Context Effects in the Formation of Deletions in *Escherichia coli*

Toni Kazic and Douglas E. Berg

Department of Molecular Microbiology and Department of Genetics, Washington University School of Medicine, St. Louis, Missouri 63110

Manuscript received January 15, 1990
Accepted for publication May 18, 1990

ABSTRACT

We have examined the frequency with which identical deletions are formed in different chromosomal contexts. A panel of six mutant *bla* genes containing palindrome/direct repeat structures were moved from pBR322 to three locations: at λ*att*, at chromosomal *lac*, and at F′*lac*. Deletion of the palindromes and one of the direct repeats results in reversion to Amp^r. The frequency of deletion for all alleles declines beyond the reduction in copy number when they are moved from the multicopy plasmid environment to a single-copy chromosome. The magnitude of the declines varies in an allele-specific and location-specific manner. Our data support the hypothesis that context can influence the frequency of mutation independent of the immediate DNA sequence.

OVER the last 15 years, a number of studies have defined characteristic DNA sequence features that help determine the position and frequency of deletions in prokaryotic genomes. Many deletions occur between short direct repeats, for example in plasmids (ALBERTINI *et al.* 1982), bacteriophage T4 (SINGER 1988), and *Escherichia coli* (FARABAUGH *et al.* 1978). In addition, the presence of palindromes between direct repeats can significantly raise deletion frequency (DASGUPTA, WESTON-HAFER and BERG 1987; GLICKMAN and RIPLEY 1984). However, neither direct repeats nor palindromes are essential for deletion, as demonstrated by several well-characterized instances (FARABAUGH *et al.* 1978; IKEPEAZU and VALENZUELA 1989; PEIJENBERG *et al.* 1989) [for reviews, see ALLGOOD and SILHAVY (1988), BALBINDER (1988) and EHRLICH (1989)].

Another line of investigation suggests that DNA conformation can influence the frequency of deletion. Thus, inclusion of a replication origin from a single-stranded bacteriophage on a multicopy plasmid elevates the frequency of several different deletions (LEVINSON and GUTMAN 1987; MICHEL, D'ALENCON and EHRLICH 1989). Similarly, reduced DNA gyrase activity has been shown to reduce deletion formation in *E. coli* (SAING *et al.* 1988). Several other mutations in DNA metabolism have been found to affect the frequency of deletion in a variety of systems (KONRAD 1977; LEACH and STAHL 1983; LEVINSON and GUTMAN 1987; LUNDBLAD and KLECKNER 1984; PEIJENBERG *et al.* 1989; SHURVINTON, STAHL and STAHL 1987).

In all of these cases, it is difficult to separate the effects of immediate DNA sequence, as measured over tens of base pairs, from those which would depend on the surrounding context of a kilobase or more. Small, multicopy plasmids differ from large, single copy chromosomes in many ways, including superhelicity, replication frequency, and the arrangement of expressed genes. Transient changes in DNA conformation which arise as a result of these processes could have a significant impact on the probability that particular mutations arise. This hypothesis could also apply to regions of the *E. coli* chromosome: the two strands may differ in the frequency of gene expression (BREWER 1988), and the chromosome is divided into about 50 topologically distinct domains (SINDEN and PETTIJOHN 1981; WORCEL and BURGI 1972).

Do different portions of the bacterial chromosome offer distinct environments to a particular sequence? We approached this question by systematically examining the effects of chromosomal context on the mutation frequency of a panel of deletion alleles. These are a set of palindrome/direct repeat structures in the β-lactamase gene (*bla*) of pBR322. Deletion of these structures is detected by reversion to Amp^r, and is dependent upon their position in the *bla* gene (DAS GUPTA, WESTON-HAFER and BERG 1987). Here we place several mutant *bla* alleles in the same position in *lacZ*, then move these *lacZ::blaX* genes to three different locations: to an F′*lac*, to a λ*lac* integrated in the *E. coli* chromosome at *att*λ, and to the *E. coli* chromosomal *lac* gene. We find that the reversion frequency for these alleles varies in both an allele-specific and a location-specific manner. Further, some alleles are tolerated at only some locations. These data support the hypothesis that context can affect the frequency of mutation, and suggest these locations are distinct biochemical environments.

MATERIALS AND METHODS

Materials: Enzymes were purchased from New England Biolabs or Boehringer-Mannheim and used according to instructions. Bicinchoninic acid (BCA) was from Pierce

