Oxidation of Thymine to 5-Formyluracil in DNA Promotes Misincorporation of dGMP and Subsequent Elongation of a Mismatched Primer Terminus by DNA Polymerase*

Aya Masaoka, Hiroaki Terato, Mutsumi Kobayashi, Yoshihiko Ohyama and Hiroshi Ide‡

From the Department of Mathematical and Life Sciences, Graduate School of Science, Hiroshima University, Higashi-Hiroshima 739-8526, Japan

*This work was supported by Grants-in-Aid from Ministry of Education, Science, and Culture of Japan (H.I.) and by JSPS Research Fellowships for Young Scientists (A.M.).
‡To whom correspondence should be addressed. Phone. & Fax: +81-824-24-7457; E-mail: ideh@hiroshima-u.ac.jp.

Running title: Mutagenic Potential of 5-Formyluracil
Summary
5-Formyluracil (fU) is a major oxidative thymine lesion generated by ionizing radiation and reactive oxygen species. In the present study, we have assessed the influence of fU on DNA replication to elucidate its genotoxic potential. Oligonucleotide templates containing fU at defined sites were replicated in vitro by *Escherichia coli* DNA polymerase I Klenow fragment deficient in 3’-5’ exonuclease. Gel electrophoretic analysis of the reaction products showed that fU constituted very weak replication blocks to DNA synthesis, suggesting a weak to negligible cytotoxic effect of this lesion. However, primer extension assays with a single dNTP revealed that fU directed incorporation of not only correct dAMP but also incorrect dGMP, albeit much less efficiently. No incorporation of dCMP and dTMP was observed. When fU was substituted for T in templates, the incorporation efficiency of dAMP ($f_A = V_{max}/K_m$) decreased to 1/4-1/2 depending on the nearest neighbor base pair and that of dGMP ($f_G$) increased 1.1- to 5.6-fold. Thus, the increase in the replication error frequency ($f_G/f_A$ for fU vs. T) was 3.1- to 14.3-fold. The misincorporation rate of dGMP opposite fU ($pK_a = 8.6$) but not T ($pK_a =10.0$) increased with pH (7.2-8.6) of the reaction mixture, indicating the participation of the ionized (or enolate) form of fU in the mispairing with G. The resulting mismatched fU:G primer terminus was more efficiently extended than the T:G terminus (8.2- to 11.3-fold). These results show that when T is oxidized to fU in DNA, it promotes both misincorporation of dGMP at this site and subsequent elongation of the mismatched primer, hence potentially mutagenic.
Faithful replication of DNA is essential for maintaining genetic integrity of living organisms. High fidelity of DNA replication is achieved by two cellular functions that involve discrimination of correct vs. incorrect nucleotides by DNA polymerases (1, 2) and postreplication mismatch repair (3). The overall error frequency of DNA replication is one in $10^8$ to $10^{10}$ base pairs when they function properly. Fidelity of DNA replication also relies on the structural integrity of DNA itself that serves as a template for the newly synthesized strand. A number of endogenous and exogenous agents have been identified to induce structural deterioration of DNA (4). Among them, reactive oxygen species generate a very complicated spectrum of DNA damage (5, 6). These lesions are mostly restored by the base excision repair pathway both in prokaryotic and eukaryotic cells, but if left unrepaired, they arrest DNA synthesis or direct misincorporation of nucleotides during DNA replication, hence exerting deleterious effects on cells (7, 8). Replication blocks and nucleotide misincorporation have been related to lethality and mutation of cells, respectively, until recently. However, this concept is now challenged by the discovery of numerous error-prone and error-free DNA polymerases that can bypass the blocking lesions (9).

Although past several years have witnessed the discovery of novel lesion replicating DNA polymerases (mentioned above) as well as remarkable progress in understanding the molecular basis for the nucleotide discrimination mechanism by DNA polymerases (10, 11), the assessment of genotoxic effects of structurally diverse oxidative DNA damage largely relies on experimental data obtained from defined lesions (7, 8, 12, 13). We have been studying the response of DNA polymerases to the encountered oxidative thymine (13-17) and other base (18-21) lesions. Oxidative thymine lesions formed by ionizing radiation, Fenton-type reactions and photosensitized reaction have been best characterized among the four DNA bases and can be classified into four subgroups depending on their structural features. The first group includes C5-C6 saturation products such as thymine glycol (5,6-dihydroxy-5,6-dihydrothymine), 5,6-dihydro-5-hydroxymethylthymine. 5,6-Dihydrothymine belongs to this group though it is a reduction product formed by ionizing radiation. The second group
is ring fragmentation products such as a urea residue and its analogues. The response of DNA polymerases to the first and second groups has been clarified fairly well by in vitro and in vivo studies (13-15, 22, 23). The third group includes 5-hydroxy-5-methylhydantoin, a ring contraction product. The ability of this lesion to block DNA replication has been recently demonstrated by in vitro DNA polymerase reactions using a defined oligonucleotide template (24). The fourth group contains methyl oxidation products such as 5-hydroxymethyluracil and 5-formyluracil (fU)†. Several lines of evidence indicate that 5-hydroxymethyluracil is neither a replicative block nor mutagenic (25, 26), hence being an innocuous lesion.

The genotoxic potential of fU belonging to the fourth group has been assessed in this (27, 28) and other (29, 30) laboratories. In our previous approach, we synthesized 5-formyl-2′-deoxyuridine 5′-triphosphate (fdUTP) and studied its incorporation into DNA by DNA polymerases. fdUTP efficiently substituted for dTTP and to a much less extent for dCTP. Moreover, the pH-dependent variation of the substitution efficiency for dCTP suggested involvement of an ionized (or enolate) form of fU as a key intermediate responsible for the mispairing with template G. Such a mutation mechanism involving ionized bases (thymine and 5-bromouracil (BrU)) was originally suggested by Lawley and Brookes (31). Later, the pH-dependent variation of the replication error frequency due to ionization of 5-BrU and 5-fluorouracil was experimentally demonstrated by Yu et al. (32), providing the conceptional basis of the previous studies (27, 28). Translesion bypass and nucleotide incorporation at the site of template fU were also studied (29, 30) using oligonucleotides containing the site specific lesion (33). Consistent with very efficient substitution of fdUTP for dTTP in our study, DNA polymerase readily passed through the fU site in the template. However, to our surprise, the primer extension study showed that fU directed incorporation of dCMP as well as correct dAMP, implying the formation of an fU:C mispair during DNA replication. The discrepancy between the two studies concerning the base pairing capacity of fU might have originated from several reasons. First, although overall Watson-Crick geometry of a newly formed base pair plays a dominant
role in nucleotide selection by DNA polymerases (1, 2), this process can also be
affected by base pairing symmetry whether a X:Y base pair is formed from X
(template):Y (dNTP) or X (dNTP):Y (template) (34-36). Secondly, the sequence
context can affect the selection of dNTP opposite the template lesion (7). Thirdly, the
base ionization mechanism somehow does not hold when fU is present in template
DNA.

In view of the potential influences of base pairing asymmetry, the sequence
context, and deviations from the base ionization mechanism mentioned above, we have
prepared oligonucleotide templates containing site specific fU following the previously
reported phosphoramidite method (33) and reexamined the base pairing capacity of
template fU with the four possible nearest neighbor base pairs. The results show that
fU in the template directs misincorporation of dGTP in a pH-dependent manner,
supporting our previous results obtained by the analysis of fdUTP incorporation.
EXPERIMENTAL PROCEDURES

Chemicals and Enzymes——Ultra-pure dATP, dGTP, dCTP and dTTP (purity > 99.3%) were purchased from Amersham Pharmacia. 5-Formyl-2’-deoxyuridine (fdU) was synthesized as described previously (27). The phosphoramidite monomer of protected 5-(1,2-dihydroxyethyl)-2’-deoxyuridine was synthesized following the reported procedure (33). $[\gamma^{32}\text{P}]\text{ATP}$ (110 TBq/mmol) was purchased from Amersham Pharmacia. 

*Escherichia coli* DNA polymerase I Klenow fragment (Pol I Kf), Pol I Kf deficient in 3’-5’ exonuclease [Pol I Kf (exo-)], and T4 polynucleotide kinase were obtained from New England Biolabs, and *Penicillium citrium* nuclease P1 and calf intestine alkaline phosphatase were from Boehringer Mannheim.

Oligonucleotides——Oligonucleotides comprising normal components were synthesized by the standard phosphoramidite method and purified by reversed phase HPLC. Oligonucleotides containing fU were prepared following the reported procedure (33). First, oligonucleotides containing 5-(1,2-dihydroxyethyl)uracil, a precursor of fU, were synthesized by the standard phosphoramidite chemistry using a phosphoramidite monomer of protected 5-(1,2-dihydroxyethyl)-2’-deoxyuridine, and purified by reversed phase HPLC. The oligonucleotides containing 5-(1,2-dihydroxyethyl)uracil was treated by sodium periodate to convert 5-(1,2-dihydroxyethyl)uracil to fU. After treatment, crude oligonucleotides were desalted by passing through a Sephadex G-10 column and further purified by reversed phase HPLC. The sequences of oligonucleotides used in the present study are listed in Table I.

