Escherichia coli DNA polymerase III holoenzyme is the replicative enzyme primarily responsible for the duplication of the E. coli chromosome. This process occurs with high accuracy, less than $10^{-9}$ to $10^{-10}$ errors being committed per base pair per round of replication. As a first step in understanding the mechanisms responsible for the high fidelity of this process, we have purified the polymerase III $\alpha$ catalytic subunit, free of exonuclease activity, and analyzed its fidelity in vitro. We employed a newly developed gap-filling assay using the N-terminal 250 bases of the lacI gene as a forward mutational target. When synthesizing across this target, $\alpha$ subunit produced mutations at a frequency of 0.6% DNA sequencing revealed that the mutants created in vitro consisted mainly of frameshift mutations, although some base substitutions were also observed. The frameshifts, occurring at more than 120-fold above the background, consisted largely of $-1$ deletions. Among them, about 80% were the deletion of a purine template base with a pyrimidine 5’-neighbor. These results suggest that the $\alpha$ subunit (i) has a relatively low ability to extend from misincorporated bases, accounting for the low level of observed base substitutions, and (ii) has a relatively high capability of extension after misalignment of a misincorporated base on the next (complementary) template base, accounting for the high level of frameshift mutations. This model is supported by an experiment in which $\alpha$ subunit was required to initiate DNA synthesis from a terminal mispair in a sequence context that allowed slippage on the next template base. Among the products of this reaction, frameshifts outnumbered base pair substitutions by greater than 70-fold. A comparison to in vivo mutational spectra suggests that the pol III accessory factors may play a major role in modulating the fidelity of DNA synthesis.

DNA replication in prokaryotic and eukaryotic organisms is a highly accurate process, only one error being committed per $10^9$ to $10^{10}$ nucleotides incorporated (1). This high fidelity of DNA replication is achieved by at least three critical steps (reviewed in Refs. 2 and 3): (i) the insertion fidelity of the DNA polymerase, which selects correct over incorrect nucleotides via template-directed discrimination, (ii) the $3' \to 5'$ (proofreading) exonuclease ease of the DNA polymerase, which removes incorrectly inserted bases before DNA synthesis continues; and (iii) postreplicative DNA mismatch repair, which corrects errors resulting from DNA replication errors by incising the newly synthesized strand.

DNA polymerase III holoenzyme is the main enzyme responsible for replication of the Escherichia coli chromosome, including the high fidelity of this process (4–6). DNA polymerase III holoenzyme is a dimeric polymerase that simultaneously replicates the leading and lagging strands (4–7). It consists of 10 different subunits, $\alpha$, $\epsilon$, $\theta$, $\gamma$, $\delta$, $\delta'$, $\chi$, $\psi$, and $\beta$ with a presumed overall composition (10). The tightly bound $\alpha$, $\epsilon$, and $\theta$ subunits form the pol III core. The $\alpha$ subunit, encoded by the dnaE gene, is the catalytic DNA polymerization subunit (8). The $\epsilon$ subunit, encoded by the dnaQ gene, provides the $3' \to 5'$ proofreading activity (9). The function of the $\theta$ subunit is as yet unknown. The fine structure and dynamics of the pol III holoenzyme are under active investigation (6). Postreplicative DNA mismatch repair is performed by the mutH, mutL, mutS, and mutU gene products (see Ref. 10, for review).

The relative contributions of base selection, proofreading, and DNA mismatch repair to in vivo fidelity have been estimated from mutant frequencies and sequenced mutation spectra in strains defective in the respective pathways (3, 11). For example, using mismatch repair-defective strains, it was estimated that pol III holoenzyme replicates DNA at an average accuracy of $10^{-7}$ errors per base pair per round of replication, of which proofreading may contribute an approximate factor of 10$^{-2}$ (3, 12). Postreplicative mismatch repair improves overall fidelity by 200- to 300-fold (3, 11), yielding the final mutation rate of $10^{-9}$ to $10^{-10}$.

To gain insights into the mechanisms by which pol III holoenzyme achieves its high fidelity, we have initiated an in vitro investigation of its fidelity. Despite the fact that this enzyme is one of the best characterized replication complexes, its fidelity has not been studied in detail (13–15). In the current study, we focus on the simplest unit, the $\alpha$ catalytic subunit. One previous study on the fidelity of $\alpha$ subunit has been reported, in which steady-state kinetic assays at a limited number of nucleotide sites were performed to determine nucleotide misincorporation rates by this enzyme (14). Here, we analyze the fidelity of $\alpha$ subunit using a newly developed in vitro forward mutation assay that uses the N-terminal region of the lacI gene as a sequence target. This target allows a variety of mutations to be recovered, including base substitutions, frameshifts, deletions, and duplications. We show that this assay $\alpha$ subunit produces relatively few base pair substitutions, but a high level of single-base frameshift mutations. The implications of these findings for the overall fidelity mechanism used by pol III holoenzyme are discussed.

