Evidence for Two Mechanisms of Palindrome-Stimulated Deletion in Escherichia coli: Single-Strand Annealing and Replication Slipped Mispairing

Malgorzata Bzymek¹ and Susan T. Lovett

Department of Biology and Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02454-0110

Manuscript received January 31, 2001
Accepted for publication March 19, 2001

ABSTRACT

Spontaneous deletion mutations often occur at short direct repeats that flank inverted repeat sequences. Inverted repeats may initiate genetic rearrangements by formation of hairpin secondary structures that block DNA polymerases or are processed by structure-specific endonucleases. We have investigated the ability of inverted repeat sequences to stimulate deletion of flanking direct repeats in Escherichia coli. Propensity for cruciform extrusion in duplex DNA correlated with stimulation of flanking deletion, which was partially sbcd dependent. We propose two mechanisms for palindrome-stimulated deletion, SbcCD dependent and SbcCD independent. The SbcCD-dependent mechanism is initiated by SbcCD cleavage of cruciforms in duplex DNA followed by RecA-independent single-strand annealing at the flanking direct repeats, generating a deletion. Analysis of deletion endpoints is consistent with this model. We propose that the SbcCD-independent pathway involves replication slipped mispairing, evoked from stalling at hairpin structures formed on the single-stranded lagging-strand template. The skew of SbcCD-independent deletion endpoints with respect to the direction of replication supports this hypothesis. Surprisingly, even in the absence of palindromes, SbcD affected the location of deletion endpoints, suggesting that SbcCD-mediated strand processing may also accompany deletion unassociated with secondary structures.

In vivo, large DNA palindromes are intrinsically unstable sequences (reviewed in Leach 1994) and are selectively removed from the genome. Large inverted repeats are also unstable in the yeast Saccharomyces cerevisiae (Gordenin et al. 1993). In addition, spontaneously occurring genetic rearrangements in Escherichia coli are often associated with inverted repeat sequences (Galas 1978; Glickman and Ripley 1984). Systematic analysis in Escherichia coli shows that inverted repeats stimulate deletion at nearby direct repeat sequences (Foster et al. 1981; Albertini et al. 1982; Glickman and Ripley 1984; Singer and Westlye 1988; Weston-Hafer and Berg 1989; Pierce et al. 1991; Sinden et al. 1991).

The potential to form DNA secondary structures is the basis of sequence instability at palindromes. Two types of structures can be formed: intrastrand pairing at palindromes results in hairpin formation in single-stranded DNA (ssDNA) molecules and cruciform extrusion from double-stranded DNA (dsDNA; Figure 1). When DNA becomes single stranded, as during replication or repair, formation of hairpin structures at inverted repeats is favored. In contrast, there is a kinetic barrier to cruciform formation from dsDNA (Courey and Wang 1983). The pathway for cruciform formation is believed to initiate with unpairing of the duplex, followed by nucleation of base pairing within the arm structures that can extend by branch migration (Figure 1; Sullivan and Lilley 1986). Crosslinking studies have detected structures consistent with cruciform secondary structures formed by inverted repeat sequences in vivo (Zheng et al. 1991). Genetic experiments also argue that cruciforms do form in vivo (Davison and Leach 1994).

Processing of large palindromic DNA sequences in Escherichia coli is mediated by the products of the sbcC and sbcd genes (Leach 1994). The sbcd genes were originally identified as antirecombination factors (Lloyd and Buckman 1985) and were later shown to promote palindrome-associated inviability (Chalker et al. 1988; Gibson et al. 1992). In vitro, SbcCD possesses ssDNA endonuclease and dsDNA exonuclease activities and can cleave hairpin structures (Connelly and Leach 1996; Connelly et al. 1997, 1998, 1999). In vivo, SbcCD appears to introduce double-strand breaks at long palindromic sequences; after cleavage, chromosomal integrity can be restored by RecA-dependent homologous recombination between sister chromosomes (Leach et al. 1997; Cromie et al. 2000).

The mechanisms by which palindromes stimulate nearby deletion may be complex. One model for the mechanism of palindrome deletion has been replication slipped mispairing (Balbinder et al. 1989; Weston-
A

hairpin

B

cruciform

Figure 1.—DNA structures assumed by palindrome sequences. (A) Hairpin formation in ssDNA. (B) Cruciform extrusion in dsDNA. Initial melting of DNA allows nucleation of DNA pairing to form arms of the cruciform, which can be extended by branch migration. Inverted repeats with interruption of pairing at the center of the repeats (as F14S) are less likely to form cruciform structures than perfect palindromes (as F14C) because of a decreased probability of this nucleation of arm pairing.

Hafer and Berg 1989) where, after stalling at a hairpin, the nascent strand misaligns at a direct repeat of several nucleotides that flanks the palindromic element. Because of the relative single-strandedness of the lagging-strand template during replication, formation of hairpin structures by inverted repeats should be more prevalent on the lagging strand. Indeed, in yeast, large palindromes are deleted at a higher rate in the presence of mutations that perturb replication on the lagging strand (Gordenin et al. 1992; Ruskin and Fink 1993). In support of this, some experiments suggest that palindrome deletions do occur more often on the lagging strand than on the leading strand in E. coli (Trinh and Sinden 1991; Pinder et al. 1998), although another analysis failed to see substantial differences (Weston-Hafer and Berg 1991). A lagging-strand bias would implicate hairpin structures in ssDNA as the deletion-prone intermediates; however, a different study correlated deletion rate with cruciform formation in dsDNA (as F14C) because of a decreased probability of this nucleation of arm pairing.

