Dissecting the Fidelity of Bacteriophage RB69 DNA Polymerase: Site-Specific Modulation of Fidelity by Polymerase Accessory Proteins

Anna Bebenek,*,†,[1] Geraldine T. Carver,* Holly Kloos Dressman,*,[2] Farid A. Kadyrov,* Joseph K. Haseman,‡ Vasiliy Petrov,§ William H. Konigsberg, Jim D. Karam§ and John W. Drake*,[3]**

**Laboratory of Molecular Genetics and ‡Biometry Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709-2233, †Institute of Biochemistry and Biophysics, Polish Academy of Science, 02-106 Warsaw, Poland, §Department of Biochemistry, Tulane University Health Sciences Center, New Orleans, Louisiana 70112-2699 and **Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06510-3219*

Manuscript received May 9, 2002
Accepted for publication August 6, 2002

ABSTRACT

Bacteriophage RB69 encodes a replicative B-family DNA polymerase (RB69 gp43) with an associated proofreading 3′ exonuclease. Crystal structures have been determined for this enzyme with and without DNA substrates. We previously described the mutation rates and kinds of mutations produced *in vivo* by the wild-type (Pol^+ Exo^+) enzyme, an exonuclease-deficient mutator variant (Pol^+ Exo^-), mutator variants with substitutions at Tyr^{567} in the polymerase active site (Pol^M Exo^+), and the double mutator Pol^M Exo^-. Comparing the mutational spectra of the Pol^+ Exo^- and Pol^+ Exo^+ enzymes revealed the patterns and efficiencies of proofreading, while Tyr^{567} was identified as an important determinant of base-selection fidelity. Here, we sought to determine how well the fidelities of the same enzymes are reflected *in vitro*. Compared to their behavior *in vivo*, the three mutator polymerases exhibited modestly higher mutation rates *in vitro* and their mutational predilections were also somewhat different. Although the RB69 gp43 accessory proteins exerted little or no effect on total mutation rates *in vitro*, they strongly affected mutation rates at many specific sites, increasing some rates and decreasing others.

THE DNA polymerase (gp43, encoded by gene *43*) of the T4-related phages is responsible for the replication of the ~170-kb phage genome (Karam and Konigsberg 2000). Lacking DNA mismatch repair, these phages achieve the high fidelity of replication, $\sim 2 \times 10^{-8}$/bp, through the combined polymerase (Pol) and exonuclease (Exo) activities of gp43 (Dressman *et al.* 1997; Bebenek *et al.* 2001). Like T4 gp43 and the gp43-like DNA polymerases of some archaeons, gp43 of the T4-related coliphage RB69 is a member of the B-family (Polα-like) polymerases, which include the eukaryotic replicative polymerases -α, -δ, and -ε (Braithwaite and Ito 1993). The crystal structure of RB69 gp43 shows five distinct domains: N (N terminal), palm (seat of the Pol catalytic center), fingers, thumb, and Exo (seat of the exonuclease catalytic activity; Wang *et al.* 1997). The recently reported structures of three B-family polymerases from archaea show marked similarities to RB69 gp43 (Hopfner *et al.* 1999; Rodriguez *et al.* 2000; Hashimoto *et al.* 2001). Two additional RB69 gp43 crystallographic structures are now available: a complex with DNA with the primer terminus occupying the Exo site (Shamoo and Steitz 1999) and a ternary complex with partially double-stranded template-primer DNA and an incoming dNTP (Franklin *et al.* 2001). The three structures provide a basis for testing structure-function relationships in the fidelity of this polymerase.

As with other DNA polymerases, the Pol catalytic site of RB69 gp43 is composed of residues in the palm domain, thumb domain (which binds primer-duplex DNA), and fingers domain (which contains most of the side chains that bind the incoming nucleotide; Franklin *et al.* 2001). In unliganded gp43, the enzyme displays an open conformation. After binding a primer-template and a nucleotide, the resulting ternary complex displays a rotated fingers domain and changes in the conformation of the template DNA. This liganded state is referred to as the closed conformation (Franklin *et al.* 2001). Nearly all crystallographic structures for polymerases in such ternary complexes show the same rotation of the fingers domain toward the polymerase catalytic center (reviewed in Kunkel and Bebenek 2000). When the fingers close down completely, a tight binding pocket is formed around the nascent base pair. The geometry of this pocket closely accommodates only a base pair

[1]*Present address:* Institute of Biochemistry and Biophysics, Polish Academy of Science, 02-106 Warsaw, Poland.

[2]*Present address:* Department of Genetics, Box 3054, Duke University Medical Center, Durham, NC 27710.

[3]*Corresponding author:* Laboratory of Molecular Genetics, Room E-344, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709-2233.
E-mail: drake@niehs.nih.gov

with proper geometry and thus discriminates against mispairs. Therefore, although hydrogen bonding contributes strongly to base selection, geometry alone can provide reasonably correct nucleotide selection (Goodman 1997; Moran *et al.* 1997). In addition, the polymerase is envisioned to test the geometry of the nascent base pair through specific hydrogen bonding between the enzyme and the pyrimidine O2 or purine N3 atoms in the minor groove of duplex DNA. Chemistry to produce primer extension depends strongly on the hydrogen-bonding potential of the 3′-terminal nucleotide (Morales and Kool 2000) and primer extension may provide the signal for transferring the primer terminus to the Exo catalytic site for proofreading (Johnson 1993). Both structural and biochemical studies suggest that the switch from primer extension to proofreading involves a conformational transition in the enzyme from a closed to an open (or Exo) mode (Doublié *et al.* 1998; Shamoo and Steitz 1999).

Tyr^{567} in the crystal structure of closed RB69 gp43 is positioned near the Pol catalytic site at the end of helix P of the fingers domain. This tyrosine is conserved in all B-family polymerases and is a part of the highly conserved region III, KXXXNXYG (Wang *et al.* 1997). The closed RB69 gp43 ternary structure suggests that Tyr^{567} makes a water-mediated hydrogen bond with a minor-groove acceptor on the template base of the terminal primer-template base pair. Thus, Tyr^{567} is proposed to be a part of the Pol binding pocket and to be an important residue for sensing proper geometry in a newly formed base pair (Franklin *et al.* 2001). We recently reported that substitution of Tyr^{567} with any of several amino acid residues having smaller side chains dramatically increases base-pair substitution errors *in vivo* but has little effect on frameshift mutagenesis (Bebenek *et al.* 2001). Kinetic measurements *in vitro* showed that the Y567A substitution (Pol^M enzyme) increases the rate of formation of some mispairs, but at the same time reduces some rates of mispair extension.

In the same study we examined the fidelity of an RB69 gp43 whose Exo activity was inactivated by two alanine substitutions at critical metal-coordinating Asp residues. The Exo^- mutant exhibited the expected strongly increased mutation rate and promoted several kinds of mutations, including base-pair substitutions and additions and deletions of single base pairs. The Pol^M and the Exo^- mutators each increased mutation rates *in vivo* by ~500-fold. We also observed that the Pol^M Exo^- double mutator increased average mutation rates by ~2300-fold (Bebenek *et al.* 2001). This result sharply contradicts the expectations of one simple model in which the Pol and Exo fidelity contributions operate independently, where the outcome would be a multiplicative rather than an additive increase. This observation and other evidence indicate that the Pol and Exo functions of RB69 gp43 interact with each other, but the mechanism of this interaction remains unexplained.

In the studies reported here, we explored the extent to which the fidelities of the same RB69 gp43 mutant enzymes measured *in vitro* mirror those determined previously *in vivo*. We used the well-described M13mp2 *lacZ*α *in vitro* system (Bebenek and Kunkel 1995). This system monitors all 12 possible base mispairs and many kinds of base addition and deletion mutations in a variety of nucleotide sequence contexts and has been used to characterize the fidelities of numerous wild-type and mutant DNA polymerases (Kunkel and Bebenek 2000). Our study had two major aims. First, we sought to determine how well the rates and patterns of mutagenesis ascertained *in vitro* reflect the same parameters ascertained *in vivo*. This important comparison, which has been performed only infrequently with other polymerases, assesses the advantages and drawbacks of the two types of approach to measuring mutation rates and specificities. Second, we sought to measure the effects of the polymerase accessory proteins upon fidelity. Such measurements have been performed a few times before but only for the overall rate of mutation in a reporter gene or for rates of reversion at a small number of sites, rather than for rates at a large number of specific DNA sites.

Gp43 is poorly processive (Das and Fujimura 1979; Newport *et al.* 1980) except in the presence of the phage-induced sliding-clamp (a gp45 trimer), clamp loader (a 4:1 complex of gp44:gp62), and the ssDNA-binding protein (SSB, gp32; Young *et al.* 1992; Nossal 1994). In contrast to our studies *in vivo*, the *in vitro* approach allowed us to compare the replication fidelities of different RB69 gp43 mutants acting either alone or in the presence of gp32, gp44/gp62, and gp45.

Here we report that the fidelities of the wild-type and mutator RB69 gp43's measured *in vitro* only partially resemble their fidelities measured *in vivo*. Error rates are generally somewhat higher *in vitro* than *in vivo*. Both the Pol^+ Exo^- and Pol^M Exo^+ enzymes displayed mutator activities *in vitro* and, as *in vivo*, the Pol^M Exo^- enzyme displayed only modestly higher mutator activity than that of either of its components. The Pol^+ Exo^- polymerase produced diverse base substitutions, single-base deletions, and large deletions. The Pol^M Exo^+ polymerase produced mostly base substitutions, but, unlike *in vivo*, it also produced a substantial number of single-base deletions. The double-mutator Pol^M Exo^- polymerase produced mostly base substitutions, but of types not predicted by its component mutators acting alone. In addition to producing single-base deletions, Pol^M Exo^- also produced single-base additions, again unlike either of its component mutators. While the accessory proteins exhibited little or no impact on total mutation rates, they substantially increased mutation rates at some sites while decreasing rates at other sites. In addition to offering potential insights into factors that may determine the fidelity of DNA synthesis by gp43, these results indi-

cate that fidelity analyses conducted *in vitro* may not agree fully with analyses conducted *in vivo*.