**EXPERIMENTAL PROCEDURES**

Materials—E. coli strains CSH50 ($\Delta$(pro-lac)), ara, thi, strA, F $\Delta$(proAB, lac$,Z\Delta M15$, traD36) (16), NR9099 ($\Delta$(pro-lac), recA56, ara, thi, F $\Delta$(proAB, lac$,Z\Delta M15$)) (17), MC1061 (hsdR, hsdM’, araD, $\Delta$(ara, leu), $\Delta$(lacPOZ), galU, galK, strA)) (18) as well as bacteriophage mRS65 (M13 lac $\Delta_Z$) (17) were stocks of this laboratory. Wild-type
Fidelity of E. coli DNA Polymerase III

M13mp2, M13mp2A89, M13mp2(-T70) gapped DNAs and gel-purified 15-mer oligonucleotide used for the 3' → 5' exonuclease assay were obtained from Dr. Thomas A. Kunkel, NIEHS. Plasmid pDNA EOPI, an α subunit overexpressor, was kindly provided by Dr. Charles S. McHenry (University of Colorado, Denver, CO). The host strain JM109 for plasmid pDNA EOPI has been described (19). Exo+ and exo- forms of P. fluorescens DNA polymerase were obtained from United States Biochemical Corp. (Cleveland, OH). 5-Bromo-4-chloro-3-indolyl β-D-galactoside (X-gal) was from Biosynth AG (Switzerland). Polyethylene glycol 8000 (PEG-8000) and cetyltrimethylammonium bromide (CTAB) were from Sigma. The primer for DNA sequencing was obtained from Research Genetics (Huntsville, AL). HiTrap-Heparin, Mono Q and Superase 12 columns for gel filtration were obtained from Amersham. Qiagen G-50 columns were purchased from Boehringer Mannheim Co. [γ-32P]ATP (5000 Ci/mmol, 10 mCi/ml) was from Amersham.

Overexpression and Purification of α Subunit—A fresh overnight culture of E. coli JM109 harboring pDNA EOPI (dnaE+) was diluted 1:200 in LB broth and shaken at 37°C until A600 reached 0.6. Isopropyl-1-thio-β-D-galactoside was added to a final concentration of 1 mM and shaking was continued for another 4 h. A total of 8 liters of the induced cells were spun down and resuspended in 50 ml of buffer containing 50 mM Tris-HCl (pH 7.5), 10% sucrose, and 8 mM dithiothreitol (DTT). Lysosome was added to a final concentration of 1.8 mg/ml. Lysis was achieved by incubating the cells at 37°C for 10 min on ice for 15 min. Cell debris was removed by centrifugation. Ammonium sulfate was added to the supernatant to a final concentration of 0.25 g/ml followed by stirring for 30 min. The precipitated protein was recovered by centrifugation, and the pellet was resuspended in buffer A containing 50 mM Tris-HCl (pH 7.5), 20 mM NaCl, 1 mM EDTA, 50 mM DTT, and 20% glycerol. The resuspension was dialyzed against 1 liter of buffer A for 16 h with two changes. It was then loaded onto a HiTrap-Heparin column (bed volume, 5 ml) equilibrated with buffer A. The column was washed with 25 ml of buffer A. Protein was eluted with 50 ml of buffer B containing a linear gradient (0–0.4 M) of NaCl at a flow rate of 0.1 ml/min. Peak fractions from two separate batches of HiTrap-Heparin chromatography were pooled and applied to a FPLC Superose 12 column (HR 10/30) equilibrated with buffer A containing 500 mM NaCl at a flow rate of 0.15 ml/min. All operations were done at 0–4°C. Polymerase activity assays (see Other Methods) and SDS-polyacrylamide gel electrophoresis were used to monitor α subunit during the purification.