TABLE 1

Phage and bacterial strains

Strain	Alias	Genotype	Source
A. Fundamental Strains			
W1485		F$^+$ λ$^-$ wild type	CGSC[a]
MG1655		F$^-$ λ$^-$ derivative of W1485	CGSC
ED8654		F$^-$ *lac*Y1 or *lac*Δ(I–Z)6 *gal*K2 *gal*T22 *met*B1 *hsd*R514 *trp*R55 λ$^-$ *sup*E44 *sup*F58	CGSC
MC1061		F$^-$ *ara*D139 Δ(*ara-leu*)7697 Δ*lac*X74 *gal*U *gal*K *hsd*R *rps*L *lac*$^+$	Laboratory collection
RGS1017		MC1061 *lac*$^+$	P1(W1485) × MC1061 → Lac$^+$
DG101		F$^-$ *thi*-1 *end*A1 *hsd*R17 *sup*E44 *lac*I^q *lac*ZΔM15	D. GELFAND
DG102		DG101 *lac*I$^+$	D. GELFAND
DB114		Δ*trp*E *trp*R *tna sup*E$^+$? *hfl*A1	Laboratory collection
DB1506		Δ*trp*E5 Δ(*pro*B- *lac*) Sup$^+$?/F′128 *pro*B$^+$ *lac*$^+$	Laboratory collection
DB4514	MC1061/pGPX2	MC1061/pBR322 *bla tet* Δ(*Dra*I nt3232-*Dra*I nt3251) Ω(nt3253::*Xho*I) Ω(*Pvu*II nt2068::*Xho*I)	Laboratory collection
RGS830	DG102/pQZ340	DG102/pGPX2 Ω(*Xho*I::*Sac*I) Ω(*Xho*I::*Sac*I)	Insertion of *Sac*I linkers
RGS414	JM83/pUC4K	Δ(*lac-pro*) *ara rps*L *thi* φ80d*lac*ZΔM15/*kan*$_{Tn903}$	pUC4K from J. MESSING (VIEIRA and MESSING 1982)
RGS883	DG101/pQZ343	DG101/pQZ340 Ω((*tet*A *bla*)::*kan*) Ω(*Hind*III::*Pst*I) Δ(*Pst*I-*Pst*I) Ω(*Pst*I::(*Pst*I *kan*$_{Tn903}$ *Pst*I))	Insertion of PstI linker and substitution of *kan* gene from pUC4K
DB4551	MC1061/p1022	MC1061/pBR322 *tet bla*301 Ω(nt3739::(22-bp palindrome^b)) dup(nt3731–3739)	DASGUPTA, WESTON-HAFER and BERG (1987)
DB4553	MC1061/p1522	MC1061/pBR322 *tet bla*305 Ω(nt4023::(22-bp palindrome)) dup(nt4015–4023)	DASGUPTA, WESTON-HAFER and BERG (1987)
DB4554	MC1061/p2222	MC1061/pBR322 *tet bla*303 Ω(nt3738::(22-bp palindrome)) dup(nt3730–3738)	DASGUPTA, WESTON-HAFER and BERG (1987)
DB4567	MC1061/p1090	MC1061/pBR322 *tet bla*302 Ω(nt3739::(90-bp palindrome^b)) dup(nt3731–3739)	DASGUPTA, WESTON-HAFER and BERG (1987)
DB4569	MC1061/p1590	MC1061/pBR322 *tet bla*306 Ω(nt4023::(90 bp palindrome)) dup(nt4015-4023)	DASGUPTA, WESTON-HAFER and BERG (1987)
DB4570	MC1061/p2290	MC1061/pBR322 *tet bla*304 Ω(nt3738::(90-bp palindrome)) dup(nt3730–3738)	DASGUPTA, WESTON-HAFER and BERG (1987)
P1virS		*virS*	J. SCOTT
λRB1		λgt11 *lac*Z Ω(*Eco*R1::*Sac*I) Ω(*Sac*I::*Nae*I) *cI*857 *nin*5 *S*100	R. BARSTEAD
λRGS10		λRB1 Ω(*Sac*I::(*bla tet*A)$_{pBR322}$)	Insertion of PstI fragment from pQZ340
λRGS11		λRGS10 Ω(*Sac*I::[(*bla tet*A)::*kan*]$_{pRGS343}$)	Recombination between λRGS10 and pQZ343
B. Plasmid strains			
RGS1058		MG1655/p1022	Transformation
RGS1059		MG1655/p1522	Transformation
RGS1060		MG1655/p2222	Transformation
RGS1061		MG1655/p1090	Transformation
RGS1062		MG1655/p1590	Transformation
RGS1063		MG1655/p2290	Transformation
C. Lysogens of recombinant λ phage			
RGS987		ED8654(λRGS11)/p1022	Transformation of ED8654; select Kan^r lysogens
RGS988		ED8654(λRGS11)/p1522	Transformation of ED8654; select Kan^r lysogens
RGS989		ED8654(λRGS11)/p2222	Transformation of ED8654; select Kan^r lysogens
RGS990		ED8654(λRGS11)/p1090	Transformation of ED8654; select Kan^r lysogens
RGS991		ED8654(λRGS11)/p1590	Transformation of ED8654; select Kan^r lysogens
RGS992		ED8654(λRGS11)/p2290	Transformation of ED8654; select Kan^r lysogens
RGS997	DB114(λRGS12)	DB114(λ Ω(*lac*Z4087::(*bla*301 *tet*A)) *cI*857 *nin*5 *S*100)	Induction of RGS987; select Tet^r lysogens
RGS998	DB114(λRGS13)	DB114(λ Ω(*lac*Z4089::(*bla*303 *tet*A)) *cI*857 *nin*5 *S*100)	Induction of RGS989; select Tet^r lysogens