Composition Analysis of Oligonucleotides——25T and 25F (0.2 OD) were incubated with nuclease P1 (1 unit) in a reaction buffer (60 μl) containing 21 mM sodium acetate (pH 5.3) and 1 mM ZnSO$_4$ at 37 °C for 1 h. To this solution, alkaline phosphatase buffer (30 μl) comprising 0.5 M Tris-HCl (pH 9.0) and 10 mM MgCl$_2$ and alkaline phosphatase (3 units) were added and the reaction mixture was further
incubated at 37 °C for 2 h. The sample was passed through a molecular weight cut off filter (Mw = 10000) and an aliquot of the filtrate was analyzed by HPLC equipped with a C18 WS-DNA column (4.6 x 150 mm, Wako). The sample was eluted by a gradient of methanol in 10 mM sodium phosphate buffer (pH 7.4) at a flow rate 0.8 ml/min. The concentration of methanol was 0% for a 0-5 min and 0-5% linear gradient for 5-35 min. The column temperature was maintained at 40 °C by a column oven and eluents were monitored at 280 nm.

_Treatments with Repair Enzymes_——25F was 5’-end labeled as described for the primers (see below) and annealed to the complementary strand (25COM). 25F/25COM (0.02 pmol) was incubated with AlkA (1 pmol) followed by endonuclease (Endo) IV (0.03 pmol) in a reaction buffer (10 μl) at 37 °C. Alternatively, the substrate was incubated with Endo III (0.3 pmol), formamidopyrimidine glycosylase (Fpg, 0.3 pmol), or Endo IV (0.03 pmol) in a similar manner. The procedures of the enzymatic treatments were essentially similar to those reported previously for AlkA (28, 37), Endo III (38), Fpg and Endo IV (39). After incubation, products were analyzed by 16% denaturing PAGE.

_Preparation of Template/Primer_——The primers were 5’-end-labeled using T4 polynucleotide kinase and [γ-32P]ATP. Appropriate template and primer (molar ratio = 2:1) were annealed in 20 mM Tris-HCl (pH 7.5) and 50 mM NaCl by heating the solution at 90 °C for 5 min and allowing to cool slowly to room temperature.

_Analysis of Translesion DNA Synthesis_——The annealed template/primer (0.5 pmol) containing T (24AT/P10 and 24CT/P10) or fU (24AF/P10 and 24CF/P10) at the same position was incubated with Pol I Kf (exo-) (0.05 unit) and four dNTPs (50 μM each) in a polymerase reaction buffer (15 μl) at 25 °C for 3-10 min. The polymerase reaction buffer consisted of 66 mM Tris-HCl (pH 7.5), 1.5 mM 2-mercaptoethanol and 6.6 mM MgCl2.

_Analysis of Nucleotides Incorporated Opposite fU_——To determine the nucleotide incorporated opposite fU, primer extension reactions were performed in a manner essentially similar to that described for translesion DNA synthesis except that...
the reaction mixture contained a single dNTP (50 μM) and incubation time was 5 min. The template/primer used in the analysis was 24AF/13T, 24GF/13C, 24CF/13G and 24TF/13A that contained fU at the same position and different nearest base pairs next to fU (i.e. primer terminus base pairs). For control reactions, primer extension assays were also performed using template/primers (24AT/13T, 24GT/13C, 24CT/13G and 24TT/13A) that contained T in place of fU. Using these template/primers containing fU or T, kinetic parameters of nucleotide incorporation opposite fU and T were also determined. For dATP incorporation, the dATP concentration was 0.05-1 μM and the amount of Pol I Kf (exo-) was 0.002 unit, while for dGTP incorporation, those were 10-80 μM and 0.03 unit. The incubation time was 5 min for both dATP and dGTP incorporation assays. Under these conditions, the extent of primer elongation was essentially proportional to the reaction time and the unit of Pol I Kf (exo-) used. The initial velocity of the reaction (V) (average of two experiments) was calculated as the percent of the extended primer per min per 0.03 unit of Pol I Kf (exo-). $K_m$ and $V_{max}$ values were evaluated from a hyperbolic curve fitting program.

The pH effect on the incorporation of dATP and dGTP opposite template fU or T was determined by varying the pH (pH 7.2-8.6) of the polymerase reaction buffer as described above. The template/primer (24AF/13T and 24AT/13T, 0.5 pmol) was incubated with Pol I Kf (exo-) (0.03 unit) and dATP or dGTP (20 μM) at 25 °C for 5 min. For a wide pH range (pH 6.9-9.3), GTA buffer (buffering capacity pH 3.5-10) was used in place of the Tris buffer for the DNA polymerase reaction. The composition of GTA buffer was 3,3-dimethylglutaric acid, Tris, and 2-amino-2-methyl-1,3-propanediol (17.3 mM each), and pH was adjusted by adding HCl or NaOH.

Extension of Mismatched Primer Terminii——Template/primers (0.5 pmol) containing a correctly paired (24AT/14TA and 24AF/14TA) or mismatched (24AT/14TG and 24AF/14TG) primer terminus were incubated with Pol I Kf (exo-) (0.05 unit) and dCTP (10 μM) in the polymerase reaction buffer (15 μl) at 25 °C for 5 min. Alternatively, kinetic parameters of the reaction (average of two experiments) were determined by varying the dCTP concentration between 0.01-1 μM for the
correctly paired primer terminus (24AT/14TA and 24AF/14TA) or 0.01-20 μM for the mispaired primer terminus (24AT/14TG and 24AF/14TG). The reaction was performed as described above except that the amount of Pol I Kf (exo') was 0.015 unit. The parameters for mismatch extension with templates 24CT and 24CF were also determined in the same manner.

**Electrophoresis**—DNA polymerase reactions were terminated by adding loading buffer containing 0.1% xylene cyanol, 0.1% bromophenol blue, 20 mM EDTA and 95% formamide. The sample was boiled and subjected to 16% denaturing polyacrylamide gel electrophoresis (PAGE). Electrophoresis was performed at 1800 V and gel was autoradiographed at -80 °C overnight. The radioactivity of the separated bands was quantified on a phosphor imager Fuji Bas 2000.
RESULTS

Nucleoside Composition of Oligonucleotides—To ensure the validity of phosphoramidite method used in the present preparation of oligonucleotides containing fU, pilot oligonucleotides containing T (25T) and fU (25F) in the same sequence were prepared. 25T and 25F were digested by nuclease P1 and alkaline phosphatase and the nucleoside composition was analyzed by HPLC. Digestion of 25T resulted in the HPLC peaks of dC, dG, dT and dA with an expected molar ratio (8:6:4:7) (data not shown). In the HPLC analysis of the digested 25F, two extra peaks were observed in addition to the four normal nucleosides (Fig. 1 A). The first peak eluted at 15.1 min was readily identified as 5-formyl-2’-deoxyuridine (fdU) by comparison with the retention time of authentic fdU. The peak at 32.4 min (indicated by *) was an unknown product. When authentic fdU was incubated under the same conditions as those used for oligonucleotide digestion, fdU was partially converted to this product (Fig. 1 B). The formyl group of fdU is fairly reactive and forms adducts with nucleophiles (40). Accordingly, this product is most likely an adduct between fdU and a nucleophilic molecule present in the reaction buffer or enzyme preparations. The long HPLC retention time of the product relative to fdU and retention of the UV absorption around 280 nm were also consistent with the adduct formation of the exocyclic formyl group of fdU. The fU moiety of fdU is known to be degraded by strong base and oxidizing reagents, giving rise to ring fragmentation products (27, 41) and 5-carboxyuracil (42), respectively. If such products were formed during the preparation of 25F, they might be present as contaminated lesions in 25F. However, the retention of the UV absorption (around 280 nm) of the product was inconsistent with the ring fragmentation products bearing no chromophores. The retention time of the product (32.4 min) much longer than fdU (15.1 min) in the reversed phase HPLC column also contradicted the expected very short retention time of 5-carboxy-2’-deoxyuridine bearing a negative charge of a carboxylate ion. Thus, the product was not the ring fragmentation products or 5-carboxy-2’-deoxyuridine. Moreover, when the amount of authentic fdU converted
to the putative adduct was taken into account, the corrected molar ratio of nucleosides in 25F agreed with the expected value (dC:dG:dT:dA:fdU = 8:6:3:7:1).

We also attempted to identify the structure of the unknown product by mass spectrometric analysis after isolating the product by HPLC. However, the attempt was unsuccessful because of the lack of the apparent molecular ion (M⁺) in the spectrum. In an alternative approach, 25F (as a duplex) was digested by several DNA repair enzymes with different damage specificities. Consistent with the previous reports (28, 37), the treatment with AlkA followed by Endo IV resulted in incision of 25F at the fU site (data not shown). In contrast, neither Endo III, Fpg, nor Endo IV incised 25F², supporting the absence of base damage other than fU in 25F. On the basis of the results from the composition analysis and the treatment with repair enzymes, we concluded that fU was successfully incorporated into oligonucleotides in the present procedure of synthesis.

