Proofreading of DNA Polymerase η-dependent Replication Errors*

Received for publication, September 29, 2000, and in revised form, November 21, 2000
Published, JBC Papers in Press, December 11, 2000
DOI 10.1074/jbc.C000690200

Katarzyna Bebenek‡, Toshiro Matsuda‡, Chikahide Masutanis‡, Fumio Hanakoš‡, and Thomas A. Kunkel‡‡

From the ‡Laboratory of Molecular Genetics and Structural Biology, NIEHS, National Institutes of Health, Research Triangle Park, North Carolina 27709, and §Institute for Molecular and Cellular Biology, Osaka University and CREST Japan Science and Technology Corporation, 1-3 Yamada-oka, Suita, Osaka 565-0871, Japan, and the ‡Laboratory of Physical and Chemical Research (RIKEN), Wako-shi, Saitama 351-0198, Japan

Human DNA polymerase η, the product of the skin cancer susceptibility gene XPV, bypasses UV photoproducts in template DNA that block synthesis by other DNA polymerases. Pol η lacks an intrinsic proofreading exonuclease and copies DNA with low fidelity, such that pol η errors could contribute to mutagenesis unless they are corrected. Here we provide evidence that pol η can compete with other human polymerases during replication of duplex DNA, and in so doing it lowers replication fidelity. However, we show that pol η has low processivity and extends mismatched primer termini less efficiently than matched termini. These properties could provide an opportunity for extrinsic exonuclease(s) to proofread pol η-induced replication errors. When tested in human cells, extracellular extracts of pol η-induced replication infidelity was found to be modulated by changing the dNTP concentration and to be enhanced by addition of dGMP to a replication reaction. Both effects are classical hallmarks of exonucleolytic proofreading. Thus, pol η is ideally suited for its role in reducing UV-induced mutagenesis and skin cancer risk, in that its relaxed base selectivity may facilitate efficient bypass of UV photoproducts, while subsequent proofreading by extrinsic exonuclease(s) may reduce its mutagenic potential.

The UmuC/DinB superfamily of DNA polymerases includes human DNA polymerase η (pol η), 1 the product of the XPV (Rad30A) skin cancer susceptibility gene (1, 2). Human pol η has the ability to efficiently copy cis-syn thymine-thymine dimers in template DNA (1). Mutations in XPV that inactivate pol η (1, 2) render cells hypermutable by UV radiation (3–7) and defective in replicating DNA containing UV photoproducts (Ref. 8 and references therein). These facts demonstrate an important role for pol η in modulating UV-induced mutagenesis and in reducing the risk of human skin cancer. Genetic analysis reveals that human pol η inserts incorrect nucleotides opposite undamaged (9, 10) and damaged (10) template bases more efficiently than most other DNA polymerases. Moreover, human pol η lacks an intrinsic proofreading exonuclease activity (9), and its base substitution error rates when copying undamaged DNA are much higher than are those of most other eukaryotic polymerases, whether they have proofreading activity or not (9). We (9) and others (10) have suggested that this generally relaxed discrimination ability during DNA synthesis may be critical to the ability of human pol η to bypass certain DNA lesions that impede synthesis by other DNA polymerases (1, 8, 11, 12).

Current models suggest that pol η competes with other replicative polymerases for 3'-OH termini at a replication fork (reviewed in Refs. 13–15). Given the intrinsically low fidelity of pol η, mechanisms may exist to prevent errors by pol η from reducing the accuracy of chromosomal replication. We previously suggested (9) two obvious error correction mechanisms, exonucleolytic proofreading of pol η mistakes by a separate exonuclease(s) and post-replication DNA repair of mismatches generated by pol η. To test the proofreading hypothesis, here we examine the effects of pol η on the accuracy of replication of double-stranded DNA catalyzed by the replication apparatus in extracts of human cells. For this we used the SV40 replication system, whose protein requirements are similar to those of human chromosomal replication (16). We first establish replication conditions under which pol η reduces replication fidelity, suggesting that pol η can indeed compete with other polymerases during semiconservative DNA replication. We then demonstrate that pol η has two intrinsic biochemical properties, low processivity and slow mismatch extension, that could allow a separate exonuclease to compete for mismatched primer termini at the replication fork. Finally, we demonstrate that pol η-induced replication infidelity depends on the dNTP concentration and is increased in the presence of a deoxynucleoside monophosphate, both classical hallmarks of exonucleolytic proofreading.

