Ionization of Bromouracil and Fluorouracil Stimulates Base Mispairing Frequencies with Guanine*

(Received for publication, November 17, 1992, and in revised form, March 23, 1993)

Hong Yu, Ramon Eritja, Linda B. Bloom, and Myron F. Goodman

From the Department of Biological Sciences, University of Southern California, Los Angeles, California 90089-1340 and the 
Department of Molecular Genetics, Center for Research and Development, Jordi Girona 18-26, 08034 Barcelona, Spain

To test whether ionized base pairs influence polymerase-catalyzed misinsertion rates, we measured the efficiency of forming 5-bromouracil (B), 5-fluorouracil (F), and thymine base pairs with guanine and adenine as a function of pH using avian myeloblastosis reverse transcriptase when B, F, and T were present as dNTP substrates, misincorporation efficiencies opposite G, normalized to incorporation of C opposite G, increased by about 20-, 13-, and 7-fold, respectively, as reaction pH increased from 7.0 to 9.5. Incorporation efficiencies to form the correct base pairs, B-A and F-A, normalized to T-A, decreased by 4- and 8-fold, respectively, with increasing pH. The effects of pH on misincorporation efficiencies were about 10-fold greater when B, F, and T were present as template bases. The relative misincorporation efficiencies of G opposite template B, F, and T, normalized to incorporation of A opposite B, F, and T, increased by about 430-, 370-, and 70-fold, respectively, as pH was increased from 6.5 to 9.5, while correct incorporation of A opposite template B and F decreased about 10-fold over the same pH range. Plots depicting incorrect and correct incorporation efficiencies versus pH were fit to a pH titration equation giving the fraction of ionized base pairs in mutagenesis. A few examples of structures that have been observed include: G-T wobble (6-8); A-C protonated wobble (9-11); G(syn)-A(anti), G(anti)-A(syn), and G(anti)-A(anti) with both bases as favored tautomers (6, 12-14). The structure of 5-bromouracil (B) and 5-fluorouracil (F) base pairs with A and G has been determined by NMR. B-G and F-G base pairs exist in a pH-dependent equilibrium between ionized and wobble structures (15, 16), while B-A and F-A are pH-independent Watson-Crick structures (17-19). Ionization of B and F allows them to adopt a Watson-Crick structure when forming mispairs with G (15, 16). The observation of 5-bromouracil in an anionic form in a B-G mispair agrees with an earlier proposal by Lawley and Brookes (20) and more generally with the ideas proposed by Ramsey Shaw and co-workers (21) on the involvement of ionized base pairs in mutagenesis. In this study, we varied pH over a range of 6.5-9.5 and measured the kinetics of forming mispairs involving B, F, and T with G and correct pairs with A. In the lower portion of the pH range the halouracils and T are predominantly electrically neutral, while in the higher range, the halouracils as triphosphates, or inside of the DNA molecule paired with G, are mainly ionized (anionic form) (15, 16). We asked whether the efficiencies of mispair formation increased with pH and whether the efficiencies of correct pair formation decreased with pH. The main objective is to determine whether the rate of forming halouracil-G mispairs in the polymerase active cleft is proportional to the concentration of halouracil anions present either on the template strand or as dNTP substrates.

* This research was supported by National Institutes of Health Grant GM21422, North Atlantic Treaty Organizational Collaborative Research Grant 900554, and National Institutes of Health Postdoctoral Research Fellowship GM15034 (to L. B. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 The abbreviations used are: dNTP, deoxyribonucleoside triphosphate; AMV, avian myeloblastosis virus; RT, reverse transcriptase; FuUTP, fluoroxydouridine triphosphate; BrdUTP, bromodeoxyuridine triphosphate; B, 5-bromouracil; F, 5-fluorouracil; U present either on the template strand or as dNTP substrates.
An increase in halobacterial base mispairing frequencies and a concomitant decrease in correct base pairing frequencies with increasing pH would be consistent with ionization models of mutagenesis (20, 21) and inconsistent with models requiring disfavored tautomers to form H-bonded base mispairs (1-4).

**EXPERIMENTAL PROCEDURES**

**Enzymes—**AMV RT (specific activity, 62,700 units/mg) was purchased from U. S. Biochemical Corp. One unit is that amount of enzyme required to incorporate 1 nmoI of labeled dTMP into nucleic acid product in 10 min at 37 °C.

**Chemicals—**Nonradioactive dNTP substrates were purchased from Pharmacia/LKB Technology Inc., except dUTP which was a product of Sierra Biochemicals (Tucson, AZ). Radioactive [γ-32P] ATP (4500 Ci/mmol) was purchased from ICN Radiochemicals, Inc.

**Primer-Templates—**For assays in which dUTP and BrdU dUTP were present as substrates for AMV RT, single-stranded M13 DNA template-primer, isolated from wild type M13 bacteriophage grown in *E. coli* strain JM103, was annealed with primer p5 (5′ AT- TAAATCCCTTGGCCG 3′) and p14 (5′ AAACGGGTAAAAATACGT 3′) (22). For assays in which B and F were present in the template strand, the 26-mer oligonucleotides (5′ CATCACCXGGAAACCTGCTGAGTCGTGACTGGGA 3′, where X is B, F, or T) were used together with primer p7 (5′ TCCCAAGCTGACCCTGGAG 3′). All synthetic oligonucleotides except for B- and F-containing oligonucleotides were synthesized on an Applied Biosystems DNA synthesizer by Lynn Williams (Comprehensive Cancer Center, University of Southern California, Los Angeles, CA) using standard 2-cyanoethyl phosphoramidites and purified by polyacrylamide gel electrophoresis.