Translesion DNA Synthesis at the fU Site——To clarify whether fU present in template DNA constitutes a replication block, 24AF and 24AT containing fU and T, respectively, at the same site (4 nucleotides beyond the primer terminus) were primed by ³²P-labeled P10, and the templates were replicated by Pol I Kf (exo⁻) for up to 10 min. The resulting products were analyzed by denaturing PAGE (Fig. 2). After 3 min of incubation, the primer annealed to the undamaged template 24AT was almost completely extended to a fully replicated product (lane 2). Similarly, the primer annealed to 24AF containing fU was mostly extended to fully replicated and one nucleotide shorter products after 3 min of incubation (lane 6). In addition, very weak bands also appeared at and one nucleotide prior to the fU site, indicating a pause of DNA synthesis at these sites. Quantification of the arrested and bypassed products showed that 91% of the original primer was extended beyond the fU site at 3 min. This result indicates that fU in template DNA allows efficient translesion DNA synthesis. Similar results were obtained with templates 24CF and 24CT (data not shown).

Nucleotides Incorporated opposite fU——Since efficient translesion DNA synthesis occurred at the fU site, the nucleotide incorporated opposite this lesion was analyzed by a primer extension assay. In this assay, primers (13T, 13C, 13G, 13A)
that were one nucleotide shorter than the template fU site were annealed to appropriate templates (24AF, 24GF, 24CF, 24TF) and the primers were extended by Pol I Kf (exo') in the presence of a single dNTP (50 μM) at 25 °C for 5 min. These template/primers (24AF/13T, 24GF/13C, 24CF/13G, and 24TF/13A) contained four different nearest neighbor base pairs in the primer terminus. Control experiments were also performed under the same conditions using template/primers (24AT/13T, 24GT/13C, 24CT/13G, 24TT/13A) that contained T instead of fU. The reaction products were analyzed by denaturing PAGE. Fig. 3A shows PAGE data obtained for 24AF/13T and 24AT/13T containing an A (template):T (primer) pair at the primer terminus. According to the band intensity of the extended product (14-mer), dAMP was most efficiently incorporated opposite fU (lane 6) as well as T (lane 1), with a preference for T. The bands indicative of misincorporation of dGMP opposite fU (lane 7) and T (lane 2) were also observed, but the incorporation was more efficient for fU than T. In the presence of dCTP and dTTP, extended products were not observed over the background for both templates containing fU (lanes 8 and 9) and T (lanes 3 and 4), showing that the misincorporation frequency of dCMP and dTMP was below the detection limit under these conditions. Essentially similar results were obtained with other template/primers containing G:C (Fig. 3B), C:G (Fig. 3C), and T:A (Fig. 3D) as the nearest neighbor base pairs. Accordingly, fU directed incorporation of correct dAMP and to a less extent incorrect dGMP but not pyrimidine nucleotides (dCMP and dTMP).

Parameters of dAMP and dGMP Incorporation———For quantitative analysis of the nucleotide incorporation efficiency and the sequence context effect, kinetic parameters of dAMP and dGMP incorporation opposite fU and T were determined by the gel fidelity assay under standing start conditions (43). The experiments were performed as described under experimental procedures using a set of template/primer employed above (Nucleotides Incorporated opposite fU). Table II summarizes the parameters ($V_{max}$ and $K_m$) and efficiencies ($f = V_{max}/K_m$) of dAMP incorporation (average of two experiments). Although there were variations depending on the nearest
neighbor base pair, the $V_{\text{max}}$ values for dAMP incorporation were consistently higher for T than fU. Conversely, the $K_m$ values for T were consistently lower than for fU. Consequently, the incorporation efficiency of dAMP ($f_A$) opposite fU was reduced to $1/4$-$1/2$ of that opposite T. The efficiency difference between T and fU was not large but significant, showing that conversion of T to fU in template DNA slows down incorporation of the correct nucleotide dAMP. Table III summarizes the parameters ($V_{\text{max}}$ and $K_m$) and efficiencies ($f_G = V_{\text{max}}/K_m$) of dGMP misincorporation together with the replication error frequencies ($f_{\text{RE}} = f_G/f_A$) (average of two experiments). With the same nearest neighbor base pair, the $V_{\text{max}}$ values for dGMP misincorporation were consistently higher for fU than T. However, the $K_m$ values for fU showed no systematic variations. Despite these variations, misincorporation of dGMP was favored for fU over T by 1.1- to 5.6-fold as judged from relative $f_G$. Granting competitive incorporation of dAMP and dGMP at the same site (T or fU) of DNA, the replication error frequency ($f_{\text{RE}} = f_G/f_A$) was calculated for individual template/primers using the data in Tables II ($f_A$) and III ($f_G$). The values of $f_{\text{RE}}$ for fU were consistently in a $10^{-4}$ range, whereas those for T were in a $10^{-5}$ range. Thus, the increase in $f_{\text{RE}}$ due to the substitution of fU for T was 3.1- to 14.3-fold (Fig. 4A). These increases arose from reduced $f_A$ (Table II) and increased $f_G$ (Table III). Comparison of the parameters ($V_{\text{max}}$ and $K_m$) in Tables II and III also indicated that discrimination of the nucleotide at the fU site originated from $V_{\text{max}}$ and $K_m$. The averaged discrimination factors for $V_{\text{max}}$ and $K_m$ were 16 and 360, respectively, which were calculated from the ratio of the parameters for dAMP vs. dGMP incorporation. Therefore, the contribution of $K_m$ was much greater than that of $V_{\text{max}}$.

**pH Effects of dAMP and dGMP Incorporation opposite fU**——We have previously reported pH-dependent misincorporation of fdUTP opposite template G by DNA polymerase and have pointed out the importance of an ionized (or enolate) form of fU in fU:G mispair formation (27). To ask whether this mispairing scheme also held for fU in template DNA, the pH effect on the dGMP misincorporation was analyzed. Template/primers containing fU (24AF/13T) or T (24AT/13T) were incubated with Pol
I Kf (exo-) and a single dNTP (dGTP or dATP, 20 μM) at pH 7.2-8.6. The percentage of the extended primer resulting from incorporation of dAMP or dGMP was determined by PAGE analysis (Fig. 5). Incorporation of dAMP was virtually unaffected by the pH change and was less efficient for fU than T (Figs. 5A and 5C). In contrast, incorporation of dGMP opposite fU showed a clear pH-dependence and the amount of dGMP increased with increasing pH (Figs. 5B and 5D). Although dGMP incorporation opposite T was also pH-dependent, the increase with pH was extremely small (Figs. 5B and 5D). The pH effect on dGMP misincorporation opposite fU was further analyzed in a wider pH range (pH 6.9-9.3) using GTA buffer in place of Tris buffer (Fig. 5D inset). The plot of the efficiency of dGMP misincorporation against pH showed a sigmoidal curve reminiscent of a pH titration, though the efficiency was somewhat different between the GTA and Tris buffer systems. Since the pK_a values of fU and T were 8.6 and 10.0, respectively, these results strongly suggest that an acid-base equilibrium of fU (Fig. 6A) is involved in the misincorporation of dGMP and the ionized (or enolate) form of fU forms a mispair with incoming dGTP during DNA synthesis (Fig. 6B left).

Extension of Matched and Mismatched Primer Termini———The primer extension assay described above revealed that fU in template DNA directed incorporation not only of correct dAMP but also of incorrect dGMP, albeit less efficiently, during DNA synthesis, thereby giving rise to matched (fU:A) and mismatched (fU:G) primer termini. It is known that mismatched primer termini are extended less efficiently than matched termini by DNA polymerases and constitute a barrier for erroneous replication of DNA. Thus, the extension of primer termini containing fU:A and fU:G pairs was examined and the results were compared to those of T:A and T:G pairs. Fig. 7 shows gel data when matched (24AT/14TA (T:A pair) and 24AF/14TA (fU:A pair)) and mismatched (24AT/14TG (T:G pair) and 24AF/14TG (fU:G pair)) primer termini were extended by Pol I Kf (exo- ) in the presence of dCTP, a nucleotide to be incorporated following the primer termini. The matched primer termini containing T:A and fU:A pairs were elongated with comparable efficiencies (lanes 2 and
4). Although extension of the mismatched primer termini containing T:G (lane 6) and fU:G (lane 8) pairs were less efficient than matched ones, the extension of the fU:G terminus was clearly preferred over T:G. For quantitative comparison of the extension efficiencies, kinetic parameters ($V_{\text{max}}$ and $K_m$) for dCMP incorporation were determined by the gel fidelity assay under standing start conditions (43) using 24AT, 24AF, 24CT and 24CF (average of two experiments) (Table IV). Irrespective of the undamaged (24AT and 24CT) or damaged (24AF and 24CF) templates, the matched primer termini containing fU:A or T:A pairs were efficiently extended, with a slight preference of T:A (ca. 1.2-fold as the extension efficiency ($f_C$)). With the matched termini, $K_m$ values of extension (i.e. dCMP incorporation) were comparable to those of dAMP at the previous site (Table II), whereas the corresponding $V_{\text{max}}$ values were several fold lower, presumably due to the difference in the incorporated nucleotide (C or A) or the sequence context. The extension of the mismatched primer termini containing fU:G and T:G were inefficient and the extension efficiency ($f_C$) was two or three orders of magnitude lower than that of the corresponding matched termini (Table IV). For both T and fU, discrimination of the matched and mismatched termini exclusively originated from $K_m$. Interestingly, the mismatched primer termini containing the fU:G pair was extended with significantly higher efficiencies than those containing the T:G pair. The mismatch extension frequency ($f_{\text{EX}} = f_C(\text{mismatched terminus})/f_C(\text{matched terminus})$ for the same template) was in a $10^{-2}$ range for fU, whereas that for T was in a $10^{-3}$ range. The increase in $f_{\text{EX}}$ associated with the substitution of fU for T was 8.2-fold (24AF vs. 24AT) and 11.3-fold (24CF vs. 24CT) (Fig. 4B). Accordingly, conversion of T to fU in template DNA promotes not only misincorporation of dGMP (Fig. 4A) but also elongation of the resulting mismatched primer termini (Fig. 4B).
DISCUSSION

