

Exonucleolytic proofreading of leading and lagging strand DNA replication errors

JOHN D. ROBERTS, DAVID C. THOMAS, AND THOMAS A. KUNKEL*

Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709

Communicated by I. Robert Lehman, January 31, 1991

ABSTRACT We have asked whether exonucleolytic proofreading occurs during simian virus 40 origin-dependent, bidirectional DNA replication in extracts of human HeLa cells. In addition, we have compared the fidelity of leading and lagging strand DNA synthesis. In a fidelity assay that scores single-base substitution errors that revert a TGA codon in the *lacZα* gene in an M13mp vector, providing an excess of a single dNTP substrate over the other three dNTP substrates in a replication reaction generates defined, strand-specific errors. Fidelity measurements with two vectors having the origin of replication on opposite sides of the opal codon demonstrate that error rates for two different A·dCTP and T·dGTP mispairs increase when deoxyguanosine monophosphate is added to replication reaction mixtures or when the concentration of deoxynucleoside triphosphates is increased. The data suggest that exonucleolytic proofreading occurs on both strands during bidirectional replication. Measurements using the two simian virus 40 origin-containing vectors suggest that base substitution error rates are similar for replication of the leading and lagging strands.

The replication of eukaryotic chromosomes is a complicated process requiring high fidelity. Studies in higher eukaryotes with the simian virus 40 (SV40) system (1–4) suggest that the enzymology at the replication fork is asymmetric—i.e., the leading and the lagging strands are synthesized by a different complement of proteins (5–7). A current model (8) suggests that the leading strand is replicated by DNA polymerase δ (pol δ), an enzyme containing an associated 3' \rightarrow 5' exonuclease activity (4), while the lagging strand is replicated by DNA polymerase α (pol α), which has generally been isolated free of exonucleolytic activity (4). We previously reported the development of an assay with which one can examine the fidelity of a human cell replication complex during bidirectional, semiconservative DNA synthesis on a circular DNA molecule containing the SV40 origin of replication and the *lacZα* gene from M13mp2 as a mutational target (9). A similar system has been described that uses the *supF* gene as the mutational target (10). We found that the replication apparatus is significantly more accurate than the purified pol α -primase complex isolated from HeLa cells (9). Given the high fidelity of the replication complex, we have now used this assay to examine the replication apparatus for the ability to proofread errors during replication. In addition, we have manipulated both reaction conditions and the DNA template to examine the base substitution fidelity of leading and lagging strand synthesis.

MATERIALS AND METHODS

Replication Extracts and Large Tumor Antigen (T Antigen). The HeLa (S3) cell cytoplasmic extract was prepared as described (11, 12). SV40 large T antigen either was purified as described (12) from HeLa cells infected with an adenovirus recombinant [R284 (13)] containing the T antigen gene under

the control of the adenovirus major late promoter or was purchased from Molecular Biology Resources (Milwaukee, WI).

Replication Reactions. Standard 25- μ l reaction mixtures contained 30 mM Hepes (pH 7.8), 7 mM MgCl₂, 4 mM ATP, 200 μ M each CTP, GTP, and UTP, 100 μ M each dATP, dGTP, dTTP, and [α -³²P]dCTP (4000 cpm/pmol), 40 mM creatine phosphate, 100 μ g of creatine phosphokinase per ml, 15 mM sodium phosphate (pH 7.5), 40 ng of either M13mp2SVA89 ori right or M13mp2SVA89 ori left replicative form DNA (123 pmol of nucleotide), \approx 1 μ g of T antigen, and \approx 75 μ g of protein from a HeLa cytoplasmic extract. After incubation for 6 hr at 37°C, the reaction was stopped as described (14), 10 μ l was removed and added to 0.5 ml of ice-cold 10% trichloroacetic acid, and the acid-insoluble products were collected on a glass fiber filter and analyzed for ³²P by liquid scintillation counting. The rest of the DNA products were collected by precipitation at -20° C in a final vol of 136 μ l in the presence of 20 μ g of tRNA/1 M ammonium acetate/50% isopropanol. The DNA pellets were dried, resuspended in 100 μ l of TE (10 mM Tris·HCl, pH 8.0/1 mM Na₄EDTA), and further purified by extraction with phenol (2 \times 100 μ l) followed by diethyl ether (2 \times 500 μ l). The DNA was then precipitated as described above (no additional tRNA added), dried, and resuspended in TE. Restriction enzymes were purchased from New England Biolabs and used as described (9). The *Dpn* I-treated, replicated DNA was then used to transfect *Escherichia coli* strain NR9162, a *mutS* form of MC1061, by electroporation with a Bio-Rad Gene Pulser. These cells were plated within 30 min with CSH50 as the β -galactosidase α -complementation strain to yield 2000–10,000 M13mp2SV plaques per plate. Revertant (blue) plaques were scored and purified as described, and single-stranded DNA was prepared and sequenced as necessary.