MATERIALS AND METHODS

RB69 DNA polymerases and accessory proteins: The production and purification of wild-type RB69 gp43 (Pol^+ Exo^+), D222A/D327A gp43 (Pol^+ Exo^-), Y567A gp43 (Pol^M Exo^+), and Y567A/D222A/D327A gp43 (Pol^M Exo^-) have been described (Bebenek *et al.* 2001). The RB69 accessory proteins gp45 and gp44/62 were prepared as described (Yeh 1998; Yeh *et al.* 1998). T4 gp32, which can support RB69 DNA replication (Borjac 1997), was prepared as described by Kadyrov and Drake (2001).

Gap-filling synthesis and forward-mutation assay: Gap-filling reactions were performed with phage M13mp2 *lacZ*α gapped substrates prepared as described by Bebenek and Kunkel (1995). Reactions with polymerase alone (25 μl) contained 1.6 nm DNA substrate, 25 mm Tris acetate, 10 mm magnesium acetate, 150 mm potassium acetate, 2 mm dithiothreitol, and 40 nm gp43. Reactions were started by adding all four dNTP's to final concentrations of 1 mm each. The mixtures were incubated for 15 min at 37° and were terminated by adding ethylenediamine tetraacetic acid (EDTA) to 15 mm.

Reactions with polymerase plus accessory proteins were conducted in the same mixture supplemented with 1 mm ATP, 6 nM gp44/gp62 complex, 80 nM gp45 (as a monomer) and 5.6–11.2 μm gp32. Reaction mixtures were preincubated for 30 sec at 37° to allow the accessory proteins to assemble on the gapped DNA substrate before polymerase was added to 20 nm. This mixture was incubated at 37° for another 30 sec to allow the polymerase to bind to the nascent replication complex and primer extension was then started by adding all four dNTP's to final concentrations of 1 mm each. After 10 min at 37° reactions were terminated by adding EDTA to 15 mm. A 20-μl sample from each reaction was analyzed on a 0.8% agarose gel to verify the extent of gap filling. The remaining 5 μl was diluted with 100 μl of distilled water and centrifuged twice through Microcon 30 columns, the DNA being recovered in a volume of 10–50 μl. Small portions (1–5 μl) of this DNA were subsequently electroporated into competent *Escherichia coli* MC1061 cells that were then plated to form M13 plaques. All plaques were counted and putative mutant plaques were scored as white or light blue against a background of dark-blue nonmutant plaques. Apparent M13 *lacZ*α mutants were picked and replated to verify their phenotypes and single isolated mutant plaques were resuspended in 100 μl of water. A representative set of mutants from each collection was sequenced to obtain mutational spectra.

Mutation-frequency and spectral data were usually compiled from several reactions, frequently with several transfections per reaction. Mainly because of small sample sizes, mutation frequencies usually varied by a fewfold (typically twofold or less) among samples for a given polymerase. Where appropriate, we call attention to mutation-frequency differences that we judge to be not strongly reliable due to underlying variation or small numbers of mutants scored.

DNA sequencing: Prior to sequencing a mutant, the entire *lacZ*α gene was amplified by the polymerase chain reaction (PCR) using two 20-mer primers for PCR purchased from Research Genetics (Birmingham, AL). The forward primer was 5′-TAAGGGATTTTGCCGATTTC, the reverse primer was 5′-CAGTTTGAGGGGACGACGAC, and the length of the final product was 642 bp. The PCR consisted of 30 cycles of 1 min at 94°, 1 min at 58°, 1 min at 72°, with a final extension time of 10 min at 72° using Taq large-fragment polymerase (Display System Biotech TAQFL from PGC Scientifics). PCR products were purified with the QIAGEN (Valencia, CA) PCR purification kit. Sequencing was performed using the same primers and an ABI Prizm 377 automatic sequencer and dRhodamine terminator cycle sequencing kits (PE Applied Biosystems, Foster City, CA). Each mutation was identified by sequencing in both directions. When mutational spectra were constructed, only mutations known from historical experience to produce a mutant phenotype were included. Although silent mutations appearing singly are exceedingly rare, they are more common in mutants that also bear a phenotypically detectable mutation. The detectable base substitutions comprise those described by Bebenek and Kunkel (1995) as well as three others discovered by us (retaining the criterion that such a mutation must be observed as a single event at least twice): G → C at 178, T → C at 183, and C → A at 189 using the numbering system of Figure 1.

Statistical methods: To determine whether the distribution of mutations across the 274 *lacZ*α bases displayed in the mutational spectra was similar in reactions with or without the polymerase accessory proteins, a chi-square statistic was calculated on the basis of the observed distribution of the mutations. However, because of the low expected mutation frequencies at each site, this statistic was not compared to standard chi-square tables. Instead, using the approach suggested by Piegorsch and Bailer (1994), a Monte Carlo simulation study was performed, conditional upon the observed totals of mutations from reactions with or without accessory proteins as well as upon the total number of mutations at each site. The overall significance of the difference between reactions with or without accessory proteins with regard to the observed pattern of mutation frequencies was determined by comparing the observed chi-square statistic to chi-square statistics from 5000 simulated data sets. Pairwise comparisons of mutation frequencies at each site were made by two-sided Fisher's exact tests.

RESULTS

Fidelities of RB69 DNA polymerase mutants: Having characterized the fidelities of several variant DNA polymerases *in vivo* (Bebenek *et al.* 2001), we proceeded to conduct parallel analyses *in vitro* by measuring the accuracy of DNA replication during the filling of a 407-nucleotide (nt) single-stranded gap in bacteriophage M13mp2 RF DNA. The gap contained the mutation-reporter *lacZ*α as a template for DNA synthesis (Bebenek and Kunkel 1995). This forward-mutation assay allows the detection of many base substitutions, most single-base additions and deletions, and many more complex mutations. Mutants are detected as light-blue or colorless M13 plaques against a background of dark-blue nonmutant plaques. The gap was filled by either the wild-type or a mutant RB69 gp43 in reactions either with or without the gp43 accessory proteins. T4 gp43 is poorly processive in the absence of the accessory proteins (Das and Fujimura 1979; Newport *et al.* 1980). Because RB69 gp43 differs from its T4 counterpart at ~39% of its residues (Wang *et al.* 1995), we compared the processivities of the two proteins *in vitro*. Both enzymes displayed identical limited processivity, the predominant products being about 20 nt long but with some products of 100–140 nt (data not shown).

For each of the four gp43 constructs studied here,

TABLE 1

Mutation frequencies (MF) and rates (μ) for wild-type and mutant RB69 DNA polymerases *in vitro* and *in vivo*

Polymerase	APs[a]	Total plaques	Mutant plaques	Correction factor[b]	MF[c] × 10^4	$\mu_{lacZ\alpha} \times 10^{4}$[d]	$\mu_{rI} \times 10^{4}$[e]
Pol^+ Exo^+	−	54,411	52	0.88	8.4		
	+	53,197	22	0.95	3.9		
	±	107,608	74	0.90	6.2	?	0.043
Pol^+ Exo^-	−	91,799	276	0.95	28.6		
	+	82,775	219	0.96	25.5		
	±	174,574	495	0.96	27.1	35	22
Pol^M Exo^+	−	32,621	181	0.99	54.9		
	+	96,336	648	1.04	69.6		
	±	128,957	829	1.03	65.9	100	21
Pol^M Exo^-	−	57,952	883	1.08	164.1	263	
	+	71,394	1,556	1.13	246.6	401	
	±					263–401	74–120

[a] Reaction conducted without (−) or with (+) accessory proteins (APs), followed by the data combined for both conditions (±) in those cases where the difference with and without APs is too small to be significant.

[b] Factor based on sequencing. Occasional light-blue mutants had no sequence change within the 293-bp *lacZ*α sequence, while some mutants had multiple detectable mutations.

[c] Mutation frequency = (mutant plaques)(correction factor)/(total plaques).

[d] [(mutator MF) − (wild-type MF)] divided by 0.6 to adjust for the loss of mutational heteroduplexes (Bebenek and Kunkel 1995). Because there is no significant signal above the historical background of this assay for the Pol^+ Exo^+ gp43, no rate can be estimated.

[e] From Bebenek *et al.* (2001) for the T4 *rI* reporter sequence *in vivo*, which is similar in size to the *lacZ*α reporter sequence used here (see text).

Table 1 describes mutation rates, Table 2 describes the kinds of mutations produced, and Figure 1 shows the mutational spectra. In the spectra, certain sites are highlighted because the accessory proteins either enhanced or inhibited mutagenesis at those sites; these are necessarily mutation-rich sites where significant differences between numbers of mutations produced with and without accessory proteins can be recognized. Error rates for specific classes of mutations can be uniquely well estimated in the *lacZ*α system because previous studies have established which changes at each site yield detectable mutations (Bebenek and Kunkel 1995). Thus, error rates for each specific kind of mutation can be normalized to the number of informative targets, and these rates are provided in Tables 5 and 6.

Below, we show results suggesting that the mutations in the Pol^+ Exo^+ (wild-type gp43) reaction products were produced during the growth of the M13 phage vector and the processing of the M13 DNA template and not by the RB69 polymerase itself. This background contributes modestly to mutations generated by the Pol^+ Exo^- polymerase, slightly to those produced by the Pol^M Exo^+ polymerase, and insignificantly to those produced by the Pol^M Exo^- polymerase. Significant contributions will be noted at the appropriate places. When we estimated such contributions, we pooled the Pol^+ Exo^+ results with and without accessory proteins because most of the mutations preceded the reactions *in vitro*.