fU is one of the major oxidative thymine lesions found in DNA and nucleoside that were exposed to ionizing radiation (44-46), Fenton type reactions (46, 47), photosensitized reactions (48, 49), and peroxy radicals (50). The yield of fU in Fenton type reactions and γ-irradiation is comparable to those of 8-oxoG (47) and 5-hydroxypyrimidines (45) that are known as major mutagenic oxidative base lesions (51-55). Bacterial (28, 37, 56) and mammalian (57, 58) cells contain repair enzyme or activity that excises fU from damaged DNA, implying potential genotoxic influences of this lesion in vivo. Direct incorporation of fdUTP into permeated E. coli cells resulted in a small but significant increase in chromosomal lacI mutation with G:C→A:T transitions being most preferred (59). fdU was also mutagenic to Salmonella typhimurium when added to the culture medium (44). In the present study, we have assessed the genotoxic potential of fU in template DNA by utilizing in vitro DNA replication reactions. The product analysis of translesion synthesis revealed that fU constituted very weak blocks to DNA synthesis (Fig. 2). Thus, unlike other thymine lesions such as thymine glycol, urea residues, and 5-hydroxy-5-methylhydantoin (13-15, 24), fU will exert a weak to negligible cytotoxic effect due to inhibition of DNA replication in vivo. Conversely, fU was shown to be a potentially mutagenic lesion based on the following results. First, the substitution of T by fU promoted misincorporation of incorrect dGMP (1.1-5.6 times as f_0) and at the same time retarded incorporation of correct dAMP (1/4-1/2 as f_a), hence leading to 3.3- to 14.3-fold increases in the replication error frequency (f_re) relative to T (Fig. 4A). Secondly, the resulting mismatched primer terminus containing an fU:G pair was more readily extended (8.2-11.3 times as f_ex) than that containing a T:G pair (Fig. 4B) (see also the discussion below on f_ex). This step will affect the probability that a genome DNA molecule is replicated to completion, and thereby scored as mutation. According to the present data, fU is moderately mutagenic, but for more quantitative estimation of the
mutation frequency of fU, it is necessary to consider the influence of repair and the property of replicative DNA polymerases.

The mismatch extension frequency \( f_{\text{EX}} \) in the absence of proofreading depends explicitly on the binding constant of DNA polymerase to matched vs. mismatched template/primer DNA as well as on the concentrations of the template/primer DNA and next-correct dNTP (60). Accordingly, to evaluate the intrinsic mismatch extension frequency \( f_{\text{EX(int)}} \), possible differential binding of DNA polymerase to matched and mismatched primer termini need to be taken into account under standing start conditions, i.e. differential binding to T:A vs. T:G and fU:A vs. fU:G termini in this study. The relationship between \( f_{\text{EX}} \) and \( f_{\text{EX(int)}} \) has been formulated in equation (1) (60), where \([D_r] \) and \([D_w] \) are the concentrations of template/primer DNA having correctly and incorrectly paired primer termini (\( D_r \) and \( D_w \)), respectively, and \( K_r \) and \( K_w \) are the equilibrium constants for dissociation of polymerase-\( D_r \) and polymerase-\( D_w \) complexes, respectively.

\[
\frac{f_{\text{EX}}}{f_{\text{EX(int)}}} = \frac{1+[D_r]/K_r}{1+[D_w]/K_w} \quad (1)
\]

In equation (1), it is generally assumed that the affinity of DNA polymerase for a correctly paired terminus is similar to or higher than that for an incorrectly paired terminus (\( K_w \geq K_r \)). The values of \([D_r] \) and \([D_w] \) are both 33 nM in this study. When Pol I (exo-) has comparable affinities (\( K_r = K_w \)) for the matched and mismatched termini (i.e. T:A and T:G, and fU:A and fU:G), \( f_{\text{EX}} \) in Table IV approximately represents the intrinsic value. According to equation (1), the largest discrepancy between \( f_{\text{EX}} \) and \( f_{\text{EX(int)}} \) occurs when \( K_w \) is much higher than \([D_w] \) (\( K_w >> 33 \text{ nM} \)). In this case, equation (1) can be transformed into equation (2) by approximation.

\[
\frac{f_{\text{EX}}}{f_{\text{EX(int)}}} = \frac{1+[D_r]/K_r}{1+[D_w]/K_w} \quad (2)
\]

Although the \( K_r \) values of Pol I (exo-) are not known, those for *Avian myeloblastosis* reverse transcriptase (AMV RT) and *Drosophila melanogaster* DNA polymerase α have been estimated as 5 nM and 20-50 nM, respectively (60). Granted that the \( K_r \) value of Pol I (exo-) is in a similar range (5-50 nM), the \( f_{\text{EX}} \) values in Table IV are subjected to up to 7.6-fold reduction. However, the correction factor given by equation (2) is
presumably similar for the T and fU templates since the affinities (K_r) of Pol I (exo-) for
the structurally resembling T:A and fU:A termini are likely comparable. These
considerations suggest that the relative difference in f_EX(int) for the T vs. fU templates
remains similar to that shown in Fig. 4B, though the absolute value of f_EX(int) may be
lower than that obtained experimentally in this study (f_EX).

Concerning the mispairing mechanism of fU, we have previously suggested
participation of the ionized (or enolate) form of fU based on the pH-dependent
misincorporation of fdUTP opposite template G (27). Privat and Sowers also proposed
a similar mispairing scheme on the basis of pK_a measurement of fdU and related
nucleosides (61). Consistent with this mechanism, the efficiency of dGMP
misincorporation opposite template fU increased around the pK_a value of fU (pK_a =
8.6), whereas the corresponding increase for T (pK_a = 10.0) was much smaller than
that for fU (Figs. 5B and 5D). The result obtained for T also agrees with a small
increase in the base substitution frequency of Pol I (exo-) in this pH range (62). Unlike
thymine bearing an electron donating methyl group, fU has an electron withdrawing
formyl group that promotes ionization of fU in an acid-base equilibrium (Fig. 6A).
According to the pK_a values, the fraction of the ionized form of fU increases form 4%
to 50% in the pH range of 7.2-8.6, but that of thymine is virtually negligible (0.2% to
4%). Ionized fU can form a base pair with incoming dGTP through two hydrogen
bonds (Fig. 6B left). The base pair formed between ionized fU and dGTP essentially
assume Watson-Crick geometry (or B form geometry) and can fits into the active site of
DNA polymerase. Since the geometric recognition is key to discrimination of correct
vs. incorrect nucleotides by DNA polymerases (1), this geometry probably promoted
misincorporation of dGMP opposite fU. Participation of base ionization promoted by
electron withdrawing substituents has been demonstrated in the mispairing of 5-
halogenated uracils (BrU and 5-fluorouracil) with G (32). Thus, fU and 5-halogenated
uracils share a common mutation mechanism. Although participation of a rare enol
tautomer of fU (Fig. 6A) in the mispairing with G can not be fully ruled out, recent
NMR studies show that the tautomeric equilibrium between keto and enol forms of fU
is not significantly affected by oxidation of the methyl group of T to the formyl group (63, 64). Therefore, involvement of the enol form of fU is unlikely in the mispairing with G. It is assumed that after dGMP incorporation, the resulting fU(ionized):G pair in Watson-Crick geometry shifts to wobble geometry (fU(keto):G) due to the acid-base equilibrium (Fig. 6B right). However, a certain fraction of the base pair will still exist as an fU(ionized):G pair whose geometry can again promote incorporation of the next nucleotide. Probably this is the reason why the mismatched primer terminus containing an fU:G pair was more efficiently extended than that containing a T:G pair. Although there is no experimental evidence that directly shows the equilibrium between fU(keto):G and fU(ionized):G base pairs in duplex DNA, the presence of such an equilibrium has been demonstrated for BrU(keto):G and BrU(ionized):G pairs in a duplex oligonucleotide by the nuclear magnetic resonance (NMR) study (65).

