapidly and at low superhelical densities ($\sigma = -0.038$). Moreover, we have shown that this sequence can exist as a cruciform in >50% of plasmid topoisomers in living *E. coli* RS2 (*topA10*), and that the formation of cruciforms leads to an increase in the deletion frequency (38, 53). Within the CAT gene in HB101, the F14C inverted repeat exists as a cruciform at low levels (≈0.2%) in plasmids in living *E. coli* cells (53). The presence of the cruciform favors the deletion of this sequence by a mechanism that may involve slipped mispairing stabilized by a long hairpin stem (38). A 2.4-fold difference in the deletion frequency of F14C inverted repeats was observed in HB101 cells upon reversal of the CAT gene (Table 3). However, in contrast to the situation with the 17-bp inserts, the frequency of deletion was higher in the reverse orientation (pBR523F14C) in RecA$^-$ strain HB101 than in the natural orientation. In strain RR1, which is a RecA$^+$ derivative of HB101, there may be no difference in the deletion frequencies between pBR325F14C and pBR523F14C. Only precise deletions were observed in either orientation.

DISCUSSION

Reversal of the CAT gene containing the asymmetric palindromic p1 or p11 insert in pBR325 reverses the misaligned intermediates (Fig. 2) that can form in the leading or lagging strand (Fig. 3). As shown previously, when the CAT gene containing the p1 insert was reversed, a 20-fold reduction in the Cm^r reversion frequency was observed (44). This was consistent with slippage stabilized by the hairpin occurring preferentially in the lagging strand (Fig. 3). The data in Table 1 extend this observation, showing that the actual deletion frequency for the p1 insert in the lagging strand compared with that for the leading strand is more than 50-fold. Both the frequencies of Cm^r reversion and the distribution of deletions and other mutational events were similar in pBR523p1, pBR325n1, and pBR523n1 (Table 1). Deletion in pBR523p1 may occur through the formation of an unpaired DNA loop between the misaligned direct repeats that is similar to that for the misaligned intermediates in pBR523n1 and pBR325n1 (44).

Since the dramatic differences in the frequency and spectrum of Cm^r reversion upon reversal of the CAT gene have been observed only with the p1 construct, the p11 insert was designed to further characterize this phenomenon. The p11 insert, like p1, has direct repeats that are asymmetric with respect to the inverted repeat (Fig. 1). The misalignment stabilized by a three-way junction can occur in either the leading or the lagging strand, depending on the orientation of the insert and the orientation of the CAT gene. Therefore, with the p11 insert, it is possible to analyze the frequency of deletion in four constructs, i.e., those with the insert in either orientation in both pBR325 and pBR523. The two constructs in which the misalignment could be stabilized by a three-way junction in the lagging strand, pBR325p11W and pBR523p11C, had higher reversion frequencies than the other two constructs in which an unpaired loop would form during