Fidelity of the Pol^+ Exo^+ RB69 DNA polymerase: The wild-type RB69 and T4 DNA polymerases were previously shown to be so accurate in copying *lacZ*α that their mutation frequencies ($3–5 \times 10^{-4}$) were at or slightly below the historical background frequency ($5–7 \times 10^{-4}$) for unfilled template DNA (Dressman *et al.* 1997). This background frequency simply reflects the intrinsic average mutation rate per base pair for phage M13 replication (Drake 1991), but it can hinder analyses of the fidelities of very accurate polymerases. A similar result was encountered here: In reactions either with or without accessory proteins, the mutation frequency was similar to or indistinguishable from the historical background (Table 1). We do not consider the apparent difference between the mutation frequency without ($\sim 8 \times 10^{-4}$) and with ($\sim 4 \times 10^{-4}$) accessory proteins to be reliable because of the small numbers of mutants scored. The kinds of mutations observed (Table 2) and their spectra (Figure 1A) are indistinguishable in the presence or absence of accessory proteins, as expected from mutations that already existed in the template DNA prior to the reactions conducted *in vitro*. Among the 43 sequenced mutants, 26 were C → T transi-

TABLE 2

Kinds of mutations produced *in vitro* by wild-type and mutant RB69 DNA polymerases without and with accessory proteins

	Pol^+ Exo^+		Pol^+ Exo^-		Pol^M Exo^+		Pol^M Exo^-	
Mutation	−AP	+AP	−AP	+AP	−AP	+AP	−AP	+AP
Total	22	21	159	177	147	133	154	172
G → A	0	0	8	8	7	15	54	56
C → T	13	13	45	58	17	8	12	21
A → G	0	0	1	4	2	0	0	0
T → C	1	0	13	24	76	72	37	68
G → T	3	1	15	7	1	5	2	0
G → C	0	0	7	10	2	2	4	0
C → A	0	0	1	6	1	0	0	3
C → G	0	0	1	2	0	0	0	0
A → T	0	0	1	6	2	1	2	1
A → C	0	0	1	1	0	0	1	0
T → A	1	0	0	6	2	3	8	3
T → G	0	0	0	5	0	3	0	1
Transitions	14	13	67	94	102	95	103	145
Transversions	4	1	26	43	8	14	17	8
−1	2	3	40	32	37	23	17	14
+1	0	1	0	3	0	1	11	4
−1 in runs of								
1 base	0	0	18	7	6	4	9	2
2 bases	0	1	12	6	10	8	8	8
3 bases	0	2	2	2	6	7	6	5
4 bases	2	1	7	12	7	2	3	1
5 bases	0	1	1	8	8	3	2	2
−(2–7)	0	1	11	1	0	0	1	0
−(13–436)	2	2	12	3	0	0	4	1
Other	0	0	3	1	0	0	1	0

Only mutations with a detectable phenotype are included. −AP, +AP: without or with accessory proteins, respectively. Of the 14 deletions of 2–7 bases, 8 occurred in runs or between repeats. The "Other" mutations consisted of one addition of 3 bases and four complex mutations. Deletions of 13–436 bases are described in Table 3 and the complex mutations are described in Table 4.

tions, 5 were single-base deletions, and 5 were larger deletions, a distribution similar to one described previously (Kunkel 1985).

Fidelity of the Pol^+ Exo^- RB69 DNA polymerase: A pair of alanine substitutions at two catalytic aspartic-acid residues (D222A/D327A) in the exonuclease domain of RB69 gp43 completely inactivates the exonuclease function (Abdus Sattar *et al.* 1996; Franklin *et al.* 2001). This Exo^- defect had little effect (<10% reduction) on DNA synthesis and phage production *in vivo* (Bebenek *et al.* 2001).

The exonuclease-deficient RB69 gp43 is a mutator polymerase *in vitro*, exhibiting a *lacZ*α mutation rate of 35×10^{-4} that is only a little larger (1.6-fold) than the *rI* forward mutation rate *in vivo* of 22×10^{-4} (Table 1). The accessory proteins do not significantly affect the total mutation rate. The Pol^+ Exo^+ background mutation frequency contributes ~23% of the mutations observed in the two Pol^+ Exo^- collections.

The mutations produced by the Pol^+ Exo^- polymerase *in vivo* are about 50% transitions, 25% transversions, and 25% additions or deletions of single base pairs (Bebenek *et al.* 2001). This polymerase produces roughly similar proportions of mutations *in vitro* and, in addition, some larger deletions. The specific kinds of base substitutions produced *in vivo* and *in vitro* are

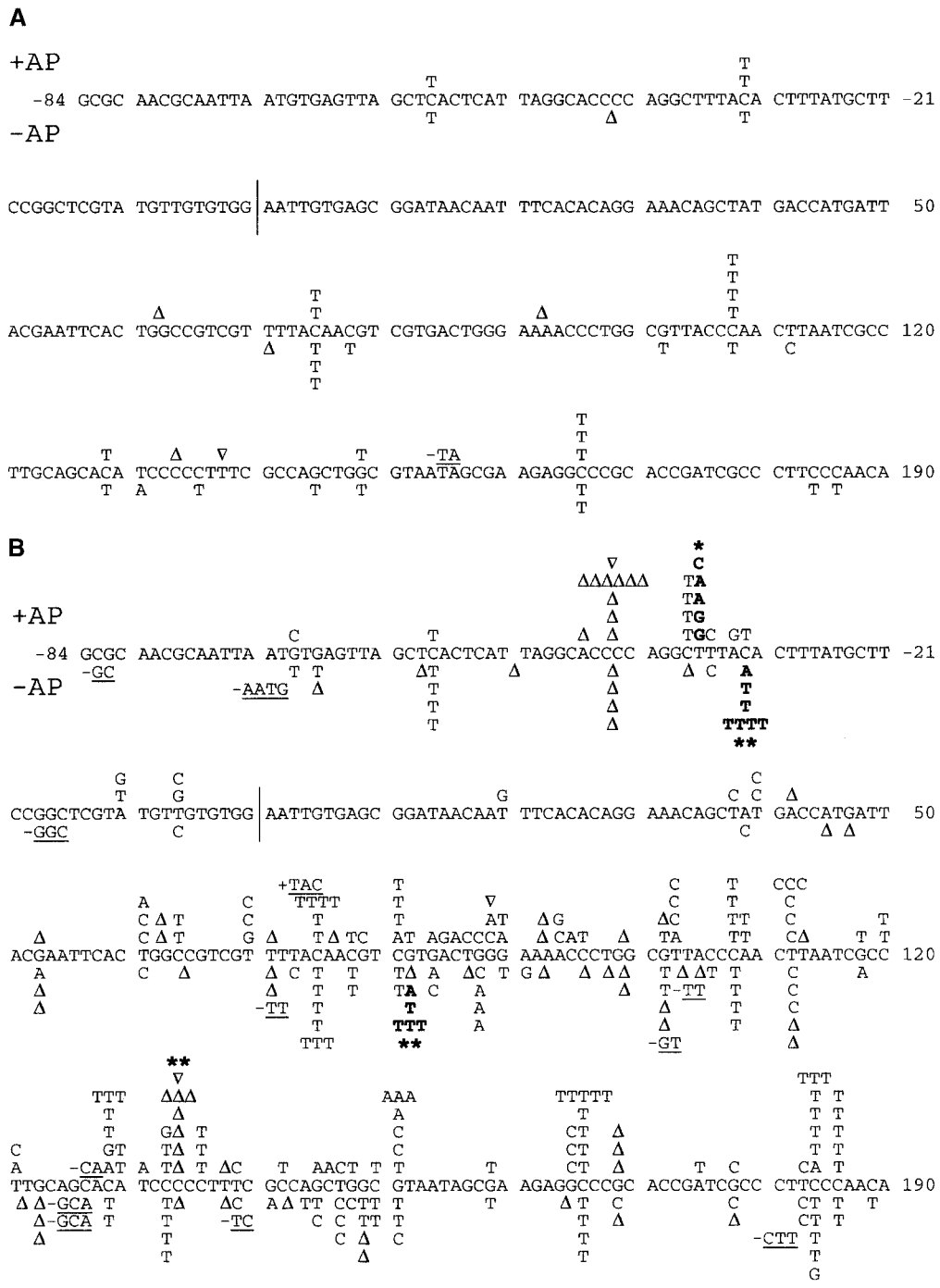

FIGURE 1.—Mutational spectra for RB69 DNA polymerases with and without accessory proteins. The 5′ → 3′ sequence of the viral template strand of the *lacZ*α sequence in M13mp2 is shown from position −84 through +190 where +1 is the first transcribed base, the vertical line separates − and + bases, and the initiation codon is coded by ATG at 39–41; the last seven bases are not shown because they generated no mutations in this study. The direction of DNA synthesis is leftward from 190. +AP indicates the presence of accessory proteins and −AP indicates their absence during the reaction. Capital letters indicate base substitutions. The deletion of a single base is indicated by Δ and the addition of a single base by ∇ except that the addition of a base unlike a template base is indicated by an arrow pointing between the two bases between which the insertion occurred and with the inserted base indicated below the arrow. The deletion or addition of two or more bases is indicated by − or + before the underlined capital letters. Larger deletions are described in Table 3 and complex mutations in Table 4. Significant asymmetries between the spectra with and without accessory proteins are indicated by ** for $P < 0.05$ or * for $P < 0.10$ and by displaying the more frequent kind of mutation in boldface. (A) Pol^+ Exo^+. (B) Pol^+ Exo^-. (C) Pol^M Exo^+. (D) Pol^M Exo^-. Note that the mutation shown as −TCG below the first 10 bases in the second line of template sequence in this spectrum could result from deletions of either TCG or CGT from the fugal sequence TCGT.

also similar (Tables 2, 5, and 6) except that single-base additions were favored *in vivo* whereas single-base deletions predominate *in vitro*.

The only significant effect of the background contribution is to reduce the number of polymerase-induced C → T transitions about twofold (from 45 and 58 to 24 and 32 without and with accessory proteins, respectively) and to reduce the numbers of transitions by about one-third (from 67 and 94 to 45 and 67 among similar numbers of mutations). The background contributed most of the C → T mutations at positions −32, 75, 108, and 166.

