JOURNAL OF BACTERIOLOGY, Jan. 1991, p. 315–318
0021-9193/91/010315-04$02.00/0

Vol. 173, No. 1

Limits to the Role of Palindromy in Deletion Formation

KATHLEEN WESTON-HAFER[1] AND DOUGLAS E. BERG[1,2]*

Departments of Molecular Microbiology[1] and Genetics,[2] Box 8230, Washington University School of Medicine, St. Louis, Missouri 63110

Received 2 July 1990/Accepted 16 October 1990

We tested the effect of palindromy on deletion formation. This involved a study of reversion of insertion mutations in the pBR322 *amp* gene at a site where deletions end either in 9-bp direct repeats or in adjoining 4-bp direct repeats. Inserts of palindromic DNAs ranging from 10 to more than 26 bp and related nonpalindromic DNAs were compared. The frequency of deletions (selected as Amp^r revertants) was stimulated by palindromy only at lengths greater than 26 bp. The 4-bp direct repeats, one component of which is located in the palindromic insert, were used preferentially as deletion endpoints with palindromes of at least 18 bp but not of 16 or 10 bp. We interpret these results with a model of slippage during DNA replication. Because deletion frequency and deletion endpoint location depend differently on palindrome length, we propose that different factors commit a molecule to undergo deletion and determine exactly where deletion endpoints will be.

Short repeated sequences are commonplace in bacterial DNAs. They specify important structural features in rRNAs and tRNAs, often serve as sites of regulation of gene expression, and may also help maintain the overall organization of DNA domains in the bacterial nucleoid (8, 9, 12, 18). Repetitious sequences can also cause instability, however. DNA sequencing showed that the endpoints of spontaneous deletions occur preferentially in short direct repeats (1, 6, 17). In addition, many deletions are associated with palindromic DNAs, and it has been tempting to imagine that even inverted repeats of just a few base pairs can provoke deletion events (1, 7, 14).

Tests based on the reversion of insertion mutations have been used to examine deletion formation more closely. We found that deletion frequencies increased with palindrome length in the range of 22 to 90 bp, up to 18,000-fold at some sites, as little as 8-fold at others (5). These outcomes indicated that both palindromy and other subtle features of DNA sequence affect the deletion process. Palindromes may stimulate deletion events by allowing hairpin structures to form in single-stranded DNAs that then provoke slippage errors during DNA synthesis (Fig. 1). Given the abundance and importance of short or imperfect inverted repeats in bacterial DNAs, it became interesting to test directly the effects of palindrome length on deletion events.

We exploited a set of insertions at a site in the pBR322 *amp* gene. Here we report that changes in palindrome size from 10 to 22 bp did not significantly affect deletion frequencies, measured as the formation of Amp^r revertants. Deletions at this site can end in 9-bp direct repeats of *amp* sequence that bracket the insert or in 4-bp repeats, one copy of which is in the insert (Fig. 2). Here we report that the 4-bp repeat is used preferentially for palindromes of at least 18 bp but not for palindromes of 16 or 10 bp. This implies different thresholds for effects of palindromy on the local probability of deletion and the exact location of deletion endpoints.

MATERIALS AND METHODS

General procedures. Media, bacterial growth, plasmid DNA extractions, recombinant DNA cloning, and restriction endonuclease analyses have been described elsewhere (10, 11). DNA sequences were determined by the chain termination method (13) adapted for double-stranded DNA (19). The primer used to sequence new insertion mutant alleles (near pBR322 positions 3730 to 3740) and representative revertant alleles corresponds to pBR322 positions 3677 to 3693 (5′-GGCGAGTTACATGATCC). Synthetic oligonucleotides were made on an Applied Biosystems 380A DNA synthesizer.

Bacterial strains. *Escherichia coli* K-12 derivative MC1061 (4) was the host for both plasmid constructions and reversion tests; GM119 (Dam$^-$) (2) was used for DNAs to be digested with *BclI*. All plasmids used are derivatives of pBR322 (Amp^r Tet^r).

Amp^s insertion mutant alleles. Palindromes of 18 and 20 bp at site A10 were made as follows. Plasmid DNA containing inserts of a segment consisting of 45-bp inverted repeats bracketing the *kan* gene from Tn*5* at site A10 was digested with *BclI* and either 2 or 4 bp were removed by digestion with mung bean nuclease prior to religation to generate palindromes of 20 and 18 bp, respectively (16).

The 12-bp palindrome at site A10 was made by oligonucleotide mutagenesis. pBR322 DNA was digested with *PvuI* (pBR322 position 3735) and religated with complementary oligonucleotides designed to create a 12-bp palindrome

```
5′   CGCTGTCTAGACAGCTCCGAT
     TAGCGACAGATCTGTCGAGGC
```

(*XbaI* site underlined). The 16-bp palindrome was made by digesting the 12-bp palindrome with *XbaI* and filling in with the Klenow fragment of DNA polymerase I before religation. To make the 10-bp palindrome at site A10, the 12-bp palindrome was digested with *XbaI*, followed by limited digestion with mung bean nuclease and religation. All new palindromic inserts were confirmed by DNA sequencing.