3'-Exonuclease Activity Assay—A 5'-32P-labeled 15-mer oligonucleotide was annealed to M13mp2 single-stranded DNA to generate a mismatched substrate with an Affleckol G 3'-terminal dT (24). The oligonucleotide was complementary to nucleotides 106–119 of the lacZa gene with an A-G mispair located at position 105. Annealing was done in 150 mM sodium chloride, 15 mM sodium citrate, by heating the mixture at 70°C for 10 min and then allowing it to cool down to room temperature. The DNA:oligonucleotide ratio was 1.5 to 1. Following hybridization, the product was purified through a Sephadex G-50 column. Exonuclease reactions (25 μl final volume) were started by adding 10 pmol of α subunit or 12 pmol of pl Klenow fragment into a mixture containing 50 mM Tris-HCl (pH 7.8), 2 mM DTT, 10 mM MgCl2, and 300 ng of terminally-mismatched DNA. After incubation at 37°C for 10 and 60 min, 5 μl aliquots were removed and mixed with stop solution (20 mM EDTA, 95% formamide, 0.05% bromphenol blue, 0.05% xylene cyanol FF). The samples were then heated at 95°C for 5 min and run on a 20% polyacrylamide-6 M urea gel and analyzed by PhosphorImager (Molecular Dynamics, Sunnyvale, CA). Exonuclease activity was calculated as the rate of disappearance of the starting 15-mer.

Construction of mRS65 Gapped DNA—Gapped molecules were prepared essentially according to the method described by Kunkel and Soni (20). 15-kb DNA of plasmid mRS65 was isolated by Qiagen column. The double-stranded DNA was digested with restriction endonucleases MluI and SpHl, which cut at positions –60 and +379 of lacZa, respectively, producing two fragments of size 439 and 8040 base pairs (bp). The 8040-bp fragment was separated from the small fragment by PEG precipitation using 6.5% PEG-8000 in the presence of 0.5 mM NaCl and resuspended in TE buffer (pH 8.0, 0.1 mM EDTA). Using these conditions, the 8040-bp preparation does not contain detectable 439-bp fragment. 0.38 pmol of 8040-bp fragment in TE buffer were heated at 70°C for 5 min; 0.25 pmol of single-stranded circular mRS65 DNA were added, and heating was continued for another 2 min. NaCl of 1 M was added to a final concentration of 0.3 M, and heating continued at 65°C for 5 min. The product was desalted using a Microcon-30 centrifugation device and analyzed via agarose gel electrophoresis. The gapped DNA obtained through this procedure did not contain any detectable single-stranded DNA.

Gap-filling DNA Synthesis and Mutant Frequency/ Error Rate Determination—The mutant frequency was determined using the method described by Nakamura and Soni (22). RF I DNA of phage mRS65 was isolated by Qiagen column. The polyethylene glycol 8000 (PEG-8000) treatment was used for gap-filling assays. Reactions were initiated by addition of α subunit to a mixture (total volume, 20 μl) containing 20 mM Tris-HCl (pH 7.5), 8 mM DTT, 10% glycerol, 10 mM MgCl2, 1 mM each of the four dNTPs and 50 fmol of gapped DNA. For the gap-filling reaction with mp2 DNAs, 28 pmol of α subunit were added. For gap-filling reaction with mRS65 DNA, 5.4 pmol of α subunit was used. The reaction was run at 30°C for 1 h and terminated by adding EDTA to a final concentration of 15 mM. A portion of the reaction mixture was run on an 8% agarose gel to confirm completeness of gap-filling. Another portion of the reaction mixture was diluted in sterile water and transfected into competent cells of strain MC1061 by electroporation. Competent cells were prepared by electroporation (25) at an approximate cell concentration of 5.4 × 108/ml, which was mixed with 3 ml of soft agar containing 2.5 mg of X-gal and 0.5 ml of indicator strain NR9099 and plated on minimal medium plates (20). The plates were incubated at 37°C for 24 h. Mutants were scored as colorless (or light-blue) plaques in case of mp2 DNA and as (light and dark) blue plaques in case of mp2A89, mp2(T70), and mRS65. The mutant frequency was calculated by dividing the number of mutant plaques by the total number of plaques. The error rate per nucleotide polymerized in the lab target was calculated by dividing the mutant frequency (after subtraction of the control frequency) by 0.55 (the average --strand expression factor, see legend to Fig. 3) and then by the approximate number of nucleotides in the target (220 for the combined mutations or the frameshift mutations and 100 for base pair substitutions per column).

Terminal-mispair Utilization Experiments—A gapped mp2 DNA molecule containing a mismatched primer was prepared as described by Kunkel and Soni (20), except that the primer fragment was obtained by agarose gel purification, instead of PEG precipitation. The gapped primer template (gap size 366 nucleotides) contained a 3'-terminal T (template) (22). DNA synthesis by α subunit (10.8 pmol) or by E. coli DNA polymerase I Klenow fragment (exo-) (0.5 unit) was performed as described for gap-filling assays, above. The product DNA was electroporated into strain MC1061, followed by plating on strain CSH50 in the presence of X-gal and isopropyl-1-thio-β-D-galactopyranoside and scoring for plaque-color phenotype (dark-blue, light-blue, or colorless) (22).