Strain	Alias	Genotype	Source
RGS999	DB114(λRGS14)	DB114(λ Ω(*lac*Z4088::(*bla*302 *tet*A)) *cI*857 *nin*5 *S*100)	Induction of RGS990; select Tet^r lysogens
RGS1000	DB114(λRGS15)	DB114(λ Ω(*lac*Z4092::(*bla*306 *tet*A)) *cI*857 *nin*5 *S*100)	Induction of RGS991; select Tet^r lysogens
RGS1015	DB114(λRGS16)	DB114(λ Ω(*lac*Z4091::(*bla*305 *tet*A)) *cI*857 *nin*5 *S*100)	Induction of RGS988; select Tet^r lysogens
RGS1016	DB114(λRGS17)	DB114(λ Ω(*lac*Z4090::(*bla*304 *tet*A)) *cI*857 *nin*5 *S*100)	Induction of RGS992; select Tet^r lysogens
RGS1046		MG1655(λRGS12)	Induction of RGS997; select Tet^r lysogens
RGS1047		MG1655(λRGS13)	Induction of RGS998; select Tet^r lysogens
RGS1048		MG1655(λRGS14)	Induction of RGS999; select Tet^r lysogens
RGS1049		MG1655(λRGS15)	Induction of RGS1000; select Tet^r lysogens
RGS1050		MG1655(λRGS16)	Induction of RGS1015; select Tet^r lysogens
RGS1051		MG1655(λRGS17)	Induction of RGS1016; select Tet^r lysogens
D. F′*lac* strains			
RGS1018		RGS1017(λRGS12)	Select Tet^r lysogens
RGS1019		RGS1017(λRGS13)	Select Tet^r lysogens
RGS1021		RGS1017(λRGS15)	Select Tet^r lysogens
RGS1022		RGS1017(λRGS16)	Select Tet^r lysogens
RGS1023		RGS1017(λRGS17)	Select Tet^r lysogens
RGS1025		RGS1021 Ω(*lac*Z4092::(*bla*306 *tet*A)) λ?	Select Temp^r Tet^r; score Lac$^-$
RGS1026		RGS1023 Ω(*lac*Z4090::(*bla*304 *tet*A)) λ?	Select Temp^r Tet^r; score Lac$^-$
RGS1027		RGS1018 Ω(*lac*Z4087::(*bla*301 *tet*A)) λ?	Select Temp^r Tet^r; score Lac$^-$
RGS1030		RGS1022 Ω(*lac*Z4091::(*bla*305 *tet*A)) λ?	Select Temp^r Tet^r; score Lac$^-$
RGS1031		RGS1019 Ω(*lac*Z4089::(*bla*303 *tet*A)) λ?	Select Temp^r Tet^r; score Lac$^-$
RGS1071		DB1506 F′128-21 Ω(*lac*Z4092::(*bla*306 *tet*A))	P1(RGS1025) × DB1506 → Tet^r
RGS1072		DB1506 F′128-22 Ω(*lac*Z4090::(*bla*304 *tet*A))	P1(RGS1026) × DB1506 → Tet^r
RGS1073		DB1506 F′128-23 Ω(*lac*Z4087::(*bla*301 *tet*A))	P1(RGS1027) × DB1506 → Tet^r
RGS1076		DB1506 F′128-24 Ω(*lac*Z4091::(*bla*305 *tet*A))	P1(RGS1030) × DB1506 → Tet^r
RGS1077		DB1506 F′128-25 Ω(*lac*Z4089::(*bla*303 *tet*A))	P1(RGS1031) × DB1506 → Tet^r
RGS1089		MG1655/F′128-21 Ω(*lac*Z4092::(*bla*306 *tet*A))	RGS1071 × MG1655 → Min$^+$ Tet^r
RGS1090		MG1655/F′128-22 Ω(*lac*Z4090::(*bla*304 *tet*A))	RGS1072 × MG1655 → Min$^+$ Tet^r
RGS1091		MG1655/F′128-23 Ω(*lac*Z4087::(*bla*301 *tet*A))	RGS1073 × MG1655 → Min$^+$ Tet^r
RGS1092		MG1655/F′128-24 Ω(*lac*Z4091::(*bla*305 *tet*A))	RGS1076 × MG1655 → Min$^+$ Tet^r
RGS1093		MG1655/F′128-25 Ω(*lac*Z4089::(*bla*303 *tet*A))	RGS1077 × MG1655 → Min$^+$ Tet^r
E. Chromosomal *lac* alleles			
RGS1133		MG1655 Ω(*lac*Z4092::(*bla*306 *tet*A))	P1(RGS1089) × MG1655 → Tet^r
RGS1135		MG1655 Ω(*lac*Z4087::(*bla*301 *tet*A))	P1(RGS1091) × MG1655 → Tet^r
RGS1136		MG1655 Ω(*lac*Z4091::(*bla*305 *tet*A))	P1(RGS1092) × MG1655 → Tet^r
RGS1137		MG1655 Ω(*lac*Z4089::(*bla*303 *tet*A))	P1(RGS1093) × MG1655 → Tet^r

[a] Coli Genetic Stock Center
[b] DASGUPTA, WESTON-HAFER and BERG (1987).

Chemical Co.; 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) was from Boehringer-Mannheim. All other fine chemicals were from Sigma. Microtiter dishes (no. 3072) were from Falcon. Nitrocellulose was from Schleicher and Schuell.

General methods: Standard microbiological and recombinant DNA methods were employed (MILLER 1972; MANIATIS, FRITSCH and SAMBROOK 1982). Cells were grown at 30° in LN broth and on plates (BERG, WEISS and CROSSLAND 1980) with fresh 5 μg/ml tetracycline (Tet) or 100 μg/ml ampicillin (Amp) as appropriate. Methods for P1 transductions (DUNCAN, ROCKSTROH and WARNER 1978) and transformations (DAVIS, BOTSTEIN and ROTH 1980) have been described. Relative plasmid copy number was determined essentially as described (EHRET and MATZURA 1988), except that concentration was estimated by visual comparison of serially diluted extracts to diluted standards. Values were converted to numbers of copies by assuming 18 copies/cell

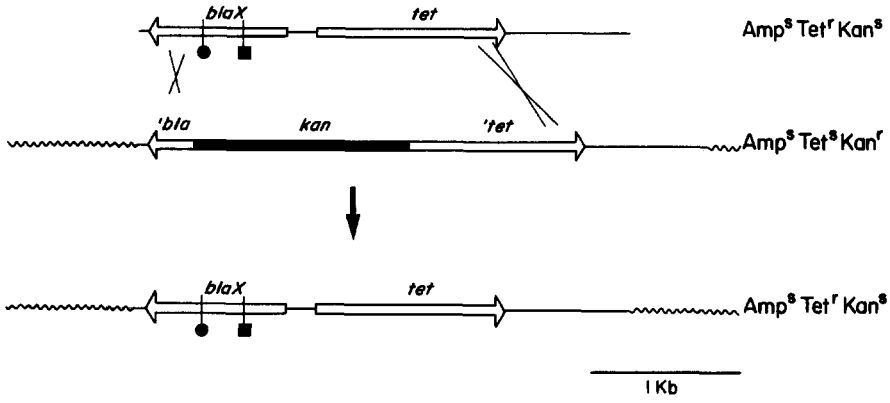

FIGURE 1.—Construction of palindromic *lac* alleles. The basic maneuver is illustrated. The donor replicon carries *blaX tet*A, and the recipient carries a *kan* insertion spanning *bla* and *tet*A. The surrounding recipient DNA is indicated by the wavy lines. Homologous recombination produces *blaX tet*A progeny, which are retrieved by selection for recipient chromosomes carrying Tet^r. (●) Position of A10 and A22 insertions; (■) position of A15 insertion.