The 10-bp nonpalindromic inserts at site A10 were made by oligonucleotide mutagenesis as described for the 12-bp palindrome. pBR322 DNA was digested with *PvuI* and ligated to the double-stranded oligonucleotide

```
5′   CGCTGTCNNNNNCTCCGAT
     TAGCGACAGNNNNNGAGGC
```

where N represents an equimolar mixture of G, A, T, and C.

* Corresponding author.

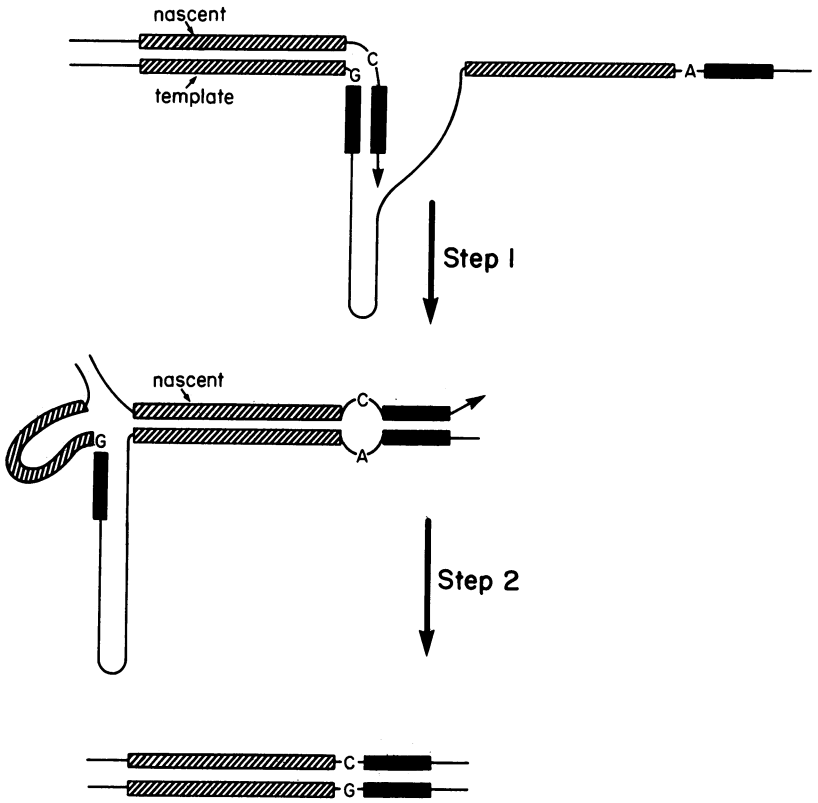

FIG. 1. DNA synthesis error model of deletion formation. Hatched boxes, 9-bp direct repeats; filled boxes, 4-bp direct repeats. Replication is shown 5′ (left) to 3′. Step 1: DNA synthesis is interrupted as the nascent strand stalls within the hairpin. Branch migration results in displacement of the stalled nascent strand and reformation of the complete hairpin. The stalled nascent strand slips and pairs with the second copy of the direct repeat, and the first copy of the repeat loops out. Step 2: Any unpaired bases at the end of the nascent strand are trimmed and synthesis resumes, thereby fixing the deletion. The deleted strand is copied during the next round of replication (reprinted from reference 16 with permission).

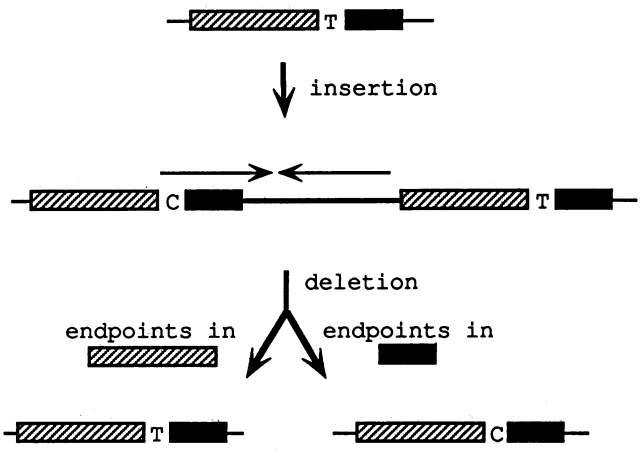

FIG. 2. Location of deletion endpoints is reflected in the sequence of revertants. Hatched and filled boxes are used as in Fig. 1. Horizontal arrows, Palindromic insert (note that one of the 4-bp repeats is part of the palindrome). In the A10 site shown here, endpoints in the 9-bp repeat (hatched) result in a T (wild type) at pBR322 position 3740 in revertants; endpoints in the 4-bp repeats (filled) result in a C at position 3740. These are distinguished by colony hybridization or, for the 3730Tpal set of alleles, by restriction analysis (see text; adapted from reference 16 with permission).