EXPERIMENTAL PROCEDURES

Materials—Materials for the SV40 replication fidelity assay were from previously described sources (17–20), and exonuclease-deficient human pol η was described earlier (1, 9).

SV40 Replication Fidelity Measurements—SV40 origin-dependent replication reactions were performed using extracts of human TR6 and HeLa cells as described previously (18, 20). When analyzed by agarose gel electrophoresis, the products generated in the reactions listed in Table I were similar to those seen in earlier studies (16–18, 21). Analysis of the lacZ mutant frequency of the replication products by introduction into an E. coli lacZ a-complementation host strain and plating to score wild-type (blue) and mutant (colorless and light blue) plaques was performed as described previously (20).

Procesivity Analysis—Reactions (25 μl) contained 40 mM Tris-HCl (pH 8.0), 10 mM dithiothreitol, 6.5 μg of BSA, 60 mM KCl, 2.5% glycerol, 10 mM MgCl2, 1 mM dNTPs, 5 mM 30-mer template primed at a 1:2 to 1 molar ratio with a 5'-32P-labeled 20-mer oligonucleotide and 0.005 nM pol η. Ten-μl aliquots were removed after 5, 15, or 30 min at 37 °C and analyzed by electrophoresis in a 16% polyacrylamide gel in parallel with products of sequencing reactions on the same template. Product bands were quantified by phosphorimaging.

Kinetic Analysis of Mismatch Extension—Reactions (25 μl) were as above except that they contained 200 mM template primed at 1:2 to 1 molar ratio with a 5'-32P-labeled primer, 2 nM pol η, and either dATP or
Proofreading of DNA Pol \( \eta \)-dependent Replication Errors

### Table I

<table>
<thead>
<tr>
<th>Condition</th>
<th>Extract Tag</th>
<th>Pol ( \eta )</th>
<th>Extract + pol ( \eta )</th>
<th>Relative M.F.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total plaques</td>
<td>Mutants</td>
<td>M.F.</td>
<td>Total plaques</td>
</tr>
<tr>
<td>Extract-tag</td>
<td>10,101</td>
<td>8</td>
<td>8</td>
<td>50</td>
</tr>
<tr>
<td>Experiment 1. Pol ( \eta )-dependent replication infidelity (100 ( \mu )M dNTPs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Replicated (TK6)</td>
<td>10,137</td>
<td>16</td>
<td>16</td>
<td>50</td>
</tr>
<tr>
<td>Replicated (HeLa)</td>
<td>36,089</td>
<td>20</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>100 ( \mu )M dNTPs</td>
<td>13,570</td>
<td>6</td>
<td>50</td>
<td>8,483</td>
</tr>
<tr>
<td>1000 ( \mu )M dNTPs</td>
<td>19,510</td>
<td>21</td>
<td>11</td>
<td>50</td>
</tr>
<tr>
<td>Experiment 2. Stimulating polymerization by increasing the dNTP concentration</td>
<td>10 ( \mu )M dNTPs</td>
<td>25,256</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>100 ( \mu )M dNTPs</td>
<td>13,570</td>
<td>6</td>
<td>50</td>
<td>8,483</td>
</tr>
<tr>
<td>1000 ( \mu )M dNTPs</td>
<td>19,510</td>
<td>21</td>
<td>11</td>
<td>50</td>
</tr>
<tr>
<td>Experiment 3. Inhibiting exonuclease by adding dGMP (100 ( \mu )M dNTPs)</td>
<td>2 ( \mu )M dGMP</td>
<td>12,474</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Experiment 4. Replication in the presence of a dNTP pool imbalance</td>
<td>14,054</td>
<td>11</td>
<td>8</td>
<td>50</td>
</tr>
</tbody>
</table>

