## Marked infidelity of human immunodeficiency virus type <sup>1</sup> reverse transcriptase at RNA and DNA template ends

(mutagenesis/recombination/retrovirus/DNA polymerase)

PREMAL H. PATEL AND BRADLEY D. PRESTON\*

Laboratory for Cancer Research, Rutgers University Coliege of Pharmacy, Piscataway, NJ 08855-0789

Communicated by Allan H. Conney, July 8, 1993

ABSTRACT Human immunodeficiency virus type <sup>1</sup> (HIV-1) is genetically highly variable. This is attributed to the error-prone nature of HIV-1 replication and its proclivity for recombination. During replication and recombination, reverse transcriptase (RT) must polymerize DNA to the <sup>5</sup>' ends of multiple RNA and DNA template termini while converting HIV-1 RNA to double-stranded DNA. We have determined the fidelity of HIV-1 RT in vitro during polymerization to the <sup>5</sup>' ends of HIV-1 long terminal repeat DNA template sequences and to the end of <sup>a</sup> partial HIV-1 genomic RNA template that mimics a recombination intermediate. HIV-1 RT readily extended recessed DNA primers to form full-length blunt-end DNA-DNA and DNA-RNA duplexes. In addition, HIV-1 RT catalyzed high yields of products with one to four extra nucleotides at the <sup>3</sup>' ends of the nascent DNAs. These products were formed processively via a nontemplated mechanism that is highly specific for the addition of purine nucleotides  $(A > G)$  $\gg$  T  $\geq$  C). Thus, HIV-1 RT is extremely unfaithful at both DNA and RNA template ends, introducing errors (extra nucleotides) in one out of every two or three nascent strands processively polymerized. This error rate is 1000 times higher than for HIV-1 RT-catalyzed errors at internal template positions. Blunt-end additions were also catalyzed by other retroviral RTs at relative rates of HIV-1  $\approx$  Moloney murine leukemia virus > avian myeloblastosis virus. These data suggest a potentially important mechanism for retroviral mutation mediated by nontemplated blunt-end addition of purines prior to forced copy-choice recombination.

Acquired immunodeficiency syndrome (AIDS) results from a chronic infection by the human immunodeficiency virus type <sup>1</sup> (HIV-1; ref. 1). HIV-1 is a highly variable retrovirus that continuously alters its genetic composition during the course of infection (2, 3). This propensity for mutation underlies the alarmingly adaptive nature of HIV-1, which permits evasion of immunosurveillance, rapid emergence of drug-resistant variants, virus spread, and continued disease progression (4). An understanding of the mechanisms of HIV-1 mutation is of critical importance in the development of antiviral strategies (3).

The high mutation frequency of HIV-1 is attributed to the error-prone nature of retroviral replication (2, 3). Retroviruses convert their single-stranded genomic RNA to linear duplex DNA in <sup>a</sup> series of complex steps involving two template transpositions (jumps) and RNA- and DNAtemplated DNA polymerization catalyzed by virus-encoded reverse transcriptase (RT; for review, see ref. 5). RT exhibits low fidelity during both RNA- and DNA-templated DNA polymerization in vitro (3, 6, 7). In addition, RT mediates strand-transfer reactions that lead to recombination (8, 9). It is likely that RT-catalyzed polymerization errors and recombination are major determinants of HIV-1 variability (2, 3, 10).

A unique aspect of retrovirus replication is that RT must polymerize complementary strands of four different linear template ends  $(5)$ : (i) the RNA 5' R region just prior to the first jump,  $(ii)$  the RNA primer-binding site just prior to the second jump, (iii) the DNA U3 long terminal repeat (LTR) during completion of minus-strand synthesis, and  $(iv)$  the DNA U5 LTR during completion of plus-strand synthesis. In addition, retroviral RNAgenomes are often broken (11), and RT copies these random RNA template endsjust prior to recombination by a forced copy-choice mechanism (10, 11). Thus, HIV-1 RT encounters multiple template ends during each cycle of replication, and these ends presumably must be copied efficiently to sustain viral growth.

