DNA Sequence Effects on Single Base Deletions Arising During DNA Polymerization in Vitro by Escherichia coli Klenow Fragment Polymerase

Fei Jun Wang and Lynn S. Ripley

Department of Microbiology and Molecular Biology, UMD-New Jersey Medical School, Newark, New Jersey 07103

Manuscript received September 28, 1993
Accepted for publication November 29, 1993

ABSTRACT

Most single base deletions detected after DNA polymerization in vitro directed by either Escherichia coli DNA polymerase I or its Klenow fragment are opposite Pu in the template. The most frequent mutations were previously found to be associated with the consensus template context 5'-PyTPu-3'. In this study, the predictive power of the consensus sequence on single base deletion frequencies was directly tested by parallel comparison of mutations arising in four related DNAs differing by a single base. G, a deletion hotspot within the template context 5'-TTGA-3', was substituted by each of the 3 other bases. Previous studies had shown that deletions opposite the G were frequent but that deletions opposite its neighboring A were never detected. Based on the predictions of the consensus, the substitution of T for G should produce frequent deletions opposite the neighboring A due to its new 5'-TTTA-3' template context. This prediction was fulfilled; no deletions of this A were detected in the other templates. The consensus further predicted that deletions opposite template C would be lower than those opposite either A or G at the same site and this prediction was also fulfilled. The C substitution also produced a new hotspot for 1 bp deletions 14 bp away. The new hotspot depends on quasi-palindromic misalignment of the newly synthesized DNA strand during polymerization; accurate, but ectopically templated synthesis is responsible for this mutagenesis. Mutations templated by quasi-palindromic misalignments have previously been recognized when they produced complex sequence changes; here we show that this mechanism can produce frequent single base deletions. The unique stimulation of misalignment mutagenesis by the C substitution in the template is consistent with the singular ability of C at that site to contribute to extended complementary pairing during the DNA misalignment that precedes mutagenesis.

In vitro studies of DNA polymerization fidelity provide one approach to identifying a subset of mechanisms that may contribute to spontaneous mutagenesis. The specificity of polymerization errors and their dependence upon experimentally manipulable factors such as DNA sequence and DNA polymerase provide important clues to their underlying causes. Moreover, demonstration of dependence (or independence) of mutagenesis in vivo and in vitro on the same manipulable factors may provide a basis for deciding which spontaneous mutations are caused directly by in vivo polymerization errors (Ripley 1990).

In vitro studies using a wide variety of DNA polymerase enzymes and several different DNA templates have been used to generate mutants which have now been sequenced in substantial numbers (Kunkel 1985a, b; de Boer and Ripley 1988; Kunkel et al. 1989; Mo et al. 1991). The studies reveal that multiple mechanisms contribute to polymerization dependent frameshift mutagenesis in vitro and that the relative importance of these mechanisms to the spectra of mutants produced by different polymerases varies widely.

Studies in this laboratory have focused on the frameshifts produced by Escherichia coli DNA polymerase I and its Klenow fragment while copying single-stranded templates designed to permit the specific recovery of frameshifts in either the +1 or −1 directions (de Boer and Ripley 1988; Papanicolaou and Ripley 1989, 1991; Revich and Ripley 1990). In vitro studies of polymerization-dependent frameshifts in the −1 direction, led to the description of three distinguishable frameshift mechanisms. Two of the three mechanisms rely on misaligned DNA pairing. One misalignment mechanism implicates misalignment of directly repeated DNA sequences. In vitro the sequence changes characteristic of this mechanism are multibase deletions deleting one copy of the repeat and bases between the repeats. This mechanism is like that proposed by Stresinger and co-workers (1966) and has been implicated repeatedly as a major contributor to spontaneous frameshift mutagenesis in vivo (Stewart and Sherman 1974; Farabaugh et al. 1978; Prinnow et al. 1981; Ripley et al. 1986). The second misalignment mechanism implicates misalignment of palindromic sequences (Papanicolaou and Ripley 1989, 1991). In vitro the DNA sequence changes characteristic of this mechanism are frequently
complex frameshifts (substitution/deletion/insertion changes) which are precisely predicted by templated, but misaligned DNA synthesis, followed by realignment of the DNA to the template. Similar frameshift sequences attributable to these misalignment classes have been found in vivo and are implicated in spontaneous mutagenesis (Ripley 1982; de Boer and Ripley 1984; Hampsey et al. 1988). Misalignment mutagenesis in its simplest form might be imagined to be primarily a function of DNA sequence and to be largely insensitive to DNA polymerase. Experimentally, this appears not to be the case. Both in vitro and in vivo, misalignment mutagenesis specificity and frequency is strongly influenced by the DNA polymerase (Ripley and Shoemaker 1983; Ripley et al. 1983; de Boer and Ripley 1984; Kunkel 1985a,b; Fix et al. 1987; Papanicolaou and Ripley 1989, 1991; Bebeneke et al. 1990; Masker and Crissey 1993; Schaaper 1993).