**Synthesis of B- and F-containing Oligonucleotides—**An important matter when non-natural base analogues are introduced in DNA templates is to ensure that DNA chemical synthesis does not affect the integrity of the base analogue. When we prepared oligonucleotides containing B and F for NMR studies using phosphate triester methodology (17), a side reaction was observed with 5-bromodeoxyuridine containing phosphoramidites (Biogenex, CA) on a 392 DNA synthesizer (Applied Biosystems) using commercially available 5′-O-DMT-F and -B and PAC-protected templates except for B- and F-containing oligonucleotides were synthesized on an Applied Biosystems DNA synthesizer by Lynn Williams (Comprehensive Cancer Center, University of Southern California, Los Angeles, CA) using standard 2-cyanoethyl phosphoramidites and purified by polyacrylamide gel electrophoresis.

**Nucleotide Misinsertion Rate Dependence on pH**

- **Reaction Conditions—**Procedures of primer 5′-end labeling, primer-template annealing, gel electrophoresis, autoradiography, and densitometry integration were performed as described (22, 27). To remove the buffer in primer-template solution, the annealed primer-template was precipitated by adding 0.1 volume of 5 M ammonium acetate and 2 volumes of cold ethanol, and the DNA pellet was resuspended with H2O (the same volume. The primer-template/enzyme solution was prepared by adding 5 μl of AMV RT (0.45 units/μl, 42 nm) to 85 μl of primer-template solution (70 nm). All kinetic measurements were carried out using a "running start" reaction (22, 27), where nucleic acid incorporation was examined at a target template site located 2 bases beyond the 3′ end of the initial primer 3′ terminus. The reactions were started by mixing equal volumes (3 μl) of primer-template/enzyme solution with dNTP solution (150 mM NaCl, 16 mM MgCl2, 1 mM dithiothreitol, and 120 mM Tris-HCl) containing a constant (saturating) concentration of the running start nucleotide and a variable concentration of dNTP for insertion at the target template site. To verify that the reaction pH was not changed significantly at high dNTP concentrations, at each pH between 6.5 and 9.5, in increments of 0.5 pH unit, we measured pH both in the presence and absence of 5 mM dTTP, dCTP, or UTP at 37 °C; at the highest and lowest pH, the reduction was less than 0.1 pH unit in the presence of 5 mM bromodeoxyuridine triphosphate. The reactions were carried out at 37 °C for 2 min and terminated by the addition of 1 μl of 100 mM HCl in 95% formamide. The reaction conditions were chosen such that product accumulation was linear with time. Labeled primer molecules extended by 0, 1, and 2 or more nucleotides were separated by polyacrylamide gel electrophoresis, and the fraction of each product was determined by integrating band intensities (22, 27).

**Curve Fitting—**The pH rate profiles were fit by assuming that the observed rate, Robs, set equal to either Vmax/Km or to the misinsertion efficiency, Iimm, given by a population weighted average of the rates for insertion of the protonated and ionized species as shown in Equation 1.

\[ R_{obs} = \frac{K_m}{V_{max}} + \frac{[H^+]^2}{K_a} \]

where subscripts NH and N represent protonated and ionized species, respectively, and X represents the mole fraction. The mole fractions of protonated and ionized species can be expressed in terms of [H+], determined by measuring the pH under reaction conditions, and the dissociation constant for the acid (Ka) as follows.

\[ R_{obs} = \frac{[H^+]^2}{K_a} \]
Observed rates as a function of pH were fit by Equation 2a using a nonlinear regression routine contained in SigmaPlot\textsuperscript{a}, Jandel Scientific (Corte Madera, CA) to estimate values for the parameters \( K_a \), \( R_{on} \), and \( R_{off} \). The data conform reasonably well to the generated titration curves (see Figs. 4 and 6 and 7). However, an accurate estimate for \( pK_a \) values cannot be obtained because the data exhibit insufficient curvature in the high pH region for incorrect base pairs and low pH regions for correct base pairs. Sufficient curvature is needed to make an accurate extrapolation to minimum and maximum \( R_{on} \) and \( R_{off} \) values, and to estimate the inflection point (\( pK_a \) value).

A rough estimate for \( pK_a \) values was made by fitting the data to a general sigmoidal function, Equation 3, using the nonlinear regression curve-fitting routine in SigmaPlot\textsuperscript{a}.

\[
R_{obs} = \frac{R_{max} - R_{min}}{1 + (\frac{pH}{pK_a})^b} + R_{min} \quad \text{(Eq. 3)}
\]

where \( R_{obs} \) are the observed values for \( f_{obs} \), \( R_{max} \), and \( R_{min} \) are the asymptotic maximum and minimum rate values, respectively, pH is the measured pH, \( pK_a \) is the calculated \( pK_a \) (pH value at the inflection point), and \( b \) is an adjustable slope parameter, where \( b < 0 \) gives a rising sigmoidal curve. The \( pK_a \) values obtained were 8.7 ± 0.2, 9.1 ± 0.8, and 11 ± 8 for insertion of F, B, and T opposite G, respectively (see Fig. 4b, inset), for an average value of \( b = -11.8 ± 3.8 \). For the reciprocal measurement having a halouracil base present on the template strand, the \( pK_a \) values are 8.6 ± 0.4 for insertion of dGMP opposite B and 8.2 ± 0.1 for insertion of dGMP opposite F. The \( pK_a \) for insertion of dGMP opposite T is 8.6 ± 0.1, indistinguishable from insertion of dGMP opposite B. It should be emphasized that the calculated \( pK_a \) values use data obtained at lower pH to generate expected data at higher pH and, therefore, must be regarded as no more than rough approximations.