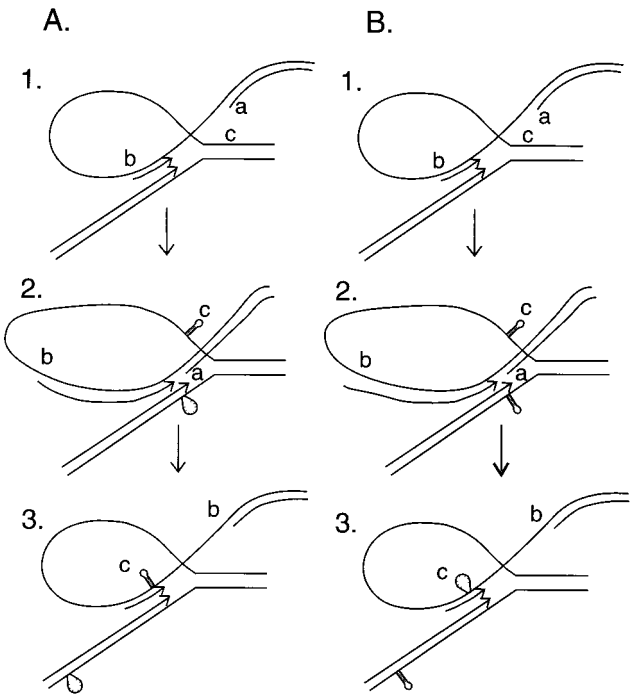

FIG. 3. Reversing the CAT gene changes the strand in which a slipped misalignment stabilized by DNA secondary structure occurs. (A) Possible misalignment events occurring in the leading or lagging strand in pBR325p11W. 1, during concurrent DNA replication, the inverted repeat, present at c, will be replicated first in the leading strand. a and b mark the ends of two Okazaki fragments. 2, during synthesis of the inverted repeat in the leading strand, a misalignment involving an unpaired loop is indicated to demonstrate the misalignment that would lead to deletion in this strand (see Fig. 2). The inverted repeat in the lagging strand is single stranded and has the potential to form a hairpin structure as shown. 3, following completion of one Okazaki fragment and reinitiation of a new Okazaki fragment, the inverted repeat is replicated in the lagging strand. Slipped misalignment of the direct repeats in the p11 mutation insert could be stabilized by a hairpin forming a perfect three-way junction as shown in Fig. 2. (B) With the CAT gene in the reversed orientation in pBR523p11W, the slipped misaligned structures that can form in the leading strand now form in the lagging strand. Misalignment stabilized by the hairpin would occur in the leading strand (structure no. 2), while the misalignment containing the unpaired loop would occur in the lagging strand (structure no. 3). Although the inverted repeat could form a hairpin in the lagging strand as shown, the misaligned intermediate necessitates that a loop be formed, as shown in Fig. 2. Since misalignment stabilized by a hairpin can increase the frequency of deletion compared with a loop (45), if misalignment during replication occurred preferentially in either the leading or the lagging strand, then one orientation of the CAT gene should exhibit a higher deletion frequency when the hairpin-stabilized three-way junction exists in the strand in which the preferential deletion occurs (44). If the case of the p1 and p11 mutation inserts, preferential deletion occurs when the three-way junction occurs in the lagging strand.

misalignment, giving rise to deletion in the lagging strand. Rather than the 50-fold difference observed for p1, the preferential deletion frequencies for the p11 insert were 10.1 and 3.2 for the two orientations of the insert in the CAT gene. These results confirm and support the previous conclusion that slipped misalignment stabilized by DNA secondary structure leads to deletion and can occur preferentially in the lagging strand during DNA replication (44).

The results with the p11 mutation insert are significant since they demonstrate that the preferential deletion associated with formation of a three-way junction in the lagging strand is independent of the orientation of the CAT gene, of the DNA sequence flanking the inverted repeat, and of the occurrence of the three-way junction in the transcribed or nontranscribed strand. With the p1 insert, the formal possibility exists that the preferential deletion of a hairpin-stabilized three-way junction is dependent on the orientation of the CAT gene with respect to the direction of replication. While reversal of the CAT gene preserves the relationship of the mutation insert to the flanking DNA sequence, there is an effect on the mutation process, as evidenced from the general 2-fold decrease in reversion frequencies observed with the asymmetrical inserts. When this is taken into account, there is a 5- to 6.4-fold preference for deletion in pBR325p11W and pBR523p11C. The preferential deletions in pBR325p11W and pBR523p11C, therefore, indicate the dependence on placing the three-way junction in the lagging strand and demonstrate that the orientation of the CAT gene is not a major determinant. Reversing the p11 insert within the CAT gene of either pBR325 or pBR523 changes the sequence context flanking the insert. Although this will likely influence the three-dimensional structure of the three-way junction (28) and thus influence its interaction with DNA polymerase, a preference for deletion of the three-way junction in the lagging strand was still apparent when pBR325p11W was compared with pBR325p11C and when pBR523p11C was compared with pBR523p11W (Table 2). Since the preferential deletion associated with the three-way junction is not limited to the transcribed or nontranscribed strand, this phenomenon is not likely the result of transcriptional coupled DNA repair events (15, 25).

The data in Table 1 confirm the initial observation that mutation inserts without an inverted repeat do not show large differences in Cm^r reversion frequencies upon reversal of the CAT gene. Reversal of the CAT gene which contained the n1 or n2 nonpalindromic inserts resulted in, at most, a twofold difference in the Cm^r reversion frequencies, and the percentages of deletions relative to other mutations were not markedly different. Although the p4 insert contains an inverted repeat, it is a symmetric palindromic insert, and reversal of the CAT gene does not change the misaligned intermediates that can form in either the leading or the lagging strand (44, 45). The results with the p4 insert demonstrate that there is little difference in reversion frequencies when misalignment can be stabilized by a three-way junction in both the leading and the lagging strands.