DNA Templates. The double-stranded M13mp2 DNA molecules used as templates were M13mp2SVA89 ori left, which has been described (9), and M13mp2SVA89 ori right, which contains the SV40 origin region (carried on an \approx 200-base-pair *Hind*III/*Sph* I fragment described previously) inserted into the *Bsu*36I site within the *lacZα* gene on the right side of the opal codon target. The procedure for inserting the SV40 origin was the same as described for the original construction of M13mp2SV (9).

RESULTS

Opal Codon Reversion Assay. We developed an assay with extracts of human HeLa cells (9) to examine the fidelity of bidirectional, semiconservative DNA replication of a circular M13mp2 DNA molecule containing the SV40 origin of replication and the *lacZα* gene as a mutational target. This fidelity assay scores single-base substitution errors that revert a TGA codon to restore α -complementation of β -galac-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: SV40, simian virus 40; T antigen, large tumor antigen; pol α , DNA polymerase α ; pol δ , DNA polymerase δ .
*To whom reprint requests should be addressed.

tosidase activity (scored as a blue plaque phenotype). This termination codon yields a colorless-plaque phenotype when the products of a replication reaction are introduced into bacterial host cells, which are then plated on indicator plates to monitor β -galactosidase activity. Base substitution errors are detected as blue plaques. The reversion frequency—i.e., the proportion of blue to total plaques—reflects the base substitution error rate.

Providing an excess of a single dNTP substrate over the other three dNTPs in a replication reaction generates defined, strand-specific errors. For example, excess dGTP leads to T-dGTP transition mispairs at position 89 in one strand of the *lacZ α* sequence at the TGA opal codon or at position 87 in the opposite strand (Fig. 1A). Likewise, excess dCTP results in transition mispairs of A-dCTP at position 89 in one strand and at position 87 in the other (Fig. 1A). The location of the error in any given revertant can be determined by DNA sequence analysis. A role for exonucleolytic proofreading on both strands can be examined by adding deoxynucleoside monophosphates, which prevent excision of a nucleotide at the primer terminus (15, 16), or by increasing the concentration of deoxynucleoside triphosphates, thus increasing the rate of polymerization of the next correct nucleotide onto a mispaired template-primer and decreasing the probability of excision (17–19).

Replication reactions were performed with excess dGTP or dCTP in order to force T-dGTP and A-dCTP mispairs (Fig. 1A). These mispairs have been shown to be effectively proofread by purified enzymes (19–22). To examine the effect of monophosphates, we chose to add dGMP to the reaction mixture, based on its ability to effectively inhibit exonucleolytic proofreading with purified enzymes (21, 23–25).

Both the pool imbalances and the addition of 1 mM dGMP slightly reduced replication efficiency (measured as total pmol of dNTP incorporated; data not shown). When aliquots of each reaction mixture were analyzed by agarose gel electrophoresis (9), the replication products were found to be

mostly monomer circular DNA molecules, with some higher molecular weight DNA present. Diagnostic restriction endonuclease digestion (9) demonstrated that the monomer circles were primarily the result of a single round of replication.

The remaining reaction products were digested with *Dpn* I to eliminate unreplicated molecules (26), and the newly replicated DNA was used to transfect competent host cells to score opal codon reversion frequencies (Table 1). Replication using equimolar concentrations of all four dNTPs yielded a reversion frequency only slightly above the reversion frequency of unreplicated DNA. Even when dGTP (Exp. 1, 200 μ M dGTP) or dCTP (Exps. 2 and 3, 200 μ M dCTP) was provided in 20-fold excess over the other dNTPs, the reversion frequency of replicated DNA was not significantly above that of unreplicated DNA. These results demonstrate that replication is highly accurate for single-base substitution errors.

Evidence for Exonucleolytic Proofreading During Replication. The addition of dGMP led to an increase in the reversion frequency for both pool bias conditions (Table 1). At 1 mM dGMP, the effect was ≥ 8.6 -fold for the dGTP pool bias (Exp. 1) and ≥ 28 -fold for the dCTP pool bias (Exp. 2). The effect of 2 mM dGMP with a dCTP pool bias was even greater, ≥ 80 -fold (Exp. 3). Since nucleoside monophosphate-induced infidelity has been taken as strong evidence for exonucleolytic proofreading with DNA polymerases (see ref. 27 and, for review, ref. 28), the simplest interpretation of these data is that proofreading is also occurring during bidirectional replication. Consistent with the interpretation that the dGMP effect was due to end-product inhibition of exonucleolytic proofreading activity, addition of 1 mM guanosine to a replication reaction mixture had no effect on reversion frequency (Exp. 3).