**RESULTS**

5-Bromouracil and 5-fluorouracil are predicted to form "stable," multiple H-bonded base mispairs with G (Fig. 1). Based on NMR spectroscopy of double-stranded oligonucleotides, mispaired species involving ionized and wobble B, G, and F-G were observed in pH-dependent equilibrium (15, 16). The triply H-bonded structures predicted to occur when B and F are in unfavorable enol tautomeric conformations (Fig. 1) were not observed. It is possible that disfavored tautomers participate in forming base mispairs during catalysis while ionized base pairs, observed by NMR, are present in the final DNA product. Our objective was to investigate whether ionization of B and F influenced deoxynucleotide misinsertion rates. The kinetics of formation of B-G and F-G mispairs and B-A and F-A correct base pairs were measured as a function of pH, with B, F, and T present as dNTP substrates (Fig. 2a) or as DNA template bases in the reaction (Fig. 2b).

Insertion kinetics were carried out in a "running start" reaction (22, 27) in which 5' labeled primers were extended by addition of two nucleotides prior to reaching a template target site G or A (Fig. 2a) or T, B, or F (Fig. 2b). The running start substrates, dATP or dTTP, were held at constant (saturating) concentrations, and the [dNTP] for insertion at the target site was varied. Primer extension dependence on pH and [dNTP] are shown for correct insertion of C and misinsertion of B, and T opposite template G (Fig. 3). The velocity to insert a nucleotide opposite the target G site is proportional to the ratio of integrated band intensities at the target (C-G, B-G, or T-G) compared with the previous (A-T) site, \( I_{T}/I_{T-1} \) (27, 28). The nucleotide insertion efficiency, \( V_{max}/K_m \), is obtained by dividing the \( V_{max}/K_m \) ratio for misinsertions by \( (V_{max}/K_m)_{cor} \) for correct insertions (Equation 4).

\[
f_{obs} = \frac{(V_{max}/K_m)_{w}}{(V_{max}/K_m)_{h}} \quad \text{(Eq. 4)}
\]

The nucleotide insertion fidelity is the reciprocal of \( f_{obs} \).

Insertion of 5-Bromo-2'-deoxyuridine, 5-Fluoro-2'-deoxyuridine, and Thymidine Nucleotides Opposite Template G and A as a Function of pH—The relative velocity for insertion of B opposite G is given by the ratio of adjacent integrated band intensities, \( f(B.G)/I(A.T) \) (Fig. 3b). The relative velocity for insertion of B opposite G was considerably more sensitive to pH changes than the insertion of either T opposite G, \( I(T.G)/I(A.T) \) (Fig. 3c), or C opposite G, \( I(C.G)/I(A.T) \) (Fig. 3a).

Changes in pH appeared to have a much more pronounced effect on values of relative \( V_{max} \) than \( K_m \) (data not shown).

---

**Fig. 1. Models for U^X.G mispairing.** a, three forms of a halouracil base: conversion of keto to ionized (anionic) form increases with pH, while favored keto and disfavored enol tautomeric forms are present in pH-independent equilibrium. X represents either CH, B, or F. b, three models of U^X.G mispairs involving different forms of U^X: ionized (anionic) form of U^X, favored keto tautomer of U^X as part of a wobble base pair, and disfavored tautomer (enol) form of U^X.
Relative \( V_{\text{max}} \) misinsertion values increased 5–50-fold with increasing pH, with the halouracils present either on the template or as dNTP substrates. Apparent \( K_m \) values for misinsertion of dGMP opposite template halouracils bases showed a 2-fold decrease between pH 6.5 and 9.5, in a range of 2.5 to 1 mM, from low to high pH. When the halouracils were present as deoxynucleotide substrates, misinsertion \( K_m \) values varied between 400 \( \mu \)M and 1.2 mM. For correct insertions opposite template halouracils (dAMP opposite template G or A target site), there were no clear systematic changes in \( K_m \) and \( V_{\text{max}} \); the range of \( K_m \) values was 15–150 \( \mu \)M.

Effect of pH on Insertion Efficiencies and Fidelity—\( V_{\text{max}}/K_m \) and \( f_{\text{ins}} \) values for insertion of C, B, F, or T opposite template G in the pH range 6.5–9.5 are given in Table I, and the values corresponding to insertion of T, B, and F opposite template A are shown in Table II. \( V_{\text{max}}/K_m \) and \( f_{\text{ins}} \) are plotted as a function of pH in Fig. 4. The data for the B·G, F·G, and T·G mismatches exhibited qualitatively similar “S-shaped” pH profiles predicted by the pH titration equation (Fig. 4a, solid curves; Equation 2a). Insertion efficiencies, \( V_{\text{max}}/K_m \), increased by about 10- and 14-fold for B·G and F·G mismatches, and about 4-fold for T·G mismatches as the pH was raised from 6.5 to 9.5 (Table I).

In contrast to the increase in \( V_{\text{max}}/K_m \) for insertion of 5-substituted deoxynucleotides as the pH was increased from 6.5 to 9.5, the shapes of \( V_{\text{max}}/K_m \) curves for correct pairs C·G and T·A (Fig. 4c) appeared as typical enzyme activity profiles, having a maximum in activity at about pH 7.0, and showed a decrease in activity as the pH was increased from 7 to 9.5. Although the general shapes of the pH profiles for B·A and F·A appeared qualitatively similar to the C·G and T·A curves, the nucleotide analogue insertion efficiencies exhibit a much steeper decrease for pH values above 7. From pH 7.0 to 9.5, B·A and F·A insertion efficiencies decreased by about 9- and 15-fold, respectively, while C·G and T·A efficiencies decreased by only about 2–3-fold (Tables I and II). The pH profile for insertion of C opposite G should provide a fairly accurate indication of the pH activity profile of the enzyme since C contains no ionizable groups in the pH 7–9 range, and G is roughly 17% ionized at pH 9 (\( K_c \) of dGMP = 9.7, Ref. 29). We suggest that ionization contributes to the steep decrease in insertion rates of B and F opposite A because the anionic forms of B and F should not form stable base pairs with A. Also, less efficient binding of the B and F ionized triphosphates to the triphosphate binding site of the polymerase, resulting from the additional negative charge, could also cause a decreased rate of insertion of B and F nucleotides opposite A with increasing pH (Table II).