The deletion frequencies observed for the p11 series constructs were significantly higher than those observed for corresponding deletion events involving hairpin or unpaired loop structures for the n1, p1, and p4 inserts (Table 1). There may be three reasons for this. First, the hairpin-stabilized misalignment intermediate in p11 consists of a perfect three-way junction which may promote misalignment and deletion at higher frequency. In contrast, the DNA secondary-structure intermediate involved in deletion of the p1 construct contained a G A mismatch at the base of the three-way junction (44). Second, the p11 insert contains 7-bp direct repeats compared with 6-bp direct repeats for the p1 and n1 insert. While our previous analysis did not indicate a significant effect of the lengths of direct repeats (between 5 and 8 bp) on the deletion frequencies (45), Pierce et al. (29) demonstrated a logarithmic relationship between the length of the direct repeat and deletion of sequences between direct repeats. Thus, the longer direct repeats may increase the deletion frequency. Third, the particular structures of the p11 three-way junctions, when the insert is in either orientation, may be more conducive to extension by DNA polymerase than that for the p1 insert. One explanation for the smaller difference in deletion frequencies between the three-way junction and unpaired loop intermediate is that the 7-bp direct repeats may significantly increase the deletion frequency of the misaligned intermediates containing either the

unpaired loop or the three-way junction, in part lessening the stabilizing influence of the three-way junction.

The molecular intermediates that lead to deletion in the leading and lagging strands are similar or identical in the symmetrical n1, n2, p4, and F14C inserts. Consequently, there should be little or no difference in the deletion frequencies. However, deletion frequencies of the short n1, n2, and p4 constructs were reduced about twofold upon reversal of the gene from its original orientation. The relative orientations of transcription and replication may be responsible for the difference in deletion frequencies in these short inserts. In pBR325, the direction of transcription and leading-strand replication are toward each other. In pBR523, replication, which is about 10-fold faster than transcription, and transcription occur in the same direction. In pBR325, the replication fork may be slowed by an oncoming transcription complex, while in pBR523, the DNA polymerase complex may proceed at a normal rate (11). The higher deletion frequency in the n1, n2, and p4 inserts in the natural orientation in pBR325 may be the result of the slower progression and/or transient pausing of the replication fork. A lower rate of replication may increase the probability that secondary structures may form in the lagging strand and be responsible for the increased mutation frequency. Transient pausing can also promote mutagenesis as demonstrated by a number of investigators (1–3, 18, 27). Moreover, the presence or absence of transcriptional activity can influence the end points of deletion (48). Therefore, transcription-induced pausing or slowing of replication may lead to an increased probability for slipped misalignment or, in the case of p4, DNA secondary formation in the lagging (or leading) strand.

The results for reversal of the CAT gene containing the F14C inverted repeat in HB101 were the opposite of those for the short constructs, in that lower deletion frequencies were observed for pBR325. An approximately 2.4-fold increase in deletion frequency was observed upon reversal of the CAT gene containing the F14C inverted repeat in HB101. The F14C inverted repeat can form a stable cruciform structure in living cells that provides a substrate for mutagenesis (38, 53). Slipped mispairing may occur between flanking direct repeats stabilized by a hairpin with a short unpaired loop at the base of the hairpin stem (see Fig. 5 of reference 38). A slightly lower deletion frequency in pBR325 might result from a reduction in negative supercoiling between the convergent macromolecular complexes, which would reduce the probability of cruciform formation (20, 31, 53). In pBR523, the DNA between the transcription and replication complexes (which are moving in the same direction) would not lose as much negative supercoiling as in the convergent arrangement. Therefore, more cruciforms may exist in pBR523F14C than in pBR325F14C. This differential in supercoiling may not be important for deletion of the short inserts since they do not likely form stable supercoil-dependent cruciform structures in DNA.