In addition to misalignment-based frameshifts, E. coli PolI and Klenow polymerase frequently produce 1-bp deletions at specific DNA contexts whose characteristics are inconsistent with frameshift mechanisms depending on misaligned pairing stabilized by either direct repeats or quasi-palindromes (de Boer and Ripley 1988). The deletions are most frequently seen opposite template purines, but not all template purines are prone to deletions. Only template purines adjacent to 5' Ts were frequently mutated. These mutations do not depend on repetition of the purine. The average deletion frequencies opposite template Pu's which are part of a run of identical bases are not higher than frequencies opposite Pu's that are not part of runs (de Boer and Ripley 1988). This shows that primer template misalignments, stabilized by the kind of slipped mispairing hypothesized by Streisinger et al. (1966) and regularly invoked to explain the loss of a base pair in a run of identical bases does not regularly stimulate this mutagenesis.

We initially proposed that the context might be error prone because the template sequence could assume a DNA configuration which would promote the ability of the polymerase to bypass the purine and move on to the adjacent 5'T template site. Evidence generally consistent with this idea was suggested by studies of the specificity of acridine-induced frameshift mutagenesis due to the presence of acridine during in vitro DNA polymerization and molecular modeling of sequence-specific, crystalline nucleic acid structures associated with acridines into the context of a replicating DNA molecule (Revich and Ripley 1990; Berman et al. 1992). However, an alternative model has been proposed to explain 5' context effects on the basis of initiation of mutagenesis through misincorporation (Kunkel 1990). In this model, the first step of mutagenesis is the incorporation of a nucleotide opposite the deletion site that is non-complementary to that site, but is instead complementary to the 5' nucleotide. Subsequent realignment of this misincorporated base and its extension produce the frameshift.

Despite their substantial differences, the two models are difficult to distinguish experimentally. For example, it appears that in some cases slippage is likely to occur first. Slippage allowing incorporation of a base complementary to the 5' adjacent template base which is postulated to subsequently "slip back" to the normal alignment thereby accounting for higher frequencies of base substitutions complementary to the 5' base at the preceding site (Kunkel and Soni 1988). The observation that slippage may precede apparent mispairing qualifies the generality of any experiment employing forced misincorporations to produce predicted frameshifts in support of the model that requires misincorporation first (Bebeneke and Kunkel 1990). Experimentally forced misinsertion experiments clearly demonstrate that realignments between primer and template occur but cannot address the issue of whether misinsertion precedes misalignment or whether instead misalignment precedes complementary incorporation. It is presently unclear whether at a particular site or group of sites, misincorporation or misalignment is the first step of mutagenesis.

Our previous studies derived a preferred DNA context for E. coli DNA PolI-induced 1-bp deletions by developing a consensus sequence from the overall spectrum of frameshifts arising in approximately 120 bp of DNA. In this study, the direct manipulation of a single DNA base was used to test the predictive power of the consensus sequence on mutagenesis at one site. A specific 1-bp deletion context, 5'TTGA3' was selected for study. In this sequence deletions opposite G are frequent. This report describes the result of substituting the G with each of the other bases (A, T or C) on the frequency and specificity of frameshifts arising during in vitro polymerization. The substitutions are variously predicted to disrupt an existing hotspot and/or create novel hotspots. This measurement examines polymerization errors opposite four different bases embedded within an otherwise constant DNA context.