The next-nucleotide effect was examined with replication reaction mixtures that contained a consistent 20-fold dNTP pool imbalance. In addition, the concentration of each of the dNTPs was increased 5-fold—e.g., low dNTP reaction mix-

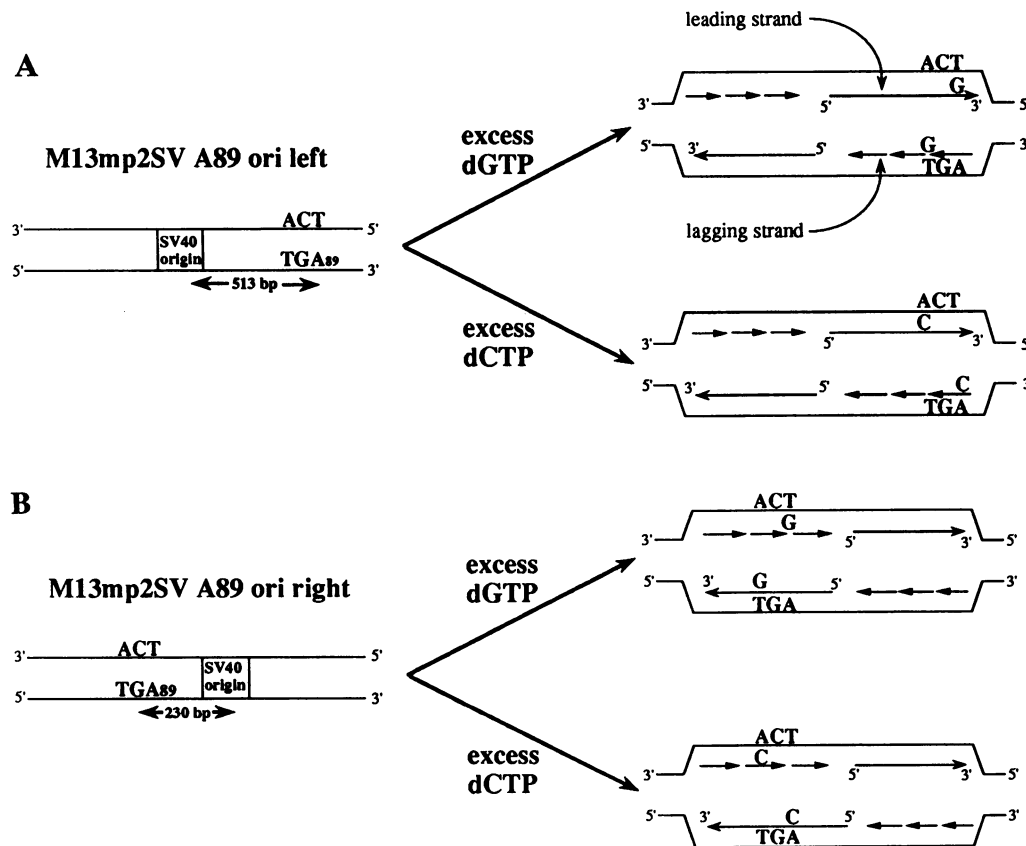


FIG. 1. Representation of leading and lagging strand synthesis across the opal codon with the SV40 origin of replication on either the left (A) or the right (B) of the *lacZ α* gene. Shown are the transition mispairs that might be generated at a single position within the opal codon on each strand in the presence of excess dGTP or dCTP. Semiconservative DNA synthesis at a replication fork requires that one strand be copied by the leading strand apparatus and one strand be copied by the lagging strand apparatus. The polarity of the DNA strands at the opal codon target remains the same regardless of the position of the origin of replication. Thus, the assignment of leading and lagging strands depends only on the position of the origin and the assumption that the replication forks move at approximately the same rate from the origin. The distance from the opal codon to the site of the first incorporated nucleotide during initiation of DNA replication is 513 nucleotides in the ori left vector and 230 nucleotides in the ori right construct. The size of the entire molecule is 7398 base pairs (bp).