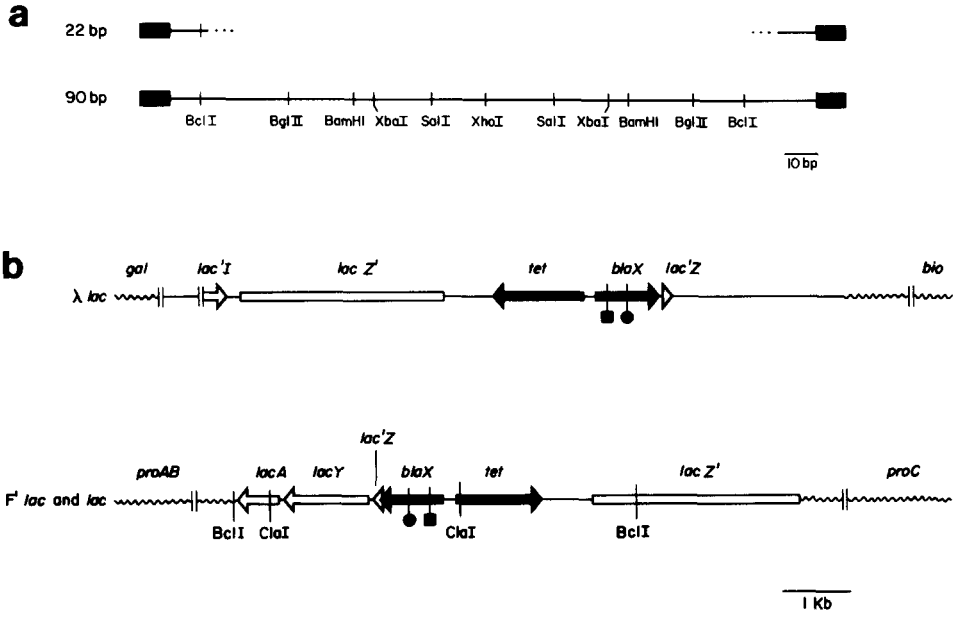

FIGURE 2.—The structure of the palindromic alleles. Panel a shows the basic insertion: a palindrome of either 22 or 90 bp is flanked by generic 9-bp direct repeats (boxes). Reversion to Amp^r requires deletion of the palindrome and one of the direct repeats. Restriction sites in the palindromes are marked. Panel b shows the extended context in λ*lac* and in F'*lac* and chromosomal *lac*. DNA replication fork movement is from left to right in the figure. The open arrows represent *lac* sequences, while the filled arrows mark genes from pBR322. The thin line is for pBR322 or λ sequences, and the wavy lines for the surrounding *E. coli* sequences. Transcription proceeds toward the arrowheads. (●) Position of A10 and A22 insertions; (■) position of A15 insertion. Several useful restriction sites are marked.

for pBR322 (COVARRUBIAS *et al.* 1981). Southern blotting was carried out as described (Schleicher and Schuell 1987).

Strain construction: The *E. coli* K-12 and bacteriophage strains employed in this work are listed in Table 1. The various *bla* alleles are collectively abbreviated *blaX*. The general strategy for moving *blaX* alleles among replicons was to select for recombination between a *blaX tet* plasmid (Amp^S Tet^r Kan^S) and a phage which substitutes a *kan* gene for parts of *bla* and *tet*, including the sites of the *blaX* mutations (Amp^S Tet^S Kan^r; Figure 1). Since the recipient's drug markers are embedded in a *lacZ* gene, the recombinant *lac*::(*blaX tet*) alleles can be crossed into a *lac*$^+$ replicon. The *lac*::[(*bla tet*)::*kan*] derivative of λgt11, λRGS11, was constructed as indicated (Table 1, part A) (YOUNG and DAVIS 1983). A similar approach was used for various control constructions.

The λ phage bearing *blaX* alleles were constructed by inducing phage from cells lysogenic for λRGS11 and carrying plasmid *blaX tet* (RGS987–992), and selecting Tet^r lysogens formed by recombinant progeny phage in an *hflA* strain (DB114; RGS997–1016). Phage were induced from these strains, purified, then used to construct the MG1655 lysogens (RGS1046–1051; Table 1, part C). To obtain chromosomal *lac*::(*blaX tet*) alleles, lysogens of λ*lac*::(*blaX tet*) were heat induced (44° 10 min; RGS1018–1023), then plated at 42° on tetracycline X-gal isopropyl *S*-D-thiogalactoside (IPTG) plates. The resulting colonies proved to be mixtures of true *lac*::(*blaX tet*) recombinants and λ*lac*::(*blaX tet*) lysogens (RGS1025–1031), so P1 *virS* lysates were made and used to transduce an F'*lac* strain (DB1506) to Tet^r (RGS1071–1077). The F's were then crossed into MG1655, making phage-free, F'*lac*::(*blaX tet*) strains (RGS1089–1093; Table 1, part D). Finally, the *lac*::(*blaX tet*) alleles were transduced into the chromosome of MG1655 with P1 *virS* (RGS1133–1137; Table 1, part E).

Physical verification of alleles: The structure of all alleles was confirmed by restriction mapping of the appropriate substrates. For λ lysogens, phage were induced, passed through a *dam* host to produce the unmethylated DNA required by several of the restriction enzymes, and the DNA purified (HELMS, DUTCHIK and OLSON 1987). Strains with F'*lac* and chromosomal *lac* alleles were made *dam*, DNA was purified (SILHAVY, BERMAN and ENQUIST 1984), and Southern blots were performed. In all cases, the position and orientation of the insertion relative to the surrounding chromosome were determined (*Cla*I and *Pst*I); the presence of a palindrome restriction site shown (*Bcl*I); and the 22- and 90-bp palindromes distinguished from each other by the presence or absence of appropriate restriction sites (*Bgl*II and *Xba*I; data not shown).