* The amount of pol \( \eta \) in extracts of wild-type TK6 and HeLa cells is not known.
* M.F., mutant frequency.
* In this reaction, human pol \( \beta \) (provided by S. Wilson, NIEHS) was used.
* Reactions contained 20 \( \mu \)M dATP, TTP, dCTP and 1000 \( \mu \)M dGTP.

### Table II

<table>
<thead>
<tr>
<th>Type of error</th>
<th>Extract only error rate</th>
<th>Extract + 50 ( \mu )M pol ( \eta )</th>
<th>Increase in error rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutants</td>
<td>Error rate</td>
<td>Error rate</td>
<td></td>
</tr>
<tr>
<td>Experiment 1: Equimolar 100 ( \mu )M dNTP, 100 ( \mu )M pol ( \eta )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All 12 base substitutions</td>
<td>( \leq 6.7 \times 10^{-6} )</td>
<td>17</td>
<td>( 33 \times 10^{-6} )</td>
</tr>
<tr>
<td>T*dGTP</td>
<td>( \leq 3.4 \times 10^{-6} )</td>
<td>11</td>
<td>( 100 \times 10^{-6} )</td>
</tr>
<tr>
<td>Single base deletion</td>
<td>( \leq 1.0 \times 10^{-6} )</td>
<td>7</td>
<td>( 16 \times 10^{-6} )</td>
</tr>
<tr>
<td>Experiment 2: Excess dGTP, 50 ( \mu )M pol ( \eta )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T*dGTP on lagging</td>
<td>( 12 \times 10^{-6} )</td>
<td>21</td>
<td>( 340 \times 10^{-6} )</td>
</tr>
<tr>
<td>T*dGTP on leading</td>
<td>( 3.3 \times 10^{-6} )</td>
<td>9</td>
<td>( 220 \times 10^{-6} )</td>
</tr>
</tbody>
</table>

* From Ref. 17.
* From Ref. 21.
* These results are for a 100-fold excess of dGTP, needed to obtain a mutant frequency above the control value (taken from Ref. 19).

### Results and Discussion

Previous studies indicated that the fidelity of replication of undamaged M13mp2 DNA by extracts of human TK6 or HeLa cells is high (17–19, 21). In this study, inclusion of human pol \( \eta \) in a replication reaction reduced fidelity, as indicated by the concentration-dependent increase in the frequency of lacZ mutants among the M13mp2 products of semiconservative replication (Table I, Experiment 1). No increase in mutant frequency was observed when excess human pol \( \beta \) was included or when pol \( \eta \) was incubated with the extract in the absence of the SV40 large T-antigen that is required for replication from the SV40 origin. In fact, the mutant frequencies under both reaction conditions were within the range of frequencies \( (3–7 \times 10^{-4}) \) for control DNA that had not been replicated in the extract.

DNA sequence analysis of 28 independent lacZ mutants from the reaction containing 100 \( \mu \)M pol \( \eta \) revealed the presence of 17 single-base substitutions, 11 of which were consistent with incorporation of dGTP opposite T. One clone had a tandem double-base substitution and seven contained a single nucleotide deletion. This error specificity is remarkably similar to that of pol \( \eta \) during gap-filling synthesis (9). A 14-nucleotide insertion was also recovered and three of the lacZ mutants contained two widely separated sequence changes, circumstances not encountered in previous studies of replication fidelity. Single base substitution and frameshift error rates calculated from these data are higher than for replication in the absence of added pol \( \eta \) (Table II, Experiment 1). Overall, these data suggest that pol \( \eta \) is capable of competing with other DNA polymerases at the replication fork and generates both base substitution and frameshift mutations.