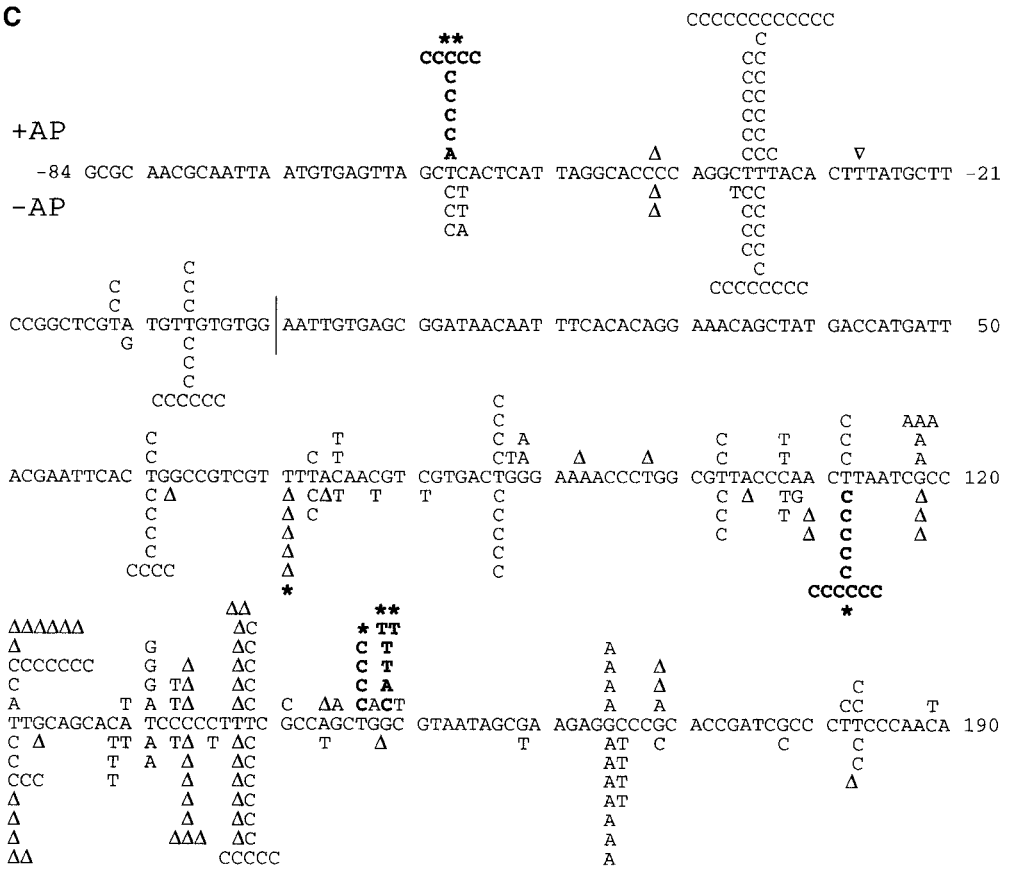

FIGURE 1.—*Continued.*

The accessory proteins altered the distributions of mutations among types in several ways. When numbers of mutations are normalized to relative mutation rates with and without accessory proteins (Table 1) and to sample sizes (Table 2), the accessory proteins modestly increased the frequencies of transitions (by 12%) and transversions (by 32%). They decreased the overall frequency of single-base deletions by 36%, but increased the frequency most (by 69%) at single-base targets while increasing the frequencies modestly ($P > 0.1$) at a four-C run and strongly (by 540%) at a five-C run. They reduced the frequency of larger deletions by 86%.

The two spectra display four specific sites (at positions −36, −32, 82, and 132–136) in which the accessory proteins significantly altered the numbers of mutations (Figure 1B). At two of these sites the mutation rate is higher in reactions with accessory proteins and at two it is lower. (When this spectrum was adjusted by subtracting the calculated contributions of the Pol$^+$ Exo$^+$ background, the resulting spectrum displayed a similar pattern.) Thus, although the accessory proteins had no effect on the gross mutation rate for this polymerase, they substantially altered mutational specificity in class- and site-specific ways.

Fidelity of the Pol^M Exo$^+$ RB69 DNA polymerase: The Y567A substitution at the Pol site is strongly mutagenic *in vivo*, generating transitions almost exclusively (Bebenek *et al.* 2001). The Pol^M Exo$^+$ enzyme is also a mutator *in vitro* (Table 1) and is sufficiently strong that the Pol$^+$ Exo$^+$ background contributes only ~10% of the mutations.

The Pol^M Exo$^+$ mutation rate *in vitro* (100×10^{-4}) is about five times higher than the corresponding rate *in vivo* (21×10^{-4}). However, as discussed later, this difference is likely to have been augmented by our use of dNTP concentrations high enough to partially inhibit proofreading. The mutational propensities of the Pol^M Exo$^+$ gp43 *in vitro* differ in three ways from those observed *in vivo* (Table 2). First, the mutator activity of this enzyme is about threefold stronger than that of the Pol$^+$ Exo$^-$ enzyme *in vitro*, whereas the two mutators were identically strong *in vivo*. Second, a substantial minority of the mutations produced *in vitro* are single-base deletions, in contrast to almost none *in vivo*. Third, the favored transition is G·C → A·T *in vivo* but T → C *in vitro*. Larger deletions were absent from the Pol^M Exo$^+$ spectrum both *in vivo* and *in vitro*.

The only significant impact of subtracting the spontaneous background from the values in Table 2 is to reduce the numbers of C → T transitions. These are reduced from 17 to 7 without accessory proteins and from 8 to 0 with accessory proteins.

The accessory proteins appear to slightly increase (by 27%) the overall Pol^M Exo$^+$ mutation frequency (Table 1), but this increase is within the variability of the individual measurements pooled to produce the values in Table 1. The accessory proteins also seem to have little effect on the proportions of mutations of various kinds (Table 2). Nevertheless, they significantly modify site-specific mutation rates at five positions (−58, 71, 112, 147, and 149), where rates were increased at three positions and decreased at two (Figure 1C). The spontaneous background has almost no effect on these two mutational spectra.

Fidelity of the Pol^M Exo$^-$ RB69 DNA polymerase: When examined *in vivo* (Bebenek *et al.* 2001), this double-mutator gp43 was only a little stronger (an estimated 2300-fold increase over the Pol$^+$ Exo$^+$ mutation rate) than its component Pol^M and Exo$^-$ mutators (each with ~500-fold increases). This enzyme exhibited a significantly higher mutation rate *in vitro* in the presence of accessory proteins (~400×10^{-4}) than in their absence (~260×10^{-4}; Table 1) and both of these values exceed the range of mutation rates ($74–120 \times 10^{-4}$) that we estimated using the *rI* system *in vivo*. The Pol$^+$ Exo$^+$ background does not contribute significantly to mutations produced by the Pol^M Exo$^-$ enzyme. Because of the high *lacZ*α background mutation frequency contributed by growth in M13, the Pol$^+$ Exo$^+$ mutation frequency could not be determined and mutator factors therefore cannot be expressed as fold increases over the Pol$^+$ Exo$^+$ value. However, as observed previously *in vivo*, the Pol^M Exo$^-$ mutation rate *in vitro* is only a fewfold higher than the sum of its component single mutations regardless of the extent to which the Pol^M Exo$^+$ rate may have been overestimated because of the high concentration of dNTP's used here (discussed below).

As previously observed *in vivo*, the Pol^M Exo$^-$ enzyme is a base-substitution mutator that produces many more transitions than transversions. While its frequencies of A → G transitions are low, as also seen with both single mutators, the frequencies of the other three transitions are not well predicted by their frequencies in the single mutators.

Approximately 14% of the mutations produced by the Pol^M Exo$^-$ enzyme *in vitro* are composed of losses or additions of single bases, compared with about 2% *in vivo*. The accessory proteins produced at most modest effects on frequencies of base substitutions, but perhaps diminished frequencies of single-base additions and of larger deletions (Table 2). On the other hand, their effects on site-specific mutation rates are striking. The accessory proteins decreased mutation rates at positions −68, −66, 87, and 89 and increased mutation rates at positions 118, 139, and 183 (Figure 1D). Because the overall mutation rate is increased ~1.5-fold by the accessory proteins, it is likely that increases occurred at many other sites but were not detectable because of the small number of mutations at most sites.

DISCUSSION

Mutational targets: The targets we used *in vivo* and *in vitro* to conduct these RB69 gp43 fidelity studies display

similarities as well as differences. The T4 *rI* target used *in vivo* contains 291 translated bases while the *lacZ*α target used *in vitro* contains 281 bases, of which 197 are transcribed and 159 are translated. The mRNA strand of T4 *rI* has the base composition A = 91, T = 94, G = 59, and C = 47 and an A·T content of 63.5%, which is close to the 64.7% calculated for the whole T4 genome (http://www.ncbi.nlm.nih.gov:80/cgi-bin/Entrez/getff?gi=15081&db=Genome). The *lacZ*α mRNA strand has the base composition A = 68, T = 68, G = 62, and C = 83, or 48.4% A·T, which is close to the 49.2% A·T for the *E. coli* genome from which it originated (Blattner *et al.* 1997) but with a modest excess of C over G. Both targets contain numerous single-base runs of various lengths that are prone to frameshift mutagenesis, and both targets contain numerous adjacent and distant repeats that are prone to deletion mutagenesis. The main advantages of the *rI* target are that it reflects biochemical realities *in vivo* and that, for the wild-type polymerase, it has the low average spontaneous mutation rate per base pair expected for a T4 sequence, $\sim 2 \times 10^{-8}$ (Drake 1991). The main advantages of the *lacZ*α target are that it differentiates between the mutabilities of each base in a base pair and that a historical record of detectable base substitutions is available.

Mutation rates *in vivo* and *in vitro*: Table 1 shows that the mutation rates of the three mutator RB69 gp43's are somewhat higher *in vitro* than *in vivo*, by 1.6-fold for the Pol^+ Exo^- enzyme, by 5-fold for the Pol^M Exo^+ enzyme, and by ~4-fold for the Pol^+ Exo^- enzyme in the presence of accessory proteins. The Pol^+ Exo^- difference is small and probably insignificant. The Pol^M Exo^- difference is uncertain because of the indirect method used to estimate the rate *in vivo* and the disparate values obtained with and without accessory proteins *in vitro*, but is likely to be significantly higher *in vitro*. The Pol^M Exo^+ difference may have been overestimated because we used dNTP's at 1 mm, which is high enough to inhibit proofreading activity to some extent. Proofreading is generally inhibited by high dNTP concentrations (such as 1 mm *vs.* 10 μm) in DNA polymerases with 3′-exonucleases such as T4 gp43 (Clayton *et al.* 1979), *E. coli* Pol I (Kunkel *et al.* 1981; Bebenek *et al.* 1990), and *E. coli* Pol III holoenzyme (Fersht 1979; Pham *et al.* 1998). However, Mathews and Sinha (1982) estimated the effective dTTP concentration at the compartmentalized replication fork to be ~220 μm, ~5-fold lower than the concentration used in this study. More importantly, the kinds of mutations produced by the Pol^M Exo^+ enzyme are not at all what might be expected from a strong contribution by a phenotypically Exo^- condition. In one of several examples, in the absence of accessory proteins, the Pol^M Exo^+ enzyme produced no deletions of ≥2 bases, whereas $23(147/159) \approx 21$ were expected from the results with the Pol^+ Exo^- enzyme. Thus, the Pol^M Exo^+ difference *in vitro* would probably decrease but not disappear at lower dNTP concentrations. In summary, the mutation rates observed with the RB69 gp43 mutators are only modestly higher *in vitro* than *in vivo*.