According to the proposed mutation mechanism for fU, it is reasonable that 5-hydroxymethyluracil, another methyl oxidation product of T, does not direct misincorporation (25, 26) since the hydroxymethyl group has electron donating nature and can not promote ionization of the base. Although a mutation mechanism involving an altered acid-base equilibrium has been previously demonstrated for 5-halogenated uracils (32), to our knowledge, fU is the first example adapting to this mechanism among oxidative DNA base lesions.

To assess the sequence context effect on the base pairing property of fU, the 3’-nearest neighbor base of template fU and the paired base (i.e. the primer terminus base pair) was systematically changed and the nucleotide incorporated opposite fU was analyzed. fU with the four possible nearest neighbor base pairs directed incorporation of dAMP and to a less extent dGMP but not dCMP and dTMP (Fig. 3). Thus, 3’-nearest neighbor base exhibited no influence concerning the type of base pairs formed from template fU. Combining the result with fdUTP (27, 28) and template fU (this study), it follows that fundamental base pairing symmetry is retained whether the fU base pairs are formed from incoming fdUTP or template fU during DNA polymerase reactions. Although fundamental base pairing symmetry held with respect to the
formation of the fU base pairs, the nearest neighbor base pair showed quantitative effects on the incorporation efficiency of dAMP and dGMP. Regardless of the correct or incorrect incorporation (Figs 4C and 4D), the sequence context effect was less pronounced for fU than T. For example, the difference in $f_G$ for T was 9.4-fold for the highest (TT) and lowest (GT) sequences, whereas the corresponding value (TF vs. AF, F = fU) was only 2.2-fold for fU (Fig. 4D). Another notable sequence context effect was a tendency of preferred incorporation of dAMP and dGMP for the sequences containing 3’-pyrimidines over 3’-purines, though the difference was not so large. In other words, incorporation of the nucleotides was favored for the primer terminus containing purines. This was common to T and fU. Presumably, stabilization of the incoming purine nucleotides through a favored purine-purine stacking interaction (purine-purine > pyrimidine-purine > pyrimidine-pyrimidine) (66) promoted their incorporation over pyrimidine nucleotides. Despite the sequence context dependent variations of $f_A$ and $f_G$, the $f_{RE} (= f_G/f_A)$ values of fU were consistently higher than those of T and were virtually independent of the nearest neighbor base pair (Fig. 4E). This result suggests that the distribution of T→C transitions induced by fU will be relatively uniform with respect to the variation of the 3’-nearest neighbor base unless heterogeneous formation or repair of fU occurs in cells. Finally, the relative increase in $f_{RE}$ associated with the conversion of T to fU was calculated. The order of the increase was A > G > C > T with respect to the 3’-nearest neighbor base (Fig. 4A), showing an inverse correlation with that (T > C > G > A) ranged by $f_{RE}$ of T (Fig. 4E). Thus, the sequence with a low replication error frequency before conversion to fU gave rise to a relatively large increase in the replication error frequency after conversion to fU.

Previously, Zhang et al. (29, 30) assessed the mutagenic potential of fU using an in vitro system similar to that used in this study. According to their experiments with Pol I Kf, Pol I Kf (exo−) and thermophilic DNA polymerases, fU directed incorporation of dCMP as well as dAMP in all cases, but not dGMP and dTMP at all. The incorporation efficiency of dCMP relative to dAMP was 0.09-0.15 for Pol I Kf, 0.06 for Pol I Kf (exo−), and 0.23-0.27 for Tth DNA polymerase. These values are
unusually high as a frequency of pyrimidine:pyrimidine mispair formation by DNA polymerases. Generally, pyrimidine:pyrimidine pairs are too small to fit into the B form helix. For this reason, formation of these mispairs is not favored by prokaryotic and eukaryotic DNA polymerases. The most common mispairs are G:T mispairs with observed frequencies between $10^{-2}$-$10^{-4}$, and the least common ones are pyrimidine:pyrimidine mispairs with frequencies between $10^{-4}$-$10^{-5}$ (1). The rate of misincorporation of dNMP opposite T by Pol I Kf and Pol I Kf (exo) also follows this rule (dGMP $>$ dCMP $>$ dTMP) (62, 67). The measurement of the melting temperature ($T_m$) of duplex oligonucleotides containing an fU:N pair (N = A, G, C, T) revealed that $T_m$ decreased in the following order: fU:A $>$ fU:G $>$ fU:T $>$ fU:C. This order indicates that the fU:C pair exerts the largest destabilization effect on DNA among the four possible base pairs and further suggests the least favored formation of this pair during DNA replication. In addition, Zhang et al. (29) found significant incorporation of dCMP but not dGMP opposite normal T (Table V), suggesting a fundamental discrepancy in the experimental set up used in the present and their studies. We repeated this experiment using the same template/primer (i.e. template 1 and primer 3 shown in Table V). However, the reported result was not reproduced and misincorporation of only dGMP was detected (Table V). Thus, T in this sequence context was not particularly prone to incorporate dCMP. We also repeated another control experiment with $T_{th}$ DNA polymerase under the reported conditions (29, 30). Template 1 (see Table V for the sequence) was annealed to a 9-mer primer (5’-TGCAGGTCG) and primer extension assays were performed in the presence of a single dNTP at 74 °C for 5 or 10 min. Although they observed incorporation of dAMP opposite T under these conditions, we did not see incorporation of any nucleotides. We believe the present result is reasonable in light of the expected low melting temperature of the 9-mer primer ($T_m$ = 30 °C). It is very likely that the template/primer dissociated at 74 °C and was unable to serve as a substrate for $T_{th}$ DNA polymerase in the present experiment. For the same reason, the reported incorporation of dCMP with the 9-mer primer by $T_{th}$ DNA polymerase (29, 30) is very unlikely to occur at such a high
temperature. In view of the inconsistencies such as the contradiction against the general preference of mispair formation by DNA polymerases and the lack of reproducibility of the certain experimental results, we believe fU directs misincorporation of dGMP but not dCMP. Therefore, whether fU is in incoming dNTP or a template, base pairing symmetry in the nucleotide selection by DNA polymerase and the mutagenesis mechanism involving ionized fU hold, though the incorporation efficiency of dAMP and dGMP varies depending on the nearest neighbor base pair.

**Acknowledgment**——We thank Akira Matsuda and Naoko Karino (Hokkaido University) for communicating the T_m data for the duplexes containing fU.
REFERENCES