MATERIALS AND METHODS

Enzymes and chemicals: In vitro DNA synthesis was catalyzed by either E. coli DNA polymerase I-Harge fragment (Klenow) or its entire form (PolI), both purchased from Boehringer Mannheim. T4 phage polynucleotide kinase was from Stratagene. [γ-32P]ATP was from Amersham.

DNA constructions: The M74 phage is a derivative of M13 mp8, previously described by de Boer and Ripley (1988). Derivatives of M74 with base substitutions at position G108 used in this study were constructed using oligonucleotide-directed mutagenesis (Amersham kit). The constructions are designated as M74-A108, M74-C108 and M74-T108. The original M74 is analogously designated M74-G108. The four DNAs differ only at their 108 site.

Oligonucleotides: M13, S20 universal sequencing primer (5'-GTAAACGACGCTGACGTTTAT'-3') from Stratagene was used
for in vitro to prime DNA synthesis for mutagenesis experiments. The T at the 3' end of this primer is complementary to the A at position 128 in M74. DNA sequencing was primed by either S20 primer or S40 universal primer (5'-GGTTT-TCCCAGTCACGAC-3') from New England Biolabs. Five oligonucleotides for sequence specific construction and probing, were synthesized using a Applied Biosystems 380 Synthesizer. The oligos used for constructions and probing were each complementary to the 108 site, covering the 98-112 region. Their sequences are shown in the Figure 1.

In vitro DNA synthesis: The S20 primer and single-stranded M74 DNAs were annealed at a template concentration of 50 nM in 50 mM of Tris-HCl (pH 7.6) and 0.1 mM NaCl. Primer was added in approximately 2-fold molar excess to template; the total volume of the annealing mixture was 60 µl. Annealing was accomplished by incubating the mixture at 65°C for 5 min, followed by 15 min at 37°C. The annealed mixture was held on ice until use. Primer extensions were carried out by adjusting the annealed template-primer mixture to 10 nM in 50 mM Tris-HCl (pH 7.6), 10 mM MgCl2, 1 mM dithiothreitol and 250 µM of each of the four dNTPs. A sample of 0.1-0.2 unit/µl DNA polymerase was added to start the synthesis; the total reaction volume was 300 µl. The extension reactions were incubated for 30 min at 37°C.

The extent of DNA synthesis was measured as a function of time. The newly synthesized DNA strand was labeled by incorporation of 32P-labeled dATP. Samples from the reaction were stopped by the addition of EDTA, and were spotted onto DE81 Whatman disks. These disks were washed in 0.3 M ammonium formate and 10 mM of sodium pyrophosphate, dried and counted. Klenow-directed DNA polymerization reached roughly to one round of extension, if all templates are used.

Mutant mapping by oligonucleotide probing: A dot-blot assay was used to quickly identify the mutants in the vicinity of the G108 site, using a protocol similar to that used for T4 DNA (RIPLEY et al. 1988). Mutants between 86 and 112 in the DNA sequence were identified by probing with two oligonucleotides one extending from 86 to 100 and another extending from 97 to 112. The oligo extending from 97 to 112 was appropriately matched to the identity of the base at 108 in the M74 template for the experiment (Figure 1). A 60-µl sample of phage stock (0.5-1 x 10^12) was deposited on a nitrocellulose filter using a dot-blot apparatus. DNA was released from phage by addition of 60 µl of 0.6 M NaOH. Filters were prehybridized and washed as previously described (RIPLEY et al. 1988). Hybridization was at 37°C overnight followed by a 10-min stringent wash at 48°C.

DNA sequencing: All of the mutants identified by the probing assay as lying in the vicinity of the 108 site were sequenced. DNA preparation and sequencing were carried out as described previously (PAPANICOLAOU et al. 1989). The sequencing primer was 5’ labeled with 32P using T4 polynucleotide kinase.

Polymerase pausing: In vitro pausing experiments were carried out as in DNA polymerization experiments used to gather mutants except that the primer was first 5’,32P-labeled using T4 polynucleotide kinase. Polymerization was carried out in a total reaction volume of 60 µl from which aliquots were removed in a time course and the reaction stopped by directly pipetting into the gel loading solution containing formamide and dye. Samples were loaded on a denaturing polyacrylamide gel and run in parallel with a DNA sequencing reaction using the same primer. Pausing pattern was visually estimated from the radioautograms.