To examine the influence of DNA secondary structure on deletion formation between fairly large, 101-bp direct repeats (Figure 2). We have extensively characterized deletion of these 101-bp repeats in the tetA gene of pBR322 (unassociated with palindromes), which occurs at a relatively high rate in the population (\(\sim 10^{-3}\)), independent of RecA (Lovett et al. 1994; Lovett and Feschenko 1996; Feschenko and Lovett 1998; Bzymek et al. 1999). The relatively high rate of deletion of these repeats should make genetic and physical analysis of palindrome effects more facile. Moreover, genetically marked versions of these repeats have been constructed to allow us to determine the endpoints of deletion within 20 nucleotides. The location of the deletion event can provide useful clues toward its molecular nature. We chose to examine the effects of a pair of related inverted repeats \(~100\) bp in length (Zheng et al. 1991): one is a perfect palindrome that more readily forms cruciform structures in duplex DNA in vivo; the other inverted repeat is identical in sequence except for permutation of the central 14 nucleotides (Figure 2). Because of the lack of perfect symmetry, the latter sequence is less likely to extrude in dsDNA as a cruciform (Zheng et al. 1991; see Figure 1B). In ssDNA both repeats form hairpin structures of equal stability (Zheng et al. 1991). Any differential effect of these inverted repeats should allow us to determine whether structures formed in ssDNA or dsDNA are critical substrates for deletion formation.

We show that the DNA palindromes greatly elevate RecA-independent deletion of direct repeats in E. coli. The perfect palindrome that readily forms cruciforms had a much greater stimulatory effect on deletion, implicating cruciform formation in dsDNA as the primary deletion-prone substrate. Our genetic analysis leads us to propose that palindromes stimulate deletion via two pathways: SbcCD-dependent cruciform cleavage/an-nealing and SbcCD-independent replication slippage on the lagging strand. An unexpected finding was that SbcCD apparently processes intermediates in direct repeat deletion, even in the absence of palindromes. An ability of SbcCD to cleave 3’ strands accounts both for our observations and for the original isolation of sbcCD mutations as cosuppressors of the recombination deficiency caused by loss of RecBCD (Lloyd and Buckman 1985; Gibson et al. 1992).

Materials and Methods

Bacterial strains and growth: All strains used are derived from the E. coli K-12 strain AB1157 [F+ thi-1 hisG4 Δ(301-proA)62 argE3 thr-1 leuB6 kdgK51 rfaD1 ara-14 lacY1 galK2 xyl-5 mil-1 tss-33 supE44 rpsL31 lacI12 arresting K+ (Bachmann 1996)] and are listed in Table 2. Experimental strains were constructed by P1 virA transduction using the indicated donors, recipient strains, and selections (Table 1). Experiments employed LB rich medium and growth at 37°C unless otherwise noted. P1 lysate preparation and transductions employed LCG medium, LB supplemented with 1% glucose and 2 mM calcium chloride. Antibiotics used were ampicillin (Ap) at 100 \(\mu\)g/ml, tetracycline (Tc) at 15 \(\mu\)g/ml, chloramphenicol (Cm) at 15 \(\mu\)g/ml, and kanamycin (Km) at 30 \(\mu\)g/ml. Minimal medium employed for certain strain constructions was 56/2 salts (Willett et al. 1969) supplemented with 0.4% glucose, 0.001% thiamine, and 0.1% essential amino acids.

Plasmids: All plasmids used are pBR322 derived and are
relatively low copy number (approximately 20/cell; Helinski et al. 1996), contain a functional copy of bla, conferring ampicillin resistance, and carry various repeated sequences in the tetA gene. Deletion formation between 101-bp perfect direct repeats in tetA was assayed using the plasmid pSTL57 (Lovett et al. 1994) and its derivatives. Deletion endpoints were determined using plasmid pSTL113 (Lovett and Feschenko 1996) and its derivatives listed in Table 2. These latter plasmids carry imperfect 101-bp tetA repeats with four silent transition mutations within the downstream repeat. A unique BgII restriction site is present between the two tetA repeats of pSTL57 and pSTL113.

Plasmids pMB302 and pMB303 were made in several steps. A synthetic oligonucleotide of sequence 5’ AATTTCGCGA was inserted into the EcoRI site of pSTL57, converting it to a new BspEI site and generating plasmid pSTL136. Into the interrepeat BgII site of pSTL136, a unique EcoRI site was added by ligation with the synthetic oligonucleotide of sequence 5’ GATCCGGCGG GAATT CCTCG AG. The resulting plasmid, pSTL298, was digested with EcoRI and ligated with EcoRI fragments F14C and F14S, whose sequences are listed in Figure 2C. Fragments F14C and F14S were excised from plasmids pF14C and pF14S, respectively, and used for transformation.

Plasmids pMB304 and pMB305 are derivatives of pSTL113 and were made by the following steps. The EcoRI site of plasmid pSTL113 was converted to a BspEI site, yielding plasmid pSTL38, by ligation of a synthetic oligonucleotide of sequence 5’ AATTTCGCGA to EcoRI-cleaved pSTL113 DNA. Plasmid pSTL39 carries an inversion of tetA with respect to the replication origin (designated the “−” orientation) and was constructed by cleavage of pSTL136 with BspEI and religation. A unique interrepeat EcoRI site was introduced into pSTL138 by ligation of the synthetic oligonucleotide of sequence 5’ GATCCGGCGG GAATT CCTCG AG into the BgII site, producing plasmid pSTL299. Plasmids pMB304 and pMB305 were constructed by ligation of the inverted repeat sequences F14C and F14S, respectively (Figure 2C), into the EcoRI site of plasmid pSTL299. Plasmid pMB306 was constructed by cleavage of pMB304 with BspEI and religation and thus carries an inversion of tetA with respect to the replication origin, designated as the − orientation. The nucleotide sequences of all resulting constructs were confirmed by DNA sequencing reactions using Sequenase 2.0 from United States Biochemical (Cleveland).