While our results have been interpreted in terms of asymmetric replication errors, we have no direct proof that replication is involved. It is possible that genetic recombination or DNA repair (which would likely involve repair replication) formally result in the observed mutations. However, we show here and previously (38) that mutations in *recA* do not influence the frequency of deletion of long inverted repeats. An asymmetry in the frequency and spectrum of certain spontaneous mutations, which is dependent on the occurrence of a strand that can form stable secondary structures in the leading or the lagging strand, is clearly demonstrated. The results with the p11 constructs, in which a lagging-strand preference was observed for both the transcribed and the nontranscribed strands, exclude the possibility that transcription-coupled repair is responsible for the observed asymmetry.

The results presented here provide further evidence that spontaneous mutations can show distinct leading- and lagging-strand preferences. A number of investigators have presented results that can also be interpreted in terms of leading- or lagging-strand replication errors. Weston-Hafer and Berg (51) did not observe a leading- or lagging-strand preference in the deletion of inverted repeats associated with direct repeats. However, their constructs contained a number of differences compared with the ones studied here, including the inability to form a reasonable three-way junction. The results described by Gordenin et al. (14), who studied Tn*5* excision in polymerase mutants of *Saccharomyces cerevisiae*, are consistent with the hypothesis that an increased opportunity for secondary-structure formation in the lagging strand increases the frequency of mutation. Roberts et al. (34, 35) have found leading- or lagging-strand asymmetries during replication in an in vitro HeLa cell system. A 20-fold preference for the introduction of frameshift mutations in the lagging strand by *n*-2-acetylaminofluorene adducts to guanine has been reported elsewhere (47). Recently, we have demonstrated that the correction of a quasipalindrome to a perfect palindrome occurs with about a 10-fold preference in the leading strand (36). This leading-strand preference is due to an interstrand switch during replication that is likely limited to the leading strand for physical reasons.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant ES-05508 from the National Institutes of Environmental Health Sciences.

We thank Adam Jaworski, M. Angela Parsons, and Christopher E. Pearson for reading the manuscript. We thank Shannon F. Kramer for technical assistance.

REFERENCES

1. **Bebenek, K., J. Abbotts, J. D. Roberts, S. H. Wilson, and T. A. Kunkel.** 1989. Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. J. Biol. Chem. **264:**16948–16956.
2. **Bebenek, K., J. Abbotts, S. H. Wilson, and T. A. Kunkel.** 1993. Error-prone polymerization by HIV-1 reverse transcriptase. J. Biol. Chem. **268:**10324–10334.

2a. **Becker, M.** Personal communication.

3. **Bierne, H., S. D. Ehrlich, and B. Michel.** 1991. The replication termination signal terB of the Escherichia coli chromosome is a deletion hot spot. EMBO J. **10:**2699–2705.
4. **Bolivar, F.** 1978. Construction and characterization of new cloning vehicles. III. Derivatives of plasmid pBR322 carrying unique EcoRI sites for selection of EcoRI generated recombinant DNA molecules. Gene **4:**121–136.
5. **Cha, T. A., and B. M. Alberts.** 1990. Effects of the bacteriophage T4 gene 41 and gene 32 proteins on RNA primer synthesis: coupling of leading and lagging strand synthesis at a replication fork. Biochemistry **29:**1791–1798.
6. **Cooper, D. N., and M. Krawczak.** 1990. The mutational spectrum of single base-pair substitutions causing human genetic disease: patterns and predictions. Hum. Genet. **85:**55–74.
7. **Cooper, D. N., and M. Krawczak.** 1991. Mechanisms of insertional mutagenesis in human genes causing genetic disease. Hum. Genet. **87:**409–415.
8. **Debyser, Z., S. Tabor, and C. C. Richardson.** 1994. Coordination of leading and lagging strand DNA synthesis at the replication fork of bacteriophage T7. Cell **77:**157–166.
9. **Drake, J. W.** 1991. A constant rate of spontaneous mutation in DNA-based microbes. Proc. Natl. Acad. Sci. USA **88:**7160–7164.
10. **Drake, J. W., B. W. Glickman, and L. S. Ripley.** 1983. Updating the theory of mutation. Am. Sci. **71:**621–630.
11. **French, S.** 1992. Consequences of replication fork movement through transcription units in vivo. Science **258:**1362–1365.
12. **Freund, A. M., M. Bichara, and R. P. Fuchs.** 1989. Z-DNA-forming sequences are spontaneous deletion hot spots. Proc. Natl. Acad. Sci. USA **86:**7465–7469.
13. **Glickman, B. W., and L. S. Ripley.** 1984. Structural intermediates of deletion mutagenesis: a role for palindromic DNA. Proc. Natl. Acad. Sci. USA **81:**512–516.
14. **Gordenin, D. A., A. L. Malkova, A. Peterzen, U. N. Kulikov, Y. I. Paviov, E.**