Table 1. Monophosphate and next-nucleotide effects on base substitution fidelity during DNA replication

| Incorrect nucleotide | dGMP | Plaques scored | | Revertant frequency ($\times 10^{-5}$) | |
|----------------------|------|----------------------------|------|--|------------------|
| | | Total ($\times 10^{-4}$) | Blue | Observed | Minus background |
| Unreplicated | — | 110 | 15 | 1.4 | |
| Replicated (=dNTPs) | — | 68 | 16 | 2.4 | 1.0 |
| Exp. 1 | | | | | |
| 200 μ M dGTP | — | 270 | 58 | 2.1 | ≤ 1.0 |
| 200 μ M dGTP | 1 mM | 120 | 124 | 10 | 8.6 |
| 1000 μ M dGTP | — | 47 | 108 | 23 | 22 |
| 1000 μ M dGTP | 1 mM | 32 | 141 | 44 | 43 |
| Exp. 2 | | | | | |
| 200 μ M dCTP | — | 290 | 29 | 1.0 | ≤ 1.0 |
| 200 μ M dCTP | 1 mM | 140 | 403 | 29 | 28 |
| 1000 μ M dCTP | — | 100 | 54 | 5.4 | 4.0 |
| 1000 μ M dCTP | 1 mM | 130 | 285 | 22 | 21 |
| Exp. 3 | | | | | |
| 200 μ M dCTP | — | 7.6 | 0 | ≤ 1.3 | ≤ 1.0 |
| 200 μ M dCTP | 2 mM | 2.7 | 22 | 81 | 80 |
| 200 μ M dCTP | dGua | 51 | 10 | 2.0 | ≤ 1.0 |
| 1000 μ M dCTP | — | 34 | 19 | 5.6 | 4.2 |

Replication reactions were carried out as described. The *DpnI*-treated, replicated DNA was then used to transfect *E. coli* strain NR9162, a *mutS* form of MC1061, by electroporation with a Bio-Rad Gene Pulser. Revertant (blue) plaques were scored and purified as described (21), and single-stranded DNA was prepared and sequenced as necessary. Our previous observation of an increase in error rate with a 20-fold pool bias (9) reinforces the present conclusion that a next-nucleotide effect occurs since the earlier experiments were carried out with a high concentration (1 mM) of both dCTP and dGTP versus 50 μ M dATP and dTTP. Exp. 1, ratio of dGTP/dATP, dTTP, and dCTP was 20:1; Exps. 2 and 3, ratio of dCTP/dATP, dTTP, and dGTP was 20:1.

tures contained 200 μ M dGTP and 10 μ M dATP, dCTP, and dTTP, whereas high dNTP reaction mixtures contained 1000 μ M dGTP and 50 μ M dATP, dCTP, and dTTP. For errors involving misincorporation of dGTP (Table 1, Exp. 1), this increase in dNTPs resulted in a ≥ 22 -fold increase in error rate. For errors involving misincorporation of dCTP (Exps. 2 and 3), the effect was a ≥ 4 -fold increase. This effect is expected of a polymerization reaction in which proofreading is operative (refs. 17 and 18; for review, see ref. 28), because, as the dNTP concentration is increased, the rate of polymerization from a terminal mispair increases, diminishing the probability that the exonuclease will excise the error prior to polymerization. Consistent with this logic, DNA sequence analysis of revertants demonstrated that only three of the four transition mispairs showed a substantial increase in frequency. For the A-dCTP mispair at position 87, no effect of dNTP concentration was seen. An examination of the neighboring DNA sequence around the opal codon (5'-CTGA⁸⁹G-3') provides a clear explanation. This is the only transition mispair where the pool bias used (high dCTP) does not provide a high concentration of the next correct nucleotide (in this case, dGTP, complementary to the template C in the minus strand at position 88).

Fidelity of Leading and Lagging Strand Replication. Previous studies (11, 29) have suggested that the rate of replication fork movement is likely to be similar in both directions from the SV40 origin. Fidelity measurements with two vectors having the origin of replication on opposite sides of the opal codon permit a determination of the fidelity of replication of the same sequence by either the leading or lagging strand apparatus (Fig. 1). The vector used for the studies described above contains the origin of replication on the left side of the

TGA codon (M13mp2SVA89 ori left). The opal codon is 513 nucleotides distant from the first nucleotides incorporated at the origin (30), a distance that is small relative to the total size of the vector (7398 base pairs). Thus, the strand containing the 5'-TGA-3' sequence is likely to be replicated as the lagging strand (Fig. 1A). Conversely, the complementary strand, containing the template sequence 3'-ACT-5', is likely to be replicated as the leading strand.

To examine the fidelity of leading versus lagging strand replication, the opposite situation was created by placing the origin of replication on the right side of the opal codon (M13mp2SVA89 ori right). Since the distance between the origin of replication and the opal codon (230 nucleotides) is again small relative to the size of the vector, it is likely that the strand containing the 5'-TGA-3' sequence will be replicated as the leading strand (Fig. 1B).