Reversion assays: Modified fluctuation tests were performed essentially as described (HAYES 1968). It seemed likely many of the reversion frequencies would be too low to measure with conventional assays. Increasing the number of replicates and measuring all strains together improves sample size and minimizes day effects, but at the cost of more complicated logistics. A microtiter dish-based assay was therefore devised. All strains were streaked from master stocks on LN tetracycline plates and incubated 22 hr at 30°. Eleven sibling colonies of each strain were then picked into individual wells of a microtiter dish containing 175 μl LN + tetracycline and grown for 24 hr at 30° in a rotating, humid box. Viable count was estimated by measurement of optical density in an ELISA reader (T. KAZIC, manuscript in preparation). Final cell densities ranged from 2×10^8 to 2×10^9/ml, depending on the strain. Within a given strain replicate cultures varied no more than twofold in cell density.

Cells were plated by one of two methods depending on the expected frequency. For reversion frequencies of $\sim 10^{-7}$ or less, cultures were plated directly on half-plates of LN + ampicillin and incubated 24–48 hr. For higher reversion frequencies, cultures were diluted using a Titertek multichannel pipettor, and aliquots from several dilutions were filtered onto nitrocellulose using a 96-well manifold (Schleicher and Schuell). The filters were laid on a film of LN broth + Amp on 15 cm LN + Amp plates and incubated 10 hr at 30°. They were then laid on nylon mesh saturated with a solution of BCA (SMITH *et al.* 1985) for several minutes to stain the microcolonies for protein and counted under a dissecting microscope at 70× to determine the number of revertants. Though up to 300 microcolonies per well can be counted, dilutions giving ~150 microcolonies were chosen for ease and reliability in counting. Control experiments showed the microcolony counts agreed well with direct plating and counting of fully grown colonies, and that frequencies of plasmid alleles determined by this method agreed well with previous measurements (DASGUPTA, WESTON-HAFER and BERG 1987).

Numerical analysis: Two trials of the entire panel of 21 strains were performed. The reversion frequencies reported are the median values for the set of 11 replicates for one of the trials. Medians, rather than means, are reported to reduce the effects of jackpots, but this prevents calculation of standard deviations or other measures of error. However, for 20 strains, these median frequencies differ by no more than twofold in duplicate trials. Two strains, RGS1047 and RGS1090, had medians which differed more than a 100-fold between the two trials as the result of runs of jackpots. To decide between the alternative medians, a third trial for these strains and several controls was performed; the value within twofold of the third trial's median is shown. Values for the control strains were within twofold of those from their previous two trials, indicating there was nothing unusual about the third trial.

An important question is whether any pair of median frequencies is statistically different. Standard tests are inapplicable here, since they either test matched samples (Wilcoxon test), assume normality (T-tests), or compare one distribution to the lumped distribution of all samples (J-median and Kruskal-Wallis tests; WINKLER and HAYS 1975). To determine which frequencies were significantly different, a two-step analysis was performed. Mann-Whitney tests on the distribution of reversion frequencies among the eleven sibs were performed for each pair of alleles. All cases of difference were significant at the $P = 0.05$ level, and most were significant at the $P = 0.01$ level. Alleles scoring as different in the Mann-Whitney test were then checked for differences in their medians, and scored as different if these differed by more than threefold. Since the Mann-Whitney test is sensitive to distributions of frequencies, not the median values, distributions which score as different can have similar medians. The results were then used to group alleles into similarity sets, where all members of a set test as not significantly different from each other. Computations were performed using RS1 v3.0 (Bolt Baranek and Newman, Inc.) on a DEC VAX 6220.

RESULTS

Rationale: The basic experiment is to place a variety of deletion-prone sequences in the same immediate context in *lacZ*, and measure the frequency of deletion when these alleles are placed in different extended contexts on single-copy chromosomes. The test sequences consist of a 22-bp or 90-bp palindrome, flanked by three different 9-bp direct repeats, in the *bla* gene of pBR322 (Figure 2, panel a) (DASGUPTA, WESTON-HAFER and BERG 1987). The longer palindrome is simply an extension of the shorter, and the three different direct repeats reflect the sites of the palindromes in the *bla* gene. These are A10 [a repeat of nucleotides (nt) 3731–3739], A22 (nt 3730–3738), and A15 (nt 4015–4023). These are abbreviated by giving the site, followed by the palindrome length: thus, A15::22 is the 22-bp palindrome, flanked by direct repeats of pBR322 nt 4015–4023, at the A15 site. All *blaX* alleles are linked to *tet*. For convenience, the palindrome/direct repeat structures will be referred to as "structures."

These sequences resemble many substrates for deletion formation (GLICKMAN and RIPLEY 1984), and offer the advantages of easy detection and previous characterization. Deletion is scored by reversion to Amp^r. Even in multicopy plasmids, restoration of Amp^r requires removal of both the palindrome and one of the two direct repeats; no pseudorevertants which might have reduced resistance have been observed (DASGUPTA, WESTON-HAFER and BERG 1987; WESTON-HAFER and BERG 1989b). Our criterion for reversion is quite stringent, and the fact that structures are present in single copy makes confusion of mutations which restore some resistance with true revertants unlikely. This does not exclude the possibility that other deletion events may occur, but helps assure the detection of a specific event. The A10 and A22 sites were chosen because the deletion frequencies of the 90-bp palindrome at these sites differ by more than four orders of magnitude, though the sites differ by only one nucleotide. In contrast, the 22-bp palindromes' frequencies differ by about sixfold at these sites. At the A15 site, reversion of the 22-bp and 90-bp palindromes differs by 40-fold.

Reversion frequencies of the alleles: Fifteen congenic alleles which place *blaX tet* in the same position in *lacZ* (nt 4302 from the start of the operon) were constructed in *E. coli* K-12 strain MG1655 (Figure 2, panel b). These differ by location: in the chromosome

TABLE 2

Median deletion frequencies

Allele	*lac*::	λ*lac*::	F′*lac*::	pBR322
A10::22	0.18	0.10	1.3	0.51
A10::90	NC[a]	580	NC	4100
A22::22	<0.001	<0.001	0.007	0.079
A22::90	NC	6.6	4.8	6.0
A15::22	<0.001	<0.001	<0.002	0.093
A15::90	0.040	1.4	0.096	0.35

Median deletion frequencies ($\times 10^{-6}$) are shown, corrected for relative plasmid copy number and assuming pBR322 is present at 18 copies/cell (see MATERIALS AND METHODS). Proceeding down the column, the relative copy numbers are: 18, 14, 24, 9, 7, and 54.