Other Methods—DNA polymerase activities used to monitor α subunit during purification were done using poly(3αD) oligo(dT) as template-primer (150 μM as nucleotide, dA:dT = 2:1) as described (23). One unit of polymerase activity is defined as the amount that catalyzes the incorporation of 1 pmol of dTMP per min. SDS-polyacrylamide gel electrophoresis was done as described (24). Protein concentrations were determined using the Bio-Rad protein assay dye reagent with bovine serum albumin as a standard. Isolation of single- and double-stranded DNA was done by a published method (25). DNA sequencing of mRS65 mutants containing a mutation in the lab gene was performed as described (26), using an 18-mer primer complementary to positions 259 to 276 of lacZa (27). Mutations were recorded when residing in the region between position +30 and +250 (3, 27).

RESULTS

Purification and Properties of α Subunit—DNA polymerase III α subunit was purified from strain JM109 harboring the dnaE-overexpressing plasmid pDNA EOPI. From 8 liters of culture we obtained 1.5 mg of α subunit. The protein was judged to be pure (>95%) by SDS-polyacrylamide gel electrophoresis (Fig. 1) and silver staining (not shown). The specific activity was 4.5 × 106 units/mg of protein, a value comparable to those published by others (8).

Since our primary goal was to perform in vitro fidelity assays with the purified α subunit, it was important to establish the level of 3'-exonuclease activity in the preparation. Contaminat-

1 The abbreviations used are: X-gal, 5-bromo-4-chloro-3-indolyl β-D-galactoside; DTT, dithiothreitol; FPLC, fast protein liquid chromatography; bp, base pair(s); PEG, polyethylene glycol.
Fidelity of E. coli DNA Polymerase III

**Fig. 1. SDS-polyacrylamide gel electrophoretic analysis of α subunit.** DNA polymerase III α subunit from the peak fractions of the final gel filtration step were denatured and electrophoresed on an 8% SDS-polyacrylamide gel (lane A). The proteins were stained with Coomassie Blue. Lane M contained protein markers; myosin (212 kDa), α2-macroglobulin (170 kDa), β-galactosidase (116 kDa), transferrin (76 kDa), and glutamate dehydrogenase (53 kDa).

**Fig. 2. Lack of 3′ → 5′ exonuclease activity in α subunit preparation.** Supercoiled 12 gel filtration (see “Experimental Procedures”) fractions 30 through 36 were used. Each reaction contained 10 pmol of α subunit. Pol I Klenow fragment reactions contained 13.2 pmol of protein. Reactions were run for the indicated time and analyzed as described under “Experimental Procedures.”

**Table I**

| Enzyme | Forward assay (× 10⁻⁶) | Base substitution (A89) (× 10⁻⁶) | Reversion assay (T70) (× 10⁻⁶) |
|--------|------------------------|-------------------------------|-----------------------------|
| None   | 4.3                    | 13                           | 5.3                         |
| α Subunit | 30                     | 1.3                           | 310                         |

*The typical background frequency for the mp2A89 reversion assay is ~2 × 10⁻⁶ (28, 32). In the present case, we obtained 8 blue plaques among 614,000 total. For the copying reaction by α subunit, we obtained 1 blue plaque among 798,000 total.*

Reversion takes place by (−1) frameshift mutation at the 5 Ts or in a limited region around it (28). Table I shows that reversion frequency was about 30-fold above the background. The results suggest that the α subunit is quite accurate for base substitutions, but relatively inaccurate for frameshift mutations. This is an unusual specificity among DNA polymerases tested (16, 22, 28–35).

A drawback of using the mp2 system to study the fidelity of α subunit is that a very large excess of enzyme over DNA was required (>400:1) to completely fill the gap. This amount is some 10-fold higher than generally required among enzymes tested (e.g., Refs. 32, 36, and 37). Further analysis (not shown) revealed that the high enzyme requirement was likely due to secondary structures in the single-stranded region of the gapped mp2 molecules, such as the palindromic lac operator, which is located about two-thirds down the gap (38, 39). For this reason, additional, more detailed studies were performed using a new gap-filling assay based on the E. coli lacZ gene, as described below.