Based on this rationale, replication reactions with the ori right vector were performed with equimolar concentrations of all four dNTPs. As with the ori left vector (Table 1), these reactions yielded low reversion frequencies (data not shown), indicating that base substitution fidelity is high for both strands. We then performed reactions under the same conditions described in Table 1 and compared the reversion frequencies with those obtained in a parallel set of reactions with the ori left vector, in both cases for A-dCTP or T-dGTP mispairs at the same position (nucleotide 89).

When a 20-fold excess of dCTP was used at a low overall substrate concentration, the reversion frequency for either vector was 1×10^{-6} (Table 2, Exp. 1). This suggests that the rate of misincorporation of dCTP opposite a template A at position 89 was low regardless of whether this template nucleotide was replicated by the leading or lagging strand replication apparatus. Addition of dGMP to these reaction mixtures increased misincorporation by 32- and 23-fold, respectively. These increases in misincorporation induced by the presence of a nucleoside monophosphate suggest that proofreading is contributing to fidelity during both leading

Table 2. Fidelity of leading and lagging strand DNA replication

| Incorrect nucleotide | 1 mM dGMP | Revertant frequency ($\times 10^{-6}$) | | Relative revertant frequency leading/lagging |
|----------------------|-----------|--|-----------|--|
| | | ori left | ori right | |
| Exp. 1 | | Lagging | Leading | |
| 200 μ M dCTP | — | 1 | 1 | 1.0 |
| 200 μ M dCTP | + | 23 | 32 | 1.4 |
| 1000 μ M dCTP | — | 12 | 3 | 0.25 |
| 1000 μ M dCTP | + | 22 | 24 | 1.1 |
| Exp. 2 | | Leading | Lagging | |
| 200 μ M dGTP | — | 5 | 4 | 1.2 |
| 200 μ M dGTP | + | 39 | 64 | 0.6 |
| 1000 μ M dGTP | — | 54 | 28 | 1.9 |
| 1000 μ M dGTP | + | 120 | 84 | 1.4 |

The background revertant frequency of unreplicated DNA (6×10^{-6}) has been subtracted. To determine which revertants could be scored as blue plaques with M13mp2SVA89 ori right, all revertants that can result from single-base changes were generated by oligonucleotide-directed mutagenesis (as described in ref. 31) and their identity was confirmed by DNA sequence analysis. The TGG codon yielded blue revertants, while codons resulting from the other single-base changes yielded colorless phenotypes. Any blue revertants recovered with this vector thus contain the TGG codon resulting from either an A89-dCTP or T89-dGTP mispair, eliminating the need to resort to DNA sequence analysis to describe error specificity. It is also possible to focus on these same two transition mispairs for the ori left vector without sequencing revertants by determining the reversion frequency for dark blue plaques only. As confirmed by sequence analysis, these contain the TGG sequence. Exp. 1, high dCTP scores A89-dCTP mispairs; Exp. 2, high dGTP scores T89-dGTP mispairs.

and lagging strand replication. Again the similarity in reversion frequencies for the same mispair at the same position suggests that, with this reaction condition, replication fidelity is similar for both the leading and lagging strands (last column in Table 2).

Reactions in which excess dCTP and a 5-fold higher overall dNTP concentration were used also yielded an increase in reversion frequency compared with the low dNTP condition for leading and lagging strand synthesis (Table 2, Exp. 1). Addition of dGMP further increased the reversion frequencies. Both effects are consistent with the suggestion that proofreading is contributing to fidelity on both strands. As for the low substrate concentration reactions, the values obtained with the two vectors at high substrate concentration and added dGMP were similar (22×10^{-6} and 24×10^{-6}). An additional experiment with excess dGTP was performed to examine the rate of misincorporation for the T-dGTP mispair at position 89 (Exp. 2). Monophosphate and next-nucleotide effects were observed for replication of both vectors and, within a factor of 2, reversion frequencies were the same with both vectors. Interestingly, with the dCTP pool imbalance in the absence of dGMP, a 4-fold difference was observed between the two vectors (compare 12×10^{-6} with 3×10^{-6}). We performed an independent repeat of this experiment and again obtained a 4-fold difference.

Specificity of Proofreading. With the ori left vector it is possible to score 10 different mismatches at the opal codon (see Table 3 and legend). To establish which of these misincorporations were responsible for the monophosphate- and next-nucleotide-mediated increases in reversion frequencies shown in Table 1, we determined the DNA sequence of revertants generated with each pool imbalance. The results demonstrate that (i) $\approx 95\%$ of the mutants contain a single-base change consistent with misincorporation of the nucleotide present in excess—i.e., either dGTP or dCTP; (ii) $\approx 94\%$ of the revertants resulted from transition base substitutions at the first or third position of the TGA codon (Table 3); (iii) monophosphate effects were observed for all four transition mismatches and next-nucleotide effects were seen for three of the four transition mismatches; and (iv) substantial quantitative differences in error rates are observed by comparing either different mismatches at the same position (the T-dGTP mispair versus the A-dCTP mispair at either position 87 or position 89) or the same mispair at different positions (the A-dCTP mismatches at two positions).