Plasmid DNA was purified by the alkaline/SDS method (Sambrook et al. 1989), QIAprep spin miniprep kit (Qiagen, Valencia, CA), or by Wizard midipreps DNA purification system (Promega, Madison, WI). These plasmids were introduced into respective strains by electroporation (Dower et al. 1988) or treatment with polyethylene glycol and Mg2+ salts (Chung et al. 1989).

Deletion assays and deletion endpoint determination: Deletion was assayed as described previously (Lovett et al. 1993) for a total of 16–64 independent isolates. Briefly, independent cultures were prepared in liquid medium, diluted, and plated to determine the number of TcR colonies compared to the total number of ApR colonies. Deletion rates were calculated by the method of the median (Lea and Coulson 1949) using the formula: deletion rate = M/N, where M is the calculated number of deletion events and N is the final average number of ApR cells in the 1-ml cultures. M is solved by interpolation from experimental determination of n, the median number of TcR cells using the formula n = M(1.24 + lnM). A 95% confidence interval was determined as described previously (Saveson and Lovett 1997).

Deletion products were selected from multiple independent cultures by plating on LB-Tc medium. Before sequencing, plasmids were retransformed, selecting ApR TcR, into a ΔrecA strain, JC10287, to separate deletion molecules from any remaining parental plasmids. Sequence determination employed dideoxy terminators and either Sequenase 2.0 from United States Biochemicals or SequiTherm Excel DNA polymerase from Epicentre Technologies (Madison, WI), using methods recommended by the manufacturers. Statistical analysis of deletion endpoint distribution was performed by contingency chi-square analysis.

Plasmid competition assays: By electroporation, 100 ng plasmid DNA of pSTL297 (with an insertion of cat into the HinII EcoRV fragment of pBR322) was transformed into strains with resident palindromic-free (pSTL57) or palindromic-containing (pMB301) deletion assay plasmids maintained with Ap selection during growth. For each point 7 × 10^7 to 3 × 10^8 electro-
The efficiency of transformation was determined by the frequency of appearance of Cm-resistant colonies (Table 5) in the transformed population.

**RESULTS**

**Effects of palindromes on RecA-independent direct repeat deletion:** We designed a plasmid-based genetic assay in which palindromic sequences were inserted between two direct repeats in tetA (Figure 2). This is an adaptation of a previously described assay (Lovett et al. 1994), in which there is an internal 101-bp duplication (base pairs 567–667) within the tetA gene of pBR322 (pSTL57, Table 2; Figure 2A). Various inverted repeat sequences (Figure 2B) were introduced between the two direct repeat sequences (pMB302, pMB303 in Table 1). Precise deletion of one of the direct repeats and the intervening sequence restores functional tetA and confers tetracycline resistance to the cell. Two types of inverted repeat sequences were examined: a “perfect palindrome,” inverted repeats with nonrepeated spacer DNA (F14C in Figure 2C), and an “interrupted palindrome,” an inverted repeat whose central region sequences contain nonpalindromic sequences (F14S in Figure 2C). The inverted repeats in F14S are identical to those in F14C except for the central region. The central region in F14S has the same base composition as the central region in F14C.

In a wild-type strain, AB1157, perfect palindrome F14C stimulated deletion formation between flanking 101-bp direct repeats by two orders of magnitude (Table 3). This stimulation was similar to that found for another perfect palindrome, 114 bp in length, of unrelated se-
Plasmids used for deletion rate and endpoint determination

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Heterology</th>
<th>Palindrome&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Orientation&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSTL57</td>
<td>–</td>
<td>None</td>
<td>+</td>
<td>Lovett et al. (1994)</td>
</tr>
<tr>
<td>pMB302</td>
<td>–</td>
<td>F14C</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td>pMB303</td>
<td>–</td>
<td>F14S</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td><strong>B.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSTL113</td>
<td>+</td>
<td>None</td>
<td>+</td>
<td>Lovett and Feschenko (1996)</td>
</tr>
<tr>
<td>pMB304</td>
<td>+</td>
<td>F14C</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td>pMB305</td>
<td>+</td>
<td>F14S</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td>pSTL139</td>
<td>+</td>
<td>None</td>
<td>–</td>
<td>Feschenko and Lovett (1998)</td>
</tr>
<tr>
<td>pMB306</td>
<td>+</td>
<td>F14C</td>
<td>–</td>
<td>This study</td>
</tr>
</tbody>
</table>

Plasmids in part A with identical direct repeats of 101 bp were used for measuring deletion rates. Plasmids in part B carried heterologous 101-bp repeats that were used for the determination of deletion endpoint intervals.

<sup>a</sup>Sequences of the various palindromic insertions are listed in Figure 2C.

<sup>b</sup>Plus designates the wild-type orientation of the tetA allele with respect to the replication origin, as in pBR322. Minus designates the reversed orientation of the tetA allele with respect to the replication origin.

Plasmide (M. Bzymek and S. T. Lovett, unpublished results). In contrast, the interrupted palindromic insert F14S (Figure 1C) did not exhibit a stimulatory effect and may have decreased the rate 5-fold relative to that observed for the control construct pSTL57 (Table 3). The decrease in rate may result from the decreased proximity of the direct repeats in this construct, known to diminish deletion rates (Bi and Liu 1994; Chedin et al. 1994; Lovett et al. 1994). In a strain unable to carry out homologous recombination due to a loss of the recA gene, the perfect palindromic F14C similarly stimulated deletion rates over 100-fold. This confirms that the palindromes stimulate a nonrecombinational deletion mechanism. Surprisingly, the absence of RecA revealed a stimulatory effect of the interrupted palindromic sequence F14S on deletion rates. This increase, of ~20-fold, was not as pronounced as that observed with perfect palindromes. This, nonetheless, is an unusual instance of a sequence rearrangement event being dependent on the absence of a functional RecA protein (see discussion).

