The Effects of dNTP Pool Imbalances on Frameshift Fidelity during DNA Replication*

(Received for publication, July 30, 1991)

Katarzyna Bebenek, John D. Roberts, and Thomas A. Kunkel†

From the Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

The use of unequal concentrations of the four deoxynucleoside triphosphates (dNTPs) in DNA polymerization reactions alters base substitution error rates in a predictable way. Less is known about the effects of substrate imbalances on base addition and deletion error rates. Thus, we examined pool bias effects on frameshift fidelity during DNA synthesis catalyzed by replicative DNA polymerases. Imbalanced pools altered the frameshift fidelity of the human immunodeficiency virus type-1 reverse transcriptase. Both mutagenic and antimutagenic effects were observed for minus-one, plus-one, and minus-two nucleotide errors, in a highly sequence-specific manner. Most of this specificity can be rationalized by either of two models. One involves frameshifts initiated by pool bias-induced nucleotide misinsertion, and the other involves pool bias-initiated template-primer slippage. Several examples of complex mutations were also recovered more than once in small mutant collections. These contained closely spaced single-base substitution and minus-one base frameshift changes. The two changes occurred at a frequency much higher than predicted if they were generated independently. This suggests that when the polymerase makes one mistake, the probability that it will make a second mistake within the next few incorporations increases significantly. Perturbation of dNTP pools also affected the frameshift fidelity of the replicative yeast DNA polymerase α. In reactions containing a low concentration of one dNTP, the error rate increased for one-nucleotide deletions at homopolymeric template nucleotides complementary to the dNTP whose concentration was low. We extended this approach to determine the frameshift fidelity of simian virus 40 origin-dependent semiconservative replication of double-stranded DNA in extracts of human cells. In reactions performed with an equal concentration of all four dNTPs, replication was highly accurate for minus-one-nucleotide errors. However, when the concentration of one dNTP was decreased, the replication error rate increased at complementary, homopolymeric template positions. This response provides an approach for describing frameshift accuracy during replication of the leading and lagging strands.

Many studies have demonstrated that the concentrations of the four deoxynucleoside triphosphates (dNTPs) in mammalian cells are not equal but can vary by more than 10-fold (for example, see Cohen et al., 1981; Bestwick et al., 1982; Leeds et al., 1985; Arecco et al., 1988). Furthermore, in both prokaryotes and eukaryotes, exposing cells to exogenous nucleosides or modifying pathways for nucleotide metabolism by genetic defects or drug treatments leads to perturbations in the concentrations of the dNTPs (Kunz, 1982). It is well documented that such perturbations alter mutation rates (Kunz, 1982; Meuth, 1984; Phear et al., 1987; Sargent and Mathews, 1987; Kunz, 1988; Sargent et al., 1989). One explanation for this comes from studies of the fidelity of DNA synthesis catalyzed by purified DNA polymerases in vitro. At least for base substitution mutations, providing unequal concentrations of the dNTP substrates in a polymerization reaction alters error rates (for review, seeechols and Goodman, 1991). Since the insertion probability reflects the ratio of the correct versus incorrect nucleotide for any given template position, dNTP substrate imbalances can be either mutagenic or antimutagenic.

Studies of polymerization fidelity have also demonstrated that, in addition to single base substitution errors, DNA polymerases may delete or add one or more nucleotides (for review, see Kunkel, 1990). Compared with studies of base substitution fidelity, less is known about the effects of dNTP pool imbalances on frameshift error rates. Nevertheless, with the advent of models for production of frameshift errors (for review, see Kunkel, 1990 and Ripley, 1990), there are reasons to expect that substrate pool imbalances could influence frameshift error rates. For example, one model (Kunkel and Soni, 1988a) posits that frameshift errors might be initiated by misinsertion of a nucleotide that, if complementary to a nearby template nucleotide, could misalign to yield a frameshift error. The specificity of pool imbalance-dependent minus-one nucleotide errors by the exonuclease-deficient large fragment of Escherichia coli DNA polymerase I is consistent with this model (Bebenek and Kunkel, 1990).

In the present study we analyzed the frameshift fidelity of polymerization reactions containing unequal concentrations of dNTPs. Our continuing interest in examining replication fidelity and models for frameshift errors, combined with the observation that polymerization by the human immunodeficiency virus (HIV-1) reverse transcriptase is error prone for frameshifts (Bebenek et al., 1989), led us to first examine the effects of substrate pool imbalances with this enzyme. We wanted to know if pool biases affected the frameshift fidelity of this replicative polymerase, and if so, if the pattern of pool bias-induced errors supported the model that nucleotide misinsertions initiate some frameshifts. We are also interested in

*This work was supported in part by the National Institutes of Health Intramural AIDS Targeted Antiviral Program. 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.

† To whom correspondence should be sent. Tel.: 919-541-2644; Fax: 919-541-7585.

The abbreviations used are: HIV-1, type 1 human immunodeficiency virus; pol α, polymerase alpha; SV40, simian virus 40; Hepes, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid.
the fidelity of semiconservative DNA replication in eukaryotes (Roberts and Kunkel, 1988; Roberts et al., 1991). Therefore, we extended these studies to examine pool bias effects on the frameshift error rate of the eukaryotic replicative DNA polymerase α and simian virus 40 (SV40) origin-dependent replication in extracts of human HeLa cells. We present evidence here that substrate pool imbalances alter frameshift fidelity in all three polymerizing systems.

EXPERIMENTAL PROCEDURES

Materials—Mutant derivatives of bacteriophage M13mp2 were obtained as follows: (a) deletion of two Ts from TTTT at positions 70–73 (Bebenek and Kunkel, 1990); (b) deletion of a C from the CC run at positions 64–65, together with a silent G → C transition at position 69, constructed by oligonucleotide-directed mutagenesis as described in Kunkel et al. (1987); (c) addition of a T to a TT run at position 103, constructed by oligonucleotide-directed mutagenesis (Kunkel et al., 1987); and (d) addition of a T to a TTTT run at positions 70–73 (from Kunkel et al., 1989). The latter mutant was used to construct (Thomas et al., 1990) a vector containing the SV40 origin of replication. Yeast DNA pol α and pol δ, the recombinant form of HIV-1 reverse transcriptase were from previously described sources (Kunkel et al., 1988; Roberts et al., 1988). HeLa cell extracts were prepared as described (Roberts and Kunkel, 1988).

Frameshift Fidelity Assays—Four reversion assays were employed to measure the frameshift fidelity, each allowing the detection of a different set of frameshift errors. For the purified polymerases, the assays use an M13mp2 DNA substrate with a 361-nucleotide single-stranded gap. The single-stranded template of each substrate contains the appropriate change (described below) was DNA containing the appropriate change (described below). These numbers have been adjusted for the addition or deletion of nucleotides to allow consistency when comparing sequences. The selectable errors for each assay are given on the right. Runs of two or more bases are underlined. The four DNA substrates were generated from the wild-type lacZα sequence as follows. In Assay A two T residues were deleted at the TTTT run at positions 70–73. In Assay B a C residue was deleted from the TTT run at positions 64–65, and a silent G → C transition was made position 69. In Assay C a T residue was added to the TT run at positions 73–74, and the SV40 origin of replication was inserted into the M13mp2 outside of the mutational target sequence as described (Thomas et al., 1990). Sites at which the changes were made are marked with an asterisk.

![Fig. 1. Target sequences for frameshift reversion assays. The viral (+) strand DNA sequence is shown. Synthesis of the complementary strand proceeds from right to left across these targets. The numbers represent the wild-type numbering system in lacZα, in which position +1 is the first transcribed nucleotide. These numbers have not been adjusted for the addition or deletion of nucleotides to allow consistency when comparing sequences. The selectable errors for each assay are given on the right. Runs of two or more bases are underlined. The four DNA substrates were generated from the wild-type lacZα sequence as follows. In Assay A two T residues were deleted at the TTTT run at positions 70–73. In Assay B a C residue was deleted from the TTT run at positions 64–65, and a silent G → C transition was made position 69. In Assay C a T residue was added to the TT run at positions 73–74, and the SV40 origin of replication was inserted into the M13mp2 outside of the mutational target sequence as described (Thomas et al., 1990). Sites at which the changes were made are marked with an asterisk.](image-url)
frequencies that are more than 100-fold higher than the background frequency for unoccupied DNA. Although the reversal frequency is similar for reactions performed with 20 versus 1,000 μM concentrations of each dNTP the frequency of revertants increased severalfold in the presence of a 50-fold excess of dGTP, dATP, and dCTP, respectively. An excess of dTTP did not cause a substantial change in the reversal frequency.

**Specificity of Frameshift Errors with HIV-1 Reverse Transcriptase**—For base substitution errors, dNTP pool imbalances can be mutagenic at some template positions and antimutagenic at others. To determine if this was so for frameshift errors, we examined the DNA sequence of independent revertants obtained under the different reaction conditions in Assay A (Fig. 1 and Table I). The DNA sequence of 58, 21, and 22 revertants was determined from reactions with equal dNTPs, excess dGTP, excess dATP, and excess dTTP, respectively. In all cases, minus-one nucleotide errors were detected. No revertants containing minus-four, plus-two or other kinds of frameshift errors were recovered.

Some revertants contained one or more base substitution errors accompanying the frameshift error. In almost all cases the base substitutions were consistent with misincorporation of the nucleotide provided in excess, demonstrating the effectiveness of the pool imbalance. In some instances, the frameshift and base substitution changes were separated by many nucleotides, suggesting that they were generated independently by the reverse transcriptase. However, some revertants contained two (or even three) mutations that were closely spaced (Table II). Five different revertants were recovered independently more than once although the total number of revertants analyzed was small. The closely spaced changes occurred much more frequently in combination with each other than alone or in combination with any other change (Table II). Since this suggests that these complex mutations may have resulted from a specialized processing (see "Discussion"), they were not included in the quantitative analysis presented next.

The distribution of minus-one base errors for each reaction condition is shown in Fig. 2. At equimolar substrate concentrations, the minus-one frameshifts are distributed throughout the target. The frequency of deletions at certain sites (e.g. A at position 76–77) is much higher than at others. Biasing the dNTP concentrations in the reaction significantly changes the distribution of errors. With excess dGTP, all the errors are loss of a template A at site 76–77. From the frequency (Table I) and specificity (Fig. 2) data, the calculated increase in error rate for this base loss, resulting from a 50-fold imbalance, is 18-fold. With excess dATP, the frameshift mutations are more randomly distributed. With excess dTTP,

**TABLE I**

| dNTP pool                     | Plaques (×10^5) | Revertant frequency (×10^-3) | Relative frequency |
|-------------------------------|----------------|-----------------------------|-------------------|
| Equimolar (20 μM)             | 220,000        | 262                         | 120               |
| Equimolar (1,000 μM)          | 300,000        | 454                         | 150               |
| Excess dGTP                   | 83,000         | 795                         | 960               |
| Excess dATP                   | 100,000        | 437                         | 440               |
| Excess dCTP                   | 960,000        | 2,726                       | 280               |
| Excess dTTP                   | 700,000        | 776                         | 110               |

**TABLE II**

| Condition          | Mutational changes | Total occurrences | Mutant frequency (×10^-3) | Observed | Predicted |
|--------------------|--------------------|-------------------|---------------------------|----------|-----------|
| Excess dTTP        | A                  | 5'-CTGGCCGT-3'    | 5/22                      | 25       | ±0.5      |
| Excess dATP        | A                  | 5'-ACAAGTCG-3'    | 4/21                      | 84       | ±5        |
| Excess dTTP        | A                  | 5'-ACAAGTCG-3'    | 2/22                      | 10       | 1.6       |
| Excess dATP        | A                  | 5'-ACAAGTTC-3'    | 3/22                      | 15       | ±0.01     |
| Excess dTTP        | A                  | 5'-CTGGCCGT-3'    | 3/21                      | 63       | ±3        |

**FIG. 2.** Spectra of one-base deletion errors by HIV-1 reverse transcriptase. Frame shift Assay A was used to detect primarily minus-one base frameshifts at nonreiterated template positions. The nucleotides shown are from position 46 to 85 in the lacZα sequence. Each open triangle represents an independent one-base deletion. Triangles below underlined nucleotides (two-base runs) indicate that one of two possible nucleotides was deleted. Of the 58 revertants sequenced, 48 were obtained from the 1,000 μM reaction and 10 from the 20 μM reaction. Although the data sets are small, there is no indication that the error specificity is different at these two dNTP concentrations. For example, 22 of 58 revertants from the 1,000 μM reaction were the loss of an A at positions 76–77 (mutant frequency, 57 × 10^-3) whereas this same mutant comprised 4 of 10 revertants from the 20 μM condition (mutant frequency, 48 × 10^-3).
position 76-77 (Fig. 2). The reversion frequency at this site decreased by 11-fold.

To facilitate discussions of models (see "Discussion"), the data are expressed as frequencies in Table III, with respect to the template base lost and its 5' neighbor. For example, a high concentration of dTTP is 14-fold antimitogenic for the loss of a template A (Table III, part A) whereas high dGTP is 11-fold mutagenic for the loss of template nucleotides whose 5' neighbor is a template C (Table III, part B).

**The Specificity of Minus-two and Plus-one Base Errors by HIV-1 Reverse Transcriptase**—Having shown that dNTP pool imbalance can be either mutagenic or antimitogenic for minus-one base errors, we next asked if similar effects could be detected for plus-one and/or minus-two base errors. For this purpose, we used a second frameshift assay (Assay B, see "Experimental Procedures") that scores plus-one and minus-two base errors at more than 50 template nucleotides. Again, polymerization by HIV-1 reverse transcriptase with equimolar dNTPs is error prone, generating a reversion frequency 1,000-fold above the background frequency for uncopied DNA (Table IV).

Imbalancing the dGTP or dATP concentration by 20-fold changed the reversion frequency less than 2-fold (Table IV). Despite these small effects on reversion frequencies, we were encouraged by the substantial pool imbalance-induced differences in the specificity of minus-one errors (above). Thus, we examined the specificity of plus-one and minus-two base errors by sequencing a small collection of revertants from each reaction condition in Assay B. The error distributions are shown in Fig. 3 and expressed as frequencies for various classes of errors in Table IV. Only minus-two or plus-one errors were recovered. At equal substrate concentrations, the proportion of minus-two base (open triangles) and plus-one base errors (closed triangles) is almost the same (Fig. 3), and the mutations are nonrandomly distributed (Fig. 3). Consistent with previous results from a forward mutation assay (Bebenek et al., 1989), at equal dNTP pools the majority of frameshifts occurs at homopolymeric runs, and the TTTT run at positions 70-73 is a mutational hot spot.

The specificity analysis revealed that substrate imbalances substantially altered error rates (Table IV). Both mutagenic and antimitogenic effects were observed. For example, with excess dGTP, a 10-fold increase in the frequency of additions of a C at positions 68-69 and 95-97 was observed, and the frequency of deletion of TT from a TTTT run at positions 70-73 increased ~3-fold. In contrast, the frequency of the latter error decreases by at least 5-fold with excess dATP. Possible explanations for these effects are discussed below.

**Pool Bias Effects on Minus-One Frameshift Errors by Yeast DNA Polymerase α**—One mechanism to explain frameshift mutations, first proposed by Streisinger (Streisinger et al., 1966; Streisinger and Owen, 1985), suggests that slippage of the two DNA strands can lead to a misaligned intermediate that ultimately is processed into a frameshift mutation. This idea agrees with observations that frameshifts both in vitro and in vivo occur more frequently in reiterated sequences. Furthermore, it has been demonstrated that the frequency of minus-one frameshift errors by eukaryotic replicative DNA polymerase α increases as the length of the homopolymeric run increases (Kunkel and Soni, 1988b; Kunkel, 1990). This correlation may be because the longer the run, the larger the number of potential misaligned intermediates and the larger the number of correct base pairs that could stabilize the intermediates.

We were interested in whether imbalanced dNTP pools affect reiterated sequence frameshift error rates by a normal cellular replicative DNA polymerase. We examined this with yeast DNA pol α, using a reversion assay that specifically

### Table III
**Specificity of one-base deletions by HIV-1 reverse transcriptase**

| Minus-one error | Revertant frequency ($\times 10^{-5}$) with |
|-----------------|------------------------------------------|
| Equal pools     | High dGTP | High dATP | High dTTP |
| A. Loss of a template | 70 | 960 | 147 | 5 |
| G               | 31 | <46 | 42 | <5 |
| T               | 19 | <46 | <21 | 10 |
| C               | 12 | <46 | 63 | 35 |
| B. Loss of a template nucleotide whose 5' neighbor is a template | |
| A               | 10 | <46 | <21 | 35 |
| G               | 17 | <46 | 84 | 5 |
| T               | 17 | <46 | 63 | <5 |
| C               | 89 | 960 | 126 | 10 |

### Table IV
**Specificity of plus-one and minus-two nucleotide errors by HIV-1 reverse transcriptase**

Assay B was used with HIV-1 reverse transcriptase as described under "Experimental Procedures." The concentration of dNTPs in DNA synthesis reactions with equal dNTP pools was 1,000 μM. Pool bias reactions contained 1,000 μM dNTP in excess and each of the three remaining nucleotides at 50 μM. The individual site reversion frequencies were calculated based on total reversion frequencies and the number of mutants at each site (Fig. 2). Forty-four revertants from the equal dNTP pools reaction, 16 revertants from reaction with excess dGTP, and 17 from reaction with excess dATP are included in the analysis. The reversion frequency of uncopied DNA was 0.3 $\times 10^{-5}$.

| Error | Position | Revertant frequency ($\times 10^{-5}$) with |
|-------|----------|------------------------------------------|
|       | dNTPs   | Equal | Excess dGTP | Excess dATP |
| All errors |          | 360 | 520 | 320 |
| Plus T | TTTT at 70-73 | 130 | ±27 | 160 |
| Minus TT | TTTT at 70-73 | 82 | 260 $\pm 18$ | |
| Plus C | CCC at 95-97 | 8.2 | 82 $\pm 18$ | |
| Plus C | CC at 68-69 | 8.2 | 82 $\pm 18$ | |
| Plus T | T at 98 | 8.2 | ±27 | 32 |

![Fig. 3. Spectra of two-base deletion and one-base addition errors by HIV-1 reverse transcriptase in Assay B. Underlined are all iterated and reiterated nucleotides in the target. Closed triangles above the line represent independent one-base additions, and open triangles below the line represent independent two-base deletions.](image-url)
detects frameshifts in runs of As, Cs, Ts, and Gs (Assay C in Fig. 1 and under "Experimental Procedures"). Reactions were performed with equal dNTP pools and under biased substrate conditions wherein the concentration of a single dNTP was decreased by 25-fold. Under all conditions, yeast pol α produced frameshift errors at a frequency that was substantially higher than the background frequency (Table V). Mutagenic responses were observed for low dTTP (13-fold) and for low dGTP (6-fold) whereas lowering the concentration of dCTP or dATP had little apparent effect on reversion frequencies.

To examine site-specific error rates, the DNAs of 97 revertants from the equal dNTP pool reaction and of 20 revertants from each dNTP biased reaction were analyzed by sequencing. This analysis revealed that all revertants contained minus-one base errors. Their distribution is presented in Fig. 4, with frequencies given in Table VI for loss of template A, C, G, and T nucleotides.

With equal pools, the mutations are nonrandomly distributed throughout the target. The frameshift error rate, expressed per nucleotide polymerized to correct for the number of reiterated and nonreiterated nucleotides, is ~6-fold higher at the four- and three-base runs than at the two-base run and nonrun sequences.

TABLE V

Pool bias effects on minus-one nucleotide errors with yeast DNA polymerase α

Assay C was used as described under "Experimental Procedures." The equimolar dNTP concentration was 1,000 μM; the low dNTP concentration was 40 μM, yielding a 25-fold pool bias. The revertant frequency of uncopied DNA was 1.1 × 10⁻⁷.

| dNTP pool   | Total Plaques | Blue Revertant frequency (× 10⁻⁴) | Relative frequency |
|-------------|---------------|-----------------------------------|--------------------|
| Equimolar   | 270,000       | 97                                | 36                 | 1.0                 |
| Low dTTP    | 53,000        | 253                               | 480                | 13                  |
| Low dGTP    | 65,000        | 142                               | 220                | 6.1                 |
| Low dCTP    | 53,000        | 25                                | 47                 | 1.3                 |
| Low dATP    | 86,000        | 31                                | 36                 | 1.0                 |

Perturbation of the dNTP pools changed the error specificity. For all four biased reaction conditions, lowering the concentration of one nucleotide increased the minus-one base deletion frequency at template positions complementary to the nucleotide which concentration has been lowered. For example, when the concentration of dTTP was lowered 25-fold, only minus-A errors at the AAAA run were observed (Fig. 4), representing a 37-fold increase in minus-one base frameshifts at this site. Reducing the concentration of dGTP resulted in a 22-fold increase in the frequency of deletions for template Cs from the two CCC runs. When the concentration of dCTP was reduced, a 12-fold increase in minus-one events occurred at template G sites. Finally, lowering the concentration of dATP led to a 3-fold increase in one-base deletions at the TTT run.

At the same time, there are sites at which the biased substrate conditions have a clearly antimutagenic effect. For example, when the concentration of dCTP is lowered, the frequency of minus-one frameshifts at the A-run and the T-run decreases by at least 5- and 4-fold, respectively. A common feature of these two homopolymeric runs is that both sequences are flanked on the 5' side by a template G, where dCTP is the correct incoming nucleotide.

Imbalanced dNTP Pools Affect Frameshift Fidelity during Semiconservative DNA Replication—We (Roberts and Kunkel, 1988) and others (Hauser et al., 1988) have shown previously that the fidelity of semiconservative DNA replication in extracts of HeLa cells is high for single-base substitutions. More recent measurements (Thomas et al., 1991) suggest that this is also the case for minus-one base framshifts. Since high frameshift accuracy may reflect the contribution of additional components of the replication complex to fidelity, it is of interest to know if a pool imbalance affects frameshift fidelity during replication. To examine this, a reversion assay was used which scores minus-one base errors in a run of five A/T base pairs (Assay D, Fig. 1, and under "Experimental Procedures").

Replication reactions in the HeLa cell extracts using the frameshift reversion substrate DNA (M13mp8SV+T70 DNA) were performed with equimolar (20 μM) dNTPs and with a 50-fold pool imbalance, in which the concentration of either dATP or dTTP was kept at 20 μM and the concentrations of the other three dNTPs were increased to 1000 μM. Based on the results with yeast pol α, we expected that the relatively low concentration of the nucleotide complementary to the bases in the homopolymeric run should increase the frequency of frameshifts within the run.

At equal dNTP pools the revertant frequency of the replicated DNA was only slightly above the background revertant
frequency of uncopied DNA (Table VII). This suggests that the replication complex is highly accurate for minus-base errors. A similar result was obtained at 100 μM equimolar dNTPs (not shown). However, when the dATP concentration was 50-fold lower than the other three dNTPs, the frequency of minus-one errors increased by 8-fold. With a 50-fold lower relative concentration of dTTP, the increase in revertant frequency was 18-fold. Of the 100 revertants recovered from the pool bias reactions, 96 produced dark blue plaques. This plaque phenotype was shown previously (Bebenek et al., 1990) to result from deletion of one of the five A:T base pairs in the homopolymeric run to yield the true wild-type sequence.

**DISCUSSION**

The results presented here demonstrate that the frameshift fidelity of DNA synthesis is affected by perturbations in the relative concentrations of the deoxynucleoside triphosphates present in a polymerization reaction. Effects are observed for synthesis by two replicative DNA polymerases and for semi-conservative DNA replication in human cell extracts. The effects can be either mutagenic or antimutagenic, depending on which dNTP is varied and on the template nucleotide sequence. Pool bias effects are not restricted to the most frequently produced minus-one errors; error rates for minus-two, plus-one, and complex changes are also altered.

Substrate pool imbalances have long been known to alter base substitution fidelity in a predictable way (for review, see Echols and Goodman, 1991). Careful examination of the sequence specificity of frameshift errors generated by HIV-1 reverse transcriptase under biased nucleotide reaction conditions (Figs. 2 and 3 and Tables III and IV) also reveals predictable patterns. In reactions performed with an excess of one nucleotide (a) the frequency of one- and two-base deletions decreases at template sites that are complementary to the nucleotide provided in excess; (b) the frequency of one-base deletions increases at sites where the nucleotide provided in excess is complementary to the template nucleotide; (c) the frequency of two-base deletions increases at sites where the nucleotide provided in excess is complementary to a adjacent template position; (d) the frequency of one-base additions decreases at sites where the nucleotide provided in excess is complementary to the next template nucleotide; and (e) the frequency of one-base additions increases at sites where the nucleotide provided in excess is complementary to the template nucleotides in which the addition occurs.

Each of these specificities is consistent with our previous suggestion (Kunkel and Soni, 1988b) that frameshift errors might be initiated by nucleotide misinsertions. In this model, misinsertion creates a mispaired template-primer that is more difficult to extend than is a correctly paired substrate. At a position where the misinserted nucleotide is complementary to a nearby template nucleotide, a realignment could occur to generate a misaligned frameshift intermediate with one or more correct terminal base pairs. Mutagenic and antimutagenic examples consistent with this model are shown in Figs. 5, A, B, and C, for minus-one, minus-two, and plus-one frameshifts, respectively. Each example represents an observation of the effects of pool imbalances shown in Figs. 2 and 3 and Tables III and IV.

The first example (Fig. 5A) displays a model for the excess dGTP-mediated 18-fold increase in the frequency of deletions of template A at the HIV-1 reverse transcriptase hot spot for minus-one base errors at positions 76–77 (Fig. 2 and Table III). This scheme is consistent with the hypothesis that misinsertion of dGTP opposite template A at position 76, followed by a realignment to form a correct C:G base pair at the adjacent template position, leads to the frameshift mutations at this site. The observation that excess dTTP in the reaction

**TABLE VII**

| Pool Concentration | Plaques Total | Blue | Revertant frequency (× 10^-5) |
|--------------------|---------------|------|-----------------------------|
| Equimolar dNTPs    | 1,700,000     | 42   | 2.5                         |
| Low dATP           | 270,000       | 54   | 20.0                        |
| Low dTTP           | 100,000       | 46   | 46.0                        |

**Fig. 5. Models for misinsertion-initiated frameshift errors.**
mixture is 11-fold antimutagenic for minus-one errors at this site is also consistent with the model. Note that excess dTTP results in a 15-fold increase in frameshift frequency at an adjacent template C where A is the 5'-neighboring nucleotide (Fig. 2, position 78).

The same logic may be used to explain the formation of minus-two and plus-one nucleotide frameshifts by HIV-1 reverse transcriptase (Figs. 5, B and C). Additional examples consistent with the model can also be found in the spectra of errors shown in Figs. 2 and 3. This model was also invoked earlier to explain the production of minus-one nucleotide frameshifts by an exonuclease-deficient form of Klenow polymerase (Bebenek and Kunkel, 1990). The present observations extend the model to a polymerase involved in replication of the HIV-1 genome. Thus, conversion of a mispaired terminus into a misaligned frameshift intermediate may be a process common to a variety of DNA polymerases. The data also extend the model to include frameshift intermediates containing two extra template nucleotides (Fig. 5B) or one extra primer strand nucleotide (Fig. 5C). Thus, at least three types of frameshift errors are affected by pool imbalances in a manner consistent with a misinsertion-initiated mechanism.

In considering alternatives for pool imbalance-mediated effects on frameshift fidelity, it is also possible that the initiating event for frameshift errors may be template-primer slippage rather than misinsertion. In such cases increasing the concentration of the next correct nucleotide to be incorporated into the misaligned DNA could be mutagenic. This effect may be caused by an increased probability of subsequent nucleotide incorporation, which stabilizes the misalignment before a return to the correct alignment can occur. Antimutagenic effects could result if excess correct nucleotide for a particular position decreased the probability of the initial template-primer slippage. Processivity of polymerization is influenced by the dNTP concentration, and, in some instances, processivity correlates with frameshift fidelity (Kunkel, 1985; Bebenek et al., 1989). Thus, it is possible that pool imbalances alter processivity in a manner that changes the probability of template-primer slippage.

Some homopolymeric nucleotide positions provide the opportunity for correct base pairing in misaligned intermediates even when misinsertion is not involved. At these sites we cannot distinguish between misinsertion-initiated and misalignment-initiated frameshift errors induced by pool imbalances. At noniterated sites, the misinsertion-initiated model provides a more plausible explanation because it allows formation of at least one terminal base pair to stabilize the misalignment and permit continued polymerization. Not all pool imbalance-mediated effects on HIV-1 reverse transcriptase frameshift fidelity can be explained by these models. For example, in reactions with excess dATP, a 5-fold increase in the frequency of minus-one base frameshifts occurs at positions where the 5'-neighboring base is a G (Table III, sixth line, compare 17 versus 84). Even more obvious are the five complex mutations shown in Table II, containing one or more base substitution errors in addition to a minus-one frameshift error. All were recovered multiple times as independent revertants in small collections. Their recovery at a frequency much greater than for reactions performed at equimolar dNTP concentrations suggests that they result from the pool imbalance.

The fact that the multiple, closely spaced changes occurred much more frequently in combination with each other than alone or in combination with any other change is particularly interesting. This suggests that if the process(es) that introduces these changes acts sequentially as polymerization proceeds through this sequence (from right to left in the sequence shown in Table II), then the presence of an error at one position may influence error rates at downstream positions. Perhaps the presence of unpaired nucleotides (as mispairs or misaligned bases) in the primer stem that is in contact with the polymerase interferes with selectivity during subsequent incorporations. A similar principle may explain what has been called "semitargeted" mutagenesis (Schaeper and Glickman, 1982; Banerjee et al., 1990), where the presence of a lesion leads to an increased mutation frequency at a nearby position.

The first three mutants listed in Table II contain the frameshift and base substitution errors at adjacent positions. Since there is no intervening nucleotide between the two errors, each can be described in two ways. For example, the second mutant can be described as containing a C → T and a minus-A change (as shown) or as containing a minus-C and an A → T change. However, this is not so for the last two mutants in Table II, since they share a common feature of having at least one intervening template nucleotide between the frameshift and the base substitution error (in the first mutant, two As and in the second mutant, a G). Interestingly, in the first of these mutants, the initial error (from right to left) is a base substitution, and the last change is a frameshift whereas the reverse is true for the other mutant. Again, if the process(es) that introduces these changes acts sequentially as polymerization proceeds through this sequence, these results suggest that the first pool bias-induced error introduced in these complex mutants can be either a frameshift or a base substitution.

The alternative to sequential introduction of the errors in these complex mutants is that two or more changes may have been introduced by a single concerted mechanism. Models of this type have been described, including one in which all the changes are templated by a distant DNA sequence (for review, see Kunkel, 1990 and Ripley, 1990). The mechanism(s) responsible for these complex errors, and whether these observations are unique to HIV-1 reverse transcriptase, remains to be determined.

Our continuing interest in describing and understanding the fidelity of eukaryotic DNA replication prompted us to examine the effect of pool imbalances on the fidelity of a eukaryotic replicative enzyme, yeast DNA polymerase α. It was from this study that the clearest pattern of pool bias effects emerged. Decreasing the concentration of a single dNTP by 25-fold increased pol α-mediated one-base deletion errors at homopolymeric template nucleotide positions complementary to that dNTP. The most substantial effect resulted from lowering the concentration of dTTP; the frequency of one-base deletions at the AAA run increased 37-fold (Table VI). Lowering the concentration of one dNTP was also antimutagenic at some sites, e.g. at run sequences in which, the 5'-neighboring template base was complementary to the dNTP whose concentration was decreased. These mutagenic and antimutagenic effects are consistent with either of the two models described above.

Studies of semiconservative DNA replication in extracts of HeLa (Roberts and Kunkel, 1988; Roberts et al., 1991) or monkey cells (Hauser et al., 1988) have demonstrated that this process is highly accurate for base substitution errors. Obtaining useful information has therefore required the use of defined reversion assays and nucleotide pool imbalances to force the replication complex to commit base substitution errors. These errors occurred in a pattern predicted by Watson-Crick base-pairing principles and the relative rates of polymerization and exonucleolytic activity on particular mispairs. This predictability allows use of defined pool biases for
quantitatively describing base substitution fidelity (Roberts and Kunkel, 1988), the contribution of exonucleolytic proofreading to fidelity and base substitution fidelity on the leading and lagging strands (Roberts et al., 1991).

Based on DNA sequence analysis of a collection of mutants obtained from replication reactions in a forward mutation assay (Thomas et al., 1991) we have found that SV40 origin-dependent DNA replication is also accurate for frameshift errors. This conclusion is reinforced by the data obtained in the more sensitive reversion assay focused on minus-one nucleotide errors in a run of five A:T base pairs, showing a low reversion frequency for replication reactions performed at equimolar dNTPs (Table VII). One objective for undertaking the pool bias experiments with polymerase α was the hope that if a predictable pattern of pool imbalance-mediated effects on frameshift fidelity could be detected with a eukaryotic replicative DNA polymerase, then pool biases also could be employed to study frameshift fidelity during SV40 origin-dependent DNA replication.

The clear pattern of pool bias effects shown for DNA polymerase α in Table VII is our first indication that this strategy may be informative. The increased revertant frequencies that result from increasing the concentrations of three of the four dNTPs indicate that just as for base substitution errors, frameshift error rates during replication are affected by pool imbalances. As discussed above, these induced errors could be initiated either by template-primer slippage or by misinsertion leading to a misalignment (as in Fig. 5). The mutagenic effect also could reflect reduced exonucleolytic proofreading when employing a high concentration of the next correct nucleotides to be incorporated after the error. Exonucleolytic proofreading has been suggested to account in part for the high frameshift fidelity of the large fragment of E. coli DNA polymerase I (Bebenek et al., 1990) and of calf thymus DNA polymerase α (Thomas et al., 1991). We are currently attempting to determine if proofreading contributes to frameshift fidelity during SV40 origin-dependent replication in extracts.

The existing information on unequal concentrations of the four dNTPs in cells (see the Introduction) and the observation that pool bias effects are not limited to studies with purified DNA polymerases but are observed for semiconservative DNA replication by a multiprotein complex suggest that pool biases may also affect frameshift mutation rates in vivo. Extrapolation from the specificity data in Fig. 4 and Table VI also suggests that it may be possible to discriminate between the fidelity of replication at different template positions or on the leading and lagging strands.

Acknowledgments—We thank Mary P. Fitzgerald and Donald Halderman for assistance in DNA sequence analysis of revertants and Roel Schaaper and James Clark for their critical evaluation of the manuscript.

REFERENCES

Arecco, A.-L., Mun, B.-J., and Mathews, C. K. (1988) Mutat. Res. 200, 165–175
Banerjee, S. K., Borden, A., Christensen, R. B., LeClerc, J. E., and Lawrence, C. W. (1990) J. Bacteriol. 172, 2105–2112
Bebenek, K., and Kunkel, T. A. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 4946–4950
Bebenek, K., Abbotts, J., Roberts, J. D., Wilson, S. H., and Kunkel, T. A. (1989) J. Biol. Chem. 264, 10948–10956
Bebenek, K., Joyce, C. M., Fitzgerald, M. P., and Kunkel, T. A. (1990) J. Biol. Chem. 265, 13878–13887
Bestwick, R. K., Moffett, G. L., and Mathews, C. K. (1982) J. Biol. Chem. 257, 9300–9304
Cohen, M. B., Maybaum, J., and Sadee, W. (1981) J. Biol. Chem. 256, 8713–8717
Echols, H., and Goodman, M. F. (1991) Annu. Rev. Biochem. 60, 477–511
Hauser, J., Levine, A. S., and Dixon, K. (1988) Mol. Cell. Biol. 8, 3267–3271
Kunkel, T. A. (1985) J. Biol. Chem. 260, 5787–5796
Kunkel, T. A. (1990) Biochemistry 29, 8003–8011
Kunkel, T. A. and Soni, A. (1988a) J. Biol. Chem. 263, 4450–4459
Kunkel, T. A. and Soni, A. (1988b) J. Biol. Chem. 263, 14784–14789
Kunkel, T. A., Roberts, R. D., and Zakour, R. A. (1987) Methods Enzymol. 154, 362–387
Kunkel, T. A., Hamatake, R. K., Motto-Fox, J., Fitzgerald, M. P., and Sugino, A. (1989) Mol. Cell. Biol. 9, 4447–4458
Kunz, B. A. (1982) Environ. Mutagen. 4, 695–725
Kunz, B. A. (1988) Mutat. Res. 200, 133–147
Leeds, J., Slabaugh, M. B., and Mathews, C. K. (1985) Mol. Cell. Biol. 5, 3443–3450
Meuth, M. (1984) Mutat. Res. 126, 107–112
Phear, G., Naibantoghi, J., and Meuth, M. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 4450–4454
Ripley, L. S. (1990) Annu. Rev. Genet. 24, 189–213
Roberts, J. D., and Kunkel, T. A. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 7064–7068
Roberts, J. D., Thomas, D. C., and Kunkel, T. A. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 3465–3469
Sargent, R. G., and Mathews, C. K. (1987) J. Biol. Chem. 262, 5546–5553
Sargent, R. G., Ji, J., Mun, B., and Mathews, C. K. (1989) Molec. Gen. Genet. 217, 13–19
Schaaper, R. M., and Glickman, B. W. (1982) Molec. Gen. Genet. 185, 404–407
Streisinger, G., and Owen, J. E. (1985) Genetics 109, 633–659
Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Tsurugio, E., and Inouye, M. (1966) Cold Spring Harbor Symp. Quant. Biol. 31, 77–84
Thomas, D. C., Roberts, J. D., Fitzgerald, M. P., and Kunkel, T. A. (1990) in Antimutagenesis and Anticarcinogenesis Mechanisms II (Kuroda, Y., Shankel, D. M., and Waters, M. D., eds) pp. 289–297, Plenum Publishing Corp., New York
Thomas, D. C., Roberts, J. D., Sabatino, R. D., Myers, T. W., Tan, C. T., Downey, K. M., So, A. G., Bambara, R. A., and Kunkel, T. A. (1991) Biochemistry 30, 11751–11759.