The relative efficiency, \( f_{\text{ins}} \), for insertion of F opposite G increased by about 24-fold, from 9.1 \( \times 10^{-3} \) at pH 6.5 to 2.2 \( \times 10^{-3} \) at pH 9.5 (Fig. 4b, Table I). F·G, B·G, and T·G mis-

---

**Fig. 2.** Primer-template sequences used to measure deoxynucleotide insertion efficiencies. a, C, B, F, or T is present as dNTP substrates in the reaction for insertion opposite an M13 template G or A target site. Both template target sites have the same two nearest neighbor bases on the 3'-side and same nearest neighbor base to the 5'-side. b, B, F, or T is present on the template, while G or A are deoxynucleotide substrates. The template strand is a synthetic oligomer (26-mer) having an identical base sequence except at the target site containing either B, F, or T. Deoxynucleotide insertion kinetics are measured by varying [dNTP].

**Fig. 3.** Gel autoradiograms showing the pH-dependent nucleotide incorporation opposite template G by AMV reverse transcriptase. A primer \( (P) \) is extended, in a running start reaction, by incorporation of two A deoxynucleotides to reach the template target site G (the running start nucleotide, dATP, is present in the assay at saturating concentration, 50 \( \mu \)M, for incorporation at the template T sites prior to the target G site). Also present in the reaction for incorporation at the target G site are variable concentrations of: c, dCTP; b, BrdUTP; c, dTTP. The reaction pH from left to right is 7.0, 7.5, and 9.0. The unextended \( 32^P \)-labeled primer is shown in the top left panel, lane 0. The [dNTP] for insertion at the target G site in lanes 1–7 is, in units of \( \mu M: a, [\text{dCTP}] = 0, 0.1, 0.2, 0.4, 0.8, 1.6, \) and 3.2; b, [BrdUTP] = 19, 37, 73, 140, 290, 560, and 1140; c, [dTTP] = 90, 190, 332, 694, 1700, 2350, and 5320.

---

Nucleotide Misinsertion Rate Dependence on pH
pairing efficiencies were roughly similar, between about 5 \times 10^{-5} and 1 \times 10^{-4} at the lower pH values, 6.5 and 7.0. There was about a 20- and 7-fold increase, respectively, in B-G and T-G values at pH 9.5 compared with pH 6.5 (pH 7.0 for the case of B; see Table I). In contrast to the plots of V_{max}/K_m, which begin to bend over above pH 8.5 (Fig. 4a), the apparent saturation for f_{ins} do not exhibit significant saturation even at pH 9.5. The apparent saturation in V_{max}/K_m profiles for the three mispairs at high pH is more likely to be caused by a decrease in the polymerase activity, as seen for C-G and T-A (Fig. 4c), than by complete ionization of the bases. The relative insertion efficiencies (Fig. 4b) may correspond more closely than V_{max}/K_m (Fig. 4a) to the degree of ionization of the bases because f_{ins} is normalized to enzyme activity, i.e., the efficiency for insertion of correct G.

To make a rough estimate of K_m values, a sigmoidal function was used to fit data for pH dependence for insertion of 5-substituted deoxyuracil derivatives opposite G (Fig. 4b, inset; see "Experimental Procedures," Equation 3). We obtained K_m values of 8.7 \pm 0.2, 9.1 \pm 0.8, and 11 \pm 8 for insertion opposite G of F, B, and T, respectively. Although these values are greater than the reported values of 7.6, 8.3, and 10 for the nucleotides F, B, and T, respectively (29, 30), they follow the expected trend. For the reciprocal measurement having a halouracil base present on the template strand, the K_m values are 8.6 \pm 0.4 for insertion of dGMP opposite B (compared with an apparent K_m of 8.6 for B-G pairs in double-stranded DNA, Ref. 16) and 8.2 \pm 0.1 for insertion of dGMP opposite F (compared with an apparent K_m of 8.3 for F-G pairs in DNA, Ref. 15). The K_m for insertion of dGMP opposite T is 8.6 \pm 0.1, indistinguishable from insertion of dGMP opposite B. However, it must be emphasized that the K_m values estimated in this way are, at best, only rough approximations because they rely on the assumption that f_{ins} data obtained in a lower pH range can be extrapolated to an experimentally inaccessible pH range.

Insertion of G and A Deoxynucleotides Opposite Template B, F, and T as a Function of pH—Oligonucleotide templates were synthesized containing either B, F, or T at a specific template target site (Fig. 2b), and the effect of pH on the kinetics of G misinsertion or A insertion was measured. The curves describing the kinetics for correct and incorrect base pair formation when B, F, or T was present on the template strand are shown in Fig. 5, where V_{max}/K_m and f_{ins} dependence on pH was fit to the pH titration equation (Equation 2a; Fig. 5, a and b, solid curves).

Differences of an order of magnitude or more were observed for the effects of pH on error rates, comparing template (Fig. 5, Table III) versus substrate halouracils (Fig. 4, Table I). The effect of pH on misinsertion rates was much more pronounced when B and F were located on the template strand. V_{max}/K_m for misinsertion of G opposite B, F, and T increased by about 35-, 67-, and 9-fold, respectively, as pH was increased from 6.5 to 9.5 (Fig. 5a, Table III). The corresponding increase in f_{ins} was about 400-fold for G-B and G-F mispairs and 70-fold for G-T mispairs (Fig. 5b, Table III).