We previously suggested that the mutagenic potential of pol \( \eta \) in human cells might be reduced if pol \( \eta \)-dependent errors were proofread (9). Several experiments were conducted to test this hypothesis. First, primer extension reactions were performed using the lacZ template and a 1000-fold molar excess of template-primer over enzyme, a condition that results in a single cycle of processive synthesis. Analysis of reaction products (Fig. 1) demonstrated that pol \( \eta \) polymerizes one to ten nucleotides per cycle of enzyme binding dissociation and that
the probability of termination of processive synthesis after each incorporation varies between about 40 and 70%. This result quantitatively confirms the earlier observation (8) that human pol η has low processivity, a property that could provide an exonuclease access to a template-primer containing a mismatch.

All DNA polymerases studied to date extend mismatched template-primer sequences less efficiently than matched termini. To determine whether this is also the case with human pol η, we performed extension reactions to obtain steady-state kinetic constants from which extension efficiencies for matched and mismatched termini were calculated. The results (Table III) indicate that pol η extended mismatched termini less efficiently than matched termini by factors of from 3-fold (T→G mismatch) to more than 100-fold (G•A mismatch). These data are consistent with the recent qualitative demonstration that human pol η extends mismatched termini less efficiently than matched termini with undamaged DNA and at sites of DNA damage (8) and with a recent report that yeast pol η also extends mismatched termini less efficiently than matched termini (22).

For polymerases having intrinsic proofreading exonucleases, slow polymerization increases the opportunity for movement of the primer terminus to an exonuclease active site for excision of a misinserted nucleotide (recently reviewed in Ref. 23). While pol η lacks an intrinsic proofreading activity (9), it is possible that editing could be performed by a separate exonuclease that is physically associated with the replication machinery, as is the case in E. coli. To determine whether replication errors induced by exonuclease-deficient pol η could be edited by an extrinsic exonuclease, replication fidelity was examined in a human extract under conditions known to modulate proofreading activity. One hallmark of proofreading is the “next nucleotide effect” (reviewed in Ref. 24). Since the probability of polymerization from a mismatch (or a misalignment) depends on the concentration of the next correct nucleotides to be incorporated, at high dNTP concentrations polymerization is favored over editing, and fidelity is reduced. This approach has already been used successfully to detect proofreading of base substitution (18, 19) and frameshift errors (19, 25) produced by the human replication complex. However, in reactions containing equimolar dNTPs, a next nucleotide effect is difficult to detect with the M13mp2 forward mutation assay, because replication fidelity is very high (17–21, 25). Thus, replication in an extract to which excess pol η was not added generated products with lacZ mutant frequencies that were similar to unreplicated DNA control values (Table I, Experiment 2). In contrast, when replication reactions were performed in the presence of pol η, fidelity decreased as the dNTP concentration was increased from 10 to 1000 μM (Table I, Experiment 2). The fact that the fidelity was higher in reactions lacking exogenous pol η indicates that the majority of errors being proofread were dependent on pol η.

A second hallmark of proofreading activity is reduced replication fidelity in the presence of a high concentration of dNMP (18–20), the product of proofreading exonucleases. In the present study, inclusion of 2 mM dGMP in a replication reaction containing 100 μM dNTPs and to which pol η had been added

### Table III

**Kinetic analysis of mismatch extension by human pol-η**

Reactions were performed as described under “Experimental Procedures.” Duplication determinations were performed using seven different concentrations of nucleotide, and kinetic constants were derived as described previously (32). The template primers used were: (a)

\[
\begin{align*}
\text{dATP} & \quad 5'-\text{XTCTTTTGGGACCGCAATGG-5'} \\
\text{dGTP} & \quad 5'-\text{AGCTCGAGACTGAGAAACCCTGCUCCACCA-3'}
\end{align*}
\]

(where X = C, T, or A) and (b)

\[
\begin{align*}
\text{dATP} & \quad 5'-\text{XTCTTTTGGGACCGCAATGG-5'} \\
\text{dGTP} & \quad 5'-\text{AGCTCGAGACTGAGAAACCCTGCUCCACCA-3'}
\end{align*}
\]

(where X = A, G, or C).