RESULTS AND DISCUSSION

M74, a derivative of M13-mp8, specifically designed for examining 1-bp deletions (DE BOER and RIPLEY 1988) was modified for these studies at the 108 site. The 3’ end of G108 was a site of frequent deletions in previous studies and is surrounded by the following context: 5’ TTGAC 3’. The G108 base was substituted with A108, C108, or T108 using oligo-directed mutagenesis. Each of the four DNAs was used as a template for in vitro DNA polymerization by E. coli DNA polymerase I Klenow fragment. After transfection of the DNA into E. coli, frameshifts that restore the correct reading frame in the α-complementing fragment of the β-galactosidase gene of M74 are detected by their blue plaques color in the
presence of IPTG and X-gal. This report describes the results of analysis of 764 blue plaques.

Mutations mapping to the vicinity of the 108 site were identified by the oligonucleotide probing assay described in the MATERIALS AND METHODS and Figure 1. The mapping identified the mutants with altered DNA sequences between bases 86 and 112 of M74. The validity of the mapping was initially established using approximately 120 sequenced frameshifts (~30 mutants from each of the 4 templates) selected for their blue plaque phenotype. Application of the mapping protocol to the sequenced mutants produced neither false negatives nor false positives.

Each mutant identified by probing to lie between 86 and 112 in the sequence was sequenced. A summary of the data is presented in Table 1. Most of the sequenced frameshifts were single base deletions. The fraction of mutants mapping to the DNA segment 86–112 ranged from 17 to 21% in the various templates and is similar to the 16% found in our previous studies of the unmodified [C108] template (PAPANICOLAOU and RIPLEY 1989). The single base deletion spectrum for each template is presented in Figure 2. The sequences of additional mutant types found are shown in Figure 3. The sequences of Deletions (Del) and Complex frameshifts (Cmplx) are shown in Figure 3. Base pair changes are indicated by position number and sequence change in the template strand.

Single base deletions as does E. coli Klenow polymerase (data not shown). Comparison of the G108 and A108 spectra reveals no major differences (Figure 2). The G108 → A108 substitution has little, if any, influence on mutations at the 108 site or its immediate neighbors. This result supports the designation of TPu as a site of frequent deletion, and does not favor a specific purine.

Single base deletion spectra in T108 and C108 templates: In contrast to the purine substitution at the 108 site, pyrimidine substitutions produced spectra readily distinguishable from the mutant spectrum in the G108 template (Figure 2). In the T108 spectrum, a novel TPu junction is created at A109. Consistent with the association of frameshift hotspots with this sequence, a new hotspot is found at A109. The A109 deletion was not detected in any other template tested. The deletions at A109 represents 20% of the mutants in the 86–112 region (or 4% of all mutants) sequenced in the M74-T108 template. This frequency of deletions at A109 is comparable to the frequencies of 1-bp deletions at C108 or A108 in the G108 and A108 templates, respectively. This observation suggests that TPu context is sufficient to define a high frequency deletion site; the identity of the purine appears to be unimportant in this region. In addition, the similar frequency of A109 or A109 deletions in the A108 or T108 templates, respectively, demonstrates that repeats of two purines do not detectably stimulate deletions at this site.

E. coli DNA polymerase I also produced A109 deletions on the M74-T108 template. After polymerization by PolI, A109 was found twice among 25 blue plaques sampled from M74-T108, but was not found at all among the 64 blue plaques sequenced from the other templates.

The absence of A109 deletions in M74-C108 demonstrates that deletions opposite A109 are not enhanced by a 5°C and supports the view that the template context TPu is preferred for spontaneous single base deletion