One potential confounding factor in our analysis is impaired maintenance of palindrome-containing plasmids. The size of palindromes chosen for this analysis was below that known to reduce maintenance (Warren and Green 1985). Moreover, we did not observe any decreased stability of these plasmids as judged by decreased copy number or increased number of plasmid-free cells (data not shown). A more direct control experiment determined the ability of a plasmid lacking any

### TABLE 3

Genetic dependence of palindrome-stimulated deletion

<table>
<thead>
<tr>
<th>Genotype insert&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Deletion rate × 10&lt;sup&gt;−5&lt;/sup&gt; (C.I.)</th>
<th>pSTL57 none</th>
<th>pMB302 F14C</th>
<th>pMB303 F14S</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. 37° assays</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>5.0 (1.1–6.5)</td>
<td>230 (140–2,400)</td>
<td>1.3 (0.3–3.9)</td>
<td></td>
</tr>
<tr>
<td>ΔrecA304</td>
<td>1.1 (0.4–2.1)</td>
<td>170 (89–210)</td>
<td>21.8 (17–27)</td>
<td></td>
</tr>
<tr>
<td>sbcD::Tn10kan ΔrecA304</td>
<td>1.2 (0.7–1.8)</td>
<td>20 (13–31)</td>
<td>6.7 (4.3–9.4)</td>
<td></td>
</tr>
<tr>
<td><strong>B. 30° assays</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔrecA304</td>
<td>1.7 (0.6–7.2)</td>
<td>310 (170–360)</td>
<td>52 (22–200)</td>
<td></td>
</tr>
<tr>
<td>dnaE486 ΔrecA304</td>
<td>1.4 (0.2–2.5)</td>
<td>5,600 (4,100–26,950)</td>
<td>3,000 (2,300–4,200)</td>
<td></td>
</tr>
<tr>
<td>dnaE486 sbcD::Tn10kan ΔrecA304</td>
<td>3.8 (3.8–7.8)</td>
<td>990 (370–3,200)</td>
<td>63 (27–140)</td>
<td></td>
</tr>
</tbody>
</table>

Deletion rates between 101-bp tetA tandem repeats were measured as described in MATERIALS AND METHODS for 16–64 independent cultures grown at 37° and 30° as indicated. Assays of palindrome-containing constructs were always performed in parallel with the nonpalindromic control on 2 or more days. A 95% confidence interval (C.I.) is also indicated.

<sup>c</sup>The sequences of palindromic insertions are listed in Figure 2C.
inverted repeat (uniquely conferring chloramphenicol resistance) to supplant various noncompatible, deletion assay plasmids, either with or without palindromic DNA (Table 4). Transformation of pSTL297 (conferring Cm resistance) into a recA mutant strain carrying pSTL57 or pMB302 revealed that the palindrome-carrying plasmids competed about as well as pSTL57. Therefore, a competitive disadvantage of the palindrome plasmid relative to a newly arising deletion product lacking the palindrome cannot account for the magnitude of the hyperdeletion phenotypes in any of the genetic backgrounds used in these experiments.

**Deletion endpoint distribution:** The use of marked direct repeats allowed us to observe changes in the location of the selected deletion events stimulated by palindromic sequences (Table 5). We used derivatives of the previous deletion assay plasmids (Lovett and Feschenko 1996), in which single-base pair transition mutations in one of the directly repeated sequences designate five intervals, 20 base pairs long, within each repeat (Figure 3). Upon mutation, the deletions are silent within tetA and thus do not affect tetracycline resistance, but allow for the determination of deletion endpoints by DNA sequence analysis (Feschenko and Lovett 1998). Plasmids pMB304, -305, and -306 contain palindromic inserts listed in Figure 2C and the marked direct repeats in tetA (Table 2). To avoid loss of the products by the action of the mismatch repair system on deletion intermediates, we analyzed endpoints in strains carrying additional mutations in mutS (the mismatch recognition factor). Strains were also always mutant for recA, to avoid contribution of homologous recombination to the observed products.

The insertion of palindromic sequences significantly altered the distribution of deletion endpoints (Table 5). Deletions from plasmids containing inserts F14C (perfect palindrome) and F14S (interrupted palindrome) had very similar distributions of endpoints, although the overall rates of deletion differed by 10-fold (Table 5). In the absence of any palindrome, deletion occurs most abundantly in interval 3 and deletions in the outside intervals 1 and 5 are very rarely recovered (Feschenko and Lovett 1998). The presence of palindromic inserts dramatically stimulates deletion events recovered in intervals 1 and 5 that abut the palindrome ($\chi^2 > 19$, $P < 0.001$).

DNA replication is unidirectional in pBR322 (Helinski et al. 1996) and its derivative plasmids used in this study. Since the palindromic sequences may assume secondary structures that interfere with replication, the distribution of deletion endpoints may depend on the direction of replication through the repeats. The asymmetry of the distribution of deleted products formed from pMB304 and pMB305 (Table 5) suggested a possible effect of the direction of replication fork movement on deletion events. In both these constructs, deletion was highest in interval 5, whose replication immediately precedes the palindrome on the lagging strand. The tetA gene was inverted relative to the origin of replication for the F14C construct (giving pMB306) and deletion endpoint distribution was determined. The recovered deletion products from the reversed construct predominated in intervals 1 and 2 (Table 5), again those intervals preceding the palindrome on the lagging strand (Figure 3). This change in skew of deletion endpoints from right to left with the inversion of replication direction was statistically significant ($\chi^2 > 4$, $P < 0.05$). We also noted (Table 6) that the rate of deletion was somewhat greater in the reversed configuration (pMB306 vs. pMB304) and the basis for this is unknown.