The effect of pH on the kinetics of forming A-B, A-F, and A-T correct base pairs is shown in Fig. 5c and Table III. Unlike the case for the halouracil dNTP substrates (Fig. 4, c and d), there was no indication of a V_{max}/K_m maximum at pH 7.0. Instead, V_{max}/K_m insertion values appeared to decrease monotonically between pH 6.5 and 9.0, with A inserted opposite B more efficiently than opposite either T (1.7-fold at pH 7.0) or F (4-fold at pH 7.0). The V_{max}/K_m values decreased as a function of pH by about 6-17-fold, while the values for f_{ins} the ratio of V_{max}/K_m values for A-F and A-B compared with A-T, changed by less than a factor of 2 (Table III).

### Table I

| pH | B-G | F-G | T-G | C-G |
|----|-----|-----|-----|-----|
| V_{max}/K_m | f_{ins} | V_{max}/K_m | f_{ins} | V_{max}/K_m | f_{ins} | V_{max}/K_m | f_{ins} | V_{max}/K_m | f_{ins} |
| \( \mu M^{-1} \) | \( \mu M^{-1} \) | \( \mu M^{-1} \) | \( \mu M^{-1} \) | \( \mu M^{-1} \) | \( \mu M^{-1} \) | \( \mu M^{-1} \) | \( \mu M^{-1} \) | \( \mu M^{-1} \) | \( \mu M^{-1} \) |
| 6.5 | NQ | 3.1 (\pm 0.9) \times 10^{-4} | 9.1 (\pm 2.9) \times 10^{-5} | 3.9 (\pm 1.0) \times 10^{-4} | 2.9 (\pm 0.8) \times 10^{-5} |
| 7.0 | 2.0 (\pm 0.1) \times 10^{-4} | 4.6 (\pm 0.9) \times 10^{-5} | 7.2 (\pm 0.0) \times 10^{-5} | 1.8 (\pm 0.9) \times 10^{-5} |
| 7.5 | 5.1 (\pm 0.9) \times 10^{-4} | 1.7 (\pm 0.8) \times 10^{-5} | 4.2 (\pm 1.6) \times 10^{-5} | 1.5 (\pm 1.1) \times 10^{-5} |
| 8.0 | 7.4 (\pm 5.7) \times 10^{-4} | 2.0 (\pm 0.5) \times 10^{-5} | 2.4 (\pm 1.8) \times 10^{-5} | 6.4 (\pm 3.3) \times 10^{-5} |
| 8.5 | 1.4 (\pm 0.6) \times 10^{-4} | 3.9 (\pm 1.3) \times 10^{-5} | 3.3 (\pm 1.9) \times 10^{-5} | 1.0 (\pm 0.5) \times 10^{-4} |
| 9.0 | 2.3 (\pm 1.0) \times 10^{-4} | 6.6 (\pm 3.4) \times 10^{-5} | 4.4 (\pm 3.4) \times 10^{-5} | 1.8 (\pm 1.1) \times 10^{-4} |
| 9.5 | 1.9 (\pm 10^{-4}) | 9.2 (\pm 3.9) \times 10^{-5} | 4.4 (\pm 1.2) \times 10^{-5} | 2.2 (\pm 0.4) \times 10^{-5} |

* V_{max} = \( \bar{t}_{V_{max}}/K_m \), which is a relative value. The V_{max} and K_m values were obtained by a nonlinear least squares fit of the kinetics data to a rectangular hyperbola. Values shown with \( \pm S.D. \) are average values obtained from the number of independent experiments shown in parentheses.

### Table II

| pH | B-A | F-A | T-A |
|----|-----|-----|-----|
| V_{max}/K_m | f_{ins} | V_{max}/K_m | f_{ins} | V_{max}/K_m | f_{ins} |
| \( \mu M^{-1} \) | \( \mu M^{-1} \) | \( \mu M^{-1} \) | \( \mu M^{-1} \) | \( \mu M^{-1} \) | \( \mu M^{-1} \) |
| 6.5 | 0.21 (\pm 0.04) \times 10^{-2} | 0.17 (\pm 0.01) | 0.42 (\pm 0.15) \times 10^{-3} | 0.40 (\pm 0.14) |
| 7.0 | 0.28 (\pm 0.06) \times 10^{-2} | 0.14 (\pm 0.02) | 0.59 (\pm 0.10) \times 10^{-3} | 0.31 (\pm 0.01) |
| 7.5 | 0.22 (\pm 0.04) \times 10^{-2} | 0.13 (\pm 0.02) | 0.35 (\pm 0.10) \times 10^{-3} | 0.26 (\pm 0.03) |
| 8.0 | 0.16 (\pm 0.01) \times 10^{-2} | 0.11 (\pm 0.01) | 0.27 (\pm 0.04) \times 10^{-3} | 0.18 (\pm 0.01) |
| 8.5 | 0.13 (\pm 0.03) \times 10^{-2} | 0.08 (\pm 0.02) | 0.15 (\pm 0.04) \times 10^{-3} | 0.11 (\pm 0.03) |
| 9.0 | 0.07 (\pm 0.01) \times 10^{-2} | 0.07 (\pm 0.01) | 0.08 (\pm 0.02) \times 10^{-3} | 0.08 (\pm 0.01) |
| 9.5 | 0.03 (\pm 0.02) \times 10^{-2} | 0.04 (\pm 0.02) | 0.04 (\pm 0.02) \times 10^{-3} | 0.05 (\pm 0.02) |

* Numbers in parentheses are the numbers of independent experiments done at that pH.

* f_{ins} = (V_{max}/K_m)_X/ (V_{max}/K_m)_C, where X = C, B, F, or T.

* NQ, not quantified; autoradiogram band intensities were barely detectable. V_{max}/K_m < 7 \times 10^{-5}, f_{ins} < 3 \times 10^{-5}.
Nucleotide Misinsertion Rate Dependence on pH

**FIG. 4.** Deoxynucleotide incorporation efficiencies as a function of pH when B, F, and T are present as dNTP substrates for AMV RT. a, efficiencies of incorporating B (V), F (O), and T (□) opposite template G; b, relative misincorporation efficiencies, $f_{in}$ (see Equation 4 in text for B-G (V), F-G (O), and T-G (□) mispairs); c, efficiencies of correctly incorporating C opposite template G (V) and T opposite template A (O); d, efficiencies of “correctly” incorporating B (V) and F (O) opposite template A. The symbols represent the experimental data (Tables I and II). The solid “S-shaped” curves in panels a and b were obtained by a non-linear least squares fit of the data to a pH titration equation (see Equation 2a and “Curve Fitting” under “Experimental Procedures”). The solid spline curves in panels c and d were drawn by direct connection of the data points. In the inset, a sigmoidal function was used to fit data for pH dependence for insertion of 5-substituted deoxyuracil derivatives opposite G (see “Curve Fitting” under “Experimental Procedures”). Data were fit using the sigmoidal function in Equation 3 and the nonlinear regression curve-fitting routine in Sigma Plot by Jandel Scientific. When the data are analyzed in this way, pK values of 8.7 ± 0.2, 9.1 ± 0.8, and 11 ± 8 are obtained for insertion opposite G of F, B, and T, respectively.

**DISCUSSION**

According to recent NMR data, a variety of base mispairs in duplex DNA involve anionic (15, 16) and cationic (9-11, 31) structures. The base analogues bromouracil and fluorouracil were observed to form ionized mispairs with G, in equilibrium with neutral wobble structures (15, 16). In this paper, we used AMV RT to measure nucleotide misinsertion kinetics as a function of pH to determine whether ionization of F and B may be involved in base mispairing during polymerization.

The transition from keto to ionized form of the halouracil base increases with pH. Since the equilibrium between keto ↔ enol forms is independent of pH, an increase in ionized form is accompanied by a concomitant decrease in the neutral enol tautomeric form (Fig. 1a). Therefore, an observation of increased base mispairing frequencies with pH would be consistent with formation of an ionized base mispair during DNA synthesis, while a reduction in mispairing frequencies would be consistent with involvement of the neutral enol tautomer.

A large increase in nucleotide misinsertion efficiencies occurred when pH values were increased from 6.5 to 9.0. Enhanced misinsertion efficiencies were observed in reciprocal experiments with: (i) B, F, and T as dNTP substrates inserted at a defined template position containing G (Fig. 4, a and b; Table I); (ii) dGTP as substrate for insertion opposite either B, F, or T at a defined template site (Fig. 5, a and b; Table III). Low misinsertion efficiencies, $f_{in}$ ~ $5 \times 10^{-5}$ to $1 \times 10^{-4}$, were observed for insertion of F, B, and T opposite G at pH 6.5-7.0. At pH 9.5, the error rates were about $9.2 \times 10^{-4}$ for B-G, $2.2 \times 10^{-3}$ for F-G, and $5.7 \times 10^{-4}$ for T-G, an increase of about 20-, 24-, and 7-fold, respectively (Fig. 4b, Table I).

At pH 6.5, misinsertion efficiencies for T and F deoxynu-
Nucleotide Misinsertion Rate Dependence on pH

a. b.

Fig. 5. Deoxynucleotide incorporation efficiencies as a function of pH when B, F, and T are present on the DNA template strand. a, efficiencies of incorporating G deoxynucleotide opposite template B (V), F (O), and T (□); b, relative misincorporation efficiencies, \( f_{\text{mis}} \), for G-B (V), G-F (O), and G-T (□) mispairs; c, efficiencies of incorporating A deoxynucleotide opposite B (V), F (O), and T (□). The solid curves were obtained by a non-linear least squares fit of the data to a pH titration equation for the fraction of ionized base as a function of pH (see "Experimental Procedures").

cleotides opposite G were 8.5 to 9.1 \( \times 10^{-3} \) (Table I). However, incorporation of B opposite template G (pH 6.5) was beneath the level of detection (Table I), using concentrations of BrdUTP as high as 1.1 mM. At pH 7.0, in order to observe B(primer)-G mispairs it was necessary to have [BrdUTP] \( \geq 290 \) \( \mu \)M (Fig. 3b). These data suggest the possibility that B-G may form wobble pairs less readily than T-G. It is possible that the relatively efficient incorporation of B opposite template G at pH 6.5 may reflect, in part, a greater degree of ionization of F (pK\(_a\), 7.6, Ref. 30) compared with B (pK\(_a\), 8.3; Ref. 30).

The pH-dependent increase in misinsertion efficiencies was more than 10-fold larger when the halouracil was present as a template rather than dNTP substrate base (Fig. 5b, Table III). A comparison of Table I (U\(^X\)-G) with Table III for (G-U\(^X\)) shows that the effect of pH on \( V_{\text{max}}/K_a \) for insertion was about 3-fold greater when B, F, and T were present as template bases compared with dNTP substrates. The more pronounced effect of pH on G-U\(^X\) misinsertion efficiencies (\( f_{\text{mis}} \), Table III) was caused by a steeper decrease (17-fold for A-B, 6-fold for A-F, and 10-fold for A-T from pH 6.5 to 9.5) in A-U\(^X\) correct insertion rates compared with an approximate 2-fold decrease in insertion of C deoxynucleotide opposite template G (Table I). At pH 6.5, G-F, G-B, and G-T mispair efficiencies, \( f_{\text{mis}} \), were roughly 1.4–2.7 \( \times 10^{-5} \). The mispair efficiencies increased to about 1 \( \times 10^{-4} \) for G-F, 6 \( \times 10^{-3} \) for G-B, and 1 \( \times 10^{-3} \) for G-T at pH 9.0–9.5, an increase of roughly 400-fold for G-B and G-F and about 70-fold for G-T. These data are in agreement with the observation that B-G mispairs are formed more easily by E. coli polymerase I as pH was increased between 7.0 and 9.0, and that the increase was greater when B was present on the template (32). Perhaps the lower misincorporation efficiencies when the halouracils were present as dNTP substrates were caused by a less efficient binding of the B and F ionized triphosphates to the triphosphate binding site of the polymerase.

If base ionization is involved in the formation of mispairs during catalysis by polymerase, then the efficiencies either to misinsert a halouracil deoxynucleotide opposite template G or to misinsert G opposite a template halouracil should resemble a pH titration curve for the fractional amount of ionic species in solution as a function of pH. Satisfactory fit to a
titration curve was demonstrated in plots of $V_{\text{max}}/K_m$ to form $U^\text{X}$-G(template) and G-$U^\text{X}$ (template), and in plots for $f_n$ (Fig. 4, a and b, and Fig. 5, a and b). The relative misinsertion efficiency, $f_n$, given by Equation 4, is obtained by dividing misinsertion $V_{\text{max}}/K_m$ values by corresponding correct base pair insertion efficiencies. Division by correct insertion efficiencies normalizes $f_n$ by canceling out pH-dependent changes in polymerase activity. Plots for $f_n$ as a function of pH can be fit to a titration curve (Figs. 4b and 5b; see “Experimental Procedures,” Equation 2a). A rough estimate of $pK_a$ values was obtained by fitting the data to a sigmoidal function (Fig. 4b, inset; see “Experimental Procedures,” Equation 3). For insertion of F, B, and T opposite template G, $pK_a$ values were 8.7 ± 0.2, 9.1 ± 0.8, and 11 ± 8, respectively, and the values were 8.6 ± 0.4, 8.2 ± 0.1, and 8.6 ± 0.1 for insertion of G opposite template B, F, and T, respectively.

In contrast to the pH dependence for the efficiencies of forming $U^\text{X}$-G mispairs, a plot of $V_{\text{max}}/K_m$ for insertion of C opposite G and T opposite A appears to resemble a “typical” enzyme activity versus pH profile (Fig. 4c). There was at most a 2.5-fold reduction in polymerase activity for insertion of C opposite G between pH 7.0 and 9.5. There was less than a 3-fold reduction in activity for T opposite A in the same pH range (Fig. 4c). To explain why it is that T-G shows a pH profile (Fig. 4b) but not T-A, it is important to note that when misinsertion efficiencies are very low as for T-G (10^{-4}; Table I), having a small percent of ionized T present (~10%) can greatly increase the misincorporation efficiency. However, in the case of T-A where correct incorporation efficiencies are high (~1, Table II), the presence of a small amount of ionized T has only a negligible effect on $V_{\text{max}}/K_m$ values for incorporating T opposite A.

The steeper (~10-fold) decline in B-A(template) and F-A(template) insertion activities with increasing pH (Fig. 4d, Table II) compared with the relatively shallow 2-fold decline for T-A(template) (Fig. 4c, Table II) may have resulted from a greater degree of ionization of fluorine- and bromine-substituted nucleotides (FUDP, $pK_a$ 7.6; Ref. 30) and BrUDP ($pK_a$ 8.3; Ref. 30) compared with dTMP ($pK_a$ 10.0; Ref. 29). We suggest that three factors may be contributing to a reduction in the rate of incorporating B and F opposite template A at high pH: (i) ionized halouracils are unable to form two H-bonds with A; (ii) ionized halo-dNTPs might bind with reduced affinity at the triphosphate binding site; (iii) the polymerase is less active at high pH. The reduction in insertion efficiency for F opposite A with increasing pH correlates with a decrease in $T_m$ with increasing pH for poly(A-F) (30).

When B, F, and T were present on the template strand, insertion of the correct base (A) decreased monotonically with increased pH (Fig. 5c). B was most efficient in directing insertion of A, while T and F were less efficient. Differences in the rates of correct base pair formation, A-B > A-T > A-F, correlate with $T_m$ values (19, 33). It is possible that substituents that increase hydrophobicity at the 5-position (bromine > methyl > fluorine) (34) may aid in stabilizing the polymerase-DNA dNTP complex. Thus, stimulation in nucleotide insertion rates might result from increased base stacking (22).

An important conclusion drawn from this study is that AMV RT catalyzed ionized B-G and F-G Watson-Crick mispairs in preference to neutral wobble structures. This conclusion is based on the qualitative similarity of $U^\text{X}$-G misinsertion rate dependence on pH to a pH titration profile. Note that the concentrations of ionized and keto forms of the bases are in pH-dependent equilibrium with each other (Fig. 1a). If AMV RT-catalyzed rates were similar for both ionized and keto forms, then mispair formation would be independent of pH. The observation that $U^\text{X}$-G mispairs rates showed a marked increase with pH is strong evidence that AMV RT favors formation of the ionized Watson-Crick structure. Other enzymes, however, may behave differently. An ability of some enzyme clefts to accommodate both wobble and ionized structures might account for the interesting observation of Driggers and Beattie (32) that changes in pH had little effect on B-G mispairing with E. coli polymerase III holoenzyme and T4 polymerase, although excision of mispairs by the potent proof-
Nucleotide Misinsertion Rate Dependence on pH

15943

reading exonucleases of polymerase III and T4 polymerase could be responsible for suppressing differences in pH-dependent mispair formation.

We cannot completely eliminate the possibility that ionization of an amino acid side chain in the enzyme active site contributed to the observed fidelity dependence on pH. However, the absence of a titration-like curve for incorporating C opposite G supports the idea that DNA base ionization is the major cause of base mispairs during DNA synthesis. Thus, in light of the recent NMR, x-ray, and enzymatic studies on base mispair structures, it no longer seems reasonable to require involvement of imino and enol tautomers in spontaneous mutagenesis, although a role for disfavored tautomers cannot be formally eliminated.

Acknowledgment—We thank Steven Creighton for making significant contributions to data analysis and for making numerous constructive suggestions concerning presentation of data.

REFERENCES

1. Watson, J. D., and Crick, F. H. C. (1963) Cold Spring Harbor Symp. Quant. Biol. 18, 123-131

2. Watson, J. D., and Crick, F. H. C. (1965) Nature 171, 964-967

3. Freese, E. (1969) J. Mol. Biol. 1, 87-105

4. Topal, M. D., and Fresco, J. R. (1976) Nature 263, 285-289

5. Straszewski, P., and Tamm, C. (1990) Angew. Chem. Int. Ed. Engl. 29, 36-57

6. Patel, D. J., Kozlowski, S. A., Ikuta, S., and Itakura, K. (1984) Fed. Proc. 43, 2663-2670

7. Hunter, W. N., Brown, T., Anand, N. N., and Kennard, O. (1986) Nature 320, 582-585

8. Hunter, W. N., Brown, T., Kneale, G., Anand, N. N., Rabinovich, D., and Kennard, O. (1987) J. Biol. Chem. 262, 9962-9970

9. Sowers, L. C., Fazakerley, G. V., Kim, H., Dalton, L., and Goodman, M. F. (1986) Biochemistry 25, 3983-3988

10. Wang, C., Gao, H., Geffrey, B. L., and Jones, R. A. (1991) J. Am. Chem. Soc. 113, 5486-5488

11. Bouliard, Y., Cognet, J. A., Gabarro, A. J., Le, B. M., Sowers, L. C., and Fazakerley, G. V. (1992) Nucleic Acids Res. 20, 1933-1941

12. Kan, L. S., Chandrasegaran, S., Pulford, S. M., and Miller, P. S. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 4263-4265

13. Brown, T., Hunter, W. N., Kneale, G., and Kennard, O. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 2492-2496

14. Brown, T., Leonard, G. A., Booth, E. D., and Chambers, J. (1989) J. Mol. Biol. 207, 455-457

15. Sowers, L. C., Eritja, R., Kaplan, B., Goodman, M. F., and Fazakerley, G. V. (1988) J. Biol. Chem. 263, 4794-4801

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

17. Fazakerley, G. V., Sowers, L. C., Eritja, R., Kaplan, B. E., and Goodman, M. F. (1987) J. Biomol. Struct. Dyn. 5, 629-666

18. Kremer, A. B., Mikita, T., and Beadley, G. P. (1987) Biochemistry 26, 391-397

19. Sowers, L. C., Eritja, R., Kaplan, B. E., Goodman, M. F., and Fazakerley, G. V. (1987) J. Biol. Chem. 262, 15436-15442

20. Lawley, P. D., and Brooks, P. (1982) J. Mol. Biol. 4, 216-219

21. Sowers, L. C., Ramsay Shaw, B., Veigl, M. L., and Sedwick, W. D. (1987) Mutat. Res. 177, 201-218

22. Mendelman, L. V., Bossalis, M. S., Petruska, J., and Goodman, M. F. (1989) J. Biol. Chem. 264, 14415-14423

23. Goldman, D., and Kalman, T. I. (1983) Nucleosides & Nucleotides 2, 175-187

24. Schulhof, J. C., Moklo, D., and Teoulle, R. (1987) Nucleic Acids Res. 15, 397-416

25. Vu, H., McCollum, C., Jacobson, K., and Theisen, P. (1990) Tetrahedron Lett. 31, 7269-7272

26. Fernandez, F. D., Palom, Y., Ikuta, S., Pedroso, E., and Eritja, R. (1990) Nucleic Acids Res. 18, 5729-5734

27. Bossalis, M. S., Petruska, J., and Goodman, M. F. (1987) J. Biol. Chem. 262, 14989-14995

28. Randall, S. K., Eritja, R., Kaplan, B. E., Petruska, J., and Goodman, M. F. (1987) J. Biol. Chem. 262, 6664-6670

29. Dunn, D. B., and Hall, R. H. (1975) in Handbook of Biochemistry and Molecular Biology (Fasman, G. D., ed) 3rd Ed., Vol. 1, pp. 76-206, CRC Press, Inc., Boca Raton, FL

30. Massoulie, J., Michelson, A. M., and Fochon, F. (1966) Biochim. Biophys. Acta 114, 16-26

31. Sowers, L. C., Fazakerley, G. V., Eritja, R., Kaplan, B. E., and Goodman, M. F. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 5434-5438

32. Driggers, P. H., and Beattie, K. L. (1986) Biochemistry 25, 1729-1735

33. Boulard, Y., and Gumpert, R. L. (1985) Nucleic Acids Res. 13, 8665-8684

34. Hansch, C., Leo, A., Unger, S. H., Kim, K. H., Nisakarzani, D., and Lien, E. J. (1973) J. Med. Chem. 16, 1207-1216