Perkins, and M. A. Resnick. 1992. Transposon Tn5 excision in yeast: influence of DNA polymerase alpha, delta, and epsilon and repair genes. Proc. Natl. Acad. Sci. USA **89:**3785–3789.
15. **Hanawalt, P. C.** 1994. Transcription-coupled repair and human disease. Science **266:**1957–1958.
16. **Inselburg, J.** 1974. Replication of colicin E1 plasmid DNA in minicells from a unique replication initiation site. Proc. Natl. Acad. Sci. USA **71:**2256–2259.
17. **Krawczak, M., and D. N. Cooper.** 1991. Gene deletions causing human genetic disease: mechanisms of mutagenesis and the role of the local DNA sequence environment. Hum. Genet. **86:**425–441.
18. **Kunkel, T. A.** 1990. Misalignment-mediated DNA synthesis errors. Biochemistry **29:**8003–8011.
19. **Kunkel, T. A.** 1992. DNA replication fidelity. J. Biol. Chem. **267:**18251–18254.
20. **Liu, L. F., and J. C. Wang.** 1987. Supercoiling of the DNA template during transcription. Proc. Natl. Acad. Sci. USA **84:**7024–7027.
21. **Lovett, M. A., L. Katz, and D. R. Helinski.** 1974. Unidirectional replication of plasmid ColE1 DNA. Nature (London) **251:**337–340.
22. **Luria, S. E., and M. Delbruck.** 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics **28:**491–511.
23. **Maki, H., S. Maki, and A. Kornberg.** 1988. DNA polymerase III holoenzyme of Escherichia coli. IV. The holoenzyme is an asymmetric dimer with twin active sites. J. Biol. Chem. **263:**6570–6578.
24. **McHenry, C. S.** 1988. DNA polymerase III holoenzyme of Escherichia coli. Annu. Rev. Biochem. **57:**519–550.
25. **Mellon, I., and P. C. Hanawalt.** 1989. Induction of the Escherichia coli lactose operon selectively increases repair of its transcribed DNA strand. Nature (London) **342:**95–98.
26. **Ohshima, A., S. Inouye, and M. Inouye.** 1992. In vivo duplication of genetic elements by the formation stem-loop DNA without an RNA intermediate. Proc. Natl. Acad. Sci. USA **89:**1016–1020.
27. **Papanicolaou, C., and L. S. Ripley.** 1991. An in vitro approach to identifying specificity determinants of mutagenesis mediated by DNA misalignments. J. Mol. Biol. **221:**805–821.
28. **Pearson, C. E., M. T. Ruiz, G. B. Price, and M. Zannis-Hadjopoulos.** 1994. Cruciform DNA binding protein in HeLa cell extracts. Biochemistry **33:**14185–14196.
29. **Pierce, J. C., D. Kong, and W. Masker.** 1991. The effect of the length of direct repeats and the presence of palindromes on deletion between directly repeated DNA sequences in bacteriophage T7. Nucleic Acids Res. **14:**3901–3905.
30. **Prentski, P., F. Karch, S. Iida, and J. Meyer.** 1981. The plasmid cloning vector pBR325 contains a 482 base-pair-long inverted duplication. Gene **14:**289–299.
31. **Rahmouni, A. R., and R. D. Wells.** 1992. Direct evidence for the effect of transcription on local DNA supercoiling in vivo. J. Mol. Biol. **223:**131–144.
32. **Ripley, L. S.** 1982. Model for the participation of quasi-palindromic DNA sequences in frameshift mutation. Proc. Natl. Acad. Sci. USA **79:**4128–4132.
33. **Ripley, L. S.** 1990. Frameshift mutation: determinants of specificity. Annu. Rev. Genet. **24:**189–213.
34. **Roberts, J. D., D. Nguyen, and T. A. Kunkel.** 1993. Frameshift fidelity during replication of double-stranded DNA in HeLa cell extracts. Biochemistry **32:**4083–4090.