Amp^s alleles were sequenced, and three different nonpalindromic inserts were recovered (see Table 2).

Reversion tests. Amp^r revertant frequencies were measured in clones grown in L broth for about 32 cell doublings from the founding cell prior to plating on L agar with ampicillin (250 μg/ml) as described previously (5, 16). Each frequency reported represents the median of determinations with at least five separate single-cell clones and counts of 100 to 300 revertant colonies from each clone. The frequencies generally varied by less than a factor of three in repeated determinations.

Analysis of Amp^r revertants. Amp^r revertants formed by deletion at site A10 (changes at position 3740) were analyzed by colony hybridization. Two types of revertants (T/A or C/G at position 3740) were obtained and were distinguished with complementary oligonucleotide probes, one containing T and the other containing C at the critical position (16). The sequences of the probes used are 5′-CTCCGATCGTTGT CAG (T probe) and 5′-CTCCGATCGCTGTCAG (C probe).

RESULTS AND DISCUSSION

Palindrome length affects deletion frequency and endpoint location. To study deletion formation, we constructed a series of insertions at the A10 site in the pBR322 *amp* gene and then measured the frequency of reversion from Amp^s to Amp^r (Table 1). The effect of moderate-sized palindromes on deletion frequency at this site is illustrated by the 100-fold

TABLE 1. Effect of palindrome length on deletion formation

Palindrome length (bp)[a]	Amp^r reversion frequency	Fraction of endpoints in 4-bp repeats
10	5.5×10^{-7}	10/58 (0.17)
16	7.2×10^{-7}	12/71 (0.17)
18	9.4×10^{-7}	63/98 (0.64)
20	1.2×10^{-6}	66/95 (0.69)
22[b]	9.5×10^{-7}	51/78 (0.65)
26[b]	1.5×10^{-6}	33/35 (0.94)
32[b]	1.1×10^{-4}	67/72 (0.93)

[a] Palindrome sequences are as follows:
10 bp, ctgtc/gacag
16 bp, ctgtctag/ctagacag
18 bp, ctgtctctt/aagagacag
20 bp, ctgtctcttg/caagagacag
22 bp, ctgtctcttga/tcaagagacag
26 bp, ctgtctcttgatc/gatcaagagacag
32 bp, ctgtctcttgatcaga/tctgatcaagagacag
[b] From Weston-Hafer and Berg (16).

difference in stability of the 32- versus 22-bp palindrome ($\sim 10^{-4}$ versus $\sim 10^{-6}$, respectively). In this context, it is striking that palindromes of 10, 16, 18, 20, 22, and 26 bp all exhibited, within a factor of two, the same stability.

Prior work with palindromes of 22 to 90 bp had shown that deletions from this site tend to end in the 4-bp, not in the 9-bp, direct repeats and had suggested that this unusual preference was due to palindromy (16). The data in Table 1 also show that the distribution of deletion endpoints between the 9- and 4-bp direct repeats was strongly affected by palindrome length, when less than 26 bp. In particular, with palindromes of 22, 20, and 18 bp, two-thirds of deletion endpoints were in the 4-bp repeats, whereas with palindromes of 16 or 10 bp, only one-sixth of deletion endpoints were in the 4-bp repeats. Thus, the placement of deletion endpoints is markedly more sensitive to differences in the length of short palindromes than is the frequency of deletion.

Deletion of short nonpalindromic inserts. To test the idea that for very short palindromes the inverted repeat structure per se is not important as a determinant of either the frequency of deletion or the location of deletion endpoints, we constructed three nonpalindromic variants of the 10-bp insert, differing only in the second set of 5 bp (Table 2, column 1). We found a 50-fold variation in reversion frequency among the four alleles at this site (Table 2); the palindromic insert was actually less deletion prone than the other three alleles. In addition, the use of the 4-bp direct repeats as deletion endpoints ranged from 5 to 17%, independent of palindromy (Table 2). Thus, for deletions of 10-bp inserts, palindromy does not significantly influence the frequency of deletion or the precise location of deletion endpoints.