lacI in Vitro Forward Mutation Assay—Phage mRS65 is a derivative of phage M13 containing, in addition to the lacZ gene, the entire adjacent lac gene (M13lac”Za”) (17, 39). Since the lacI gene encodes the lac repressor, α-complementation by this vector depends on the state of the lacI gene. When plated on α-complementation host strain NR9099 (which is lacI−) in the presence of X-gal, mRS65 will produce colorless plaques, but any lacI− derivative will produce a blue plaque. Thus, this phage provides a forward mutational target where blue mutants can be readily observed against a background of colorless plaques. The outline for the lacI in vitro assay system is shown in Fig. 3. We created an mRS65 gapped molecule containing a 439-nt gap that includes the N-terminal portion (nucleotides −60 to +379) of the lacI gene. This region of lacI was chosen because it contains the so-called lacI° region (+30
through +240), which has a high density of detectably mutable sites and has been the target of many in vivo mutagenesis studies (3, 11, 40–43). There are a number of advantages to this lac in vitro assay for measurement of DNA synthesis errors. First, smaller amounts of enzyme are sufficient for complete gap-filling (see “Experimental Procedures”), presumably due to less interference by secondary structures. Second, the lac N-terminal region has been used extensively as a target for studies of mutagenesis in vivo, more than 4,000 lac mutants having been sequenced to date (3, 11, 40–43). Results obtained in vitro with E. coli DNA replication proteins can therefore be evaluated against this in vivo data base. Third, since mutants are selected as blue plaques against a background of colorless plaques, the scoring of forward mutants is facilitated.

Mutant Frequencies and Mutant Specificity in the lac in Vitro Assay System—The products of lac gap-filling reactions by α subunit were analyzed on agarose gels (Fig. 4). The band representing the gapped DNA clearly shifted to the position of double-stranded RFII DNA. The reaction products were transferred into competent cells. The mutant frequencies are shown in Table II. The frequency for uncopied DNA was 14 x 10^{-4}, while for copied DNA the frequency was 113 x 10^{-4}, indicating that α subunit made errors at a level readily detectable above the background.

To learn the precise nature of the replication errors, we sequenced 100 independent lac mutants produced by α subunit. Our sequencing effort was restricted to the lacd site of the gap (nucleotides +30 to +250). A mutation in this region was found for close to 60% of the mutants. The resulting corrected mutant frequency showed a 12-fold increase above the background (Table II). Both base substitutions and frameshifts were recovered. The mutational spectrum is presented in Fig. 5, and the results are tabulated in Table III. Base substitutions occurred at a frequency only about 2-fold above background (Table III). Therefore, α subunit did not make many base substitution errors, as was also surmised from the mp2A89 reversion assay (Table I). Frameshifts outnumbered base substitutions 7.8:1 (after subtraction of the background, see Table III). This is in contrast to the control, where base pair substitutions exceeded frameshifts by more than 10-fold. Thus, frameshifts were increased some 130-fold above the background. Based on these data, the overall error rate (per nucleotide) can be calculated to be 1/20,000. The specific rate for base substitutions is 1/91,000; for frameshift errors 1/25,000. Thus, on a per nucleotide basis, frameshifts are about 3.5-fold more frequent than base pair substitutions. However, it is likely that the calculated base substitution rate is an overestimate (see “Discussion”) and the difference between the two classes of mutations may be larger.

Almost all frameshifts were 1-base deletions, although a few (−2) frameshifts were also observed. Among the −1 deletions, deletion of G was the most frequent event, followed by deletion of A. Overall, deletion of a purine base was 4-fold more frequent than of a pyrimidine. This is not due to a bias within the target sequence, since it contains an almost equal number of purines and pyrimidines (116 versus 103). In Tables IV and V, we have further analyzed the DNA sequence dependence of the (−1) frameshifts. In general, there is no good correlation between the deletion frequency and the length of runs of identical bases (Table IV). For example, the deletion of G occurs at similar frequency whether the G is solitary or part of a GG or GGG sequence. The nearest neighbor analysis of Table V shows that for deletion of G, all 5’ neighbors (20/20) were pyrimidines. Among the 3’ neighbors, C was preferred.

Terminal Mismatch Utilization by α Subunit—We also investigated the ability of α subunit to continue synthesis from a mismatched 3’ terminus (the immediate product of a misincorporation error). We used a gapped mp2 substrate containing a terminal (template)C mismatch at position 103 of the lacZα gene (20, 22), which has been used to investigate terminal mispair utilization by a number of different enzymes (20, 22, 28, 32). One advantage of the T-C mismatch at position 103 is that its DNA sequence context permits distinction between two types of mismatch utilization (see Table VI): direct extension of the terminal C, generating a phage with the G^{103} mutation (a light-blue plaque phenotype), or extension after misalignment of the terminal C on the next template G (providing a correctly paired terminus), thus creating a (−1) frameshift mutant (colorless plaque) (20, 22). The data in Table VI show that uncopied DNA yielded no colorless plaques and only a few light-blue phages, demonstrating the poor in vivo expression of nonextended mispairs (the majority of plaques being dark-blue, representing the T-containing wild-type template strand). However, extension with either α subunit or the exonuclease-deficient form of E. coli DNA polymerase I Klenow fragment yielded a high percentage of mutant phase. Interestingly, in the case of α subunit, the plaques were in large majority colorless, whereas in the case of the Klenow fragment they were in the majority light-blue (as observed before (32)).
lacZ strain, under the sequence 

\[ \text{lacZ} \]

in vitro, and hence undesirable. If necessary, a nucleosome may be added to the lacZ probe. It is also possible to use a probe containing a random sequence (e.g., poly[dG]).