[a] Not constructed.

TABLE 3

Sets of alleles of similar frequency

	Location				
Set no.	Chromosome	λ*lac*	F′*lac*	pBR322	Mean
1				A10::90	4100
2		A10::90			580
3		A22::90	A22::90	A22::90	5.80
4		A15::90	A10::22		1.35
5	A15::90			A10:22	0.40
				A15::90	
				A15::22	
6	A10::22	A10::22	A15::90	A22::22	0.11
7			A22:22		0.01
8	A15::22	A15::22	A15::22		<0.002
	A22::22	A22::22			

at *lac* (at 8′ on the genetic map), in a λ*lac* prophage (17′), and in an F′*lac* episome. Stable *lac* A10::90 or A22::90 alleles could not be isolated, either by direct curing of prophage, by transduction, or by spontaneous or UV-induced recombination between the appropriate F′*lac::(blaX tet)* and MG1655 chromosomes. An apparent F′ A10::90 allele was shown not to be inserted in *lac* on Southern blotting, and could not be constructed by spontaneous recombination between F′*lac::*[(*bla tet*)*::kan*] and the plasmid allele. The median reversion frequencies for the test *bla* alleles are compared to the same insertions carried on pBR322 in MG1655 in Table 2. These vary from approximately 10^{-3} to less than 10^{-9}. Two patterns to the variation are particularly apparent.

The 22- *vs.* 90-bp palindromes: The 22-bp palindromes usually delete less frequently than the homologous 90-bp palindrome. The differences range from ~0.3-fold (at the A15 site in pBR322) to nearly 10^{-4}-fold (at the A22 site in integrated λ*lac*). Most 22-bp structures revert about 10^{-2}-fold less frequently than the corresponding 90-bp structures. About half of the 22-bp alleles have frequencies less than the sensitivity of the assay ($\sim 10^{-9}$). These data are consistent with previous observations on the effect of palindrome length on deletion (DASGUPTA, WESTON-HAFER and BERG 1987; EGNER and BERG 1981; FOSTER *et al.* 1981).

Frequencies at locations other than pBR322: For most alleles, placement on larger, single-copy chromosomes reduces the frequency of reversion relative to pBR322, often by one to two orders of magnitude. This is greater than the differences in copy number among the replicons. The magnitude of the effect is both insertion-specific and location-specific. A22::90 (line 4) is relatively insensitive to locus, while A10::22 (line 1) is slightly more sensitive. A15::22 is fairly consistently reduced (line 5), while the remainder show wide variation (A10::90, line 2; A22::22, line 3; and A15::90, line 6). Both 22- and 90- bp palindromes fall into all categories. Only three alleles show no reduction: the A10::22 in F′*lac*, and the A22::90 and A15::90, both in λ*lac*.

Statistically significant differences in reversion frequencies: These observations are reinforced when the alleles are grouped into sets of similar frequency as described in the MATERIALS AND METHODS. These sets, along with the mean reversion frequencies for each set, are displayed in Table 3. Alleles with 22-bp palindromes fall mostly into sets six to eight. The insensitivity of the A22::90 structure to different locations is highlighted (set three), while most of the remainder show drops in reversion frequency when moved from pBR322: plasmid values are found exclusively in sets one to six.

DISCUSSION

The data presented are consistent with the hypothesis that the deletion frequency of a structure can be changed by the surrounding context. Thus, reversion frequencies are usually depressed when constructs are placed on large, single-copy chromosomes beyond that expected from the drop in copy number. The extent of the reduction, and the consistency of the reduction for all locations, is structure-specific. A15::22 is equally reduced in all non-pBR322 locations, others are somewhat more variable (A10::22 and A22::90), while the remainder vary widely. These results also suggest there are some limits to the change in frequency contributed by the extended context: shorter palindromes still delete infrequently, compared to longer ones, when placed at the other locations. These observations are consistent with earlier data collected from single-copy replicons in a variety of extended contexts (EGNER and BERG 1981; FOSTER *et al.* 1981).

Implications for mechanisms of deletion formation: The yield of deletions is clearly determined by the substrate's configuration and availability, the nature of the enzymatic steps for individual deletions, and the fraction of reactions going to completion (GLICKMAN and RIPLEY 1984; PAPANICOLAOU and RIPLEY 1989; SINGER 1988). That several superficially

different alleles show the same reversion frequencies suggests these alleles exist in biochemically similar environments. Reversion frequency could reflect differences in the frequency and mechanism with which unusual DNA conformations are processed. One consequence of the replication slippage hypothesis (GLICKMAN and RIPLEY 1984; STREISINGER *et al.* 1966; STREISINGER and OWEN 1985; WESTON-HAFER and BERG 1989a) is that the deletion frequency should rise when the number of opportunities to synthesize DNA are increased. This is consistent with the general decline in reversion seen in single-copy replicons.

Alternatively, the magnitude of the frequency may depend directly on the formation of particular deletion-prone DNA conformations *in vivo*. Our results suggest that different locations can significantly affect substrate formation or availability. By grouping alleles into statistically similar sets, as in Table 3, one can predict which alleles are similar substrates *in vivo*. If the probability of fork stalling and slippage (KAGUNI and CLAYTON 1982) increases with longer half-lives of nonduplex DNA conformations, deletion is likely to be more frequent, independent of the final copy number. Such deletion-prone conformations could include, but are by no means limited to, base-paired palindromes extruded from the plane of the chromosome. Local superhelical density has been shown to affect the frequency of palindrome extrusion in small plasmids (LILLEY 1981), and gene expression has been shown to influence DNA supercoiling in pBR322 (LODGE, KAZIC and BERG 1989; WU *et al.* 1988). Preliminary experiments indicate changes in supercoiling affect the viability of some palindrome-containing plasmids (T. KAZIC, unpublished results).