<table>
<thead>
<tr>
<th>Mutant assays</th>
<th>G108</th>
<th>A108</th>
<th>C108</th>
<th>T108</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total plaques examined (× 10⁶)</td>
<td>1.6</td>
<td>1.5</td>
<td>1.8</td>
<td>3.7</td>
</tr>
<tr>
<td>Total mutants (blue plaques)</td>
<td>170</td>
<td>184</td>
<td>239</td>
<td>171</td>
</tr>
<tr>
<td>Total mutant frequency (× 10⁶)</td>
<td>106</td>
<td>123</td>
<td>133</td>
<td>46</td>
</tr>
<tr>
<td>Total mutants in 86–112 region</td>
<td>34</td>
<td>35</td>
<td>54</td>
<td>37</td>
</tr>
<tr>
<td>Single base deletions in 86–112 region</td>
<td>31</td>
<td>31</td>
<td>49</td>
<td>35</td>
</tr>
<tr>
<td>Other mutants in 86–112 region</td>
<td>1 Del 1</td>
<td>1 Del 4</td>
<td>1 Del 1</td>
<td>1 Del 2</td>
</tr>
<tr>
<td>2 C₉₅ → T</td>
<td>1 C₉₅ → T</td>
<td>1 Del 5</td>
<td>1 Del 6</td>
<td></td>
</tr>
<tr>
<td>1 C₉₅ → T</td>
<td>1 Del 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 C₉₅ → T</td>
<td>1 Del 7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 C₉₅ → T</td>
<td>1 Del 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 C₉₅ → T</td>
<td>1 Del 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction of single base deletions in 86–112 region among total blue plaques (%)</td>
<td>18</td>
<td>17</td>
<td>21</td>
<td>20</td>
</tr>
</tbody>
</table>

* Average plaques per plate were 0.5–1 × 10⁶.
* Determined by oligonucleotide probing.
* The sequences of Deletions (Del) and Complex frameshifts (Cmplx) are shown in Figure 3. Base pair changes are indicated by position number and sequence change in the template strand.
DNA Contexts of Polymerase Errors

Figure 2.—Spectra of 1-bp deletions in templates G₁₀₈, A₁₀₈, T₁₀₈, and C₁₀₈. The template strand is illustrated. Each of the four templates differs by only a single base located at position 108. Mutants were mapped to this region by oligonucleotide hybridization. Each mutant mapping to the region was sequenced. The position of each deletion mutant is indicated by a ■. When a deleted base occurs in the midst of identical bases, DNA sequence does not identify which base has actually been deleted. In this figure the ■ is consistently placed at the 5' most position in the template where the deletion could have occurred.

sites over the more general PyPu designation, consistent with our earlier studies.

Deletions at C₁₀₈ are more rare than are A₁₀₈ or G₁₀₈ deletions, consistent with the absence of a purine at the deletion site. The run of three consecutive Ts in T₁₀₈ precludes assignment of a 1-bp deletion frequency specifically to T₁₀₈. However, 1-bp deletions at T₁₀₈ are probably not higher than at C₁₀₈ because the sum of mutants at sites between 106 and 108 in both the C₁₀₈ and T₁₀₈ templates is the same and the calculated frequency is 2-fold lower in the T₁₀₈ template (8/35 × 0.46 × 10⁻⁴) than in the C₁₀₈ template (8/49 × 1.33 × 10⁻⁴) (see Table 1). This example further supports the conclusion of our earlier studies (DE BOER and RIPLEY 1988) that modest monotonic base run are not hotspots for 1-bp deletions during in vitro DNA polymerization by Klenow fragment DNA polymerase.

A novel frameshift hotspot: In vitro polymerization on the M74-C₁₀₈ template produced a novel single base deletion hotspot at A₉₄, 14 bp away from the C₁₀₈ site.
Despite the distance, the hotspot clearly depends on the C108 context because the deletion is infrequent in each of the spectra from other templates. The new hotspot is accounted for by a quasi-palindromic DNA misalignment whose frequency is influenced by the C108 site. Figure 4 demonstrates how a quasi-palindromic fold-back misalignment explains this specific single base deletion. The explanation accounts for both the precise sequence of the mutant and accounts for the specific association of the hotspot with the M74-C108 template. Only the C108 template increases the complementary pairing participating in the DNA misalignment that precedes mutagenesis.

We have previously shown that quasi-palindromic misalignments account for complex DNA sequence changes arising during *in vitro* DNA polymerization.
reactions. We had predicted that this mechanism occasionally produced 1-bp deletions as well, based on the observation that a single specific 1-bp deletion site \([G_{15}]\) in M74 occurred at different frequencies during polymerization by *E. coli* DNA polymerase I or its Klenow fragment. This hypothesis was based on the synthesis of two observations: (1) the mutant sequence was consistent with a local quasi-palindromic misalignment mechanism and (2) mutations depending on quasi-palindromic misalignment often display different frequencies in comparisons of the two enzymes while single base deletions usually do not (PAPANICOLAOU and RIPLEY 1989, 1991). Thus, we examined the influence of the template sequence changes at the 108 site on DNA polymerase pausing. As can be seen in Figure 5, the DNA sequence changes at 108 produced no detectible change in pausing under the conditions of the mutagenesis experiments. Thus, the mechanism by which *C*,\(_{108}\) promotes misalignment does not appear to include an influence of *C*,\(_{108}\) on DNA polymerase processivity and is thus probably due to the additional complementary pairing during misalignment provided by the *C*,\(_{108}\) base change.