These modest differences in mutation rates cannot be adequately explained by low polymerase processivity, because including the accessory proteins markedly affected rates at numerous specific sites. We offer five possible explanations for these observations.

i. Additional undiscovered replication accessory proteins might act *in vivo* to increase fidelity. This conjecture is mildly inconsistent with the striking lack of mutator mutations in genes other than those encoding the polymerase and its accessory proteins (Drake and Ripley 1994). Note, however, that the functions of roughly half of the ≈280 T4 genes remain undescribed (Kutter *et al.* 1994).
ii. Conditions *in vitro* might be inadequate to support optimal fidelity. For instance, the cationic composition of our reaction mixture is dominated by acetate, whereas it is likely to be glutamate *in vivo*. In addition, concentrations of anionic and cationic species and even replication proteins *in vivo* are not well described for the putative metabolic compartments in which T4 (and presumably RB69) DNA replication occurs (Mathews and Allen 1983; Greenberg *et al.* 1994).
iii. As suggested by Pham *et al.* (1998), a replication fork with coupled synthesis of the leading and lagging strand might be intrinsically more accurate than half a fork. Gap filling *in vitro* certainly differs significantly from coupled, helicase-driven leading-strand synthesis and primer-dependent lagging-strand synthesis *in vivo*.
iv. A DNA polymerase might replicate a noncognate template (with which it did not coevolve) less accurately than a cognate template.
v. The *lacZ*α reporter sequence may, for unknown reasons, be intrinsically more mutable than the *rI* reporter sequence.

Two other comparisons have been made of forward-mutation systems *in vivo vs. in vitro*. In one, Fujii *et al.* (1999) studied the *E. coli rpsL* gene moved to a plasmid. Whereas the mutation frequency of this ~500-base reporter gene was similar to that of *lacZ*α when on M13, it was far lower ($\sim 0.007 \times 10^{-4}$) when plasmid borne, but close to that expected for the mutation rate of *E. coli*, $\sim 5 \times 10^{-10}$/bp (Drake 1991). Rolling-circle replication by an *in vitro* mixture containing the Pol III holoenzyme generated an *rpsL* mutation frequency of $\sim 2 \times 10^{-4}$, a 300-fold higher frequency that must reflect at the least the lack of DNA mismatch repair *in vitro*. The predominant mutations produced *in vitro* (in the absence of mismatch repair) were additions and deletions of base pairs, whereas base-pair substitutions predominated *in vivo*. The base-substitution mutational spectra were very different *in vivo* and *in vitro* on the same reporter se-

quence, but the frameshift mutational spectra were rather similar. In the other comparison, the *E. coli lacI* gene was the reporter sequence (Schaaper and Dunn 1987). In this case, the relevant rate *in vivo* is in a genetic background defective in DNA mismatch repair, in which case fidelity is determined primarily by the holoenzyme. Applying the method of Drake (1991) yields a mutation rate *in vivo* of 6.6×10^{-6}. The rate *in vitro* using the Pol III holoenzyme and combining the results for 10 and 50 μM dNTP's was $\sim 900 \times 10^{-6}$ based on a small and uncertain difference between the mutation frequency with and without enzyme (Pham *et al.* 1998). Even higher rates were obtained *in vitro* when proofreading was absent because only the α (polymerase) subunit of Pol III was used (Mo and Schaaper 1996) or because synthesis was conducted with the holoenzyme using a proofreading-impaired mutant together with 1 mM dNTP's, which also impairs proofreading (Pham *et al.* 1998); the rate with α alone was $16{,}600 \times 10^{-6}$ and with proofreading-impaired holoenzyme was $18{,}700 \times 10^{-6}$, or ~20-fold higher than that with the proofreading-proficient holoenzyme. While the ratio of base substitutions to frameshift mutations was similar *in vivo* and with the wild-type holoenzyme *in vitro* at low dNTP concentrations, both of the conditions *in vitro* that impaired proofreading enhanced frameshift mutagenesis by about an order of magnitude more than they enhanced base-substitution mutagenesis. Thus, in both the *rpsL* and the *lacI* studies, mutation rates were far higher *in vitro* than *in vivo*, in contrast to the small differences seen in this study.

Mutation specificities: The base-pair substitution hotspots produced by all four of the RB69 gp43 constructs *in vivo* tended to occur at certain specific G·C-rich 6-mers and especially at GG/CC dimers (Bebenek *et al.* 2001). This is a particularly striking tendency because of the low G·C content of the T4 genome. Like most of the *E. coli* genome, the *lacZ*α target is roughly half (G + C) but with some bias toward C. *In vitro*, the three spectra generated by mutator gp43's included 47 different sites with four or more base substitutions. These sites were taken to be much more mutable than average. (In this analysis, the Pol^+ Exo^- spectrum was first adjusted for the contribution from the Pol^+ Exo^+ background.) Unlike the hotspots produced *in vivo*, these 47 sites revealed no bias for either a G or a C to occupy either or both neighboring bases ($\chi^2 = 3.4$, $P \approx 0.4$). However, these 47 sites comprised 0 A, 22 T, 15 G, and 10 C residues. This distribution is very different from a random sample of *lacZ*α bases ($\chi^2 = 24.5$, $P < 0.0001$): The number of hotspots at a G or a C is close to the expectations from random sampling, while the number at an A is much lower and the number at a T is much higher. We examined the distribution of T residues neighboring these hotspots but observed little bias ($\chi^2 = 4.65$, $P \approx 0.2$). Thus, *lacZ*α hotspots generated by the mutator RB69 gp43's studied here occur preferentially at T sites, but not preferentially at TT or TTT sites. Wang *et al.* (2002) have recently described a process by which *E. coli* Pol II produces an excess of mutations at A·T-rich sites, but it is unclear whether this applies to the present system because the number of hotspots at A sites is relatively low while the number at T sites is relatively higher.

While the sequence determinants of base-substitution hotspots are usually either unknown or poorly understood, one model has been particularly helpful. In this model, forward misalignment of the primer terminus by one base, followed by correct incorporation, followed by realignment generates a base mispair that is then a candidate for extension (Kunkel and Soni 1988; Boosalis *et al.* 1989; reviewed in Bebenek and Kunkel 2000). Several of our base-substitution hotspots (particularly in Pol^M Exo^+ and Pol^M Exo^- backgrounds) contained mutations of the form XYY → XXY (such as CTT → CCT), which can be a signature of base substitution initiated by transient misalignment. However, we did not perform the tests that are required to establish the operation of this mechanism.

The sequence determinants of base addition and/or deletion hotspots are somewhat better understood. The first insight was the slippage model of Streisinger *et al.* (1966), which applies to events occurring within short homonucleotide runs or repeated short sequences. In this model, a primer base or short repeat misaligns backward or forward while retaining conventional base pairing, and primer extension then locks in the addition or deletion. Most frameshift hotspots in many spectra are of this type. In the second model, which is not confined to sequence repeats, the usual disinclination to extend a mispair can favor realignment of a mispaired primer-terminal base to achieve correct pairing provided the adjacent base is its complement, producing a readily extended primer and generating a single-base addition or deletion (with deletions strongly favored) (Kunkel and Soni 1988; Bebenek and Kunkel 1990; reviewed in Bebenek and Kunkel 2000). In the third model, active-site misalignment (which is also not confined to sequence repeats), the template base misaligns in a way that allows an incoming dNTP to pair with the next template base; if chemistry occurs and is followed by primer extension, a single-base deletion results (Kunkel 1986). This model is supported by both kinetic and structural studies (Efrati *et al.* 1997; Hashim *et al.* 1997; Ling *et al.* 2001). In this study, we observed numerous deletions of a single base in a nonrun environment (particularly of G and particularly in the Pol^+ Exo^- background). The high frequency of single-base deletions at positions 120–122 (CTT) and 136–139 (CTTT) in the Pol^M Exo^+ background might have been promoted by the misinsertion of G opposite the 5′ T. Again, however, we did not perform tests to establish the mode of frameshift mutagenesis.

We detected a total of 25 deletions of from 7 to 436

TABLE 3

Characteristics of deletion mutations produced by wild-type and mutant RB69 DNA polymerases *in vitro* without and with accessory proteins

Polymerase	APs	No. of bases deleted	Location in *lacZ*α[a]	Flanking repeats
Pol$^+$ Exo$^+$	−	436	All *lacZ*α	CTGGC
		111	(144–150)→	AGCTGGC
	+	201	→(−7)	None
		276	→(94–102)	ACCCTGGCG
Pol$^+$ Exo$^-$	−	260	→(57–58)	TC
		37	53–92	GAA
		363	→146	None[b]
		13	165–181	GCCC
		160	(97–101)→	CTGGC
		152	→(64–68)	CCGTC
		427	→(137–143)	TTTCGCC
		315	→(134–137)	CCCT
		7	−(37–26)	CTTTA
		317	→(166–170)	CCCGC
		317	→(166–170)	CCCGC
		317	→(166–170)	CCCGC
		317	→(166–170)	CCCGC
	+	81	(−15)–70	TCGT
		52	131–186	TCCC
		317	→(166–170)	CCCGC
Pol^M Exo$^-$	−	276	→(94–102)	ACCCTGGCG
		38	(−8)–30	None
		286	→85	None
		66	115–180	atcgcc[c]
	+	49	97–152	CTGGCGT

[a] Arrows from the left indicate that the deletion approaches the terminal repeat or the flanking base from the left from outside the *lacZ*α sequence, while arrows to the right indicate that the deletion extends outside the *lacZ*α sequence from the terminal repeat to the right. All listed flanking repeats are direct repeats.

[b] In this deletion, the left end has the sequence AAGGG CAATCagctgttgcccg and the right end has the sequence tcgccagCTGGCGT where lowercase letters indicate deleted bases. These two sequences are replete with repeats, including the direct repeat CAGCTG that overlaps the deletion ends, the nearly perfect palindromic reverse repeat GGGCAA(t)Cag/ctgttgccc surrounding the left end of the deletion, and the perfect palindrome cgccagCTGGCG precisely surrounding the right end of the deletion. For ways in which combinations of direct and palindromic repeats can enhance deletion formation, see Glickman and Ripley (1984).