<table>
<thead>
<tr>
<th>Template•Primer</th>
<th>Correct dNTP</th>
<th>(K_m) (μM)</th>
<th>(k_{cat(app)}) (min(^{-1}))</th>
<th>(\frac{(k_{cat})}{(K_m)}) (app)</th>
<th>(f_{cat})</th>
</tr>
</thead>
<tbody>
<tr>
<td>G•C (a)</td>
<td>dATP</td>
<td>31 ± 8.8</td>
<td>13 ± 4.4</td>
<td>420 ± 190</td>
<td>1</td>
</tr>
<tr>
<td>G•T (a)</td>
<td>dATP</td>
<td>730 ± 160</td>
<td>3.3 ± 1.1</td>
<td>4.5 ± 1.8</td>
<td>0.011</td>
</tr>
<tr>
<td>G•A (a)</td>
<td>dATP</td>
<td>440 ± 320</td>
<td>1.6 ± 0.9</td>
<td>3.5 ± 3.3</td>
<td>0.008</td>
</tr>
<tr>
<td>T•A (b)</td>
<td>dGTP</td>
<td>37 ± 8.2</td>
<td>11 ± 0.0</td>
<td>310 ± 69</td>
<td>1</td>
</tr>
<tr>
<td>T•G (b)</td>
<td>dGTP</td>
<td>85 ± 48</td>
<td>8.4 ± 0.9</td>
<td>98 ± 56</td>
<td>0.32</td>
</tr>
<tr>
<td>T•C (b)</td>
<td>dGTP</td>
<td>960 ± 110</td>
<td>3.2 ± 1.8</td>
<td>3.3 ± 1.9</td>
<td>0.011</td>
</tr>
</tbody>
</table>
corporation opposite template T, the most frequent error gen-
reflected misincorporation of dGMP. Thirty were due to misin-
reaction revealed 33 single base substitutions, 31 of which
matches (Table III) facilitates proofreading (Table I). Proof-
processive synthesis (Fig. 1) and slow extension of mis-
even if pol 
that discrimination for extension of a T
extends a mismatch for a few nucleotides (note
hot spot could directly bind the mismatch for immediate re-
results imply that pol \( \eta \)-in-
duced replication errors can be proofread.
We also examined whether pol \( \eta \) competes for primer termini
on the leading strand, the lagging strand, or both. Here we
employed a strategy used previously (19), in which excess
dGTP is included in a replication reaction to monitor misincor-
poration of dGTP during lagging strand replication of the
M13mp2 viral (+) strand or during leading strand replication
of the complementary (−) strand. A reaction containing 50 nM
pol \( \eta \) and a 50-fold molar excess of dGTP generated replication
products whose \( \text{lacZ} \) mutant frequency was strongly elevated
relative to control values (Table I, Experiment 4). Sequence
analysis of 29 independent \( \text{lacZ} \) mutants recovered from this
reaction revealed 33 single base substitutions, 31 of which
reflected misincorporation of dGMP. Thirty were due to misin-
corporation opposite template T, the most frequent error gen-
erated by pol \( \eta \) during gap-filling DNA synthesis (9). Of these,
nine were consistent with pol \( \eta \) misincorporation during repli-
cation of the leading strand and 21 reflected pol \( \eta \) misincor-
poration during replication of the lagging strand (Table II). The
error rates calculated from these data are much higher than
those observed for replication with excess dGTP in the absence
of excess pol \( \eta \) (Table II). Notably, seven independent \( \text{lacZ} \)
mutants contained the same T to C substitution at template
121. At this site, misinsertion of dGTP opposite this
mutants to promote genomic instability by an unknown mechanism.

Acknowledgments.—We thank William A. Beard and Youri Pavlov for
critical evaluation of the manuscript.

REFERENCES