**Discussion**

In order to better understand the precise mechanisms that enable E. coli to replicate its chromosome with high fidelity, we have examined the properties of the \( \alpha \) (catalytic) subunit of E. coli DNA polymerase III holoenzyme. From an overproducing strain, \( \alpha \) subunit was purified essentially free of 3' exonuclease activity. The current experiments may therefore allow an assessment of the accuracy of this enzyme without interference by the proofreading activity provided by the normally tightly associated \( \varepsilon \) subunit. Fidelity was assessed using in vitro gap-filling assays, which used either the lacZ or the lac gene as a mutational target. The 

\[ \text{lacZ} \]

assay was specifically developed to facilitate complete gap-filling synthesis by \( \alpha \), as the enzyme proved to have great difficulty synthesizing past presumed secondary structures in the lacZ template DNA. Such a synthesis problem may not be surprising for polymerases like \( \alpha \) that are accustomed to performing DNA synthesis in the presence of multiple accessory factors.

The salient result of our study is the seemingly low ability of \( \alpha \) subunit to produce base pair substitutions, while showing a high propensity for producing frameshift mutations. This discrepancy between base pair substitutions and frameshifting is seen in both the lacZ reversion assay and the lac forward assay. In the lacZ reversion assay, we observed no increase above the background (\( \pm 10^{-6} \)) for the base substitution marker, but a significant increase (60-fold) for the frameshift marker. Among the sequenced lacZ mutations, frameshifts outnumbered base substitutions by almost 8-fold (after subtraction of the background) (Table III).

It is furthermore possible that the small increase in lacZ base substitutions (about 2-fold, see Table III) does not reflect increased base substitution replication errors by the \( \alpha \) subunit. When we attempted to enhance the frequency of the C \( \rightarrow \) T transitions, the most frequent base pair substitution found in the lacZ sequence target, the lacZ gene was mutated at a reduced rate. For the purpose of this analysis, when deletion of G occurred from a GGG sequence, the 5'- or 3'-base adjacent to the run was used.

**Table VI**

Table of the nearest-neighbor analysis of minus-G frameshifts created by \( \alpha \) subunit

| Sequence | Number of mutations |
|----------|---------------------|
| 5'-A-G  | 11                  |
| 5'-T-G  | 20                  |
| 5'-G-A  | 12                  |
| 5'-G-T  | 17                  |
| 5'-G-C  | 20                  |

**Table VII**

Table of the nearest-neighbor analysis of minus-G frameshifts created by \( \alpha \) subunit

| Sequence | Number of mutations |
|----------|---------------------|
| 5'-A-G  | 11                  |
| 5'-T-G  | 20                  |
| 5'-G-A  | 12                  |
| 5'-G-T  | 17                  |
| 5'-G-C  | 20                  |

**Table VIII**

Table of the nearest-neighbor analysis of minus-G frameshifts created by \( \alpha \) subunit

| Sequence | Number of mutations |
|----------|---------------------|
| 5'-A-G  | 11                  |
| 5'-T-G  | 20                  |
| 5'-G-A  | 12                  |
| 5'-G-T  | 17                  |
| 5'-G-C  | 20                  |

**Table IX**

Table of the nearest-neighbor analysis of minus-G frameshifts created by \( \alpha \) subunit

| Sequence | Number of mutations |
|----------|---------------------|
| 5'-A-G  | 11                  |
| 5'-T-G  | 20                  |
| 5'-G-A  | 12                  |
| 5'-G-T  | 17                  |
| 5'-G-C  | 20                  |

**Table X**

Table of the nearest-neighbor analysis of minus-G frameshifts created by \( \alpha \) subunit

| Sequence | Number of mutations |
|----------|---------------------|
| 5'-A-G  | 11                  |
| 5'-T-G  | 20                  |
| 5'-G-A  | 12                  |
| 5'-G-T  | 17                  |
| 5'-G-C  | 20                  |
might be that the base substitutions reflect a low level of errors forced by damaged bases introduced by the gap construction procedures. For example, deamination of cytosine would readily yield C→T transitions independently of dNTP pool imbalances. Such a possibility is supported by our observation that transfection of a gapped DNA molecule generally yields a mutant frequency a few fold higher than the single-stranded DNA from which it derives, as was also observed by others. Thus, the apparent frequency of base pair substitution mutations by α subunit, as measured by gap-filling assay, may be essentially \( \leq 10^{-6} \), as deduced from the experiments with the A89 marker in Table I.