[c] In this deletion, the left end has the sequence ACCCAACT TAatcgcct and the right end has the sequence gatcgccCTTCC CAACA where lowercase letters indicate deleted bases. Here, both atcgcc repeats were deleted instead of the usual loss of only one repeat. Two additional repeats, CCCAAC and CCT, are separated from the deletion on one or both sides by intervening bases, but their contribution to the formation of the deletion is not obvious and their presence may be fortuitous. No reverse repeats occur near the ends of this deletion.

bases (Table 3). Of these, 18 appeared once, 1 appeared twice, and 1, a 317-base deletion, appeared five times and was a deletion frequently observed in this system (Kunkel 1985). Most of these deletions arose between direct repeats of 2–9 bases. These repeats are G·C-rich, 75% on average *vs.* 52% for *lacZ*α, an observation suggesting that distantly misaligned replication intermediates are more stable the higher the melting temperature of the misaligned segment. The 317-base deletion may have been promoted not only by the direct CCCGC repeat but also by a contribution from a short palindromic repeat (Kunkel and Soni 1988). Except for four deletions in the wild-type spectrum that may not have been generated *in vitro*, all the rest were generated by an Exo$^-$ polymerase. This observation suggests that large deletions generated *in vitro* by RB69 gp43 either are produced by the polymerase and are well proofread or are produced when the primer terminus partitions to the defective Exo site but returns to a distant site more frequently than from an Exo$^+$ site. In the second mechanism, accurate realignment of a primer terminus would require a normal Exo site, because of either the energetics of cleavage or the correct primer orientation. Perhaps the most interesting observation concerning the deletions of ≥2 bases is that their frequency was sharply reduced by the accessory proteins. For a short sequence to misalign with a distant repeat, several bases of the primer terminus must separate from the template strand and must then diffuse sufficiently to discover a distant unpaired complement. It would not be surprising if such a process were inhibited by a processivity factor and/or by gp32.

We observed four complex mutations of two types (Table 4). In one type, two mutations associated with a C-run were generated by a Pol$^+$ Exo$^-$ polymerase and correspond to the replacement of two adjacent bases by one different base. These mutations could have arisen by any of several possible coupled mispair-and-slippage errors. In the second type, two mutations generated by a Pol^M Exo$^+$ polymerase represent complicated palindrome-associated events, first recognized by Glickman and Ripley (1984), that may involve strand-switching templating errors during primer extension.

Certain kinds of mutations were produced more frequently *in vitro* than *in vivo*. These included single-base deletions by all three polymerases, larger deletions by the Pol$^+$ Exo$^-$ polymerase (although most of these were prevented by the accessory proteins), and transitions at A and T sites relative to those at G and C sites by the Pol^M Exo$^+$ polymerase (at least in part because of the hotspot biases described above). In contrast, the Pol$^+$ Exo$^-$ polymerase produced many fewer single-base additions *in vitro* than *in vivo* with or without accessory proteins. The higher mutation rates observed *in vitro vs. in vivo* cannot be explained simply by the additional classes of mutations sometimes produced *in vitro*. The litany of possible reasons for these differences is the same as that invoked above for the differences in mutation rates, but here pertaining to specific mutational pathways.

TABLE 4

Characteristics of complex mutations produced by wild-type and mutant RB69 DNA polymerases *in vitro* without and with accessory proteins

Polymerase	APs	Target bases in *lacZ*α	Sequence change	Donor
Pol^+ Exo^-	−	98–99	CCCTGG → CCCCG	None
		−(44–43)	ACCCC → ATCC	None
		133–139	CCCCTTT → GCTATTA	131–151[a]
Pol^M Exo^+	−	−(69–43)	27-mer → 11-mer	101–111[b]

[a] 131-TCCCCCTTTCGCCAGCTGGCG-151 → GCTATTA, the reverse complement of 140-CGCCAGC TGGCGTAATAGCGA-160. The sequence 140–152 is also shown because it is a 6:6 palindrome that might (or might not) have contributed to the formation of this mutation.

[b] (−82)-GCAA*C*GC*A*ATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGG*C*TTT-(−34) → GTTGGGTA ACG, the reverse complement of 92-AAA*C*CCTGG*C*GTTACCCAAC*T*TAAT*C*GC*C*TTGC-124. Italicized bases mark three imperfections in the palindrome.

The Pol^M-Exo^- interaction: The double mutator displayed two unanticipated traits *in vivo*: Its mutational specificity was not fully predicted by either of its two component mutators, and its mutator strength was only about twofold greater than the sum of the mutator strengths of its components. The same two traits are displayed *in vitro*: Both G → A transitions and single-base additions are more frequent than those with either of the single mutators, and the strength of the double mutator is again only about twofold greater than the sum of the mutator strengths of its components. In one simple view of replication fidelity, proofreading removes a constant fraction of nascent mutations of various types. This view predicts that the double mutator should be far stronger than observed, even though this cannot be specifically predicted because we lack a Pol^+ Exo^+ value. The modest increase in the strength of the double mutator does not seem to be due to the efficient extension of mispaired primer termini because the Pol^M enzyme extends mispairs at least as poorly as does the Pol^+

TABLE 5

Base-substitution rates (μ_b) for mutant RB69 DNA polymerases without and with accessory proteins

	Pol^+ Exo^-				Pol^M Exo^+				Pol^M Exo^-			
	No.		μ_b		No.		μ_b		No.		μ_b	
Template•primer mispair (opportunities)[a]	−AP	+AP	−AP	+AP	−AP	+AP	−AP	+AP	−AP	+AP	−AP	+AP
All (128)	93	137	17	19	110	109	47	68	120	153	160	278
A•dCTP (19)	1	4	1	4	2	0	6	4	0	0	≤9	≤12
T•dGTP (28)	13	24	11	16	76	72	150	200	37	68	230	570
G•dTTP (22)	8	8	9	7	7	15	18	54	54	56	420	590
C•dATP (25)	45	58	42	42	17	8	38	25	12	21	82	200
A•dGTP (17)	1	1	1	1	0	0	≤3	≤5	1	0	10	≤14
A•dATP (23)	1	6	1	5	2	1	5	3	2	1	15	10
G•dGTP (20)	7	10	8	9	2	2	6	8	4	0	34	≤12
G•dATP (25)	15	7	14	5	1	5	2	16	2	0	14	≤9
T•dTTP (16)	0	6	≤1	7	2	3	7	15	8	3	85	44
T•dCTP (23)	0	5	≤1	4	0	3	≤2	10	0	1	≤7	10
C•dTTP (17)	1	6	1	6	1	0	3	≤5	0	3	≤10	41
C•dCTP (9)	1	2	3	4	0	0	≤6	≤9	0	0	≤20	≤26

The mutation rates in Table 1 and summary sequence data in Table 2 were used to calculate mutation rates (μ_b) that are rounded to reflect the numbers of observed mutations. Because the wild-type polymerase did not generate errors above the historical background mutant frequency of uncopied DNA, its mutant frequency of 6.2×10^{-4} was subtracted from each mutator-polymerase frequency. Mutation rates are per 10^6 nucleotides incorporated and were calculated by multiplying the net mutant frequency (Table 1) by the proportion of mutants in each class (Table 2) and dividing by 0.6 (the correction factor for detecting errors in *E. coli*) and by the number of detectable sites ("opportunities") for each class of mutation (see MATERIALS AND METHODS, *Sequencing*). "≤" values were calculated as if one mutant had been detected.

[a] The number of detectable events is unambiguously defined for the specific mispairs. From 1 to 3 substitutions are detectable at 128 sites, a value we used for comparison with other published values. However, there are only 244 possible substitutions at these 128 sites, and careful readers could recalculate the rates normalized to this value.

TABLE 6

Base addition/deletion rates for mutant RB69 DNA polymerases without and with accessory proteins

Base addition or deletions (opportunities)[a]	Pol^+ Exo^-				Pol^M Exo^+				Pol^M Exo^-			
	No.		μ		No.		μ		No.		μ	
	−AP	+AP	−AP	+AP	−AP	+AP	−AP	+AP	−AP	+AP	−AP	+AP
1 base (199)	40	35	5	3	37	24	10	10	28	18	24	21
In nonrun (97)	18	7	4	1	6	4	3	3	9	2	16	5
In 2-base run (58)	12	6	5	2	10	8	10	10	8	8	20	30
In 3-base run (27)	2	2	2	1	6	7	10	20	6	5	40	40
In 4-base run (12)	7	12	14	18	7	2	30	10	3	1	40	20
In 5-base run (5)	1	8	5	30	8	3	90	50	2	2	70	90
−(2–436) bases[b]	23	4	540	70	0	0	60	80	5	1	850	230

Rates were calculated as described in the legend to Table 5.

[a] The number of detectable sites is unambiguously defined for the single-base events.

[b] The number of detectable sites is not defined for the larger deletions, and their mutation rates are therefore not normalized to numbers of opportunities and appear correspondingly larger than the other values.

enzyme (Bebenek *et al.* 2001), but several other explanations are tenable: (i) It might reflect an effect of the Y567 Pol mutation on the partitioning of the primer terminus to the Exo site; (ii) it might result from an effect of the Exo deficiency on fidelity at the Pol site; (iii) if this gp43 usually dissociated from a mispair with subsequent loss of the DNA substrate, many mutations might be lost to our measurement system; or (iv–v) conversely, if this gp43 usually stalled irreversibly at a mispair, at either the Pol or the Exo binding sites, many mutations could be lost.