Other mutant sequences in the 86–112 region: A small number of DNA sequence changes were found in the 86–112 region that were not 1-bp deletions. The full repertoire of changes previously associated with *in vitro* DNA polymerization errors were found; namely, base pair substitutions, multibase deletions and complex frameshifts.

**Base substitutions:** We have previously seen light blue plaque phenotypes associated with base substitutions. Here, we found C\(_{95} \rightarrow T\), three times; twice in the
Figure 5.—DNA polymerase pausing on C\textsubscript{108}, A\textsubscript{108}, C\textsubscript{108}, and T\textsubscript{108} templates. Transient pausing of DNA polymerase during extension reactions similar to those used to create mutations was measured. No influence of the identity of the base at the 108 site is seen on polymerase pausing in the DNA regions in which mutagenesis is measured. Panel A shows the C\textsubscript{108} template. Panel B shows pausing on the remaining templates. Dideoxy DNA sequencing patterns were used to mark the positions of pausing in both panels. The M74-C\textsubscript{108} template was used in both cases and synthesis was directed by the same labeled primer as used in the pausing experiment.

G\textsubscript{108} template and once in the A\textsubscript{108} template. The substitution creates an in-frame termination codon, and we believe that out-of-frame translational reinitiation accounts for the blue plaque phenotype (PAPANICOLAOU and RIPLEY 1989). A G\textsubscript{98} → T transversion was also found to produce a light blue phenotype in the A\textsubscript{108} template. This change does not produce a termination codon, and the molecular basis of its phenotype is unknown. Similarly, an adjacent transversion T\textsubscript{98} → A was found to produce a light blue phenotype in the G\textsubscript{108} template (PAPANICOLAOU and RIPLEY 1989).

**Deletions:** Deletions between direct repeats are common after *E. coli* DNA polymerase I extensions, but have been found more rarely after Klenow extensions. Six deletion sequences that include at least part of the 86–112 region were sequenced in this study and are shown in Figure 3. Some deletions have end points associated with direct repeats and can thus be explained by primer misalignments; for example, deletion 1, found after replication on both the G and the C templates. Deletion 2 can also be explained by this model, but the deleted sequence suggests that the misalignment from the upstream sequence to the downstream position, the direction opposite that predicted by primer misalignment during *in vitro* polymerization, but fully consistent with misalignments arising *in vivo*. Deletions 3–5 are imprecisely associated with imperfect direct repeats and deletion 6 is imprecisely connected to only a 2-bp repeat. Deletions with these characteristics have been quite frequently seen in studies of deletion mutagenesis in mammalian cells, but have been rarely seen in procaryotes (MEUTH 1990). The low numbers of these mutations still provides no real insight into how they arise. Their recovery in these experiments was completely dependent on Klenow fragment rather than PolI polymerase synthesis. Deletions recovered after PolI synthesis on these templates (data not shown) had endpoints associated with direct repeats that could account for their sequences and were like those we have previously reported (PAPANICOLAOU and RIPLEY 1989).

**Complex frameshifts:** Two complex mutations were found in the probed region. The sequences are shown in Figure 3. Mutant 1, arose in the M74-C\textsubscript{108} template and is an interesting variation on the b4 hotspot deletions associated with the M74-C\textsubscript{108} template. In the complex 1 (Cmplx 1) mutant the b4 deletion is accompanied by four non-adjacent base pair substitutions, covering a distance of 8 bp. These substitutions reflect an inversion of the entire local sequence in which they lie. The A\textsubscript{94} deletion and the inversion are explained as a single mutational event involving templated DNA synthesis initiated by a quasi-palindromic DNA misalignment. The sequence change cannot be accounted for by the fold-back misalignment described in Figure 4, but is instead explained by a strand-switch misalignment shown in Figure 6.