Contextual effects on deletion formation: We note the two locations on the *E. coli* chromosome compared here are not equivalent. The data show much larger variation in reversion frequencies for integrated λ*lac* prophages than chromosomal *lac* constructs (five orders of magnitude *vs.* three, respectively), and reversion of the A15::90 construct differs significantly at the two locations (Table 2, line 6, column 1 *vs.* column 2). Neither does F′*lac* perfectly mimic behavior at *lac*. Two structures which are well tolerated as prophages could not be obtained as *lac* insertions (A10::90 and A22::90), and one of these was not obtained on F′*lac*. It is tempting to speculate that the apparent lethality of these structures represents extreme cases of the same regional effects which produce deletions in the λ*lac* alleles.

Until recently there has been little direct evidence to suggest chromosomal context has any effect on the frequency of mutation. Surveys of transposon insertion document different frequencies in different locations, but each site considered differs in DNA sequence as well as location (for example, SHAW and BERG 1979). There is some evidence for different properties of portions of the bacterial chromosome: studies of large inversions in *E. coli* show some are deleterious (REBOLLO, FRANCOIS and LOUARN 1988), and inversions of some regions of *Salmonella typhimurium* cannot be recovered (SEGALL, MAHAN and ROTH 1988; SEGALL and ROTH 1989). Two classes of explanation for these results are relevant here. The first argues the chromosome is divided into structurally distinct domains of chromosome folding (REBOLLO, FRANCOIS and LOUARN 1988; SEGALL and ROTH 1989). The second class argues there are functionally distinct domains, determined by the orientation of highly expressed genes to replication fork movement (BREWER 1988). This second class does not account for our results, since transcription of *lac::blaX* and F′*lac::blaX* is antiparallel to fork movement while that of λ*lac::blaX* is parallel. The two sites considered here are contained within *E. coli* inversions, suggesting this is not a crucial difference. They do lie in different chromosomal domains as proposed by REBOLLO, FRANCOIS and LOUARN (1988): *lac* is outside of NDZ1, while *att*λ is within it (NDZ is nondivisible zone). While this may account for the differences, it is equally likely there are other features which distinguish these two portions of the chromosome. The method described here can be generalized to exploring any region of the chromosome given an appropriate λ*att*$^+$ *int*$^+$ phage carrying *E. coli* genes.

We thank M. BERLYN, B. DUNCAN, S. PHADNIS, S. SAWYER and K. WESTON-HAFER for helpful discussions, the reviewers for comments, and B. BACHMANN, R. BARSTEAD, D. GELFAND, J. MESSING and J. SCOTT for gifts of strains. T.K. was supported by U.S. Public Health Service training grant A107172. This work was supported by grants from the U.S. Public Health Service (GM-37138) and National Science Foundation (DMB-8608193) to D.E.B.

LITERATURE CITED

ALBERTINI, A. M., M. HOFER, 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.

ALLGOOD, N. D., and T. J. SILHAVY, 1988 Illegimate recombination in bacteria, pp. 309–330 in *Genetic Recombination*, edited by R. KUCHERLAPATI and G. R. SMITH. American Society for Microbiology, Washington, D.C.

BALBINDER, E., 1988 Use of predesigned plasmids to study deletions: strategies for dealing with a complex problem, pp. 378–391 in *DNA Replication and Mutagenesis*, edited by R. E. MOSES and W. C. SUMMERS. American Society for Microbiology, Washington, D.C.

BERG, D. E., A. WEISS and L. CROSSLAND, 1980 Polarity of Tn*5* insertion mutations in *Escherichia coli*. J. Bacteriol. **190:** 439–466.

BREWER, B. J., 1988 When polymerases collide: replication and the transcriptional organization of the *E. coli* chromosome. Cell **53:** 679–686.

COVARRUBIAS, L., L. CERVANTES, A. COVARRUBIAS, X. SOBERON, I. VICHIDO, A. BLANCO, Y. M. KUPERSTOCH-PORTNOY and F. BOLIVAR, 1981 Construction and characterization of new cloning vehicles. V. Mobilization and coding properties of pBR322 and several deletion derivatives including pBR327 and pBR328. Gene **13:** 79–89.

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.

Davis, R. W., D. Botstein and J. R. Roth, 1980 *Advanced Bacterial Genetics.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Duncan, B. K., P. A. Rockstroh and H. R. Warner, 1978 *Escherichia coli* K-12 mutants deficient in uracil-DNA glycosylase. J. Bacteriol. **134:** 1039–1045.

Egner, C., and D. E. Berg, 1981 Excision of transposon Tn*5* is dependent on the inverted repeats but not on the transposase function of Tn*5*. Proc. Natl. Acad. Sci. USA **78:** 459–463.

Ehret, M., and H. Matzura, 1988 Control of the *Staphylococcus aureus* chloramphenicol resistance plasmids pC223 and pUB112 in *Bacillus subtilis.* Nucleic Acids Res. **16:** 2045–2062.

Ehrlich, S. D., 1989 Illegitimate recombination in bacteria, pp.799–832 in *Mobile DNA,* edited by D. E. Berg and M. M. Howe. American Society for Microbiology, Washington, D.C.

Farabaugh, P. J., U. Schmeissner, M. Hofer and J. H. Miller, 1978 Genetic studies on the *lac* repressor. VII. On the molecular nature of spontaneous hotspots in the *lacI* gene of *Escherichia coli.* J. Mol. Biol. **126:** 847–863.

Foster, T. J., V. Lundblad, S. Hanley-Way, S. N. Halling and N. Kleckner, 1981 Three Tn*10*-associated excision events: relationship to transposition and role of direct and inverted repeats. Cell **23:** 215–227.

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.

Hayes, W., 1968 *The Genetics of Bacteria and Their Viruses,* pp. 181–185. Blackwell Scientific Publications, Oxford.