1. Echols, H., and Goodman, M. F. (1991) *Annu. Rev. Biochem.* **60**, 477-511
2. Johnson, K. A. (1993) *Annu. Rev. Biochem.* **62**, 685-713
3. Modrich, P. and Lahue, R. (1996) *Annu Rev. Biochem.* **65**, 101-133
4. Friedberg, E. C., Walker, G. C., and Siede, W. (1995) *DNA Repair and Mutagenesis*, AMS Press, Washington, D.C.
5. von Sonntag, C. (1987) *Chemical Basis of Radiation Biology*, Taylor & Francis, New York.
6. Breen, A. P., and Murphy, J. A. (1995) *Free Radical Biol. Med.* **18**, 1033-1077
7. Hatahet, Z., and Wallace, S. S. (1998) in *DNA Damage and Repair* (Nickoloff, J. A., and Hoekstra, M. F. eds) Vol. 1, pp. 229-262, Humana Press, Totowa, New Jersey
8. Wang, D., Kreutzer, D. A., and Essigmann, J. M. (1998) *Mutat. Res.* **400**, 99-115
9. Friedberg, E. C., Feaver, W. J., and Gerlach, V. L. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5681-5683
10. Lewis, D. A., Bebenek, K., Beard, W. A., Wilson, S. H., and Kunkel, T. A. (1999) *J. Biol. Chem.* **274**, 32924-32930
11. Steiz, T. A. (1999) *J. Biol. Chem.* **274**, 17395-17398
12. Gasparutto, D., Bourdat, A. -G., D’Ham, C., Duarte, V., Romieu, A., and Cadet, J. (2000) *Biochimie* **82**, 19-24
13. Evans, J., Maccabee, M., Hatahet, Z., Courcelle, J., Bockrath, R., Ide, H., and Wallace, S. S. (1993) *Mutat. Res.* **299**, 147-156
14. Ide, H., Kow, Y. W., and Wallace, S. S. (1985) *Nucleic Acids Res.* **13**, 8035-8052
15. Ide, H., Petruollo, L. A., Hatahet, Z., and Wallace, S. S. (1991) *J. Biol. Chem.* **266**, 1469-1477
16. Ide, H., Melamede, R. J., and Wallace, S. S. (1987) *Biochemistry* **26**, 964-969
17. Ide, H., and Wallace, S. S. (1988) Nucleic Acids Res. 16, 11339-11354
18. Ide, H., Yamaoka, T., and Kimura, Y. (1994) Biochemistry 33, 7127-7133
19. Ide, H., Murayama, H., Sakamoto, S., Makino, K., Honda, K., Nakamuta, H., Sasaki, M., and Sugimoto, N. (1995) Nucleic Acids Res., 23, 123-129
20. Shimizu, H., Yagi, R., Kimura, Y., Makino, K., Terato, H., Ohyama, Y., and Ide, H. (1997) Nucleic Acids Res. 25, 597-603
21. Suzuki, T., Yoshida, M., Yamada, M., Ide, H., Kobayashi, M., Kanaori, K., Tajima, K. and Makino, K. (1998) Biochemistry, 37, 11592-11598
22. Hayes, R. C., Petrullo, L. A., Huang, H. M., Wallace, S. S., and LeClerc, J. E. (1988) J. Mol. Biol. 201, 239-246
23. Maccabee, M., Evans, J. S., Glackin, M. P., Hatahet, Z., and Wallace, S. S. (1994) J. Mol. Biol. 236, 514-530
24. Gasparutto, D., Ait-Abbas, M., Jaquinod, M., Boiteux, S., and Cadet, J. (2000) Chem. Res. Toxicol. 13, 575-584
25. Levy, D. D., and Teebor, G. W. (1991) Nucleic Acids Res., 19, 3337-3343
26. Mi, L. -J., Mahl, E., Chaung, W., and Boorstein, R. J. (1997) Mutat. Res. 374, 287-295
27. Yoshida, M., Makino, K., Morita, H., Terato, H., Ohyama, Y., and Ide, H. (1997) Nucleic Acids Res. 25, 1570-1577
28. Terato, H., Masaoka, A., Kobayashi, M., Fukushima, S., Ohyama, Y. Yoshida, M., and Ide, H. (1999) J. Biol. Chem. 274, 25144-25150
29. Zhang, Q. -M., Sugiyama, H., Miyabe, I., Matsuda, S., Saito, I., and Yonei, S. (1997) Nucleic Acids Res. 25, 3969-3973
30. Zhang, Q. -M., Sugiyama, H., Miyabe, I., Matsuda, S., Kino, K., Saito, I., and Yonei, S. (1999) Int. J. Radiat. Biol. 75, 59-65
31. Lawley, P. D., and Brookes, P. (1962) J. Mol. Biol. 4, 216-219
32. Yu, H., Eritja, R., Bloom, L. B., and Goodman, M. F. (1993) J. Biol. Chem. 268, 15935-15943
33. Sugiyama, H., Matsuda, S., Kino, K., Zhang, Q., -M., Yonei, S., and Saito, I. (1996) *Tetrahedron Lett*. **37**, 9067-9070
34. Carroll, S. S., Cowart, M., and Benkovic, S. J. (1991) *Biochemistry* **30**, 804-813
35. Ahn, J., Werneburg, B. G., and Tsai, M. -D. (1997) *Biochemistry* **36**, 1100-1107
36. Einolf, H. J., Schnetz-Boutaud, N., and Guengerich, F. P. (1998) *Biochemistry* **37**, 13300-13312
37. Masaoka, A., Terato, H., Kobayashi, M., Honsho, A., Ohyama, Y., and Ide, H. (1999) *J. Biol. Chem.* **274**, 25136-25143
38. Asagoshi, K., Yamada, T., Okada, Y., Terato, H., Ohyama, Y., Seki, S., and Ide, H. (2000) *J. Biol. Chem.* **275**, 24781-24786
39. Asagoshi, K., Yamada, T., Terato, H., Ohyama, Y., Monden, Y., Arai, T., Nishimura, S., Aburatani, H., Lindahl, T., and Ide, H. (2000) *J. Biol. Chem.* **275**, 4956-4964
40. Terato, H., Morita, H., Ohyama, Y. and Ide, H. (1998) *Nucleosides Nucleotides* **17**, 131-141
41. Armstrong, V. W., Dattagupta, J. K., Eckstein, F., and Saenger, W. (1976) *Nucleic Acids Res.* **3**, 1791-1810
42. Cline, R. E., Fink, R. M., and Fink, K. (1959) *J. Am. Chem. Soc.* **81**, 2521-2527
43. Creighton S., Bloom, L. B., and Goodman M. F. (1995) *Methods Enzymol.* **262**, 232-256
44. Kasai, H., Iida, A., Yamaizumi, Z., Nishimura, S. and Tanooka, H. (1990) *Mutat. Res.* **243**, 249-253
45. Douki, T., Delatour, T., Paganon, F., and Cadet, J. (1996) *Chem. Res. Toxicol.* **9**, 1145-1151
46. Murata-Kamiya, N., Kamiya, H., Karino, N., Ueno, Y., Kaji, H., Matsuda, A., and Kasai, H. (1999) *Nucleic Acids Res.* **27**, 4385-4390
47. Murata-Kamiya, N., Kamiya, H., Muraoka, M., Kaji, H., and Kasai, H. (1997) *J. Radiat. Res*. **38**, 121-131

48. Decarroz, C., Wagner, J. R., Van Lier, J. E., Krishna, C. M., Riesz, P., and Cadet, J. (1986) *Int. J. Radiat. Biol*. **50**, 491-505

49. Saito, I., Takayama, M., and Kawanishi, S. (1995) *J. Am. Chem. Soc*. **117**, 5590-5591

50. Martini, M., and Termini, J. (1997) *Chem. Res. Toxicol*. **10**, 234-241

51. Shibutani, S., Takeshita, M., and Grollman, A. P. (1991) *Nature* **349**, 431-434

52. Maki, H., and Sekiguchi, M. (1992) *Nature* **355**, 273-275

53. Purmal, A. A., Kow, Y. W., and Wallace, S. S. (1994) *Nucleic Acids Res*. **22**, 72-78

54. Purmal, A. A., Kow, Y. W., and Wallace, S. S. (1994) *Nucleic Acids Res*. **22**, 3930-3935

55. Kreutzer, D. A., and Essigmann, J. M. (1998) *Proc. Natl. Acad. Sci. U.S.A*. **95**, 3578-3582

56. Bjelland, S., Birkeland, N., Benneche, T., Volden, G., and Seeberg, E. (1994) *J. Biol. Chem*. **269**, 30489-30495

57. Bjelland, S., Eide, L., Time, R. W., Stote, R., Eftedal, I., Volden, G., and Seeberg, E. (1995) *Biochemistry* **34**, 14758-14764

58. Zhang, Q. -M., Fujimoto, J., and Yonei, S. (1995) *Int. J. Radiat. Biol*. **68**, 603-607

59. Fujikawa, K., Kamiya, H., and Kasai, H. (1998) *Nucleic Acids Res*. **26**, 4582-4587

60. Mendelman, L. V., Petruska, J., and Goodman, M. F. (1990) *J. Biol. Chem*. **265**, 2338-2346

61. Privat, E. J., and Sowers, L. C. (1996) *Mutat. Res*. **354**, 151-156

62. Eckert, K. A., and Kunkel, T. A. (1993) *J. Biol. Chem*. **268**, 13462-13471

63. La Francois, C. J., Fujimoto J., and Sowers L. C. (1998) *Chem. Res. Toxicol*. **11**, 75-83
64. La Francois, C. J., Jang, Y. H., Cagin, T., Goddard, W. A. III, and Sowers, L. C. (2000) *Chem. Res. Toxicol*. **13**, 462-470

65. Sowers, L. C., Goodman, M. F., Eritja, R., Kaplan, B., and Fazakerley, G. V. (1989) *J. Mol. Biol.* **205**, 437-447

66. Saenger, N. (1984) *Principles of Nucleic Acid Structure*. Springer-Verlag, NY

67. Joyce, C. M., Sun, X. C., and Grindley, N. D. F. (1992) *J. Biol. Chem.* **267**, 24485-24500

68. Zhang, Q. -M., Miyabe, I., Matsumoto, Y., Kino, K., Sugiyama, H., Yonei, S. (2000) *J. Biol. Chem.* **275**, 35471-35477
FOOTNOTES

1The abbreviations used are: fU, 5-formyluracil; fdUTP, 5-formyluracil-2'-deoxyuridine 5'-triphosphate; fdU, 5-formyluracil-2'-deoxyuridine; BrU, 5-bromouracil; Pol I Kf, Escherichia coli DNA polymerase I Klenow fragment; Pol I Kf (exo'), Pol I Kf deficient in 3'-5' exonuclease; PAGE, polyacrylamide gel electrophoresis; endonuclease, Endo; Fpg, formamidopyrimidine glycosylase; HPLC, high-performance liquid chromatography.

2Recently, the activity of Endo III and Fpg for fU was reported by Zhang et al. (68). Our study on the kinetic parameter revealed that their activity for fU was two to three orders of magnitude lower than their intrinsic substrates (thymine glycol for Endo III and 7,8-dihydro-8-oxoguanine for Fpg). Thus, the activity of the two enzymes for fU was negligible under standard assay conditions and extremely large excess enzymes were required to detect the activity for fU (A. Masaoka, H. Terato, Y. Ohyama, and H. Ide, to be published).