It is, however, unlikely that the base substitution error rate of α is as low as the mutant frequency suggests. Sloane et al. (14) used a kinetic gel assay to determine single-site misincorporations by α subunit on an oligonucleotide template. For the four mismatches tested (A-A, A-C, A-G, and T-G), they observed misinsertion frequencies ranging from \( 1.5 \times 10^{-5} \) to \( 7.7 \times 10^{-4} \), depending on the mismatch. This range is quite typical for most DNA polymerases examined by this method (32, 44-47) and is consistent with the general substitution fidelity of non-proofreading enzymes measured by gap-filling assays (16, 22, 28–35). Therefore, the low base pair substitution frequency of α in the gap-filling reaction unlikely reflects its inability to produce such errors.

Instead, we suggest that the low base pair substitution frequency reflects the enzyme's limited ability to continue synthesis from the terminal mismatches that are created by the misincorporation event. In general, extension from mismatches is slow for most enzymes compared to correct pairs (46, 48–50). The reduced rate of extension is considered one determinant of fidelity, as it allows efficient competition by proofreading (46, 48–52) or, in case of non-proofreading enzymes, polymerase dissociation from the primer terminus. Nonextended terminal mismatches have a low probability of survival upon transfection into competent cells (20) causing the base substitution error to go undetected (see also Table VI). Kinetic experiments with bacteriophage T4 and T7 DNA polymerases (51, 52) have shown that dissociation is capable of competing effectively with mispair extension, and fidelity measurements with T7 DNA polymerase showed that this enzyme, when deprived of its proofreading factor thioredoxin, displayed an apparent antimutator effect for base pair substitutions, an effect attributed to its strongly reduced extension capability (34). Thus, it seems reasonable to assume that the isolated pol III α subunit, devoid of its supporting subunits, may have great difficulty extending mismatches (53).

1 frameshift mutations occurred at 60 to 120 times the background level in the lacI and lacZα assays, indicating that the α subunit readily makes such errors. Two general models have been proposed for the production of frameshifts during DNA replication: (Streisinger) slippage in runs (54, 55) and misincorporation plus misalignment of the misincorporated base on the next template base (22, 56). The latter can occur favorably if the misincorporated base is complementary to the next template base; the misalignment allows further synthesis to proceed from a correctly paired, although misaligned, 3’ terminus. From which pathway frameshifts arise can often be discerned by an inspection of the DNA sequence context at which they occur. Streisinger slippage is associated with runs of identical bases, the frequency increasing with the length of the run (55, 57). In contrast, frameshifts occurring via misincorporation occur readily at non-reiterated sequences in specific sequence contexts. Among the α subunit-induced frameshifts, we observed no bias in favor of frameshifts in runs (Table IV), suggesting that most of the frameshifts do not result from direct slippage. Further, a striking feature of the α subunit-induced frameshifts is that \( >90\% \) (2325) of the non-run deletions are the loss of a purine base that has a pyrimidine as its 5’-nearest neighbor (Fig. 5). This specificity has been observed for several other enzymes and is considered a hallmark of the misincorporation/realignment model (28, 56, 57). Purine-purine mispairs such as G-G or A-A are among the pairs extended the most poorly (46, 50), thus providing the greatest opportunity for the misalignment to occur on the next (complementary) pyrimidine. These combined observations are consistent with the idea that the frameshifts generated by α subunit proceed largely through the misincorporation/slippage model.

Direct evidence for the above model is provided by our experiments in which α subunit is provided with a preformed terminal mismatch (Table VI). While most other enzymes tested by this procedure (including polymerase I Klenow fragment used here) prefer (to varying degrees) direct extension to yield a base pair substitution (20, 22, 28, 32), α subunit almost exclusively favors misalignment on the next template base to yield a (–1) frameshift mutation.

One useful aspect of our current E. coli system is the possibility of comparing the in vitro error spectra with spectra of mutations observed in vivo in the same sequence target. For example, the in vivo mutation spectrum in a mismatch-repair defective mutL strain may be considered to reflect in vivo DNA replication errors. This spectrum is strongly dominated by (transition) base pair substitutions (3, 40, 58), although in one study a significant contribution (–25%) of frameshift errors was also observed (58). The latter occurred at a run of five identical bases (135–139, see Fig. 5), an event not very frequent in the α spectrum and likely representing a direct slippage mechanism. Thus, the error spectrum of the α subunit in vitro is quite different from the in vivo mutation spectrum. One possible explanation for this difference is that exonucleolytic proofreading removes the misaligned intermediates or the mispairs that promote a misalignment. However, the spectrum of mutations in mutDmutL double-mutator strains, defective in both mismatch repair and proofreading, is still dominated by base pair substitutions (3).

Thus, interestingly, significant differences exist in the fidelity behavior of the α subunit when acting by itself or when

---

2 T. A. Kunkel, personal communication.
acting as part of a greater assembly, such as the in vivo replication complex. The mechanisms underlying these differences are likely to be important for the precise replication by which the E. coli replication machinery achieves its high fidelity. It is likely that all the fidelity properties of higher order pol III assemblies, such as pol III core, pol III*, or pol III holoenzyme in our current in vitro assay system will provide further insights into this question.

Acknowledgments—We thank Dr. Charles S. McHenry for providing the α subunit overproducing vector, Dr. Thomas A. Kunkel for many helpful suggestions during the course of this work, Dr. Phuong P. Pham for providing the gapped mp2 DNA containing the A mismatches, and Drs. William C. Copeland and Dana T. Minnick for critically reading the manuscript.

REFERENCES
1. Dower, W. J., Miller, J. F., and Ragsdale, C. W. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 7160–7164
2. Echols, H. A., Lu, C., and Burgers, P. M. J. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 2189–2192
3. Schaaper, R. M. (1993) J. Biol. Chem. 268, 23762–23765
4. McHenry, C. S. (1988) Annu. Rev. Biochem. 57, 519–550
5. Kornberg, A. (1992) DNA Replication, 2nd Ed., pp. 165–182.
6. Kornberg, A., and Baker, T. A. (1978) Nature 274, 765–769
7. Kunkel, T. A., Hamatake, R. K., Motto-Fox, J., Fitzgerald, M. P., and Sugino, A. (1989) Mol. Cell. Biol. 9, 4447–4458
8. Kunkel, T. A. (1985) J. Biol. Chem. 260, 12866–12874
9. Tindall, K. R., and Kunkel, T. A. (1988) Biochemistry 27, 6008–6013
10. Bebenek, K., Abbotts, J. D., Wilson, S. H., and Kunkel, T. A. (1989) J. Biol. Chem. 264, 16948–16956
11. Schaaper, R. M. (1989) J. Biol. Chem. 264, 25557–25564
12. Fijalkowska, I. J., and Schaaper, R. M. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 2856–2861
13. Fisher, A. R., and Knill-Jones, J. W. (1981) Proc. Natl. Acad. Sci. U. S. A. 78, 4251–4255
14. Sloane, D. L., Goodman, M. F., and Echols, H. (1988) Nucleic Acids Res. 16, 6465–6475
15. Carraway, M., Rewinski, C., and Marinus, M. G. (1990) Mol. Microbiol. 4, 1645–1652
16. Kunkel, T. A. (1985) J. Biol. Chem. 260, 5787–5796
17. Schaaper, R. M., Danforth, B. N., and Glickman, B. W. (1985) Gene (Amst.) 39, 181–189
18. Huyhn, T. V., Young, H. T. V., and Davis, R. W. (1985) DNA Cloning, (Glover, D. M., ed) Vol. 1, pp. 56–110. IRL Press at Oxford University Press, Oxford, UK
19. Yanish-Perron, C., Vieira, J., and Messing, J. (1985) Gene (Amst.) 33, 103–119
20. Kunkel, T. A., and Soni, A. (1988) J. Biol. Chem. 263, 4450–4459
21. Dower, W. J., Miller, J. F., and Ragsdale, C. W. (1988) Nucleic Acids Res. 16, 6127–6145
22. Roberts, J. D., Preston, B. D., Johnston, L. A., Soni, A., Loeb, L. A., and Kunkel, T. A. (1989) Mol. Cell. Biol. 9, 466–476
23. Maki, H., Mo, J. Y., and Sekiguchi, M. (1991) J. Biol. Chem. 266, 5055–5061
24. Læmml, U. K. (1970) Nature 227, 680–685
25. Del Sal, G., Manfioletti, G., and Schneider, C. (1989) Biotechniques 7, 514–519
26. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
27. Farabaugh, P. J. (1978) Nature 274, 765–769
28. Kunkel, T. A., Hamatake, R. K., Motto-Fox, J., Fitzgerald, M. P., and Sugino, A. (1989) Mol. Cell. Biol. 9, 4447–4458
29. Kunkel, T. A. (1985) J. Biol. Chem. 260, 12866–12874