Helms, C., J. E. Dutchik and M. V. Olson, 1987 A lambda DNA protocol based on purification of phage on DEAE-cellulose. Methods Enzymol. **153:** 69–82.

Ikepeazu, E. V., and M. S. Valenzuela, 1989 A unique deletion in pBR322 DNA caused by insertion of poly(dA)·poly(dT). Plasmid **21:** 247–250.

Kaguni, L. S., and D. A. Clayton, 1982 Template-directed pausing in *in vitro* DNA synthesis by DNA polymerase α from *Drosophila melanogaster* embryos. Proc. Natl. Acad. Sci. USA **79:** 983–987.

Konrad, B. E., 1977 Method for the isolation of *Escherichia coli* mutants with enhanced recombination between chromosomal duplications. J. Bacteriol. **130:** 167–172.

Leach, D. R. F., and F. W. Stahl, 1983 Viability of λ phages carrying a perfect palindrome in the absence of recombination nucleases. Nature **305:** 448–451.

Levinson, G., and G. A. Gutman, 1987 High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. Nucleic Acids Res. **15:** 5323–5338.

Lilley, D. M. J., 1981 *In vivo* consequences of plasmid topology. Nature **292:** 380–382.

Lodge, J. K., T. Kazic and D. E. Berg, 1989 Formation of supercoiling domains in plasmid pBR322. J. Bacteriol. **171:** 2181–2187.

Lundblad, V., and N. Kleckner, 1984 Mismatch repair mutations of *Escherichia coli* K12 enhance transposon excision. Genetics **109:** 3–19.

Maniatis, T., E. F. Fritsch and J. Sambrook, 1982 *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Michel, B., E. d'Alencon and S. D. Ehrlich, 1989 Deletion hot spots in chimeric *Escherichia coli* plasmids. J. Bacteriol. **171:** 1846–1853.

Miller, J. H., 1972 *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Papanicolaou, C., and L. S. Ripley, 1989 Polymerase-specific differences in the DNA intermediates of frameshift mutagenesis. *In vitro* synthesis errors of *Escherichia coli* DNA polymerase I and its large fragment derivative. J. Mol. Biol. **207:** 335–353.

Peijenberg, A. A. C. M., P. V. Breed, S. Bron and G. Venema, 1989 Plasmid deletion formation in *recE4* and *addB72* mutants of *Bacillus subtilis.* Plasmid **21:** 205–215.

Rebollo, J.-E., V. Francois and J.-M. Louarn, 1988 Detection and possible role of two large nondivisible zones on the *Escherichia coli* chromosome. Proc. Natl. Acad. Sci. USA **85:** 9391–9395.

Saing, K. M., H. Orii, Y. Tanaka, K. Yanagisawa, A. Mura and H. Ikeda, 1988 Formation of deletion in *Escherichia coli* between direct repeats located in the long inverted repeats of a cellular slime mold plasmid: participation of DNA gyrase. Mol. Gen. Genet. **214:** 1–5.

Schleicher and Schuell, 1987 *Transfer and Immobilization of Nucleic Acids to S&S Solid Supports,* pp. 6–9. Schleicher and Schuell, Keene, N.H.

Segall, A., M. J. Mahan and J. R. Roth, 1988 Rearrangement of the bacterial chromosome: forbidden inversions. Science **241:** 1314–1318.

Segall, A. M., and J. R. Roth, 1989 Recombination between homologies in direct and inverse orientation in the chromosome of *Salmonella*: intervals which are nonpermissive for inversion formation. Genetics **122:** 737–747.

Shaw, K. J., and C. M. Berg, 1979 *Escherichia coli* K-12 auxotrophs induced by insertion of the transposable element Tn*5*. Genetics **92:** 741–747.

Shurvinton, C. E., M. M. Stahl and F. W. Stahl, 1987 Large palindromes in the λ phage genome are preserved in a *rec*$^+$ host by inhibiting λ DNA replication. Proc. Natl. Acad. Sci. USA **84:** 1624–1628.

Silhavy, T. J., M. L. Berman and L. W. Enquist, 1984 *Experiments with Gene Fusions,* pp. 137–139. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Sinden, R. R., and D. E. Pettijohn, 1981 Chromosomes in living *Escherichia coli* cells are segregated into domains of supercoiling. Proc. Natl. Acad. Sci. USA **78:** 224–228.

Singer, B. W., 1988 On the role of homologous sequences in chromosomal rearrangements. Genes Dev. **2:** 1800–1811.

Smith, P. K., R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson and D. C. Klenk, 1985 Measurement of protein using bicinchoninic acid. Anal. Biochem. **150:** 76–85.

Streisinger, G. M., and J. (E.) Owen, 1985 Mechanisms of spontaneous and induced frameshift mutation in bacteriophage T4. Genetics **109:** 633–659.

Streisinger, G. M., Y. Okada, J. Emrich, J. Newton, A. Tsugita, E. Terzaghi and M. Inouye, 1966 Frameshift mutations and the genetic code. Cold Spring Harb. Symp. Quant. Biol. **31:** 77–84.

Vieira, J., and J. Messing, 1982 The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19:** 259–268.

Weston-Hafer, K., and D. E. Berg, 1989a Palindromy and the location of deletion endpoints in *Escherichia coli.* Genetics **121:** 651–658.

Weston-Hafer, K., and D. E. Berg, 1989b Specificity of deletion events in pBR322. Plasmid **21:** 251–253.

Winkler, R. L., and W. L. Hays, 1975 *Probability, Inference and Decision,* Ed. 2. Holt, Rinehart & Winston, New York.

Worcel, A., and E. Burgi, 1972 On the structure of the folded chromosome of *Escherichia coli.* J. Mol. Biol. **71:** 127–147.

Wu, H. S., S. Shyy, J. C. Wang and L. F. Liu, 1988 Transcription generates positively and negatively supercoiled domains in the template. Cell **53:** 433–440.

Young, R. A., and R. W. Davis, 1983 Efficient isolation of genes by using antibody probes. Proc. Natl. Acad. Sci. USA **80:** 1194–1198.

Communicating editor: J. R. Roth