Influence of Mismatch Repair. It has recently been shown that a nuclear extract of HeLa cells contains a generalized system for repair of single-base mismatches (32). To examine

the influence of mismatch repair on measurements of replication fidelity with the cytoplasmic extracts used in this study, we performed a parallel analysis of mismatch repair, using extracts, reaction conditions (minus T antigen), and DNA concentrations (40 ng per 25- μ l reaction mixture) identical to those used in our replication reactions. Under these conditions, repair of single-base mismatches is inefficient (described in ref. 9 and more recent unpublished observations). This result and the readily observable increases in reversion frequencies resulting from selective increases in dNTP substrate concentrations suggest that most replication errors escape mismatch repair under the conditions used. Even under conditions in which mismatch repair has been found to be highly active [lower DNA concentrations (5 ng per 25- μ l reaction mixture)] (33), mismatch repair efficiency was not reduced by adding dGMP or by varying the concentration of dNTPs in the reaction mixtures. Thus, the monophosphate and next-nucleotide effects do not result from a diminution in mismatch repair. Furthermore, the efficiency of repair differed by no more than 2-fold for any of the four transition mismatches examined, suggesting that relative reversion frequencies reflect the error specificity of the replication complex.

DISCUSSION

The results presented here indicate that proofreading operates during replication of both strands. Furthermore, the data suggest that the base-substitution error rates for DNA synthesis during bidirectional replication are similar for the leading and lagging strands. This conclusion requires two simple assumptions: that in the presence of a pool bias, the mutations identified are the result of the misincorporation of the dNTP present in excess, and that the apparatus that replicates each strand is defined by the close proximity of the target to the origin and the approximately equal rate of movement of the replication forks in each direction from the origin of replication.

These findings are relevant to several aspects of DNA replication in eukaryotic cells. With only one exception, the results presented in Table 2 suggest that the fidelity of replication of the leading and lagging strands differ by no more than 2-fold for two mismatches at one template position. With a low concentration of dNTPs and no added dGMP, the fact that the reversion frequency values are low for the two vectors indicates that fidelity is high for both strands and that proofreading is contributing by factors of from 3- to ≥ 240 -fold to base substitution fidelity. Since the reversion frequencies are similar to the background frequency of unreplicated DNA, we cannot yet infer whether the fidelity of leading and lagging strand replication are different under these high-fidelity conditions.

If one makes the simplifying assumption that adding dGMP to the replication reaction mixture diminishes proofreading but does not affect base selectivity, then the similarity in reversion frequency values for the two different vectors under the low-fidelity conditions (Table 2; 1000 μ M dCTP or dGTP in the presence of dGMP) implies that, whatever the actual composition and architecture of the replication apparatus at the fork, it achieves about the same level of base selectivity for both the leading and lagging strands. The data also suggest that, at least for position 89, replication is less accurate for the T-dGTP mispair than for the A-dCTP mispair.

While seven of eight variables in Table 3 show less than a 2-fold difference in leading and lagging strand replication fidelity, one reaction condition (1000 μ M dCTP, no dGMP) yielded a 4-fold difference, wherein leading strand replication is possibly more accurate. An independent repeat of this experiment gave the same result. We interpret this difference with caution since the reversion frequency values were so

Table 3. Specificity of monophosphate and next-nucleotide effects for four transition mismatches at the TGA codon

| Mismatch | 200 μ M | | 200 μ M + dGMP | | 1000 μ M | |
|----------|--|--|--------------------|--|-----------------|--|
| | Revertant frequency ($\times 10^{-6}$) | Revertant frequency ($\times 10^{-6}$) | Relative effect | Revertant frequency ($\times 10^{-6}$) | Relative effect | |
| T87-dGTP | 4 | 47 | 12 | 65 | 16 | |
| T89-dGTP | 5 | 38 | 8 | 43 | 9 | |
| A89-dCTP | 1 | 23 | 23 | 46 | 46 | |
| A87-dCTP | ≤ 1 | 240 | ≥ 240 | ≤ 1 | — | |

Reactions were carried out with M13mp2SVA89 ori left DNA under conditions of a 20-fold pool imbalance with the incorrect dNTP present at the concentration given above each column. dGMP was added at 1 mM final concentration. Additional DNA sequence analysis indicated that the frequency of the transversion mismatches, G88-dGTP, C88-dCTP, T87-dCTP, T89-dCTP, A87-dGTP, and A89-dGTP, was not above the background of the assay at 200 μ M incorrect dNTP in either the presence or absence of dGMP. The background revertant frequency of unreplicated DNA (6×10^{-6}) has been subtracted.