The impact of accessory proteins on polymerase fidelity: The effect of the accessory proteins on the fidelity of DNA replication *in vitro* has been the subject of several studies:

i. The T7 DNA polymerase processivity factor is a tightly bound host protein, thioredoxin. The T7 holoenzyme is too accurate to produce a reliable signal in the *lacZ*α system, but an Exo^- holoenzyme did produce a signal (Kunkel *et al.* 1994). In reversion tests, the presence of thioredoxin decreased base-substitution frequencies by about threefold and decreased −1 and −2 base deletions substantially in nonrun contexts, but strongly increased additions of one or more bases within runs.

ii. As with T7 DNA polymerase, the mitochondrial DNA polymerase-γ accessory subunit increased fidelity for some base substitutions on synthetic oligonucleotide primer templates (Johnson and Johnson 2001). A similar result was obtained by Longley *et al.* (2001), who showed that fidelity increased, not during mispair formation but during mispair extension. In extensive tests with the *lacZ*α system, the overall fidelity of wild-type Polγ was unaffected by the accessory subunit, although the subunit promoted the specific template·primer mispair T·G with the wild-type enzyme. The overall fidelity of an Exo^- Polγ was decreased threefold by the accessory subunit, mainly by increases in the formation of the specific mispair A·A and in single-base deletions in nonruns or in runs of size 2. In runs of size 4 or 5, the accessory subunit considerably decreased the rate of single-base deletions.

iii. The fidelity of mammalian DNA polymerase-δ on synthetic oligonucleotide primer templates was decreased considerably by the inclusion of its processivity clamp, proliferating cell nuclear antigen (Mozzherin *et al.* 1996).

iv. When the fidelity of the *E. coli* Pol III α (polymerase) subunit alone (Mo and Schaaper 1996) was compared with that of a holoenzyme that was defective in proofreading because of the combination of a genetic defect in the proofreading subunit (*mutD5*) and a high concentration (1 mm) of dNTP's (Pham *et al.* 1998), there was no difference in the overall mutation rate. A small decrease in the relatively high frameshift rate (primarily single-base deletions) was offset by a tripling of the base-substitution rate.

T4 gp43 removed preformed terminal mismatches more efficiently when gp32, gp45, and gp44/gp62 were present (Sinha 1987). However, we cannot compare these results with those reported here because the high *lacZ*α background in our system conceals the specificity of the Pol^+ Exo^+ enzyme. When T4 gp43 infidelity was measured as dNTP → dNMP turnover using homopolymer primers and templates, the accessory proteins improved fidelity (Topal and Sinha 1983).

The study of the effects of accessory proteins on fidelity that is most relevant to the current report used a T4 Pol^+ Exo^- gp43 and the *lacZ*α system (Kroutil *et al.* 1998). Adding the gp45 clamp increased the forward-mutant frequency from 24 to 47 per 10^4, in apparent contrast to our observation of an insignificant decrease

from 29 to 25 per 10^4 (Table 1). The frequency of large deletions was decreased ~10-fold, a factor similar to that observed by us (Table 2). Adding the gp45 clamp increased the TGA revertant frequency from 4 to 10 per 10^5, an increase distributed irregularly over several of the possible mispairs. However, the wild-type precursor (TGG) to their TGA reporter at bases 87–89 is particularly prone to accessory-protein effects in the RB69 system (Figure 1, B, C, and D), exhibiting both increases and decreases. Adding the gp45 clamp had no significant effect on the frequency of single-base deletions in a TTTTTTT run. Kinetic parameters for misincorporation were similar with and without the accessory proteins. A procedural difference should be noted between the T4 and the RB69 studies. In the T4 studies, most of the comparisons were between reactions with and without gp45, the other accessory proteins being present throughout, whereas in the RB69 studies tabulated above the comparison was between gp43 alone and gp43 with all four accessory proteins. However, this procedural difference seems unlikely to be profound, because we also measured mutant frequencies (but did not sequence mutants) in reactions lacking only gp45. For comparison with the values in Table 1, the uncorrected mutant frequencies for the four polymerases (in the same order) were 4.4, 20, 71, and 275 per 10^4. In each case, the frequency with gp32 and gp44/gp62 present was very similar with and without gp45.

Including the SSB, the processivity clamp and the clamp-loading proteins in our reactions had little or no effect upon gross mutation rates (Table 1). Although we did not measure the impact of the RB69 accessory proteins on processivity, other studies have examined this matter with the related T4 proteins, both in the context of leading-strand and lagging-strand synthesis (Kadyrov and Drake 2001 and references therein) and during gap filling, all gaps being filled in 15 sec in their presence and none in their absence (Kroutil *et al.* 1998). In addition, the dramatic impact of the RB69 accessory proteins on site-specific mutation rates argues strongly for their participation in polymerase fidelity. There were only a few convincing effects on rates of specific types of base substitutions, additions or deletions (Tables 2, 5, and 6), such as modestly increased frequencies of T → C transitions and sharply decreased frequencies of larger deletions in the two Exo^- mutators. However, the mutational spectra (Figure 1) revealed numerous specific sites where rates were affected. The striking aspect of these effects is that site-specific rates were sometimes increased and sometimes decreased, with no obvious preference for the direction of the bias.

Computer simulations first established the statistical significance of the differences in the overall distributions of mutation frequencies with and without accessory proteins for each of the three RB69 gp43's, but did not identify the particular sites at which these differences might be most notable. (As expected and noted previously, there were no significant differences with the wild-type gp43.) Subsequent Fisher's exact tests then identified specific hotspots showing evidence of such differences. Because of the large number of sites examined, we expected occasional false positives, that is, P values below 0.10 or 0.05 that can be attributed to chance alone. Sites having few or no mutations do not contribute false positives because even the most extreme distribution of mutations possible at such sites would not be statistically significant. For each of the four polymerases studied here, between 21 and 27 sites could have produced a false-positive outcome on the basis of the frequency of mutations observed at those sites. Calculations based on the distribution of the Fisher's exact test statistic revealed that for each polymerase, we would expect to find approximately one instance of $P < 0.05$ and two of $P < 0.10$ by chance alone. This compares with the two to five instances of $P < 0.05$ and four to seven of $P < 0.10$ actually observed for each of the variant polymerases. We conclude that while it is possible that a few of the site-specific differences in mutation frequency associated with the presence or absence of accessory proteins may be chance occurrences, the majority appear to reflect real effects upon both base substitutions and base deletions. Many other such biases probably occur but are not detected with our sample sizes.

Although our evaluation of each site was based on all mutations detected at that site regardless of type, in most instances a single type of mutation accounted for the statistically significant bias. The biases occur for a variety of base substitutions (at least the transitions G → A, T → C, and C → T and at least the transversions T → A and G → T) and base deletions (at least −T, −C, and −G). For most kinds of mutations for which three or more asymmetries were recorded, the biases were in both directions. Of the 16 different biased sites we observed, none occurred more than once among the three mutator spectra. There was a clear tendency for a biased site to be flanked by a pyrimidine ($P \approx 0.01$) but C (12) and T (9) were represented among the flanking pyrimidines in nearly the same proportion (83:68) as in the *lacZ*α target. However, structural information about the nature of the interactions among DNA, polymerase, clamp, and SSB are far too sparse to provide an explanation for the bias toward flanking pyrimidines.

In summary, the RB69 DNA-replication accessory proteins have at most small overall effects on mutation rates *in vitro* but strongly inhibit or promote mutation at specific sites. (This is another instance in which conclusions drawn from small targets, as in most reversion tests, may be misleading and should be confirmed in a forward-mutation test.) Evolution has optimized the interactions among the proteins of DNA metabolism to achieve an optimal mutation rate perhaps driven by the deleterious effects of most mutations and the costs that must be paid to keep mutation rates low (Drake 1991).

The huge variety of local DNA sequences that interact with the replication proteins at and near the primer terminus results in high mutation rates at some sites and low rates at others. Our results may offer an opportunity to probe the specific roles of accessory proteins in causing such variability.

We thank Tom Kunkel and Roel Schaaper for their critical readings of this article.

LITERATURE CITED

Abdus Sattar, K. M. A., T.-C. Lin, C. Jones and W. H. Konigsberg, 1996 Functional consequences and exonuclease kinetic parameters of point mutations in bacteriophage T4 DNA polymerase. Biochemistry **35:** 16621–16629.

Bebenek, A., H. K. Dressman, G. T. Carver, S. N. Ng, V. Petrov *et al.*, 2001 Interacting fidelity defects in the replicative DNA polymerase of bacteriophage RB69. J. Biol. Chem. **276:** 10387–10397.

Bebenek, K., and T. A. Kunkel, 1990 Frameshift errors initiated by nucleotide misincorporation. Proc. Natl. Acad. Sci. USA **87:** 4946–4950.

Bebenek, K., and T. A. Kunkel, 1995 Analyzing the fidelity of DNA polymerases. Methods Enzymol. **262:** 217–232.

Bebenek, K., and T. A. Kunkel, 2000 Streisinger revisited: DNA synthesis errors mediated by substrate misaligments. Cold Spring Harbor Symp. Quant. Biol. **65:** 81–91.

Bebenek, K., C. M. Joyce, M. P. Fitzgerald and T. A. Kunkel, 1990 The fidelity of DNA synthesis catalyzed by derivatives of *Escherichia coli* DNA polymerase I. J. Biol. Chem. **265:** 13878–13887.

Blattner, F. R., G. Plunkett, III, C. A. Block, N. T. Perna, V. Burland *et al.*, 1997 The complete genome sequence of *Escherichia coli* K-12. Science **277:** 1453–1474.

Boosalis, M. S., D. W. Mosbaugh, R. Hamatake, A. Sugino, T. A. Kunkel *et al.*, 1989 Kinetic analysis of base substitution mutagenesis by transient misaligment of DNA and by miscoding. J. Biol. Chem. **264:** 11360–11366.

Borjac, J., 1997 RNA determinants of specificity of two phylogenetically related single-stranded DNA binding proteins. Ph.D. Thesis, Tulane University, New Orleans.

Braithwaite, D. K., and J. Ito, 1993 Compilation, alignment, and phylogenetic relationships of DNA polymerases. Nucleic Acids Res. **21:** 787–802.

Clayton, L. K., M. F. Goodman, E. W. Branscomb and D. J. Galas, 1979 Error induction and correction by mutant and wild type T4 DNA polymerase. J. Biol. Chem. **254:** 1902–1912.

Das, S. K., and R. K. Fujimura, 1979 Processiveness of DNA polymerases, a comparative study using a simple procedure. J. Biol. Chem. **254:** 1227–1232.

Doublié, S., S. Tabor, A. M. Long, C. C. Richardson and T. Ellenberger, 1998 Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. Nature **391:** 251–258.

Drake, J. W., 1991 A constant rate of spontaneous mutation in DNA-based microbes. Proc. Natl. Acad. Sci. USA **88:** 7160–7164.

Drake, J. W., and L. S. Ripley, 1994 Mutagenesis, pp. 98–124 in *Molecular Biology of Bacteriophage T4*, edited by J. D. Karam. American Society for Microbiology, Washington, DC.

Dressman, H. K., C.-C. Wang, J. D. Karam and J. W. Drake, 1997 Retention of replication fidelity by a DNA polymerase functioning in a distantly related environment. Proc. Natl. Acad. Sci. USA **94:** 8042–8046.