In preliminary studies, we observed that purified HIV-1 RT frequently adds nontemplated nucleotides while copying DNA template ends (12). In the present study, we examined the fidelity of HIV-1 RT in vitro during polymerization at LTR DNA template ends and at the end of <sup>a</sup> partial HIV-1 genomic RNA template. Our data show that HIV-1 RT copies both DNA and RNA template ends with extremely low fidelity, yielding large amounts of product (30-50% of molecules processively polymerized) with one to four extra nucleotides added beyond the template termini. This nontemplated polymerization is highly specific for the addition of purine nucleotides  $(A > G)$  and is associated with RTs from the lentivirus, avian sarcoma-leukosis, and mammalian type C retroviral groups. These results suggest a potentially important mechanism for retroviral mutation.

## MATERIALS AND METHODS

Materials. Virion and recombinant HIV-1 RTs were isolated as described (6, 13). The purified recombinant protein (8400 units/mg; units as defined in ref. 13) consisted of equimolar ratios of 66- and 51-kDa subunits and was biochemically indistinguishable from RT purified from HIV-1 virions (6, 13, 14). Escherichia coli DNA polymerase <sup>I</sup> large fragment was purchased from New England Biolabs. <sup>2</sup>'- Deoxyribonucleoside 5'-triphosphates (dNTPs; FPLC-pure), poly(rA) $(dT)_{12-18}$ , and Sephadex G-50 (DNA grade) were purchased from Pharmacia LKB Biotechnology. Bovine serum albumin (RIA grade) was purchased from Sigma. All other reagents were obtained in the purest available form from Fisher Scientific.

Polymerization at Template Ends. DNA·DNA primer-templates were constructed from synthetic oligodeoxyribonucleotides (HPLC-purified; Operon Technologies, Alameda, CA) with sequences corresponding to HIV-1 U3 and U5 LTR

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HIV-1, human immunodeficiency virus type 1; RT, reverse transcriptase; LTR, long terminal repeat.

<sup>\*</sup>To whom reprint requests should be sent at the present address: Section of Experimental Oncology, Room 3110b, Eccles Institute of Human Genetics, University of Utah, Salt Lake City, UT 84112.

termini (see Figs. <sup>1</sup> and 3). 5'-32P-labeled primers (15 or 24 nt long, i.e., 15-mers or 24-mers) were prepared by standard procedures (15) and hybridized to 30-mer templates (0.9:1 primer/template ratio) under conditions yielding >90% hybridization of the radiolabeled primer (6). All DNADNA primer-templates were designed with a 5'-recessed primer strand; this ensured that each duplex DNA contained <sup>a</sup> single 3'-hydroxyl group that is a substrate for polymerasecatalyzed primer extension or blunt-end addition.

Primertemplate DNA-RNA was prepared from run-off transcripts of <sup>a</sup> subclone of HIV-1 pNL4-3 proviral DNA (13). The Kpn 1-HindIII fragment of pNL4-3 (HIV-1 nt 6347-8130) was cloned into the vector  $pGEM-3zf(+)$ (Promega), and run-off RNA transcripts were generated in vitro by T7 RNA polymerase as described (13). The resultant RNA (1.8 kb long) corresponded to <sup>a</sup> fragment of the HIV-1 genomic plus-strand (portion of env). This RNA was hybridized to a synthetic [5'-32P]15-mer DNA to yield a DNA-RNA primertemplate with the <sup>3</sup>' nucleotide of the DNA primer positioned <sup>20</sup> nt from the <sup>5</sup>' end of the RNA template. RNA was handled under standard RNase-free conditions (15), and there was no detectable strand breakage (<1% per template site) during the hybridization and polymerization reactions (13).

Polymerizations of both DNA-DNA and DNA-RNA primer-templates were conducted in vitro at 30°C for 3-60 min in 10- to 100- $\mu$ l reaction mixtures containing 20 mM Tris-HCl (pH 8), 2 mM dithiothreitol, 5 mM MgCl<sub>2</sub>, 24 mM KCI, bovine serum albumin (0.1 mg/ml), <sup>2</sup> nM primer template, RNasin (1 unit/ $\mu$ l in RNA reaction mixtures; Promega), the indicated dNTPs (each at  $1-500 \mu M$ ), and 0.3-6 nM HIV-1 RT active sites. All reactions, except those confined to a single cycle, were initiated by the addition of HIV-1 RT. Single-cycle reactions were initiated by first equilibrating HIV-1 RT with primer-template in the absence of dNTPs (5 min at room temperature), followed by addition of poly(rA) $\cdot$ (dT)<sub>12-18</sub> (0.05 mg/ml) and dNTPs as described (13). All reactions were stopped by adding EDTA (7.5 mM). Products were eluted through Sephadex G-50 spin columns, separated in  $15\%$  or  $20\%$  gels by polyacrylamide/urea gel electrophoresis, visualized by autoradiography, and quantified by densitometry as described (6, 14).

## RESULTS AND DISCUSSION

Fidelity of Polymerization at DNA Template Ends. To examine polymerization at HIV-1 DNA template ends, synthetic DNA-DNA primertemplates were constructed that correspond to sequences of the nearly complete HIV-1 U3 and U5 LTR termini (Fig. 1). Incubations of HIV-1 RT with these primer-templates resulted in the formation of two major products (Fig. 1): the predicted blunt-end 24-mers and large amounts of 25-mers. Similar results were observed with viral and recombinant HIV-1 RTs at both the U3 and U5 LTR termini (data not shown). Analyses of primer-[5'-32P]templates before and after incubation with HIV-1 RT showed no evidence of contaminating templates with altered electrophoretic mobilities (<3%; data not shown). Thus, the 25-mer products did not result from polymerization on impure templates or templates modified during the incubations with HIV-1 RT.

The rates of both 24-mer and 25-mer synthesis were constant up to  $\approx 10$  min on both LTR templates, indicating that the products resulted from multiple cycles of RTprimer-template binding, DNA polymerization, RT dissociation, and binding to another primertemplate molecule. Moreover, the yields of 25-mers were exceptionally high (20-90% of nascent products). This high rate of 25-mer synthesis may result from two mechanisms: (i) efficient synthesis of products with an extra nucleotide during each cycle of processive



FIG. 1. Polymerization at the U5 LTR DNA end. Synthetic U5 LTR DNA-DNA primertemplate

[5'-32P]GTCAGTGTGGAAAAT GAAATCAGTCACACCTTTTAGAGATCGTC  $(-)$ 

was incubated with (lanes  $2-10$ ) or without (lane 1) recombinant HIV-1 RT (0.3 nM; 0.0010 unit/ $\mu$ ) in the presence of dATP, dCTP, dGTP, and dTTP (each at 500  $\mu$ M) at 30°C. Incubations were stopped after 3, 6, 9, 12, 15, 20, 25, 30, and 60 min, and the products were separated by polyacrylamide/urea gel electrophoresis, visualized by autoradiography, and quantified by densitometry. The predicted blunt-end products are 24 nt long (length markers are indicated to the left). Reactions containing synthetic U3 LTR DNA-DNA primer<br>template  $(5'-3^2P)$  TGGGAGTGAATTAGC  $(-)$  $(5' - 32P)$ TGGGAGTGAATTAGC

 $GAAGCAACCCTCACTTAATCGGGAAGGTCA$  (+) yielded a similar distribution of products (data not shown). The U3 and U5 primer template sequences used in these reactions are based on those predicted from sequences at the primer binding site and polypurine tract junctions of HIV-1 IIIB HXB2 (16).

polymerization and/or (ii) rapid enzyme cycling and distributive synthesis after forming blunt-end precursors.

To distinguish between these two mechanisms and to accurately determine the frequency of extra nucleotide addition during processive synthesis, reactions were conducted in the presence of competing  $poly(rA)$  $(dT)$  (Fig. 2). These single-cycle reactions resulted in yields of 24- and 25-mers essentially identical to those observed during the early time points with multiple enzyme cycling (compare Fig. 2, lane 3, with Fig. 1, lanes 2-4). Therefore, the products observed at the early time points in Fig. 1 likely resulted from processive



FIG. 2. Processive polymerization at LTR termini. Single-cycle reactions with HIV-1 RT (3 nM; 0.010 unit/ $\mu$ ) in the presence of a  $poly(rA)$ ·oligo(dT) trap were conducted as described (13) using U3 (lane 2) or U5 (lane 3) primertemplates in the presence of all four dNTPs (each at 500  $\mu$ M; see Fig. 1 for primer template sequences). The relative yields of 25-mers  $[nM \t25-mer/(nM \t24-mer + nM$ 25-mer)] after 10-min incubations were 0.3 and 0.5 on the U3 and U5 LTR primertemplates, respectively. A control reaction in which HIV-1 RT was first equilibrated with a poly(rA)-oligo(dT) trap and then initiated by the addition of dNTPs and the U3 primertemplate yielded no extension products and demonstrated that the trap was effective for at least 10 min (lane 1).

synthesis, and the high yield of 25-mers reflects an extremely high rate of HIV-1 RT errors during processive polymerization at template ends. From densitometric analysis, we calculate that HIV-1 RT introduces an extra nucleotide in one out of every two or three primer-template molecules polymerized in a single processive cycle. This error rate is 1000 times higher than for HIV-1 RT-catalyzed errors at internal template positions (6, 7).

Sequence Specificity of Extra Nucleotide Addition. The 25-mer products from both LTRs were sequenced by the Maxam and Gilbert method to determine the site of extra nucleotide addition. This analysis showed that the 25-mers were complementary to the template up to the 24-mer positions and revealed an extra nucleotide incorporated at the <sup>3</sup>' ends of both LTR nascent strands (data not shown). Therefore, the 25-mers were likely formed directly from blunt-end 24-mer precursors. The ability of HIV-1 RT to add nucleotides onto blunt-ends was verified by incubating HIV-1 RT with preformed blunt-end LTR sequences in the presence of all four dNTPs (Fig. 3, lane a). These reactions resulted in high yields of 25-mer products. Similar incubations with unhybridized [32P]24-mer resulted in no 25-mer formation (Fig. 3, lane b). Thus, extra nucleotide addition does not result from a loopback priming mechanism (17).

The Maxam and Gilbert sequence analyses suggested that dAMP is preferentially added as the extra nucleotide onto both LTR ends. However, ambiguities at the terminal nucleotide positions indicated that other nucleotides may also be added onto blunt-ends (data not shown). To determine the relative rates of addition of each nucleotide, HIV-1 RT was incubated with preformed blunt-end U3 LTR DNA in the presence of each dNTP individually, and the rates of 25-mer synthesis were compared (Fig. 3). Purine nucleotides were strongly favored substrates for this reaction with initial velocities of incorporation of  $dAMP(0.1 nM/min) > dGMP$  $(0.02 \text{ nM/min}) > d \text{TMP}$   $(0.002 \text{ nM/min}) > d \text{CMP}$  (<0.001) nM/min; limit of detection). This trend was observed over a range of dNTP concentrations (13–500  $\mu$ M) and on U5 as well as  $\phi$ X174 DNAs (data not shown). Thus, the specificity of extra nucleotide incorporation is not dependent on primer-template sequence.

These data demonstrate that the infidelity of HIV-1 RT at template ends is due to a high rate of processive polymerization of an extra nucleotide onto blunt-end duplexes. The preferential addition of <sup>a</sup> single dAMP residue on all primertemplates indicates that extra nucleotide addition is not due to template switching (8, 9), polymerization on discontinuous templates (18), or loopback polymerization (17), because these processes would result in the incorporation of each of the four nucleotides in a sequence-



FIG. 3. Specificity of extra nucleotide addition. Blunt-end U3 DNA  $(5'-32P)TGGGAGTGAATTAGCCCTTCCAGT (-))$  $\Gamma$ [5<sup> $\dot{\ }$ </sup>-<sup>32</sup>P]TGGGAGTGAATTAGCCCTTCCAGT (-GAAGCAACCCTCACTTAATCGGGAAGGTCA (+)

was incubated with HIV-1 RT (3 nM;  $0.010$  unit/ $\mu$ I) or without HIV-1 RT (control lanes "C") in the presence of 500  $\mu$ M dATP, 500  $\mu$ M dGTP, 500  $\mu$ M dTTP, or 500  $\mu$ M dCTP, individually. Incubations with HIV-1 RT were at 30°C for 3, 6, 9, 12, 15, and <sup>18</sup> min (except for the dTTP reactions where the 15-min point was omitted). Lanes a and b are from incubations (30°C, 10 min) of HIV-1 RT (3 nM;  $0.010$ unit/ $\mu$ l) with either blunt-end U3 DNA (lane a) or unhybridized [32P]24-mer (lane b) in the presence of all four dNTPs (each at 500  $\mu$ M). All products were analyzed by polyacrylamide/urea gel electrophoresis and quantified by densitometry.

dependent fashion and would yield products of heterogeneous length on different primertemplates. Other polymerases, including RTs from avian myeloblastosis virus and Moloney murine leukemia virus, have also been reported to add nucleotides onto DNA termini in vitro by <sup>a</sup> blunt-end addition mechanism (12, 19, 20).

Fidelity of Polymerization at RNA Template Ends. To determine the fidelity of polymerization at RNA template ends, an RNA segment corresponding to 1.8 kb of the HIV-1 genome was hybridized to <sup>a</sup> 15-mer DNA primer to form <sup>a</sup> DNA-RNA primer-template. The <sup>3</sup>' nucleotide of the DNA primer was positioned <sup>20</sup> bases from the <sup>5</sup>' RNA end. Thus, this synthetic DNA\*RNA partial duplex is similar to what HIV-1 RT will encounter as it approaches the <sup>5</sup>' end of a broken genomic RNA template.

Incubation of this primer-template with HIV-1 RT and dNTPs in vitro yielded a series of extension products ranging from 16 to 39 nt long (Fig. 4, lanes 3-7). The 35-mer corresponds to blunt-end DNA·RNA duplex. Thus, the formation of abundant products longer than 35-mers indicates that HIV-1 RT also introduces extra nucleotides at a high frequency during polymerization at RNA template ends. The yields of these longer products relative to the blunt-end 35-mer were constant at the early time points. Thus, the addition of extra nucleotides likely occurs during processive synthesis, as was observed on the LTR DNA templates (Figs. <sup>1</sup> and 2). Unlike the DNA templates, polymerization at this RNA template end also resulted in relatively high yields of products with up to four additional nucleotides (Fig. 4, lane 7). Reverse transcription with E. coli DNA polymerase <sup>I</sup> only yielded products  $\leq 35$  nt long (Fig. 4, lane 2). This is consistent with previous reports that blunt-end addition by E. coli DNA polymerase <sup>I</sup> is inefficient (19) and provides indirect evidence that the RNA template ends were homogeneous. Analyses of [32P]RNAs before and after hydrolysis with



FIG. 4. Polymerization at an RNA template end. [5'-32P]-DNA-RNA primertemplates were incubated with HIV-1 RT (3 nM; 0.010 unit/ $\mu$ l) in the presence of all four dNTPs (each at 1  $\mu$ M; lanes 3-7) or three dNTPs (each at 1  $\mu$ M) with the fourth dNTP at 500  $\mu$ M (lanes 8-11). Incubations at 30°C were stopped after 3, 6, 9, or 12 min (lanes 3-6) or 60 min (lanes 7-11). The products were analyzed by polyacrylamide/urea gel electrophoresis and quantified by densitometry. Incubations with E. coli DNA polymerase <sup>I</sup> large fragment (lane 2; 0.8 unit/ $\mu$ ); units as defined by the manufacturer) or without polymerase (lane 1) were conducted for 10 min at 30°C in the presence of all four dNTPs (each at 500  $\mu$ M). The 5' end of the RNA template sequence is displayed at the right, and the product lengths are shown at the left of the autoradiographs.

specific RNases confirmed that the RNA template ends were uniform (data not shown). Thus, extra nucleotide addition by HIV-1 RT is not an artifact of imprecise transcription initiation by T7 RNA polymerase. These data show that HIV-1 RT copies RNA template ends with very low fidelity and introduces extra nucleotides at a frequency of  $\approx 50\%$  during processive synthesis (Fig. 4, lane 3). RTs from other retroviruses also formed high yields of products with extra nucleotides (HIV-1 RT  $\approx$  Moloney murine leukemia virus RT  $>$ avian myeloblastosis virus RT; data not shown). Thus, this activity is associated with RTs from a broad range of retroviral groups.

To determine the base specificity of extra nucleotide addition on the DNA-RNA primer-template, the yields of products with extra nucleotides (>35-mers) were compared in polymerization reaction mixtures containing either low concentrations of all four dNTPs (each at  $1 \mu$ M) or low levels of three dNTPs (each at 1  $\mu$ M) plus a high level (500  $\mu$ M) of the fourth dNTP (Fig. 4, lanes 7-11). After 60 min in the presence of 1  $\mu$ M dNTPs, products >35-mers made up 30% of the total extended primers and contained as many as four extra nucleotides beyond the predicted blunt end (Fig. 4, lane 7). Incubation mixtures containing high concentrations of dATP or dGTP, but not dCTP or dTTP, gave significantly higher yields of products  $>35$ -mers (Fig. 4, lanes 8–11). The relative rates of elongation to products >35-mers were 20, 3, 1, and <sup>1</sup> for the reaction mixtures containing excess dATP, dGTP, dTTP, and dCTP, respectively (calculated as described in ref. 6). The selective incorporation of purine nucleotides argues against template switching, loopback priming, and polymerization on discontinuous templates as mechanisms of extra nucleotide addition (see above). Thus, as with DNA templates, extra nucleotides incorporated beyond RNA template ends result from preferential blunt-end addition of dAMP by <sup>a</sup> nontemplated mechanism.

Blunt-End Addition and Retroviral Mutagenesis. The proclivity of RT to add extra nucleotides during processive synthesis and at physiologic dNTP levels (Fig. 4, lanes 3-7) strongly suggests that blunt-end addition can occur during retroviral replication in vivo and may contribute significantly to retroviral mutation (Fig. 5). RT continuously encounters broken <sup>5</sup>' RNA template ends in vivo and responds by switching templates and continuing polymerization (10, 11). Blunt-end addition of one or several nucleotides at these broken ends will result in a high frequency of mismatched nucleotides at primertemplate termini after template transfer (Fig. 5, steps 1-3). The fate of these mismatches will depend both on the nature of the flanking sequences and on the ability of RT to extend the aberrant primer-template structures. Studies in vitro show that RTs, particularly HIV-1 RT, efficiently extend primer-templates with terminal mismatches (14, 21) or misaligned nucleotides (22, 23). Extension of terminal mismatches will yield base-substitution mutations (Fig. 5, step 4). In addition, extra nucleotides at broken RNA ends could result in small deletions or insertions after strand transfer, primertemplate misalignment, and subsequent extension (Fig. 5, steps <sup>5</sup> and 6). A combination of deletions, insertions, and base substitutions may also occur if several nucleotides are added onto blunt ends prior to strand transfer.



FIG. 5. Model for mutagenesis by blunt-end addition and forced copy-choice recombination. After polymerization to the <sup>5</sup>' end of a broken RNA genome (step 1), RT adds an extra nucleotide (depicted here as A) by blunt-end addition onto the nascent strand (step 2). Subsequent forced copy-choice strand transfer to the intact RNA genome copy will result in <sup>a</sup> mispair (step 3) with <sup>a</sup> probability of 0.75 (assuming that RNA breaks are random). This mispair can result in three classes of mutation. Mispair extension (step 4) and continued reverse transcription will result in base substitution errors  $(N \rightarrow T)$  in the proviral DNA plus strand. Alternatively, the extra adenosine residue may lead to misalignment of the primer-template resulting in either insertion or deletion mutations depending on the flanking template sequences (steps 5 and 6). Misalignment with upstream uridine residues (step 5a) will result in nucleotide insertions, while misalignment with downstream uridine residues (step Sb) will yield deletions; both of these pathways require RT to extend misaligned primer-templates with unpaired extra nucleotides in either the primer (step 6a) or the template (step 6b) strand. More than one nucleotide may be inserted or deleted depending on the extent of misalignment. The sequence shown surrounding the RNA break is arbitrary and was selected for illustration purposes. Thin and thick lines represent RNA and DNA, respectively;  $\nabla$  and  $\Delta$  represent insertion and deletion mutations, respectively.

The results of our studies suggest a potentially important mechanism for RT-catalyzed mutagenesis involving initial nontemplated polymerization at RNA template ends and subsequent forced-copy choice recombination. The rate of mutation by this mechanism is potentially high. Based on the high rate of retroviral minus-strand recombination (20%; ref. 10) and the high frequency of blunt-end addition during processive synthesis (30-50%; Figs. 2 and 4), we estimate that 1 out of every 10-20 viruses could mutate by this mechanism during a single cycle of replication. Studies of unintegrated proviral DNAs from infected cells in culture suggest that blunt-end addition occurs widely in retroviruses (24-28) and in the retroelement copia (29), although the frequency in vivo appears somewhat lower than in vitro (27, 28, 30, 31).

The observed preference for addition of purine residues (Figs. 3 and 4) leads to very distinct predictions about the nature and location of mutations arising by this mechanism. Base substitutions will be predominantly  $N \rightarrow T$  and  $N \rightarrow C$ substitutions in the proviral plus-strand, and small deletions or insertions will occur flanked by plus-strand thymidine or cytidine residues. Both types of mutations should occur at minus-strand recombination junctions. This model is readily testable in viral (10, 32) and cell-free (8, 13) replication systems and may account for the short nucleotide insertions observed at mutation junctions in retroviral deletion mutants (33, 34).

Note. While this paper was in preparation, Peliska and Benkovic (35) proposed a similar mechanism for HIV-1 base-substitution mutagenesis.