close to the background frequency of unreplicated DNA. Nonetheless, this result suggests that fidelity differences may exist for replication of the two strands in some situations. Such differences, although small quantitatively, may lead over time to a cumulative bias in nucleotide composition of the DNA strands. A recent investigation (34) analyzed nucleotide substitutions in β -globin sequences from six species of primates. The nucleotide composition of the two strands indicated that changes of purines to pyrimidines occurred more frequently than changes of pyrimidines to purines. Furthermore, the equilibrium nucleotide compositions of the two strands were not identical. These asymmetries led the authors to suggest that there was an inequality in mutation rates of the two strands.

The specificity information on the errors generated in the presence and absence of proofreading is interesting relative to results with purified DNA polymerases. In these replication studies, under low-fidelity conditions that may reflect diminished proofreading, the ratio of transition to transversion base substitutions at the opal codon is $\approx 20:1$. In contrast, during gap-filling synthesis reactions with the human pol α -DNA primase complex, this ratio is $\approx 1:1$. One intriguing speculation is that accessory proteins in the replication apparatus may preferentially improve discrimination against transversion errors, as has been implied for the mutT protein in *E. coli* (35–37).

The suggestion of proofreading on both leading and lagging strands has implications for the composition of the replication apparatus. A current model for SV40 DNA replication in human cell extracts posits that DNA pol α /primase is responsible for lagging strand DNA synthesis, whereas pol δ carries out DNA synthesis on the leading strand (2, 3). Since most preparations of pol α have been found devoid of proofreading exonuclease activity (4) (however, see below), the question that arises is how proofreading may occur on both strands during bidirectional replication. There are several possible explanations. Exonuclease activity associated with pol α may simply be removed or inactivated upon purification of the polymerase. Several reports are consistent with this possibility, including the purification of some preparations of pol α that contain multiple subunits, including a 3' \rightarrow 5' exonucleolytic activity (for examples, see ref. 4), and the observation that the polymerase and exonuclease activities of the *E. coli* replicative DNA polymerase III reside in distinct subunits encoded by different genes (38, 39).

Alternatively, DNA polymerase-associated proofreading activity may be modulated by other components of the replication apparatus. Cotterill *et al.* (40) identified an exonuclease activity in preparations of pol α /primase from *Drosophila melanogaster*. This exonuclease activity was cryptic when the purified pol α /primase was associated with additional subunits. Upon removal of specific subunits, polymerase-containing fractions showed a substantial increase in the ratio of exonuclease to polymerase activity. An additional possibility is that pol ϵ may have a role in proofreading on one or both strands during DNA replication. The strategy used to detect proofreading during DNA replication—i.e., a monophosphate and next-nucleotide effect—also demonstrated proofreading with pol ϵ (21) [formerly designated DNA polymerase δ II (41)]. Recent observations suggest that the equivalent of pol ϵ in *Saccharomyces cerevisiae*, DNA polymerase II (42), is involved in chromosomal replication in this organism (43). Proofreading on both strands in crude extracts of human cells is consistent with the involvement of either pol ϵ or pol δ , which also has a 3' \rightarrow 5' proofreading exonuclease (4) and has been shown to be involved in SV40 replication *in vitro* (2, 3). There is as yet no direct evidence that pol ϵ is involved in SV40 replication in the reconstituted *in vitro* systems under study.

Another model for achieving proofreading on both strands comes from fidelity studies by Perrino and Loeb (44) with purified enzymes. These workers suggested that the exonuclease associated with pol δ may proofread errors generated by pol α at a replication fork. An extension of this hypothesis might allow any exonuclease activity at the replication fork to function on either strand or both strands of DNA during replication, regardless of which polymerase is on the leading or lagging strand.

We thank Dinh C. Nguyen for expert technical assistance and James M. Clark and Akio Sugino for critical evaluation of the manuscript.