Efrati, E., G. Tocco, R. Eritja, S. H. Wilson and M. F. Goodman, 1997 Abasic translesion synthesis by DNA polymerase β violates the "A-rule." Novel types of nucleotide incorporation by human DNA polymerase β at an abasic lesion in different sequence contexts. J. Biol. Chem. **272:** 2559–2569.

Fersht, A. R., 1979 Fidelity of replication of phage ϕX174DNA by DNA polymerase III holoenzyme: spontaneous mutation by misincorporation. Proc. Natl. Acad. Sci. USA **76:** 4946–4950.

Franklin, M. C., J. Wang and T. A. Steitz, 2001 Structure of the replicating complex of a pol α family DNA polymerase. Cell **105:** 657–667.

Fujii, S., M. Akiyama, K. Aoki, Y. Sugaya, K. Higuchi *et al.*, 1999 DNA replication errors produced by the replicative apparatus of *Escherichia coli*. J. Mol. Biol. **289:** 835–850.

Glickman, B., and L. S. Ripley, 1984 Structural intermediates of deletion mutagenesis: a role for palindromic DNA. Proc. Natl. Acad. Sci. USA **81:** 512–516.

Goodman, M. F., 1997 Hydrogen bonding revisited: geometric selection as a principal determinant of DNA replication fidelity. Proc. Natl. Acad. Sci. USA **94:** 10493–10495.

Greenberg, G. R., P. He, J. Hilfinger and M.-J. Tseng, 1994 Deoxyribonucleoside triphosphate synthesis and phage T4 DNA, pp. 14–27 in *The Molecular Biology of Bacteriophage T4*, edited by J. D. Karam. American Society for Microbiology, Washington, DC.

Hashim, M. F., N. Schnetz-Boutaud and L. J. Marnett, 1997 Replication of template-primers containing propanodeoxyguanosine by DNA polymerase β. Induction of base pair substitution and frameshift mutations by template slippage and deoxynucleoside triphosphate stabilization. J. Biol. Chem. **272:** 20205–20212.

Hashimoto, H., M. Nishioka, S. Fujiwara, M. Takagi, T. Imanaka *et al.*, 2001 Crystal structure of DNA polymerase from hyperthermophilic archaeon *Pyrococcus kodakaraensis* KOD1. J. Mol. Biol. **306:** 469–477.

Hopfner, K.-P., A. Eichinger, R. A. Engh, F. Laue, W. Ankenbauer *et al.*, 1999 Crystal structure of a thermostable type B DNA polymerase from *Thermococcus gorgonarius*. Proc. Natl. Acad. Sci. USA **96:** 3600–3605.

Johnson, A. A., and K. A. Johnson, 2001 Fidelity of nucleotide incorporation by human mitochondrial DNA polymerase. J. Biol. Chem. **276:** 38090–38096.

Johnson, K. A., 1993 Conformational coupling in DNA polymerase fidelity. Annu. Rev. Biochem. **62:** 685–713.

Kadyrov, F. A., and J. W. Drake, 2001 Conditional coupling of leading-strand and lagging-strand DNA synthesis at bacteriophage T4 replication forks. J. Biol. Chem. **276:** 29559–29566.

Karam, J. D., and W. H. Konigsberg, 2000 DNA polymerase of the T4-related bacteriophages. Prog. Nucleic Acid Res. Mol. Biol. **64:** 65–96.

Kroutil, L. C., M. W. Frey, B. F. Kaboord, T. A. Kunkel and S. J. Benkovic, 1998 Effect of accessory proteins on T4 DNA polymerase replication fidelity. J. Mol. Biol. **278:** 135–146.

Kunkel, T. A., 1985 The mutational specificity of DNA polymerase β during *in vitro* DNA synthesis. Production of frameshift, base substitution, and deletion mutations. J. Biol. Chem. **260:** 5787–5796.

Kunkel, T. A., 1986 Frameshift mutagenesis by eucaryotic DNA polymerases *in vitro*. J. Biol. Chem. **261:** 13581–13586.

Kunkel, T. A., and K. Bebenek, 2000 DNA replication fidelity. Annu. Rev. Biochem. **69:** 497–529.

Kunkel, T. A., and A. Soni, 1988 Mutagenesis by transient misalignment. J. Biol. Chem. **263:** 14784–14789.

Kunkel, T. A., R. M. Schaaper, R. A. Beckman and L. A. Loeb, 1981 On the fidelity of DNA replication. Effect of the next nucleotide on proofreading. J. Biol. Chem. **256:** 9883–9889.

Kunkel, T. A., S. P. Patel and K. A. Johnson, 1994 Error-prone replication of repeated DNA sequences by T7 DNA polymerase in the absence of its processivity subunit. Proc. Natl. Acad. Sci. USA **91:** 6830–6834.

Kutter, E., T. Stidham, B. Guttman, E. Kutter, D. Batts *et al.*, 1994 Genomic map of bacteriophage T4, pp. 491–518 in *Molecular Biology of Bacteriophage T4*, edited by J. D. Karam. American Society for Microbiology, Washington, DC.

Ling, H., F. Boudsocq, R. Woodgate and W. Yang, 2001 Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. Cell **107:** 91–102.

Longley, M. J., D. Nguyen, T. A. Kunkel and W. C. Copeland, 2001 The fidelity of human DNA polymerase γ with and without exonucleolytic proofreading and the p55 accessory subunit. J. Biol. Chem. **276:** 38555–38562.

Mathews, C. K., and J. R. Allen, 1983 DNA precursor biosynthesis, pp. 59–70 in *Bacteriophage T4*, edited by C. K. Mathews, E. M. Kutter, G. Mosig and P. B. Berget. American Society for Microbiology, Washington, DC.

Mathews, C. K., and N. K. Sinha, 1982 Are DNA precursors concentrated at replication sites? Proc. Natl. Acad. Sci. USA **79:** 302–306.

Mo, J.-Y., and R. M. Schaaper, 1996 Fidelity and error specificity of the α catalytic subunit of *Escherichia coli* DNA polymerase III. J. Biol. Chem. **271:** 18947–18953.

Morales, J. C., and E. T. Kool, 2000 Functional hydrogen-bonding map of the minor groove binding tracks of six DNA polymerases. Biochemistry **39:** 12979–12988.

Moran, S., R. X.-F. Ren and E. T. Kool, 1997 A thymidine triphosphate shape analogue lacking Watson-Crick pairing ability is replicated with high sequence selectivity. Proc. Natl. Acad. Sci. USA **94:** 10506–10511.

Mozzherin, D. J., M. McConnell, M. V. Jasko, A. A. Krayevsky, C.-K. Tan *et al.*, 1996 Proliferating cell nuclear antigen promotes misincorporation catalyzed by calf thymus DNA polymerase. J. Biol. Chem. **271:** 31711–31717.

Newport, J. W., S. C. Kowalczykowski, N. Lonberg, L. S. Paul and P. H. von Hippel, 1980 Molecular aspects of the interactions of T4 coded gene 32 protein and DNA polymerase (gene 43 protein) with nucleic acids, pp. 485–505 in *Mechanistic Studies of DNA Replication and Genetic Recombination*, edited by B. M. Alberts. Academic Press, New York.

Nossal, N. G., 1994 The bacteriophage T4 DNA replication fork, pp. 43–53 in *Molecular Biology of Bacteriophage T4*, edited by J. D. Karam. American Society for Microbiology, Washington, DC.

Pham, P. T., M. W. Olson, C. S. McHenry and R. M. Schaaper, 1998 The base substitution and frameshift fidelity of *Escherichia coli* DNA polymerase III holoenzyme *in vitro*. J. Biol. Chem. **273:** 23575–23584.

Piegorsch, W. W., and A. J. Bailer, 1994 Statistical approaches for analyzing mutational spectra: some recommendations for categorical data. Genetics **136:** 403–416.

Rodriguez, A. C., H.-W. Park, C. Mao and L. S. Beese, 2000 Crystal structure of a pol α family DNA polymerase from the hyperthermophilic archaeon *Thermococcus sp.* 9E N-7. J. Mol. Biol. **299:** 447–462.

Schaaper, R. M., and R. L. Dunn, 1987 Spectra of spontaneous mutations in *Escherichia coli* strains defective in mismatch correction: the nature of *in vivo* DNA replication errors. Proc. Natl. Acad. Sci. USA **84:** 6220–6224.

Shamoo, Y., and T. A. Steitz, 1999 Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. Cell **99:** 155–166.

Sinha, N. K., 1987 Specificity and efficiency of editing of mismatches involved in the formation of base-substitution mutations by the 3′→5′ exonuclease activity of phage T4 DNA polymerase. Proc. Natl. Acad. Sci. USA **84:** 915–919.

Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita *et al.*, 1966 Frameshift mutations and the genetic code. Cold Spring Harbor Symp. Quant. Biol. **31:** 77–84.

Topal, M. D., and N. K. Sinha, 1983 Products of bacteriophage T4 genes 32 and 45 improve the accuracy of DNA polymerase during DNA replication *in vitro*. J. Biol. Chem. **258:** 12274–12279.

Wang, C.-C., L.-S. Yeh and J. D. Karam, 1995 Modular organization of T4 DNA polymerase. J. Biol. Chem. **270:** 26558–26564.

Wang, J., A. K. M. Abdus Sattar, C. C. Wang, J. D. Karam, W. H. Konigsberg *et al.*, 1997 Crystal structure of a pol α family replication DNA polymerase from bacteriophage RB69. Cell **89:** 1087–1099.

Wang, Z., E. Lazarov, M. O'Donnell and M. F. Goodman, 2002 Resolving a fidelity paradox. Why *Escherichia coli* DNA polymerase II makes more base substitution errors in AT- compared with GC-rich DNA. J. Biol. Chem. **277:** 4446–4454.

Yeh, L.-S., 1998 Specificity and interactions of the DNA polymerase accessory proteins of bacteriophage T4 and RB69. Ph.D. Thesis, Tulane University, New Orleans.

Yeh, L.-S., T. Hsu and J. D. Karam, 1998 Divergence of a DNA replication gene cluster in the T4-related bacteriophage RB69. J. Bacteriol. **180:** 2005–2013.

Young, M. C., M. K. Reddy and P. H. von Hippel, 1992 Structure and function of the bacteriophage T4 DNA polymerase holoenzyme. Biochemistry **31:** 8675–8690.

Communicating editor: S. Lovett