3The UV melting curves (plots of A$_{260}$ against temperature) were measured using the duplexes (total strand concentration 3 μM) of 25F and its complementary strand containing A, G, C, or T as the base opposite fU in 10 mM NaCl and 10 mM sodium cacodylate (pH 7.0). The T$_m$ value was evaluated from the inflection point of the melting curve. (A. Masaoka, H., Terato, Y. Ohyama, N. Karino, A. Matsuda, and H. Ide, to be published).
FIGURE LEGENDS

**FIG. 1.** **HPLC analysis of nucleosides in 25F.** A, 25F containing fU was digested by nuclease P1 and alkaline phosphatase as described under EXPERIMENTAL PROCEDURES. An aliquot of the reaction mixture was analyzed by reversed phase HPLC equipped with a C18 WS-DNA column (4.6 x 150 mm). The sample was eluted by a gradient of methanol (0% for 0-5 min and 0-5% linear gradient for 5-35 min) in 10 mM sodium phosphate buffer (pH 7.4). The flow rate was 0.8 ml/min and the monitoring wavelength was 280 nm. The attribution of the elution peak is indicated over the peak. The elution peak with an asterisk (*) was a putative fdU adduct formed by the enzymatic treatment. B, authentic fdU was treated by nuclease P1 and alkaline phosphatase, and subjected to HPLC analysis as described above. Note that fdU before incubation was eluted as a single peak (not shown) and no adduct (*) was observed.

**FIG. 2.** **Translesion DNA synthesis at the fU site.** DNA templates containing T (24AT) and fU (24AF) at the same site were primed with \(^{32}\)P-labeled P10. 24AT/P10 and 24AF/P10 (0.5 pmol) was incubated with Pol I Kf (exo-) (0.05 unit) and four dNTPs (50 \(\mu\)M each) at 25 °C, and products were analyzed by 16% denaturing PAGE. Templates (24AT and 24AF) and incubation time (3, 5, 10 min) are indicated on the top. Lanes 1 and 5, template/primer without the polymerase reaction; lane 9, 14-mer marker (\(^{32}\)P-labeled 14TA in Table I) showing the position of fU.

**FIG. 3.** **PAGE analysis of the nucleotide incorporated opposite fU.** Template/primers (0.5 pmol) containing four different primer terminus base pairs were incubated with the indicated dNTP (50 \(\mu\)M) and Pol I Kf (exo-) (0.05 unit) at 25 °C for 5 min, and the incorporated nucleotide opposite X (= T or fU) was analyzed by 16% denaturing PAGE. Template/primer used was 24AT/13T (lanes 1-4) and 24AF/13T (lanes 6-9) (A), 24GT/13C (lanes 1-4) and 24GF/13C (lanes 6-9) (B), 24CT/13G
(lanes 1-4) and 24CF/13G (lanes 6-9) (C), 24TT/13A (lanes 1-4) and 24TF/13A (lanes 6-9) (D). In panels A-D, lane 5 shows template/primer without the polymerase reaction. The arrow indicates the extended primer. The sequence surrounding T or fU (indicated by X) is shown over the gel.

**FIG. 4.** Sequence context effects on \( f_A \), \( f_G \), \( f_{RE} \), and increases in \( f_{RE} \) and \( f_{EX} \) associated with conversion of T to fU. A, Increases in the replication error frequency (\( f_{RE} \)) associated with conversion of T to fU. The ratio of the replication error frequencies (\( f_{RE} (X=fU)/f_{RE} (X=T) \)) was calculated for the same 3’-nearest neighbor base using the data in Table III. The ratio was plotted against the template sequence (3’-NX-5’, N = A, G, C, and T, X = T and fU). B, Increases in the mismatch extension frequency (\( f_{EX} \)) associated with conversion of T to fU. The ratio of the misextension frequencies (\( f_{EX} (X=fU)/f_{EX} (X=T) \)) was calculated for the corresponding templates (i.e. 24AF vs. 24AT and 24CF vs. 24CT) using the data in Table IV. The ratio was plotted against the template sequence (3’-NXG-5’, N = A and C, X = T and fU). C, Sequence context effects on the incorporation efficiencies of dAMP (\( f_A \)). D, Sequence context effects on the misincorporation efficiencies of dGMP (\( f_G \)). E, Sequence context effects on the replication error frequency (\( f_{RE} \)). In C - E, the values of \( f_A \), \( f_G \), and \( f_{RE} \) were taken from Tables II and III and plotted against the template sequence (3’-NT-5’ and 3’-NF-5’, N = A, G, C, T, F = fU).

**FIG. 5.** pH effects on the incorporation of dAMP and dGMP opposite fU. Template/primers (24AF/13T and 24AT/13T, 0.5 pmol) were incubated with Pol I Kf (exo') (0.03 unit) in the presence of dATP or dGTP (20 µM) at pH 7.2-8.6 (Tris buffer) or 6.9-9.3 (GTA buffer). Incubation was performed at 25 °C for 5 min. Incorporation of dAMP (A) and dGMP (B) opposite template fU and T (indicated over the gel) was analyzed by 16% denaturing PAGE. In panels A and B, pH of the reaction mixture was 7.2 (lanes 1 and 6), 7.7 (lanes 2 and 7), 8.0 (lanes 3 and 8), 8.3 (lanes 4 and 9), and 8.6 (lanes 5 and 10). The extended primer is indicated.
by the arrow. The percent of the extended primer resulting from incorporation of dAMP (C) and dGMP (D) was plotted against pH based on the product analysis in panels A and B. dGMP misincorporation opposite fU was also measured in a wide pH range (pH 6.9-9.3) using GTA buffer, and the result is shown in the inset of panel D. In panels C and D, incorporation opposite template T and fU is represented by open and closed symbols, respectively.

**FIG. 6.** A proposed mechanism for mispair formation between fU and G. A, an acid-base equilibrium of fU (right) involving keto and ionized (enolate) forms, and a tautomeric equilibrium of fU (left) involving keto and enol forms. B, base pairing schemes for the fU:G mispair involving ionized (or enolate) and keto forms of fU.

**FIG. 7.** PAGE analysis of the extension of mismatched primer termini containing T:G and fU:G pairs. Template/primers (0.5 pmol) containing correctly paired (24AT/14TA (T:A pair) and 24AF/14TA (fU:A pair)) and mismatched (24AT/14TG (T:G pair) and 24AF/14TG (fU:G pair)) primer termini were incubated with Pol I Kf (exo') (0.05 unit) in the presence of dCTP (10 μM) at 25 °C for 5 min. Products were analyzed by 16% denaturing PAGE. Lanes 1, 3, 5, and 7 show template/primers without the polymerase reaction, and lanes 2, 4, 6, and 8 show the extended products formed by the polymerase reaction. The base pair in the primer terminus is indicated over the gel.
| Abbreviation | Sequence<sup>a</sup>                        |
|--------------|------------------------------------------|
| 24AT         | 3′-TGGGTCA<sub>T</sub>GTCTATGATGGTGCT    |
| 24AF         | 3′-TGGGTCA<sub>AF</sub>GTCTAFGATGGTGCT   |
| 24GT         | 3′-TGGGTCA<sub>GT</sub>GTCTGTGATGGTGCT   |
| 24GF         | 3′-TGGGTCA<sub>GF</sub>GTCTGFATGGTGCT    |
| 24CT         | 3′-TGGGTCA<sub>CT</sub>GTCTGTGATGGTGCT   |
| 24CF         | 3′-TGGGTCA<sub>CF</sub>GTCTCFGATGGTGCT   |
| 24TT         | 3′-TGGGTCA<sub>TT</sub>GTCTTTGATGGTGCT   |
| 24TF         | 3′-TGGGTCA<sub>TF</sub>GTCTTFGATGGTGCT   |
| P10          | 5′-ACCCAGTCA<sub>C</sub>                   |
| 13A          | 5′-ACCCAGTCACGAA<sub>A</sub>              |
| 13G          | 5′-ACCCAGTCACGAG<sub>G</sub>              |
| 13C          | 5′-ACCCAGTCACGAC<sub>C</sub>              |
| 13T          | 5′-ACCCAGTCACGAT<sub>T</sub>              |
| 14TA         | 5′-ACCCAGTCACGATA<sub>A</sub>             |
| 14TG         | 5′-ACCCAGTCACGATG<sub>G</sub>             |
| 14GA         | 5′-ACCCAGTCACGAGA<sub>A</sub>             |
| 14GG         | 5′-ACCCAGTCACGAGG<sub>G</sub>             |
| 25T          | 5′-CATCGATAGCATCCGTACACGGCAG              |
| 25F          | 5′-CATCGATAGCATCCGFCACAGGCAG              |
| 25COM        | 3′-GTAGCTATCGTAGGCAGTGTCGTC               |

<sup>a</sup> F = 5-formyluracil
## Table II

*Parameters (V<sub>max</sub> and K<sub>m</sub>) and efficiencies (f<sub>A</sub>) of dAMP incorporation opposite template T and fU*

| Template/Primer<sup>a</sup> | V<sub>max</sub><sup>b</sup> | K<sub>m</sub> | f<sub>A</sub><sup>c</sup> |
|---------------------------|----------------|-----------|----------------|
|                           | %/min | µM | %/min/µM |
| 24AT/13T (A:T)            | 57    | 0.034 | 1700 (1.00) |
| 24AF/13T (A:T)            | 22    | 0.052 | 420 (0.25) |
| 24GT/13C (G:C)            | 37    | 0.038 | 970 (1.00) |
| 24GF/13C (G:C)            | 33    | 0.074 | 450 (0.46) |
| 24CT/13G (C:G)            | 57    | 0.023 | 2500 (1.00) |
| 24CF/13G (C:G)            | 38    | 0.058 | 660 (0.26) |
| 24TT/13A (T:A)            | 72    | 0.034 | 2100 (1.00) |
| 24TF/13A (T:A)            | 49    | 0.069 | 710 (0.34) |

<sup>a</sup> The parenthesis next to template/primer indicates the nearest neighbor base pair in the primer terminus.