**The role of replication in deletion:** To explore further the role of replication in palindrome-stimulated deletion, we determined deletion rates in recA strains with and without dnaE486, a temperature-sensitive mutation
in the polymerase subunit of DNA polymerase III, at its permissive temperature for growth, 30°C. At 30°C, deletion rates for the palindrome-carrying constructs in the control recA (dna+) strain were somewhat higher than at 37°C (Table 3). Deletions resulting from the nonpalindrome construct, pSTL57, were not noticeably affected by temperature. Addition of the dnaE486 allele did not detectably affect deletion of tandem direct repeats (pSTL57); however, palindrome-associated deletion was even further stimulated by dnaE486 (Table 3), suggesting that defects in replication enhance deletion promoted by the presence of palindromic sequences. Whereas the imperfect palindrome consistently affected deletion much less than its perfect counterpart in recA and wild-type strain backgrounds, in the presence of dnaE486, the effects of the imperfect palindrome approximated that of the perfect palindrome. The rate of these deletion events in dnaE mutants was extraordinarily high—deletions were found in almost 10% of the plasmid-bearing population and could be clearly seen in plasmid DNA isolated from this strain, unselected for the deletion events (data not shown).

Dependence on SbcD defines two pathways: The SbcCD nuclease is believed to introduce double-strand breaks (DSBs) at large palindromic sequences (Leach et al. 1997; Cromie et al. 2000) and contributes to the instability of these sequences (Chalker et al. 1988; Gibson et al. 1992). We therefore examined the role of the nuclease on the deletions promoted by palindromic DNA sequences. A mutation in sbcD specifically reduced, from 3- to 9-fold, the rate of RecA-independent palindrome-stimulated deletion (Table 3) but had no effect on nonpalindrome-associated deletion measured from pSTL57. The effect of sbcD was somewhat greater on the perfect F14C palindrome constructs compared to the interrupted F14S palindrome constructs. This dependence of palindrome-stimulated deletion on sbcD was also observed in the context of the dnaE486 mutation, where an sbcD mutation reduced deletion 5-fold for F14C and 50-fold for F14S (Table 3). In both dnaE+ and dnaE486 strains, however, a component of palindrome-stimulated deletion remained independent of sbcD. This defines two pathways for palindrome-stimulated deletion: one SbcD dependent and one SbcD independent.

SbcD effects on deletion endpoints: We determined the distribution of deletion endpoints in sbcD mutants for constructs carrying palindromic sequences F14C and F14S (Table 5). SbcD-independent deletion generates products with asymmetrical endpoint distribution and is skewed toward one of the intervals abutting the palindrome (Table 5; Figure 4, C and D). The skew to interval 5 for both F14C and F14S constructs, + orientation, was significantly different from that for the nonpalindrome construct in sbcD mutants (χ² > 12, P < 0.001). For palindrome F14C, this skew responds to the direction of replication. For F14C − orientation, the skew to the leftward intervals 1 and 2 was also significantly different from the construct lacking a palindrome (χ² > 10, P < 0.001) and the F14C + orientation (χ² > 17, P < 0.001).

This distribution (Table 5) and the rates of deletion (Table 6) can be used to derive a deletion rate in each of the five defined intervals. Rates of sbcD-dependent deletion in each endpoint interval were estimated by

### TABLE 5

<table>
<thead>
<tr>
<th>Genotype (palindrome)</th>
<th>Orientation</th>
<th>No. of products with endpoints in interval</th>
<th>Total sequenced</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. sbcD+</td>
<td>pSTL113 (none) +</td>
<td>1 5 16 6 0</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>pSTL138 (none) −</td>
<td>0 3 7 5 1</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>pMB304 (F14C) +</td>
<td>7 3 7 4 13</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>pMB305 (F14S) +</td>
<td>4 3 6 2 12</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>pMB306 (F14C) −</td>
<td>9 5 3 1 5</td>
<td>23</td>
</tr>
<tr>
<td>B. sbcD−</td>
<td>pSTL113 (none) +</td>
<td>7 4 4 3 2</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>pSTL138 (none) −</td>
<td>2 2 5 3 6</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>pMB304 (F14C) +</td>
<td>2 3 2 0 17</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>pMB305 (F14S) +</td>
<td>2 2 2 5 18</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>pMB306 (F14C) −</td>
<td>8 8 5 2 0</td>
<td>23</td>
</tr>
</tbody>
</table>

Independent deletion products were selected by tetracycline resistance from the indicated strains (STL2172 sbcD+ mutS recA or STL4768 sbcD::Tn10kan mutS recA) carrying the denoted plasmids. The plasmids contain 101-bp direct tetA repeats, with and without palindromic inserts, and in varying orientations relative to the direction of replication. The + orientation is that of natural pBR322; − orientation has the tetA gene reversed relative to ori. Four heterologies between the two repeats allows sequence analysis to determine deletion endpoints within one of five 20-bp intervals (Figure 3).
SbcD-independent deletion was asymmetric, dependent on the replication direction (Figure 4, C and D).

SbcD has little or no effect on the rates of deletion from constructs that lack palindromic DNA sequences (Table 3; Table 6). However, much to our surprise, a mutation in \( \text{sbcD} \) was found to dramatically alter the distribution of deletion endpoints determined from these constructs (Table 5; Figure 5). The reported distribution of endpoints from the \( \text{sbcD}^+ \) background revealed a cluster of endpoints toward the center of the repeat and almost none in the outside intervals 1 and 5. This pattern was observed irrespective of replication direction (Feschenko and Lovett 1998). In \( \text{sbcD} \) mutant genetic background, the hotspot for deletion vanished. Plasmids without palindromic sequences (pSTL113 and pSTL139) generated deletion with endpoints distributed more evenly among all intervals. There was a slight bias of \( \text{sbcD} \)-independent deletion from these nonpalindromic constructs toward one end of the repeat, relative to the direction of replication. For example, among the deletion products of pSTL113 \( >30\% \) of all deletion endpoints fell within interval 1, whereas 10% fell within interval 5. Reversal of the direction of replication in pSTL139 generated the reverse result (Table 5; Figure 4). (The number of products in interval 1 vs. interval 5 is significantly different in the two orientations, \( \chi^2 = 3.8, P = 0.05 \).) In addition to an increase of deletions occurring in interval 3 in both orientations \( \chi^2 > 5.5, P < 0.05 \). This may indicate that even in palindrome-free constructs, the SbcCD nuclease complex actively contributes to the formation of deletions in interval 3 and prevents the formation of products with endpoints in intervals 1 and 5.

\[ \text{Distribution of endpoints among products of SbcD-independent deletion is almost symmetric and biased to both ends of the repeats (Figure 4, A and B). In contrast, subtracting the rate calculated in \( \text{sbcD} \) mutant strain background from that obtained for the \( \text{sbcD}^+ \) strain. Distribution of endpoints among products of SbcD-dependent deletion is almost symmetric and biased to both ends of the repeats (Figure 4, A and B). In contrast,} \]

**DISCUSSION**

Palindromic DNA sequences placed between two direct repeats on a plasmid increase the rate of RecA-

**TABLE 6**

Deletion rates of deletion endpoint mapping plasmids

<table>
<thead>
<tr>
<th>Relevant genotype</th>
<th>Deletion rate ( \times 10^5 ) (C.I.) for plasmids</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pSTL113 (no palindrome) orientation +</td>
</tr>
<tr>
<td>( \text{sbcD}^+ )</td>
<td>2.6 (1.7–4.3)</td>
</tr>
<tr>
<td>( \text{sbcD}::\text{Tn}10\text{kan} )</td>
<td>1.9 (0.4–5.4)</td>
</tr>
</tbody>
</table>

Deletion rates between tandem repeats in \( \text{tetA} \) were measured as described in MATERIALS AND METHODS for 12–16 independent cultures grown at 37°. A 95% confidence interval (C.I.) is indicated. Replication orientation + is that of wild-type pBR322 with fork movement counter to the transcription direction of \( \text{tetA} \). In orientation − the \( \text{tetA} \) gene has been inverted so that replication proceeds in the same direction as \( \text{tetA} \) transcription. The strains STL2172 and STL4768, \( \text{sbcD}^+ \) and \( \text{sbcD}::\text{Tn}10\text{kan} \), respectively, carry the additional mutations \( \text{mutS}::\text{kan} \Delta \text{recA}304 \) and are described in Table 1. The sequences of palindromic insertions are listed in Figure 2C.
independent sequence rearrangement, resulting in deletion of one of the repeats and the intervening palindrome. The rate of deletion was stimulated by the presence of palindromes by one to four orders of magnitude when compared to tandemly positioned direct repeats. We compared two inverted repeat sequences that differ only in their central region. When in ssDNA, both form hairpins of similar thermodynamic stability, but have differing propensity to extrude cruciforms from dsDNA (Sinden et al. 1991; Zheng et al. 1991). Our observed stimulation of deletion rate was highly sensitive to the cruciform-forming potential of the sequence: in wild-type or recA mutant strain backgrounds, a perfect palindrome exhibited stimulatory effects from 10 to 100 times greater than those of an interrupted palindrome with lesser cruciform propensity. Therefore, in our system, deletions must form more commonly from secondary structures arising in dsDNA rather than in ssDNA that is revealed by replication or other processes.

Difficulties in replication can, however, potentiate the effects of palindromes. Our genetic analysis showed that the stimulatory effect of palindromic DNA on deletion was further amplified by defects in polymerization, afforded by a temperature-sensitive mutation in the polymerase subunit of DNA polymerase III. In this dnaEts strain background, the interrupted palindrome was almost as effective in stimulating deletion as its perfect palindrome counterpart. This defect in replication may cause the accumulation of ssDNA tracts, altered superhelical density, or other unknown effects that increase the chance of secondary structure formation by both the perfect and, especially, the interrupted palindrome.
This finding also supports the idea that it is the probability of formation of secondary structure, which is increased in a dnaE mutant, rather than any difference in the structure itself that underlies our different observations regarding the two palindrome types.

RecA may also influence the potential for secondary structure formation in certain cases. The presence of functional RecA abolished the stimulatory effect on deletion of the interrupted palindrome F14S, although not that of the perfect F14C palindrome. The F14S sequence may not form secondary structures unless encompassed, at least partially, by ssDNA; RecA coating of such ssDNA tracts may inhibit hairpin formation. Alternatively, RecA, either directly or indirectly, may help overcome a hairpin's block to replication, thereby diminishing its mutagenic effect.

Dependence on sbcD, the nuclease component of the structure-specific nuclease, SbcCD, defined two pathways by which palindromes stimulate deletion. A mutation in sbcD lowered, but did not abolish, the stimulatory effect of both the perfect and interrupted palindromes on deletion formation. The presence or absence of SbcD also dramatically influenced the distribution of endpoints for palindrome-stimulated deletion events. Our results are consistent with two mechanisms for deletion stimulated by palindromes: (1) SbcCD-dependent single-strand annealing (SSA), which is initiated by the recognition and processing of cruciform structures formed by palindromic sequences, and (2) SbcCD-independent slipped misalignment promoted by replication stalling at hairpins formed on the lagging strand.

**SbcCD-dependent deletion—single-strand annealing:**

We propose a model where SbcCD generates DSBs in the DNA that are subsequently repaired by SSA (Figure 6). Our search for physical evidence for a linearized plasmid yielded only negative results (M. Bzymek and S. T. Lovett, unpublished data); however, this intermediate may be short lived. Breakage of the chromosome by SbcCD at large palindromic sequences has been previously inferred by genetic results (Leach et al. 1997; Cromie et al. 2000). We suggest that even smaller cruciform structures (with ~50-bp arms) that extrude on double-stranded DNA may be incised on both strands by SbcCD. Resection of one strand on each side of the break followed by annealing at the two direct repeats and ligation repairs the break and accomplishes a deletion. This mechanism explains dependence of deletion on cruciform formation and fits well with our endpoint distribution data. Resection of only one strand from the break and annealing would generate products heterogeneous for deletion endpoints in intervals 1 and 5 (Figure 6). SbcCD-dependent deletion generated high levels of interval 1 and 5 products, at approximately equal ratio (Figure 5) and these were dependent on the presence of palindromes.

This mechanism of deletion formation does not require replication and SbcCD incision could occur at cruciforms that form pre- or postreplicationally. Accordingly, the observed distribution of SbcD-dependent products did not vary with the direction of replication fork progression on our plasmids. However, cruciforms broken prereplicationally should be disadvantageous, since the break cannot be repaired by recombination with the sister chromosome and would most likely be mutagenic if healed. It would seem desirable for the cell to restrict SbcCD incision to cruciforms formed in the wake of the replication fork but a mechanism for this restriction is not known.

A SSA mechanism for deletion has been convincingly demonstrated in eukaryotes (Lin et al. 1990; Maryon and Carroll 1991; Fishman-Lobell et al. 1992) and during bacteriophage infection in prokaryotes (Stahl et al. 1997; Tomso and Kreuzer 2000). The contribution of this pathway to deletion in *E. coli*, however, may be normally restricted by the rampant DNA degradative properties of the RecBCD enzyme. Transformation and healing of linear DNA molecules with terminal redundancy is inefficient unless the recipient is mutant for recBCD (Lovett et al. 1988; Luisi-DeLuca et al. 1989) or the DNA molecule possesses multiple chi sequences.
(FRIEDMAN-OHANA et al. 1998), which attenuate RecBCD degradation in vivo (DABERT et al. 1992; KUZMINOV et al. 1994). The absence of chi sequences in our plasmids increases the likelihood that the entire 4.5-kb linear plasmid is degraded within seconds of DSB formation. However, it is possible that SbcCD binds tightly to its break sites, thereby preventing their access to RecBCD. Supporting the fact that SbcD-dependent deletion events are not vulnerable to RecBCD nuclease degradation, we have found that mutation of RecBCD does not enhance SbcD-dependent palindrome-stimulated deletion (M. BZYMEK and S. T. LOVETT, unpublished results). SbcCD's exonuclease activity may also resect the 3' strand until it encounters regions of homology, upon which the two strands anneal (Figure 6). This is similar to a model for deletion at microhomologies promoted by Mre11 (PAULL and GELLERT 1998), which is structurally related to SbcD (SHARPLES and LEACH 1995). RAD50, the eukaryotic counterpart of SbcC, has been shown to promote deletion at short repeats, both with and without associated inverted repeats, in S. cerevisiae (GORDENIN et al. 1992; TRAN et al. 1995).

The SSA mechanism proposed here must be independent of RecA activity, since we observe efficient deletion in recA mutant strains of E. coli. DSB-stimulated SSA after T4 phage infection is independent of RecA as well as UvsX, the phage-encoded RecA homolog, but depends on the gp 45/46 nuclease and single-strand DNA-binding protein, gp32 (TOMSO and KREUZER 2000). For bacteriophage-λ, the annealing pathway requires λ-exonuclease and the annealing protein-β but not host RecA (STAHL et al. 1997). In S. cerevisiae SSA is independent of Rad51, a RecA homolog, but it is dependent on Rad52 for deletions between homologies shorter than 2 kb (reviewed in PAQUES and HABER 1999). Rad52 protein can anneal strands in vitro (MORTENSEN et al. 1996; SUGIYAMA et al. 1998). We cannot rule out the participation of a similar strand annealing activity in E. coli SSA events but the identity of such a factor is presently unknown.

SbcCD-independent deletion—replication slippage at palindromes: Deletion via slipped misalignment during replication has been proposed as an explanation for the loss of palindromic sequences (reviewed in LEACH 1994). The endpoint distribution of deletions stimulated by the perfect palindrome in an sbd mutant strain supports the hypothesis that hairpins, acting as replication stall sites, encourage misalignment at the neighboring direct repeats (Figure 7). For the (+) replication orientation, the F14C palindrome stimulates deletion exclusively in interval 5. Replication of interval 5 immediately precedes the palindrome on the lagging strand; therefore, polymerases stalled at a hairpin formed preferentially on the lagging-strand template would arrest in interval 5. Realignment of this strand, with no accompanying strand degradation, would result in an interval 5 deletion. In the reversed – orientation, deletion endpoints are skewed toward interval 1, again, the interval replicated immediately before the palindrome on the lagging strand. In this latter orientation, however, there appears to be a gradient of endpoints emanating from interval 1, with significant deletions induced in interval 2, and to a lesser extent, intervals 3 and 4. This may result from stalling at the hairpin in interval 1 followed by 3’ exonucleolytic processing of the nascent strand such that the final endpoints are somewhat upstream from the initial stall site (as in Figure 7). We do not know the basis for the difference in endpoint spreading in the two orientations although it is conceivable that sequences in interval 4 or 5 act as a barrier for 3’ end processing.

Our results observed with the sbd mutant strain are therefore consistent with reported lagging-strand bias for palindrome-stimulated replication slippage (TRINH and SINDEE 1991; PINDEE et al. 1998). Secondary structure formation may be more facile on the lagging strand because the template remains in an unpaired state—therefore the rate of intrastrand pairing and hairpin formation is greater than that on the leading strand.
Another possibility is suggested from in vitro results, which implicate polymerase dissociation from its template followed by reassociation as important for deletion formation (CANCEILL and EHRlich 1996; CANCEILL et al. 1999). Discontinuous synthesis on the lagging strand is a series of such polymerase dissociation/association events, and a hairpin structure may simply provide an obligatory end to an Okazaki fragment.

In our particular system, the SbcD-dependent SSA pathway seems to predominate over the SbcD-independent slippage pathway for palindrome-stimulated deletion. The inverted repeats used in our study are optimized for cruciform extrusion with GC-rich stems and AT-rich centers (ZhiENG et al. 1991). However, a similar SbcD dependence of deletion stimulation has been noted by us for an unrelated perfect palindrome, 114 bp in length, without such special features (M. Bzymek and S. T. Lovett, unpublished results). The relative contribution of these two pathways to deletion may vary, depending on the probability or nature of secondary structures formed, or the length and position of homologies. The relatively short palindromic sequences used in some early studies may not be recognized by SbcCD or not efficiently extruded as cruciforms and are hence deleted by slippage on the lagging strand but not by SbcCD-dependent SSA. In addition, the relatively long direct repeats used in our construct may favor SSA over slippage misalignment.

Palindrome-independent deletion and SbcCD processing: In the absence of any known secondary structure, deletion rates between tandem direct repeats were unaffected by a mutation in sbd. However, sbd strongly influenced deletion endpoint distribution, suggesting that SbcCD does participate in direct repeat deletion. The observed deletion hotspot in interval 3 (FESCHENKO and LOVETT 1998) is SbcD dependent. In the absence of a functional SbcD, the products with endpoints in intervals 1 and 5 are generated at the expense of products in interval 3. Our previous analysis of deletion between these tandem repeats supports a replication slippage mechanism for deletion and is not consistent with single-strand annealing (LOVETT et al. 1994; LOVETT and FESCHENKO 1996; FESCHENKO and LOVETT 1998; Bzymek and Lovett 2001). The simplest explanation for our results is that SbcCD converts intermediates that would give rise to interval 1 or 5 deletion products into intermediates that resolve in interval 3. (Initiation of the deletion event, unlike palindrome-associated SSA, would not depend on SbcCD.) Perhaps replication slipped misalignment is accompanied by cleavage of nascent strand 3′ ends by SbcCD nuclease activity, preferentially in interval 3 (Figure 8). Deletions in interval 3 also account for a portion of SbcD-dependent palindrome-associated deletion events (Figure 4) in the + orientation.

Processing of 3′ strands by SbcCD, even in the absence of secondary structures, may explain the original isolation of sbcd mutations as cosuppressors of the recombination deficiency conferred by mutations in RecBCD (LLOYD and BUCKMAN 1985; GIBSON et al. 1992). The partner cosuppressor mutation, sbcB, inactivates the major single-strand 3′ exonuclease, exonuclease I (KUSHNER et al. 1971). The explanation for this suppression has been that 3′ single-stranded ends are critical intermediates for recombination via the alternate RecF pathway (KUSHNER et al. 1971; HORI and CLARK 1973). “sbcB” alleles are nuclease deficient and dominant negative (PHILLIPS et al. 1988; RAZAVY et al. 1996) and are better suppressors than “xoa” alleles that completely inactivate the gene (KUSHNER et al. 1972; PHILLIPS et al. 1988). We have suggested that the SbcB form of ExoI binds but not does degrade single-strand DNA (VISWANATHAN et al. 2000), similar to dominant-negative alleles we have studied of the RecF exonuclease (SUTERA et al. 1999). This binding may protect the 3′ end from degradation by other 3′ exonucleases such as ExoVII. (Indeed, in some phenotypic assays, a double null ExoI−ExoVII− mutant mimics the single SbcB− mutant; VISWANATHAN et al. 2000). However, the bound but inactive SbcB protein should not be able to protect the single strand from endonucleolytic incision via SbcCD. The RecF pathway of recombination may be efficient only when the integrity of 3′ single-strand tails is preserved by further inactivation of the SbcCD nuclease.

A leading strand bias for deletion unassociated with secondary structures: In the absence of SbcCD, deletion endpoints for the nonpalindrome construct are slightly skewed to one side of the repeat and this skew responds to the direction of replication through the repeats. We might imagine that nascent strand slippage after almost complete replication of the repeat (as in Figure 8A)
would be more favorable than slippage after only a small portion of the repeat has been replicated, due to a longer heteroduplex intermediate in the former situation. If this is true and is the only factor governing deletion location, the skew of the distribution suggests that these slipped misalignments are leading strand events. In the + orientation, interval 1, the “hottest” interval for deletion, is replicated last on the leading strand; in the – orientation, interval 5, the hottest interval, is the last interval replicated on the leading strand. All experiments that implicate a lagging-strand bias for deletion in *E. coli* (Trinh and Sinden 1993; Pinder et al. 1998) concern deletion in the context of a palindromic sequence, whose transition to a hairpin structure may be more prevalent on the lagging-strand template. It is conceivable that in the absence of any secondary structures, misalignment occurs more readily during replication of the leading strand, although the molecular basis for this bias is unknown.

We thank D. Leach for providing us with the *shdD* insertion mutant and R. Sinden for the palindromic plasmids. This work was supported by U.S. Public Health Service grants T32 GM07122 (to M.B.) and ROI GM51753.

**LITERATURE CITED**


Connelly, J. C., E. S. de Leau and D. R. Leach, 1999 DNA cleavage and degradation by the ShcC protein complex from *Escherichia coli*. Nucleic Acids Res. 27: 1039–1046.


Davison, A., and D. R. Leach, 1994 The effects of nucleotide sequence changes on DNA secondary structure formation in *Escherichia coli* are consistent with cruciform extrusion in vivo. Genetics 137: 361–368.