- Kelly, T. J. (1988) *J. Biol. Chem.* **263**, 17889–17892.
- Stillman, B. (1989) *Annu. Rev. Cell Biol.* **5**, 197–245.
- Hurwitz, J., Dean, F. B., Kwong, A. D. & Lee, S.-H. (1990) *J. Biol. Chem.* **265**, 18043–18046.
- Bambara, R. A. & Jessee, C. B. (1991) *Biochim. Biophys. Acta* **1088**, 11–24.
- Downey, K. M., Tan, C. K., Andrews, D. M. M., Li, X. & So, A. G. (1988) *Cancer Cells* **6**, 1211–1218.
- Prelich, G. & Stillman, B. (1988) *Cell* **53**, 117–126.
- Focher, F., Ferrari, E., Spadari, S. & Hubscher, U. (1988) *FEBS Lett.* **229**, 6–10.
- Tsurimoto, T., Melendy, T. & Stillman, B. (1990) *Nature (London)* **346**, 534–539.
- Roberts, J. D. & Kunkel, T. A. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7064–7068.
- Hauser, J., Levine, A. S. & Dixon, K. (1988) *Mol. Cell. Biol.* **8**, 3267–3271.
- Li, J. J. & Kelly, T. J. (1985) *Mol. Cell. Biol.* **5**, 1238–1246.
- Wold, M. S., Weinberg, D. H., Virshup, D. M., Li, J. J. & Kelly, T. J. (1989) *J. Biol. Chem.* **264**, 2801–2809.
- Thummel, C., Tjian, R., Hu, S.-L. & Grodzicker, T. (1983) *Cell* **33**, 455–464.
- Wold, M. S. & Kelly, T. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2523–2527.
- Que, B. G., Downey, K. M. & So, A. G. (1978) *Biochemistry* **17**, 1603–1606.
- Kornberg, A. (1980) *DNA Replication* (Freeman, San Francisco).
- Ninio, J. (1975) *Biochimie* **57**, 587–595.
- Clayton, L. K., Goodman, M. F., Branscomb, E. W. & Galas, D. J. (1979) *J. Biol. Chem.* **254**, 1902–1912.
- Fersht, A. R. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4946–4950.
- Kunkel, T. A. & Loeb, L. A. (1979) *J. Biol. Chem.* **254**, 5718–5725.
- Kunkel, T. A., Sabatino, R. D. & Bambara, R. A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4865–4869.
- Bebenek, K. & Kunkel, T. A. (1990) *J. Biol. Chem.* **265**, 13878–13887.
- Byrnes, J. J., Downey, K. M., Que, B. G., Lee, M. Y. W., Black, V. L. & So, A. G. (1977) *Biochemistry* **16**, 3740–3746.
- Kunkel, T. A., Schaaper, R. M., Beckman, R. & Loeb, L. A. (1981) *J. Biol. Chem.* **256**, 9883–9889.
- Fersht, A. & Knill-Jones, J. W. (1983) *J. Mol. Biol.* **165**, 669–682.
- Li, J. J. & Kelly, T. J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6973–6977.
- Abbotts, J. & Loeb, L. A. (1985) *Nucleic Acids Res.* **13**, 261–274.
- Kunkel, T. A. (1988) *Cell* **53**, 837–840.
- Edenberg, H. J. & Huberman, J. A. (1975) *Annu. Rev. Genet.* **9**, 245–284.
- Hay, R. T. & DePamphilis, M. L. (1982) *Cell* **28**, 767–779.
- Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987) *Methods Enzymol.* **154**, 367–382.
- Holmes, J., Jr., Clark, S. & Modrich, P. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 5837–5841.
- Thomas, D. C., Roberts, J. D. & Kunkel, T. A. (1991) *J. Biol. Chem.* **266**, 3744–3751.
- Wu, C.-I. & Maeda, N. (1987) *Nature (London)* **327**, 169–170.
- Schaaper, R. M. & Dunn, R. L. (1987) *J. Biol. Chem.* **262**, 16267–16270.
- Bhatnagar, S. K. & Bessman, M. J. (1988) *J. Biol. Chem.* **263**, 8953–8957.
- Akiyama, M., Maki, H., Sekiguchi, M. & Horiuchi, T. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3949–3952.
- Scheuermann, R. H. & Echols, H. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7747–7751.
- DiFrancesco, R., Bhatnagar, S. K., Brown, A. & Bessman, M. J. (1984) *J. Biol. Chem.* **259**, 5567–5573.
- Cotterill, S. M., Reyland, M. E., Loeb, L. A. & Lehman, I. R. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5635–5639.
- Crute, J. J., Wahl, A. F. & Bambara, R. A. (1986) *Biochemistry* **25**, 26–36.
- Burgers, P. M. J., Bambara, R. A., Campbell, J. L., Chang, L. M. S., Downey, K. M., Hubscher, U., Lee, M. Y. W. T., Linn, S. M., So, A. G. & Spadari, S. (1990) *Eur. J. Biochem.* **191**, 617–618.
- Morrison, A., Araki, H., Clark, A. B., Hamatake, R. K. & Sugino, A. (1990) *Cell* **62**, 1143–1151.
- Perrino, F. W. & Loeb, L. A. (1990) *Biochemistry* **29**, 5226–5231.