The other complex frameshift sequenced (Cmplx 2, Figure 3) may be explained by a fold-back palindromic misalignment. The inserted bases can be accounted for by misalignment, but the misaligned synthesis provides no convincing explanation for the accompanying deletion specificity which would depend on no more than 1 bp of complementary pairing. The mutation may be explained by a base substitution G\textsubscript{94} → C followed by a similarly unexplained deletion.
DNA Contexts of Polymerase Errors

A

\[
\begin{array}{cccc}
94 & 95 & 100 & 121 \\
S' - TGATTCAGCAAGGGTGCATGCTGCCGACCCAG & 3' - GTGGGACGGGGTC \\
A - ACTAA & C - GAGGGTCAGCCGACCCAG \\
G - TCCAGATAGCCAAGTGGGACGGGGTC \\
\end{array}
\]

\[\downarrow\]

DNA strand-switch

B

\[
\begin{array}{cccc}
94 & 95 & 100 & 121 \\
S' - TGA & A - T & T & C - C \\
G - C & G & G & T \\
A - T & A & G & C \\
C - G & G & G & T \\
C - G & G & G & T \\
T - A & T - A & T - A & T \\
A - G & G & G & T \\
C - G & G & G & T \\
C - G & G & G & T \\
A - T & A - T & A - T & A \\
A - T & A - T & A - T & A \\
G - C & A - C & C & G \\
G - C & A - C & C & G \\
G - C & A - C & C & G \\
A - T & A - T & A - T & A \\
T - G & T - G & T - G & T \\
\end{array}
\]

Misaligned DNA polymerization

\[\downarrow\]

Realignment to the template

C

\[
\begin{array}{cccc}
94 & 95 & 100 & 121 \\
S' - TGA & A - T & T & C - C \\
G - C & G & G & T \\
A - T & A & G & C \\
C - G & G & G & T \\
C - G & G & G & T \\
T - A & T - A & T - A & T \\
A - G & G & G & T \\
C - G & G & G & T \\
C - G & G & G & T \\
A - T & A - T & A - T & A \\
A - T & A - T & A - T & A \\
G - C & A - C & C & G \\
G - C & A - C & C & G \\
G - C & A - C & C & G \\
A - T & A - T & A - T & A \\
C - G & A - G & A - G & T \\
\end{array}
\]

GTTGGACGGGGTC

Mutant sequence in new DNA strand

\[\downarrow\]

Mutant sequence in new DNA strand

D

\[
\begin{array}{cccc}
94 & 95 & 100 & 121 \\
S' - TGATTCAGCAAGGGTGCATGCTGCCGACCCAG & 3' - GTGGGACGGGGTC \\
A - ACTAA & C - GAGGGTCAGCCGACCCAG \\
G - TCCAGATAGCCAAGTGGGACGGGGTC \\
\end{array}
\]

Mutant sequence in new DNA strand

Figure 6.—Inversion mutagenesis mediated by polymerase strand-switching. A shows the starting template sequence containing C\textsubscript{94} which has been copied around the entire genome of M74 and is now polymerization continues via strand displacement. The displaced strand is shown in bold, the newest synthesis is shown in italics. B describes a strand-switch misalignment that brings DNA initially extended to the 108 position by correctly aligned polymerization into juxtaposition with the template that will account for the mutant sequence. C shows the ectopic DNA polymerization step that produces the mutant DNA sequence. Ectopically templated DNA is indicated in lower case. Realignment of this DNA to the template in D places the mutant sequence in its final context. A further round of replication of this mutational heterozygote (not shown) produces a complete mutant. The deletion of A at 94 is indicated by a dash, the concerted base pair substitutions, the consequence of the sequence inversion, are enlarged to emphasize their positions. The deletion and all 4 base substitutions are the consequence of a single misalignment events and require no mispairing by DNA polymerase, only the incorrect alignment of the primer on the DNA template.
CONCLUSIONS

The consensus DNA context for frequent 1-bp deletions produced during in vitro polymerization by E. coli PolI Klenow polymerase successfully predicted the effects of changing a base at a single deletion hotspot site to each of the other three bases. Most impressively, the frequency of deletions opposite template A in the contexts, 5'-TTTA-3' and 5'-TTAA-5' was similar to that observed at G in the original 5'-TTGA-3' context, but deletions opposite template C in the 5'-TTCA-3' context were reduced. Thus, the TPu junction sequence was consistently associated with high 1-bp deletion frequencies. Neither the identity of the purine (A or G) nor the theoretical opportunity for slippage in the 5'-TTAA-3' context influenced the deletion frequency. The important contribution of the 5' T on mutant frequency in the TPu context was strongly confirmed by the absence of A100 deletions in the 5'-TTCA-3' and 5'-TTGA-3' contexts.