<sup>b</sup> The percentage of the extended primer per min per 0.03 unit of Pol I Kf (exo'). The original data were obtained with 0.002 unit of Pol I Kf (exo') as described under EXPERIMENTAL PROCEDURES and were converted to the value for 0.03 unit (multiplied by 15) for comparison with the data for dGMP misincorporation at the same site (see Table III).

<sup>c</sup> f<sub>A</sub> = V<sub>max</sub>/K<sub>m</sub> for dAMP incorporation. The value in the parenthesis is the ratio of f<sub>A</sub> values for fU and T with the same nearest neighbor base pair.
### Table III

*Parameters* (*V*<sub>max</sub> and *K*<sub>m</sub>) and efficiencies (*f*<sub>G</sub>) of dGMP misincorporation opposite template *T* and *fU*, and replication error frequencies (*f*<sub>RE</sub>)

| Template/Primer | *V*<sub>max</sub><sup>b</sup> | *K*<sub>m</sub> | *f*<sub>G</sub><sup>c</sup> | *f*<sub>RE</sub><sup>d</sup> |
|-----------------|-----------------|-------------|-----------------|-----------------|
| 24AT/13T (A:T)  | 0.55            | 24          | 0.023 (1.0)     | 1.4 x 10<sup>-5</sup> |
| 24AF/13T (A:T)  | 2.5             | 30          | 0.083 (3.7)     | 2.0 x 10<sup>-4</sup> |
| 24GT/13C (G:C)  | 0.98            | 57          | 0.017 (1.0)     | 1.8 x 10<sup>-5</sup> |
| 24GF/13C (G:C)  | 1.9             | 20          | 0.095 (5.6)     | 2.1 x 10<sup>-4</sup> |
| 24CT/13G (C:G)  | 0.88            | 17          | 0.052 (1.0)     | 2.1 x 10<sup>-5</sup> |
| 24CF/13G (C:G)  | 1.6             | 15          | 0.11 (2.1)      | 1.7 x 10<sup>-4</sup> |
| 24TT/13A (T:A)  | 1.6             | 10          | 0.16 (1.0)      | 7.6 x 10<sup>-5</sup> |
| 24TF/13A (T:A)  | 3.9             | 22          | 0.18 (1.1)      | 2.5 x 10<sup>-4</sup> |

*a* The parenthesis next to template/primer indicates the nearest neighbor base pair in the primer terminus.

*b* The percentage of the extended primer per min per 0.03 unit of Pol I Kf (exo').

*c* *f*<sub>G</sub> = *V*<sub>max</sub>/*K*<sub>m</sub> for dGMP misincorporation. The value in the parenthesis is the ratio of *f*<sub>G</sub> values for *fU* and *T* with the same nearest neighbor base pair.

*d* *f*<sub>RE</sub> = *f*<sub>G</sub>/*f*<sub>A</sub> for the same template/primer, where the values of *f*<sub>A</sub> and *f*<sub>G</sub> were taken from Tables II and III, respectively.
| Template/Primer                  | \( V_{\text{max}} \) \(^b\) | \( K_m \) | \( f_C \) \(^c\) | \( f_{\text{EX}} \) \(^d\) |
|---------------------------------|-----------------|--------|--------------|------------------|
| 24AT/14TA (T:A)                 | 6.3             | 0.025  | 250 (1.0)    | –                |
| 24AT/14TG (T:G)                 | 9.6             | 17     | 0.56 (0.0022) | \( 2.2 \times 10^{-3} \) |
| 24AF/14TA (fU:A)                | 6.7             | 0.033  | 200 (0.80)   | –                |
| 24AF/14TG (fU:G)                | 10              | 2.8    | 3.6 (0.014)  | \( 1.8 \times 10^{-2} \) |
| 24CT/14GA (T:A)                 | 13              | 0.021  | 620 (1.0)    | –                |
| 24CT/14GG (T:G)                 | 16              | 16     | 1.0 (0.0016) | \( 1.6 \times 10^{-3} \) |
| 24CF/14GA (fU:A)                | 11              | 0.022  | 500 (0.81)   | –                |
| 24CF/14GG (fU:G)                | 16              | 1.8    | 8.9 (0.014)  | \( 1.8 \times 10^{-2} \) |

\(^a\) The parenthesis next to template/primer indicates the primer terminus base pair. The parameters were determined by quantifying the amount of dCMP incorporated next to the matched and mismatched primer termini.

\(^b\) The percentage of the extended primer per min per 0.03 unit of Pol I Kf (exo'). The original data were obtained with 0.015 unit of Pol I Kf (exo') as described under EXPERIMENTAL PROCEDURES and were converted to the value for 0.03 unit (multiplied by 2) for comparison with the data in Tables II and III.

\(^c\) \( f_C = \frac{V_{\text{max}}}{K_m} \) for dCMP incorporation. The value in the parenthesis is a relative value of \( f_C \) (standardized to 24AT/14TA or 24CT/14GA)

\(^d\) \( f_{\text{EX}} = \frac{f_C(\text{mismatched terminus})}{f_C(\text{matched terminus})} \) for the same template.
| Experiment      | DNA polymerase | Specificity<sup>a,b</sup> |
|-----------------|----------------|--------------------------|
|                 |                | dAMP  | dGMP  | dCMP  | dTMP  |
| %               |                |       |       |       |       |
| Zhang *et al.* (29) | Pol I Kf       | 91.5  | <0.3  | 6.8   | <0.3  |
| this study      | Pol I Kf       | 78.7  | 21.3  | ND    | ND    |
| this study      | Pol I Kf (exo<sup>-</sup>) | 74.5  | 25.6  | ND    | ND    |

<sup>a</sup> Primer extension assays were performed using the indicated DNA polymerase (0.02 unit) and template 1/primer 3 (50 fmol) in the presence of a single dNTP (100 μM) at 25 °C for 10 min as described in Ref. 29. The sequences of template 1 and primer 3 were 5'-GATCCCTCTAGAGTCGACCTGCA and 5'-TGCAGGTCGACTCTAG, respectively. The specificity of nucleotide incorporation was determined for underlined T in template 1. The amount of the incorporated nucleotide (I<sub>i</sub> (i = A, G, C, T)) was quantified by PAGE analysis and the nucleotide specificity (%) was calculated by 100 x I<sub>i</sub>/(I<sub>A</sub>+I<sub>G</sub>+I<sub>C</sub>+I<sub>T</sub>).  

<sup>b</sup> ND: Nucleotide incorporation was not detected over the background.
Masaoka et al. Fig. 2

| Template | Incubation time (min) |
|----------|-----------------------|
|          | 24AT                  |
|          | 24AF                  |
|          | M                     |
| 3        | 5                     |
| 10       | 3                     |
| 5         | 5                     |
| 10        | 3                     |

- Fully replicated product (24mer)
- fU site (14mer)
- Primer (10mer)
Masaoka et al. Fig. 4
Figure 5

Masaoka et al.
A

\[
\begin{align*}
\text{fU (enol)} & \quad \text{fU (keto)} \\
\text{fU (ionized/enolate)} & \\
\text{Tautomerization} & \quad \text{Ionization}
\end{align*}
\]

\[
(pK_a = 8.6)
\]

B

\[
\begin{align*}
\text{fU (ionized)} : G & \quad \text{fU (keto)} : G \\
(\text{Watson-Crick base pair}) & \quad (\text{Wobble base pair})
\end{align*}
\]

Masaoka et al. Fig. 6
Masaoka et al. Fig. 7

| Primer terminus | T:A | fU:A | T:G | fU:G |
|-----------------|-----|------|-----|------|
| dCTP            | –   | +    | –   | +    |

1 2 3 4 5 6 7 8
Oxidation of thymine to 5-formyluracil in DNA promotes misincorporation of dGMP and subsequent elongation of a mismatched primer terminus by DNA polymerase
Aya Masaoka, Hiroaki Terato, Mutsumi Kobayashi, Yoshihiko Ohyama and Hiroshi Ide
J. Biol. Chem. published online January 29, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M008598200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts