Proc. Natl. Acad. Sci. USA
Vol. 74, No. 11, pp. 4806–4810, November 1977
Biochemistry

2-Aminopurine-induced mutagenesis in T4 bacteriophage: A model relating mutation frequency to 2-aminopurine incorporation in DNA

(mutator/antimutator/DNA polymerase/replication errors)

MYRON F. GOODMAN*, RANDI HOPKINS*, AND WILLIS C. GORE†

* Department of Biological Sciences, University of Southern California, Los Angeles, California 90007, and † Department of Electrical Engineering, The Johns Hopkins University, Baltimore, Maryland 21218

Communicated by Martin D. Kamen, August 4, 1977

ABSTRACT **We measured the *in vivo* incorporation of 2-aminopurine into DNA of T4 bacteriophage allelic for gene 43 (DNA polymerase), mutator (*L56*), *43*$^+$, and antimutator (*L141*). The magnitude of incorporation (mol/mol of Thy) was 1/1500 in *L56*, 1/1600 in *43*$^+$, and 1/8900 in *L141*. The incorporation ratio *L56*:*43*$^+$:*L141 in vivo* was equal to that mediated by the purified DNA polymerases of these allelic phages *in vitro*. A model for 2-aminopurine-induced A·T ⇄ G·C transitions is discussed. The model is used to predict the magnitudes of replication errors (C mispairing with a template 2-aminopurine) and incorporation errors (2-aminopurine mispairing with a template C) per round of replication and to investigate the asymmetry in 2-aminopurine-induced transitions favoring the A·T → G·C pathway over G·C → A·T. We suggest that the fidelity of *L56* and *L141* DNA polymerases exemplifies one-step and two-step editing, respectively.**

Error detection and correction is a process basic to the maintenance of informational integrity in biological systems. Genetic studies (1–7) demonstrated that certain temperature-sensitive bacteriophage T4 mutants (mutators and antimutators) having single base substitutions in gene 43, the gene coding for DNA polymerase (8), can exhibit increased and decreased mutation frequencies, respectively, as measured by the revision of marker loci. Measurements of the insertion and removal of 2-aminopurine (AP) deoxynucleotides by purified mutator and antimutator DNA polymerases during *in vitro* DNA synthesis revealed that the fidelity of DNA replication can be influenced by the polymerase-associated 3′-exonuclease (9, 10) or proofreading activity (11). Similar studies on the mispairing of normal nucleotides using defined polymer templates (12, 13) further demonstrated that the insertion specificity may also influence the fidelity of the polymerization process.

AP, a base analogue of adenine, induces transition mutations through the pathway: A·T ⇄ AP·T ⇄ AP·C ⇄ G·C (14). In this paper, we address experimentally the initial and final steps in AP-induced A·T → G·C transitions. The first step is measured as AP incorporation into encapsulated DNA in T4 mutator (*L56*), wild-type (*43*$^+$), and antimutator (*L141*) backgrounds. The final step is measured as the AP-induced reversion (A·T → G·C transition) of the marker locus r_{II} *UV199* in gene 43 allelic bacteriophage (5, 6). Both AP incorporation and AP-induced transitions vary in a manner consistent with *L56* mutator or *L141* antimutator phenotypes. We will show that mutator and antimutator phage differ much more widely in AP-induced marker reversion than in AP incorporation. Because they respond differently to AP mutagenesis, the T4 phage must progress through the stages of AP-induced mutagenesis at significantly different rates. Therefore, it is possible to gain an intimate view of AP-induced mutagenesis and the role of DNA polymerase in nucleotide selection by quantitation of each step in the AP-induced A·T → G·C transition pathway in each of the gene 43 alleles.

A model of AP-induced transitions is proposed that facilitates analysis of the data by relating the population of each possible base pair in the transition pathway to the rate at which it is generated per round of replication. The experimental values for AP incorporation (population of AP·T base pairs) and AP-induced A·T → G·C transition frequencies (ratio of G·C/A·T base pairs generated at the site of the transition) are inserted into the model. The model allows us to deduce: (*i*) the absolute magnitudes of replication error frequencies defined as the misincorporation of C into DNA opposite a template AP in the three gene 43 allelic backgrounds; (*ii*) the basis of asymmetry observed in AP-induced transition frequencies in which A·T → G·C prevails over G·C → A·T; and (*iii*) the range of pool sizes of dAPTP relative to dATP which would indicate that the primary determinant of the absolute magnitude of AP incorporation is the T4 DNA polymerase as opposed to other replicative gene products such as genes 32, 41, 62, 44, and 45 (15). Finally, using a value predicted for the free energy difference between an A·T and an AP·T hydrogen-bonded base pair (16), we show that the process of error correction catalyzed during DNA synthesis by the *L56* and the *L141* DNA polymerases is consistent with a one-step proofreading (17) process for *L56* and a two-step process for *L141*.

MATERIALS AND METHODS

Bacterial and Viral Strains. Wild-type *Escherichia coli* B (CGSC no. 2) and CR63 (CGSC no. 3591) were obtained from the *E. coli* Genetic Stock Center at the Yale University School of Medicine. CR63 (λ) was obtained from M. J. Bessman. Bacteriophage T4B r_{II}*UV199* and temperature-sensitive mutants in gene 43, described in refs. 5 and 6, were the gift of J. W. Drake. Double mutants *UV199-L56* and *UV199-L141* were constructed in our laboratory.

2-[^{3}H]Aminopurine. [^{3}H]AP was prepared by ICN pharmaceuticals, Irvine, CA. The free base was purified on a Dowex formate column and desalted by adsorbing to activated charcoal.

Isolation of [^{3}H]AP Mutagenized Bacteriophage DNA. *E. coli* CR63 was grown to 2.5 × 10^8 cells per ml in M9a [M9 + 1% casamino acids (18)]. The cells were chilled, centrifuged, and resuspended in 1/10 volume of M9a containing double the glucose concentration. (Extra glucose is necessary to maintain the physiology of lysis inhibition in the presence of AP.) The artificially concentrated cells were prewarmed for 2 min at 32°. At 0 time, *UV199*, *UV199-L56*, or *UV199-L141* T4 bacteriophage were added at 5 phage per bacterium. At 1 min, pre-

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. §1734 solely to indicate this fact.

Abbreviation: AP, 2-aminopurine.

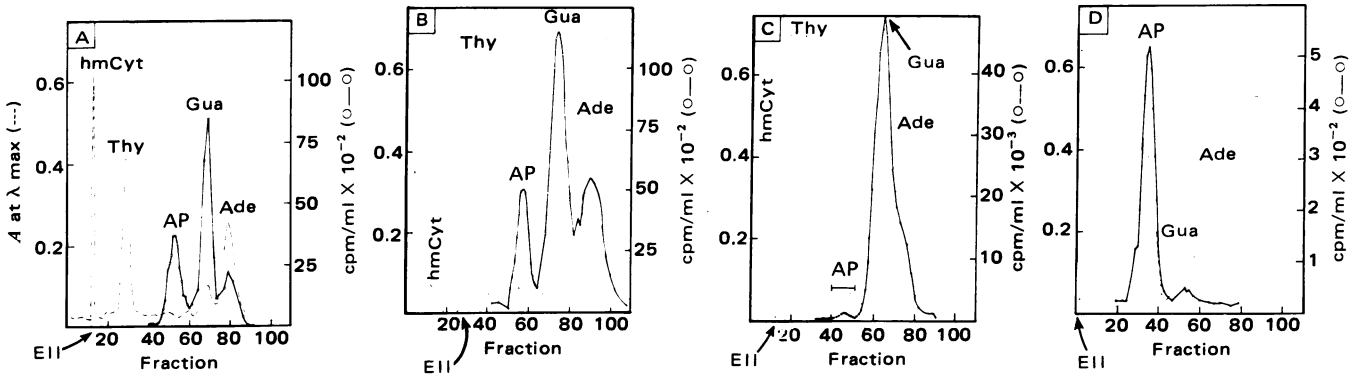


FIG. 1. Separation of free bases obtained by hydrolysis of encapsulated DNA from [^{3}H]AP-mutagenized bacteriophage. (*A*) *UV199-L56* (mutator); (*B*) *UV199-43*$^+$; (*C*) *UV199-L141* (antimutator); (*D*) chromatogram of pooled AP fractions 40–51, inclusive, from *C* with marker AP, Gua, and Ade. EII, elution buffer II (0.33 M ammonium formate, pH 9.1).

warmed growth medium was added such that the final solution contained 5×10^8 bacteria per ml, 2.5×10^9 phage per ml, 7.4 mM AP at $\sim 1.4 \times 10^8$ cpm/μmol, 0.22 mM Ade, and 0.1 mM Gua in M9a 2× glucose. The culture was aerated by shaking at 32°. At 8 min the bacteria were superinfected with a second dose of phage, bringing the total dose to 10 phage per bacterium. The lysis inhibited cells were harvested by centrifugation at 70 min. An aliquot of the supernatant medium was saved for determination of specific activity. The pellet was resuspended in $^1/_{10}$ volume of lysing mix: eggwhite lysozyme (Sigma grade I), 1 g/liter; DNase I (Sigma), 0.1 g/liter; RNase (Sigma grade I), 10 mg/liter, and excess chloroform in phage buffer (18). The lysate was dialyzed against phage buffer to remove exogenous [^{3}H]AP and unencapsulated DNA and RNA. The dialyzed lysate was titrated on permissive (*E. coli* B) and restrictive [CR63 (λ)] hosts to determine the AP-induced reversion frequency of the r_{II}*UV199* marker. The ^{3}H label did not affect the mutagenicity of AP. Bacterial debris was pelleted; phage in the supernatant were then pelleted and resuspended in $^1/_{10}$ volume of phage buffer. The phage were further purified on a CsCl step gradient (19) in which the ^{3}H label migrates with the phage band. DNA was liberated from the capsid by bringing the pooled peak phage fractions to 0.1 M in NaOH. CsCl and NaOH were removed by dialysis against 1 mM EDTA. DNA was taken to dryness and hydrolyzed to free bases by boiling for 1 hr in 70% perchloric acid (20).

Chromatography. The DNA hydrolysate was neutralized with KOH. The $KClO_4$ precipitate was centrifuged away from the free base solution. The hydrolysate was adjusted to pH 12 and loaded onto a 0.6 × 23 cm Dowex-1 (×8) formate column and eluted stepwise at 22 ml/hr; 1.5-ml fractions were collected. hmCyt was eluted with the first buffer (EI), 0.1 M ammonium formate, pH 10.0; buffer EII was 0.33 M ammonium formate, pH 9.1. The absorbance of each fraction was read at its λ max. Molar recoveries of the normal nucleotides were determined spectrophotometrically. Ade, Gua, and Thy were recovered quantitatively in ratios consistent with accepted A·T/G·C ratios for T4; recovery of hmCyt was variable. The absolute amount of AP was determined by measuring the ^{3}H label from an aliquot of each fraction in the peak in 9 ml of an aqueous scintillation fluor (9). AP specific activity was determined by adding a dilution of the growth medium to the same vials that contained the previously assayed AP fractions and then reassaying each sample. The molar amount of AP recovered from the column was computed and compared with the molar amount of Thy recovered to quantify AP·T base pairs.

Table 1. AP misincorporation and r_{II} marker reversion

Mutant	r_{II} reversion frequencies		Mean incorporation, T/AP*
	Spontaneous	AP-induced	
UV199-L56	5×10^{-5}	2×10^{-4} (6)	1515 ± 616 (3)
UV199-43$^+$	5×10^{-7}	4×10^{-5} (5)	1586 ± 113 (2)
UV199-L141	1×10^{-9}	1×10^{-7} (5)	8932 ± 3438 (2)

Reversion of the r_{II} marker *UV199* was measured as the viable titer on the restrictive host CR63 (λ) divided by the titer on the permissive host *E. coli* B. The spontaneous reversion frequencies shown are the means from five stocks grown from single plaques. In parentheses are the number of separate AP mutagenesis runs whose reversion frequencies were averaged.

* Molar ratios of T/AP in encapsulated DNA ± SD; the number of experimental determinations is given in parentheses.

RESULTS

AP Incorporation and Mutagenesis in T4 Mutator, *43*$^+$, and Antimutator Backgrounds. We measured the incorporation of AP into DNA and the rate at which that incorporation gave rise to an A·T → G·C transition (reversion at the marker locus r_{II}*UV199*) in gene 43 allelic phage. A chromatogram of hydrolyzed DNA (containing [^{3}H]AP) from each of the gene 43 alleles showed that the *UV199-L56* (mutator) incorporated 1 AP per 1200 Thy (Fig. 1*A*), *UV199-43*$^+$ 1 AP per 1700 Thy (Fig. 1*B*),‡ and *UV199-L141* (antimutator) 1 AP per 6500 Thy (Fig. 1 *C* and *D*) in their encapsulated DNA. ^{3}H label appearing in the Ade and Gua peaks was due to interconversion of AP, which is tritiated primarily at the 6 position but probably also has label at position 8. In the initial chromatogram of *UV199-L141* hydrolyzed DNA (Fig. 1*C*), it was equivocal whether the counts in the AP peak were due to contaminating radioactivity from the leading edge of the adjacent Gua peak. The fractions from the small AP peak were therefore pooled and rechromatographed on a Dowex formate column with marker Ade, Gua, and AP (Fig. 1*D*). Recovery of AP from the second column was quantitative and showed that greater than 99% of the label in the AP peak of the first column was definitely AP. As an indication of experimental reproducibility, results from several independent determinations of AP incorporation and marker reversion frequencies in each of the gene 43 backgrounds are given in Table 1. However, in the model analyses that follow, we use data from experiments performed at the same time (Fig. 1 *A–D*). Our data show that the mutator:antimutator ratio of AP incorporation *in vivo* is equal to the ratio of AP incorpora-

‡ Gottschling and Freese (21) found 1 AP per 3500 T residues in wild-type T4 grown with one-seventh the amount of AP in a purine-requiring host.

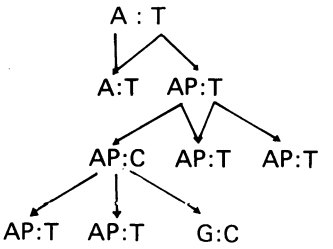


FIG. 2. The evolution of replicating base pairs along an AP-induced A·T → G·C transition mutation pathway.

tion mediated by the respective DNA polymerases *in vitro* (10). If *L56* and *43*$^+$ AP incorporation *in vivo* exactly paralleled the relative incorporation frequencies measured *in vitro;* the 20% difference between *L56* and *43*$^+$ could not be unequivocally resolved in these *in vivo* experiments. However, the *L56*:*L141* ratio of AP incorporation was easily resolved and paralleled the 5-fold difference between *L56* and *L141* AP incorporation observed *in vitro*.

A striking aspect of our results is that AP incorporation (population of AP·T base pairs) varied by a factor of 5 among the three gene 43 alleles, yet the AP-induced reversion (A·T → G·C transition) of the *UV199* marker locus in these backgrounds varied by a factor of 2000. Clearly, these phage discriminate against AP incorporation to different degrees, but they must derive their radical differences in mutability at a step in the pathway subsequent to AP incorporation. These data and the model that follows suggest that the formation of an AP·C base pair is the event at which the phage differ.

Model for AP-Induced A·T ⇄ G·C Transitions. As proposed by Freese (14), AP can pair with Thy; mispairing with C may also occur. These base pairings can lead to A·T → G·C as well as G·C → A·T transitions. Replication errors, causing A·T → G·C transitions, occur as C mispairs opposite a template AP:

$$\frac{A}{T} \rightarrow \frac{AP}{T} \xrightarrow[\text{error}]{\text{replication}} \frac{AP}{C} \rightarrow \frac{G}{C}.$$

Here AP is incorporated on a template strand opposite T. Before the mutation can occur, C must be misincorporated opposite the template AP ("replication error") during a subsequent round of replication. Incorporation errors, causing G·C → A·T transitions, occur because AP is misincorporated opposite a template C:

$$\frac{G}{C} \rightarrow \xrightarrow[\text{error}]{\text{incorporation}} \frac{AP}{C} \rightarrow \frac{AP}{T} \rightarrow \frac{A}{T}.$$

The evolution of A·T → G·C as a function of DNA doublings is shown in Fig. 2 (see also ref. 22). This picture is isomorphic to a four-state system (Fig. 3) in which each state consists of the population at any given time of one of the four base pairs—A·T, AP·T, AP·C, and G·C—present at a given locus during the transition. The state populations evolve in time in discrete steps induced by the variable n (the number of rounds of replication) according to a set of difference equations. Eq. 1 relates the population size of a given base pair to the rate at which it is formed per round of replication:

$$\begin{aligned} a_{1,n+1} &= r_{1\to1}\, a_{1,n} + r_{2\to1}\, a_{2,n} \\ a_{2,n+1} &= r_{1\to2}\, a_{1,n} + r_{2\to2}\, a_{2,n} + r_{3\to2}\, a_{3,n} \\ a_{3,n+1} &= r_{2\to3}\, a_{2,n} + r_{3\to3}\, a_{3,n} + r_{4\to3}\, a_{4,n} \qquad [1] \\ a_{4,n+1} &= r_{3\to4}\, a_{3,n} + r_{4\to4}\, a_{4,n} \end{aligned}$$

The transition rate constants, $r_{i\to i}$ and $r_{i\to j}$, governing intrastate and interstate transitions, respectively, are normalized so that their sum leaving each state equals 2 ($r_{i\to i} + \Sigma r_{i\to j} =$ 2). Because the template bases T, C, and AP are common to several states (Fig. 3), the 10 rate constants are related by 7 independent equations[§] so that Eq. 1 can be written in terms of just three independent variables, $r_{1\to2}$, $r_{2\to3}$, and $r_{4\to3}$:

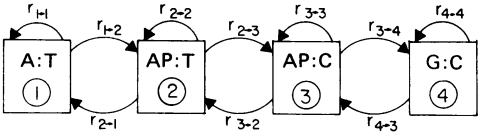


FIG. 3. Closed-loop model for generating base pair populations for AP-induced bidirectional A·T ⇄ G·C transitions. States 1–4 represent the population of A·T, AP·T, AP·C, and G·C base pairs, respectively. Interstate and intrastate transition rates are designated by a set of rate constants $r_{i\to j}$ and $r_{i\to i}$, respectively.

$$\begin{aligned} a_{1,n+1} &= (2 - r_{1\to2})\, a_{1,n} + (1 - r_{1\to2})\, a_{2,n} \\ a_{2,n+1} &= r_{1\to2}\, a_{1,n} + (r_{1\to2} + 1 - r_{2\to3})\, a_{2,n} \\ &\qquad + (1 - r_{2\to3})\, a_{3,n} \qquad [1a] \\ a_{3,n+1} &= r_{2\to3}\, a_{2,n} + (r_{2\to3} + r_{4\to3})\, a_{3,n} + r_{4\to3}\, a_{4,n} \\ a_{4,n+1} &= (1 - r_{4\to3})\, a_{3,n} + (2 - r_{4\to3})\, a_{4,n} \end{aligned}$$

in which $a_{i,n}$ is the population of the ith state (i = 1–4), (e.g., $a_{3,n}$ is the number of AP·C base pairs present after n rounds of replication have occurred). To facilitate analysis, we assume that DNA doublings occur synchronously and that all DNA strands, parentals and daughters, serve as templates in ensuing rounds of replication. We have used the number of phage equivalent units of encapsulated DNA per infected cell to extrapolate the number of rounds of replication; n = 5–7. Mutation frequencies are affected by the number of rounds of replication, the efficiency of packaging DNA into phage capsids, and the kinetics of DNA synthesis (23). Varying these parameters in the model over a physiologically reasonable range does not alter our conclusions significantly.

A·T → G·C Transition Pathway. In relating our data to the model, we have assumed that the preponderance of AP is incorporated opposite T so that the molar ratio AP/T in T4 DNA is equal to the ratio of AP·T/A·T base pairs. Because the r_{II} marker *UV199* reverts via an AP-induced A·T → G·C transition (24), we have assumed that the r_{II} reversion frequency (r^+ revertant/r_{II} phage) after AP-induced mutagenesis is equal to the relative G·C/A·T state populations (state 4/state 1) at the locus. It has also been assumed that the ratio of state 2 (AP·T)/state 1 (A·T) populations at the *UV199* locus is similar to the ratio of AP·T/A·T base pairs averaged over the entire genome. The population of A·T base pairs (state 1) at the *UV199* locus is nearly 1. The experimentally derived values have been substituted into Eq. **1a**. By using the estimates of A·T (state 1), AP·T (state 2), and G·C (state 4) populations at the transition site, the AP·C (state 3) population and the rate of each transition between states have been computed.

Limiting rates for A·T → G·C transitions appear in Table 2. Limiting rates in G·C → A·T transitions appear in Table 3. During the AP-induced reversion of *UV199*, the replication error frequency, $r_{2\to3}$ (substrate C pairing with template AP), at that locus is predicted to be 4 times larger in the mutator background and 100 times smaller in the antimutator background compared to *43*$^+$ (Table 2). With spontaneous reversion rates (Table 1) as an approximate indicator of the frequency of mispairing of C opposite template A, an AP on the template stimulates the rates of C misincorporation roughly 10^4 times

[§] $r_{2\to3} = 1 - r_{3\to2}$, $r_{2\to1} = 1 - r_{1\to2}$, $r_{4\to3} = 1 - r_{3\to4}$, $r_{2\to2} = r_{3\to2} + r_{1\to2}$, $r_{3\to3} = r_{2\to3} + r_{4\to3}$, $r_{1\to1} + r_{1\to2} = 2$, $r_{4\to4} + r_{4\to3} = 2$.

Table 2. Model-based intermediate steps—A·T → G·C pathway

Mutant	Interstate transition frequencies $r_{1\to2}$ (A·T → AP·T)	$r_{2\to3}$ (AP·T → AP·C)	Replication error AP·C / AP·T
UV199-L56	9.5×10^{-4}	2.7×10^{-1}	1.6×10^{-1}
UV199-43$^+$	6.3×10^{-4}	8.7×10^{-2}	4.2×10^{-2}
UV199-L141	1.6×10^{-4}	8.3×10^{-4}	4.1×10^{-4}

Eq. 1a is used to generate computer solutions for base pair populations and for the rate-limiting interstate transition frequencies. In these solutions, it was assumed that there were six rounds of replication. We start with a single A·T base pair at $n = 0$.

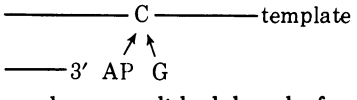
in the *L56* background and 10^6 times in the *43*$^+$ and *L141* backgrounds.

We observe from Table 2 that $r_{2\to3}$ (AP·T → AP·C) is larger than the rate of incorporating AP opposite template T. This would be expected if the AP·T pairing frequency were limited by a lack of available dAPTP in the triphosphate pool or if other gene products contribute to fidelity (e.g., see refs. 15 and 25).

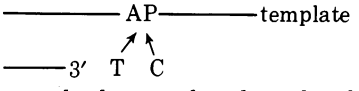
G·C → A·T Transition Pathway. The observation (26, 27) that AP induces A·T → G·C in preference to G·C → A·T transitions suggested that replication errors are more likely to occur than incorporation errors. In Table 3, we have analyzed a hypothetical G·C → A·T transition, assuming, for illustrative purposes, that the G·C → A·T transition frequencies are 4-fold lower than the AP-induced A·T → G·C *UV199* reversion frequencies for the three gene 43 alleles (see Table 2).

The frequency of incorporation errors, $r_{4\to3}$, is predicted to be less than the replication error frequency by a factor of 2800 for mutator, 5500 for *43*$^+$, and 26,000 for antimutator. Thus, the formation of an AP·C bond seems to depend strongly on whether AP serves as a substrate or is situated on the template. These large predicted differences in AP·C pairing frequencies could be due to a combination of two factors.

(*i*) Based on the increased stability of G·C relative to AP·T bonds, it should be easier to incorporate a substrate C opposite a template AP than a substrate AP opposite a template C because, in the latter case, the AP·C bond is formed in competition with a G·C bond.



This is less frequently accomplished than the formation of the AP·C bond in competition with the AP·T bond.



(*ii*) Conversion to the deoxynucleoside triphosphate substrate is less efficient for AP than for the normal nucleotides.¶ Because AP has a far greater probability of pairing with T than with C, most of the available AP will enter the template at potential sites for replication errors. Thus, errors in incorporation (AP opposite C) would be much less likely to occur when dAPTP substrate concentrations are limiting.

DISCUSSION

We have measured the first and last steps in AP-induced A·T → G·C transitions: the incorporation of AP into DNA and the resulting reversion of the r_{II} marker *UV199* in the three T4 backgrounds allelic for gene 43. It was found that the approximately 5:1 ratio of mutator to antimutator AP incorporation *in vivo* is strikingly similar to that mediated *in vitro* by their purified gene 43 products (10). Hence, the presence of a full complement of replicative gene products *in vivo* does not seem to cause significant distortion of the relative fidelities of DNA replication observed *in vitro* for the purified gene 43 products.

¶ R. Hopkins, D. R. Smith, and M. F. Goodman, unpublished data.

Table 3. Model-based intermediate steps—G·C → A·T pathway

Mutant	Hypothetical G·C → A·T* reversion frequencies, A·T / G·C	Interstate transition frequencies, $r_{4\to3}$, (G·C → AP·C)	Incorp. error, AP·C / G·C
UV199-L56	5×10^{-5}	9.1×10^{-5}	5.7×10^{-5}
UV199-43$^+$	1×10^{-5}	1.4×10^{-5}	7.7×10^{-6}
UV199-L141	2.5×10^{-8}	3.3×10^{-8}	1.6×10^{-8}

Eq. 1a is used as in Table 2 and we start with a single G·C base pair at $n = 0$.

* The "hypothetical" data are obtained by assuming that AP-stimulated A·T → G·C transition rates are 4-fold larger than G·C → A·T transition rates.

A simple model has been used to translate the AP·T state population (molar ratio AP/T in DNA) and the G·C state population (reversion frequency of the r_{II}*UV199* marker) into rates per round of replication of each intermediate step in the AP-induced A·T → G·C transition pathway (Fig. 3). In Table 2 it is shown that the replication error (misincorporation of C opposite a template AP) is, as expected, the source of the great differences in AP-induced transition frequencies among the three gene 43 alleles. These values were calculated (Eq. 1a) by assuming that the measured AP incorporation is representative of the *UV199* locus.

The predicted replication error frequency ($r_{2\to3}$) of 27% in mutator seems extremely high. However, preliminary measurements of AP·C populations in *43*$^+$ and *L56* backgrounds based on heterozygotes at the *UV199* locus captured in single-burst experiments substantiate the predicted populations.¶ Koch (28) showed that C is incorporated opposite a template AP 6–20 times more frequently if the base pair immediately preceding it was G·C than if it were A·T. Our prediction that $r_{2\to3}$ is very large for the marker *UV199* may suggest that this locus resides adjacent to a G·C base pair. Perhaps a marker adjacent to an A·T base pair would show an $r_{2\to3}$ 6–20 times smaller than that extrapolated for *UV199*. Even so, $r_{2\to3}$ would still be very high. Thus, the presence of an AP on the template stimulates misincorporation of C at that locus by several orders of magnitude irrespective of nearest neighbor effects.

In the analysis of the G·C → A·T transition (Table 3), it was assumed that AP induces 4-fold higher A·T → G·C than G·C → A·T transition frequencies. This 4-fold difference is a conservative estimate (26, 27). Yet, even if G·C → A·T transition frequencies were equal to A·T → G·C transition frequencies, the basic conclusion remains that the rate of formation of an AP·C base pair is at least 3 orders of magnitude greater when AP is situated on the template than when C is on the template. It should be feasible to test these model predictions in an *in vitro* DNA-synthesizing system by using suitably constructed polynucleotides containing AP or C in the template strand.

An analysis of *in vivo* and *in vitro* rates of AP incorporation enables one to propose a mechanism for *L56* and *L141* DNA polymerase error correction and to offer a quantitative speculation concerning the effects of other replicative gene products on the absolute level of mispaired nucleotides in DNA. A cur-

rent hypothesis (16, 17, 29) put forth to explain mutation rates *in vivo* imposes a two-step selection against improperly paired nucleotides. In the first step, an incorrect nucleotide is inserted at a frequency, f, not less than $f_0 = \exp(-\Delta G/RT)$, in which δG is the free energy difference between matched and mismatched bases at the 3'-OH growing point. Observed mutation frequencies are on the order of $(f_0)^2$ (29). The second step—e.g., 3'-exonucleolytic proofreading—operating with a precision f_0 would result in an overall error rate $\sim(f_0)^2$. How do the different values of AP incorporation measured *in vivo* and *in vitro* for the gene 43 alleles compare with the above predictions? From an analysis of *in vitro* frequencies of incorporation of AP (10), Galas and Branscomb (16) estimated a value of 1.3 kcal/mol for the free energy difference between AP·T and A·T base pairs. An enzyme having only a single opportunity to discriminate against AP, when dAPTP and dATP concentrations are not rate-limiting and are equal (10), could incorporate AP at a frequency $\geq\exp(-1.31/0.6) = 11.5\%$. Given a second opportunity for discrimination, the minimum AP incorporation rate would be reduced to 1.3%. Thus, the 10% AP incorporation frequency for *L56* (10) implies that the mutator polymerase is behaving, for the most part, as a one-step discriminator. In contrast, the antimutator polymerase *L141* (2.9% AP incorporation) closely resembles an ideal two-step discriminator. Because these polymerases insert AP in the initial discrimination step at similar frequencies, ~12%, the net differences in error rates are attributable to large differences in the polymerase-associated 3'-exonuclease activities as originally shown by Muzyczka *et al.* (9). The *UV199* spontaneous reversion frequency is also consistent with a one-step error correction process in the mutator and an approximately two-step error correction process in the antimutator.

Can the proofreading properties of the DNA polymerases account for the absolute magnitudes of AP incorporation *in vivo* or are other replication proteins such as genes 32, 41, 62, 44, and 45 (15) also required to maintain genetic integrity as suggested by the observations that gene 32 mutants can increase spontaneous mutation rates (25)? This question can be addressed, in part, by considering the relative concentrations of dAPTP and dATP in the acid-soluble nucleotide pool. Setting f_0(dAPTP/dATP) and f_0^2(dAPTP/dATP) approximately equal to the observed AP·A incorporation ratios for *L56* and *L141*, respectively, we find that the dAPTP pool size must be no greater than $^1/_{100}$–$^1/_{200}$ of the dATP pool size in order for the polymerase, acting independently of other gene products, to account for the observed AP incorporation frequencies. Assuming that the *UV199* locus is not a "hot spot" for AP incorporation, the observation that the frequency of replication errors (AP·T → AP·C) exceeds the rate of incorporating AP opposite T (A·T → AP·T; Table 2) would strongly imply that dAPTP pool sizes were rate limiting in the *in vivo* systems, suggesting in turn that DNA polymerases, by themselves, might account for most of the exclusion of single nucleotide mismatches in DNA.

Note Added in Proof. Preliminary data suggest that the dAPTP pool size in *43*$^+$ infected cells during mutagenesis is less than 4% of the dATP pool size.¶

We are grateful to Drs. John Petruska, Bruce Alberts, and David Galas for valuable insights into all aspects of this work, to David Smith for computing assistance, and to Edith Dean for preparing the manuscript. This work was supported by National Institutes of Health Grants GM21422 and CA17358 to M.F.G.

1. Speyer, J. F. (1965) *Biochem. Biophys. Res. Commun.*, **21**, 6–8.
2. Speyer, J. F., Karam, J. D. & Lenny, A. B. (1966) *Cold Spring Harbor Symp. Quant. Biol.* **31**, 693–697.
3. Speyer, J. F. & Rosenberg, D. (1968) *Cold Spring Harbor Symp. Quant. Biol.* **33**, 345–350.
4. Drake, J. W. & Allen, E. F. (1968) *Cold Spring Harbor Symp. Quant. Biol.* **33**, 339–344.
5. Drake, J. W., Allen, E. F., Forsberg, S. A., Preparata, R. M. & Greening, E. O. (1969) *Nature* **221**, 1128–1131.
6. Allen, E. F., Albrecht, I. & Drake, J. W. (1970) *Genetics* **65**, 187–200.
7. Freese, E. B. & Freese, E. F. (1967) *Proc. Natl. Acad. Sci. USA* **57**, 650–657.
8. DeWaard, A., Paul, A. V. & Lehman, I. R. (1965) *Proc. Natl. Acad. Sci. USA* **54**, 1241–1248.
9. Muzyczka, N., Poland, R. L. & Bessman, M. J. (1972) *J. Biol. Chem.* **247**, 7116–7122.
10. Bessman, M. J., Muzyczka, N., Goodman, M. F. & Schnaar, R. L. (1974) *J. Mol. Biol.* **88**, 409–421.
11. Brutlag, D. & Kornberg, A. (1972) *J. Biol. Chem.* **247**, 241–248.
12. Hershfield, M. S. (1973) *J. Biol. Chem.* **248**, 1417–1423.
13. Nossal, N. G. & Hershfield, M. S. (1973) in *DNA Synthesis In Vitro*, eds. Wells, R. D. & Inman, R. B. (University Park Press, Baltimore, MD), pp. 47–62.
14. Freese, E. (1959) *J. Mol. Biol.* **1**, 87–105.
15. Alberts, B., Morris, C. F., Mace, D., Sinha, N., Bittner, M. & Moran, L. (1975) "Systems for replication of DNA *in vitro*," *ICN-UCLA Symposium on Molecular and Cellular Biology*, eds. Goulian, M., Hanawalt, P. & Fox, C. F. (W. A. Benjamin, Inc., Reading, MA), pp. 241–269.
16. Galas, D. & Branscomb, E. (1977) *J. Mol. Biol.*, in press.
17. Hopfield, J. J. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 4135–4139.
18. Clowes, R. C. & Hayes, W. (1968) *Experiments in Microbial Genetics* (Blackwell Scientific Publications, Oxford), pp. 187–188.
19. Thomas, C. A., Jr. & Abelson, J. (1966) in *Procedures in Nucleic Acid Research*, eds. Cantoni, G. & Davis, D. (Harper and Row, New York), pp. 553–561.
20. Bendich, A. (1957) *Methods in Enzymology*, eds. Colowick, S. P. & Kaplan, N. O. (Academic Press, Inc., New York), Vol. III, pp. 715–723.
21. Gottschling, H. & Freese, E. (1961) *Z. Naturforsch. Teil B* **16**, 515–519.
22. Rudner, R. (1960) *Biochem. Biophys. Res. Commun.* **3**, 275–280.
23. Drake, J. W. (1970) *The Molecular Basis of Mutation* (Holden-Day, San Francisco, CA), p. 44.
24. Champe, S. P. & Benzer, S. (1962) *Proc. Natl. Acad. Sci. USA* **48**, 532–546.
25. Koch, R. E., McGraw, M. K. & Drake, J. W. (1976) *J. Virol.* **19**, 490–494.
26. Freese, E. B. (1968) *Mutat. Res.* **5**, 299–301.
27. Osborn, S., Phillips, S. & Funk, F. (1967) *J. Mol. Biol.* **26**, 437–447.
28. Koch, R. E. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 773–776.
29. Topal, M. D. & Fresco, J. R. (1976) *Nature* **263**, 285–289.