Differences in the frequency of deletion of the same palindrome from different locations had been explained by nonrandom pausing and loss of processivity by DNA polymerases, as seen in vitro (3, 15). We propose that such a sequence- and conformation-dependent effect is reflected in the distribution of endpoints seen here: the preferential use of the 4-bp repeats during deletion of 18-bp and larger palindromes versus the preferential use of the 9-bp repeats during deletion of 16- and 10-bp palindromes and of nonpalindromic DNA. Given this threshold (18 bp versus 16 bp), the finding that palindrome length affects deletion frequency only with palindromes longer than 26 bp implies that the local frequency of DNA synthesis errors and their exact locations are controlled differently. One interesting possibility is that deletion events are provoked by local perturbations in the balance of replication proteins, for example, an undersupply of single-stranded DNA-binding protein (Ssb) and that the chance of such an imbalance is also sequence and conformation dependent. Thus, the local probability of slippage could be increased by palindromy only with palindromes longer than 26 bp, while the chance that DNA lacking Ssb protein provokes slippage after copying the first 4 bp becomes significant with hairpin structures of at least 16 bp.

ACKNOWLEDGMENTS

We thank L. Henry for skilled technical assistance.

This research was supported by National Science Foundation grant DMB-8608193, Public Health Service grant GM-37138 from the National Institutes of Health, and Department of Energy grant DEFG02-89ER60862 to D.E.B.

REFERENCES

1. **Albertini, A. M., M. Hoffer, M. P. Calos, and J. H. Miller.** 1982. On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. Cell **29**:319–328.
2. **Arraj, J. A., and M. G. Marinus.** 1983. Phenotypic reversal in *dam* mutants of *Escherichia coli* K-12 by a recombinant plasmid containing the *dam*$^+$ gene. J. Bacteriol. **153**:562–565.
3. **Bedinger, P., M. Munn, and B. A. Alberts.** 1989. Sequence-specific pausing during *in vitro* DNA replication on double-stranded DNA templates. J. Biol. Chem. **264**:16880–16886.
4. **Casadaban, M. J., and S. N. Cohen.** 1977. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138**:179–207.
5. **DasGupta, U., K. Weston-Hafer, and D. E. Berg.** 1987. Local DNA sequence control of deletion formation in *Escherichia coli* plasmid pBR322. Genetics **115**:41–49.
6. **Farabaugh, P. I., U. Schmeissner, M. Hoffer, and J. Miller.** 1978. Genetic studies on the lac repressor. VII. On the molecular

TABLE 2. Effect of palindromy on deletion formation

Insert sequence	Insert palindromy	Amp^r reversion frequency	Fraction of endpoints in 4-bp repeats
CTCCGATCGctgtc**gacag**CTCCGATCG	+	7.1×10^{-7}	10/58 (0.17)
CTCCGATCGctgtc**attaa**CTCCGATCG	–	1.2×10^{-6}	9/58 (0.16)
CTCCGATCGctgtc**gaggc**CTCCGATCG	–	8.7×10^{-6}	3/56 (0.05)
CTCCGATCGctgtc**tcgcc**CTCCGATCG	–	3.5×10^{-5}	7/61 (0.11)

nature of spontaneous hotspots in the *lacI* gene of *Escherichia coli*. J. Mol. Biol. **126**:847–863.
7. **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.
8. **Gold, L., and G. Stormo.** 1987. Translation initiation, p. 1302–1307. *In* F. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
9. **Landick, R., and C. Yanofsky.** 1987. Transcription attenuation, p. 1276–1301. *In* F. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
10. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Nag, D. K., U. DasGupta, G. Adelt, and D. E. Berg.** 1985. IS*50* mediated reverse transposition: specificity and precision. Gene **34**:17–26.
12. **Noller, H. F., and M. Nomura.** 1987. Ribosomes, p. 104–125. *In* F. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
13. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
14. **Schaaper, R. M., B. N. Danforth, and B. W. Glickman.** 1986. Mechanisms of spontaneous mutagenesis: an analysis of the spectrum of spontaneous mutation in the *Escherichia coli lac*I gene. J. Mol. Biol. **189**:273–284.
15. **Weisman-Shomer, P., D. K. Dube, F. W. Perrino, K. Stokes, L. A. Loeb, and M. Fry.** 1989. Sequence specificity of pausing by DNA polymerases. Biochem. Biophys. Res. Commun. **164**: 1149–1159.
16. **Weston-Hafer, K., and D. E. Berg.** 1989. Palindromy and the location of deletion endpoints in *Escherichia coli*. Genetics **121**:651–658.
17. **Weston-Hafer, K., and D. E. Berg.** 1989. Specificity of deletion events in pBR322. Plasmid **21**:251–253.
18. **Yang, Y., and G. F.-L. Ames.** 1990. The family of repetitive extragenic palindromic sequences: interaction with DNA gyrase and histonelike protein HU, p. 211–225. *In* M. Riley and K. Drlica (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
19. **Zagursky, R. J., K. Baumeister, N. Lomax, and M. L. Berman.** 1985. Rapid and easy sequencing of large linear double-stranded DNA and supercoiled plasmid DNA. Gene Anal. Tech. **2**:89–94.