35. **Roberts, J. D., D. C. Thomas, and T. A. Kunkel.** 1991. Exonucleolytic proofreading of leading and lagging strand DNA replication errors. Proc. Natl. Acad. Sci. USA **88:**3465–3469.
36. **Rosche, W. A., T. Q. Trinh, and R. R. Sinden.** 1995. Unpublished data.
37. **Sinden, R. R., and R. D. Wells.** 1992. DNA structure, mutations, and human genetic disease. Curr. Opin. Biotech. **3:**612–622.
38. **Sinden, R. R., G. Zheng, R. G. Brankamp, and K. N. Allen.** 1991. On the deletion of inverted repeated DNA in Escherichia coli: effects of length, thermal stability, and cruciform formation in vivo. Genetics **129:**991–1005.
39. **Sinha, N. K., C. F. Morris, and B. M. Alberts.** 1980. Efficient in vitro replication of double-stranded DNA templates by a purified T4 bacteriophage replication system. J. Mol. Biol. **255:**4290–4303.
40. **Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita, E. Terzaghi, and M. Inouye.** 1966. Frameshift mutations and the genetic code. Cold Spring Harbor Symp. Quant. Biol. **31:**77–84.
41. **Sutcliffe, J. G.** 1978. Complete nucleotide sequence of the Escherichia coli plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. **43:**77–90.
42. **Tomizawa, J.-I., Y. Sakakibara, and T. Kakefuda.** 1974. Replication of colicin E1 plasmid DNA in cell extracts. Origin and direction of replication. Proc. Natl. Acad. Sci. USA **71:**2260–2264.
43. **Trinh, T. Q.** 1991. The role of DNA secondary structure in deletion and duplication between direct repeats in E. coli: differential mutagenesis between the leading and lagging strand. Ph.D. Dissertation, University of Cincinnati, Cincinnati, Ohio.
44. **Trinh, T. Q., and R. R. Sinden.** 1991. Preferential DNA secondary structure mutagenesis in the lagging strand of replication in *E. coli*. Nature (London) **352:**544–547.
45. **Trinh, T. Q., and R. R. Sinden.** 1993. The influence of primary and secondary DNA structure in deletion and duplication between direct repeats in Escherichia coli. Genetics **134:**409–422.
46. **Ussery, D. W., R. W. Hoepfner, and R. R. Sinden.** 1992. Probing DNA structure with psoralen in vitro. Methods Enzymol. **212:**242–262.
47. **Veaute, X., and R. P. Fuchs.** 1993. Greater susceptibility to mutations in lagging strand of DNA replication in Escherichia coli than in leading strand. Science **261:**598–600.
48. **Vilette, D., M. Uzest, S. D. Ehrlich, and B. Michel.** 1992. DNA transcription and repressor binding affect deletion formation in Escherichia coli plasmids. EMBO J. **11:**3629–3634.
49. **Wells, R. D., and R. R. Sinden.** 1993. Defined ordered sequence DNA, DNA structure, and DNA-directed mutation. Genome Anal. **7:**107–138.
50. **Weston-Hafer, K., and D. E. Berg.** 1989. Palindromy and the location of deletion endpoints in Escherichia coli. Genetics **121:**651–658.
51. **Weston-Hafer, K., and D. E. Berg.** 1991. Deletions in plasmid pBR322: replication slippage involving leading and lagging strands. Genetics **127:**649–655.
52. **Wu, C. A., E. L. Zechner, A. J. Hughes, Jr., M. A. Franden, C. S. McHenry, and K. J. Marians.** 1992. Coordinated leading- and lagging-strand synthesis at the Escherichia coli DNA replication fork. J. Biol. Chem. **267:**4064–4073.
53. **Zheng, G., T. Kochel, R. W. Hoepfner, S. E. Timmons, and R. R. Sinden.** 1991. Torsionally tuned cruciform and Z-DNA probes for measuring unrestrained supercoiling at specific sites in DNA of living cells. J. Mol. Biol. **221:**107–129.