The large increase in single base deletions occurring at A94 associated with the C106 template is well explained by misaligned synthesis mediated by a quasi-palindromic DNA misalignment. The increased mutation frequency may be ascribed to extended complementarity to mediate misaligned pairing in the C106 template, a feature expected to increase the concentration of misaligned DNA intermediates and/or to increase the ability of the polymerase to extend them. It is reasonable to propose that the low frequencies of A94 deletions in the other templates are due to less efficient formation or processing of the same misalignment. Indeed, quasipalindromic misalignments should be more seriously considered as potential explanations for 1-bp deletions in contexts that are not TPu. For example, the loss of one base, opposite G96-97 [5'-AGGTG-3'] might depend on a fold-back misalignment to the complementary DNA sequence opposite positions 127-130 [5'-CACT-3']. Such distant relationships should be readily uncovered using templates bearing base substitutions in the suspected template regions.

Palindromic misalignments have been previously demonstrated to produce complicated DNA sequence changes during polymerization reactions mediated by E. coli Klenow fragment (Papanicolaou and Ripley 1989) and Yeast PolII (Kunkel et al. 1989). Here the same class of misalignments is shown to produce 1-bp deletions in contexts that are not TPu. For example, the loss of one base, opposite G96-97 [5'-AGGTG-3'] might depend on a fold-back misalignment to the complementary DNA sequence opposite positions 127-130 [5'-CACT-3']. Such distant relationships should be readily uncovered using templates bearing base substitutions in the suspected template regions.

It has been recently noticed, that during polymerization on damaged DNA templates, the Klenow polymerase produces specific double mutations whose frequencies are far higher than would be predicted on the basis of the occurrence of independent single mutations (D. Feig and L. Loeb, personal communication). In some cases, the sequences of these double mutants are consistent with the idea that an initial error during polymerization produces a new strand bearing a change which would have an improved ability to form a fold-back misalignment that would template the second change. The substantial increase in single base deletions due to a single base change at C106 observed in this study suggests that the frequency of secondary misalignment could be substantially enhanced by a single base change. Moreover, potentially large contributions by misalignment-mediated mutagenesis mechanisms may be expected during in vitro polymerization on damaged templates in view of the fact that DNA damage is expected to interfere with processive DNA polymerization and the demonstration that DNA polymerase pause sites correlate strongly to the most frequent misalignment mutations (Papanicolaou and Ripley 1991).

Altogether this study has refined our view of the mechanisms contributing to 1-bp deletions during polymerization by the E. coli Klenow polymerase. One mechanism is clearly associated with TPu template contexts and fails to operate at a detectable level in a CPu context at the same site. Although these studies cannot per se address whether this context preference reflects misincorporation and/or misalignment properties of Klenow DNA polymerase at this site, it provides the contextual rules that must be explained by context-specific polymerase behavior. A broad exploration of DNA sequence effects on Klenow fragment-mediated misincorporation and mismatch extension reactions by Joyce et al. (1992) demonstrated large effects of context on polymerization behavior. For example, Klenow polymerase directed highly variable, site-specific patterns of misinsertions at template G sites, but usually directed the rapid misinsertion of A at template A sites. Perhaps parallel measurements of the influence of context on both the biochemical and mutational characteristics of polymerase activity could be used to define the site-specific polymerization events associated with polymerase-mediated single base deletions.

Some of the templates used in the work reported were constructed in our laboratory by Zuwei Qian and Wei Li. We thank Catherine Papanicolaou for her advice. Support for this work is provided by grant CN-50 from the American Cancer Society and by grant MCB-9305475 from the National Science Foundation.

LITERATURE CITED