We gratefully acknowledge Cherylene Schauber for purification of recombinant HIV-1 RT; George Klarmann for preparing the RNA template; Sandy Price for technical assistance; Diana Lim for graphic arts; George Klarmann, Cherylene Schauber, Nancy Garvey, and Joe Dougherty for critical reading of the manuscript; and Jim Fyfe and Paula Sherman for helpful comments and communicating unpublished data. This work was supported by grants from the National Institutes of Health (R29 CA48174 and P30 ES05022) and the Leukemia Research Foundation. B.D.P. is a recipient of a Junior Faculty Research Award from the American Cancer Society (JRFA-245) and a Henry Rutgers Research Fellowship from Rutgers University.

- 1. Fauci, A. S. (1988) Science 239, 617-622.
- 2. Katz, R. A. & Skalka, A. M. (1990) Annu. Rev. Genet. 24, 409-445.
- 3. Preston, B. D. & Garvey, N. (1992) Pharm. Technol. 16, 34–51.<br>4. Richman, D. D. (1991) Immunodefic. Rev. 2, 315–318.
- Richman, D. D. (1991) Immunodefic. Rev. 2, 315-318.
- 5. Coffin, J. M. (1990) in Virology, eds. Fields, B. N., Knipe,

D. M., Chanock, R. M., Melnick, J. L., Roizman, B. & Shope, R. E. (Raven, New York), pp. 1437-1500.

- 6. Preston, B. D., Poiesz, B. J. & Loeb, L. A. (1988) Science 242, 1168-1171.
- 7. Roberts, J. D., Bebenek, K. & Kunkel, T. A. (1988) Science 242, 1171-1173.
- 8. Luo, G. & Taylor, J. (1990) J. Virol. 64, 4321–4328.<br>9. DeStefano, J. J., Mallaber, J., M., Rodriguez-Rodri
- 9. DeStefano, J. J., Mallaber, L. M., Rodriguez-Rodriguez, L., Fay, P. J. & Bambara, R. A. (1992) J. Virol. 66, 6370-6378.
- 10. Hu, W.-S. & Temin, H. M. (1990) Science 250, 1227-1233.
- 11. Coffin, J. M. (1979) J. Gen. Virol. 42, 1-26.<br>12. Patel, P. H. & Preston, B. D. (1991) J. Acquir
- Patel, P. H. & Preston, B. D. (1991) J. Acquired Immune Defic. Syndr. 4, 344 (abstr.).
- 13. Klarmann, G. J., Schauber, C. A. & Preston, B. D. (1993) J. Biol. Chem. 268, 9793-9802.
- 14. Perrino, F. W., Preston, B. D., Sandell, L. L. & Loeb, L. A. (1989) Proc. Natl. Acad. Sci. USA 86, 8343-8347.
- 15. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab. Press, Plainview, NY).
- 16. Ratner, L., Fisher, A., Jagodzinski, L. L., Mitsuya, H., Ruey-Shyan, L., Gallo, R. C. & Wong-Staal, F. (1987) AIDS Res. Hum. Retroviruses 3, 57-69.
- 17. Komberg, A. & Baker, T. A. (1992) DNA Replication (Freeman, New York), pp. 105-106.
- 18. Clark, J. M. (1991) Gene 104, 75-80.
- 19. Clark, J. M., Joyce, C. M. & Beardsley, G. P. (1987) J. Mol. Biol. 198, 123-127.
- 20. Clark, J. M. (1988) Nucleic Acids Res. 16, 9677-9686.
- 21. Mendelman, L. V., Petruska, J. & Goodman, M. F. (1990) J. Biol. Chem. 265, 2338-2346.
- 22. Roberts, J. D., Preston, B. D., Johnston, L. A., Soni, A., Loeb, L. A. & Kunkel, T. A. (1989) Mol. Cell. Biol. 9,469-476.
- 23. Bebenek, K., Abbotts, J., Roberts, J. D., Wilson, S. H. & Kunkel, T. A. (1989) J. Biol. Chem. 264, 16948-16956.
- 24. Van Beveren, C., van Straaten, F., Galleshaw, J. A. & Verma, I. M. (1981) Cell 27, 97-108.
- 25. Olsen, J. C. & Swanstrom, R. (1985) J. Virol. 56, 779–789.<br>26. Colicelli, J. & Goff. S. P. (1988) J. Mol. Biol. 199, 47–59.
- 26. Colicelli, J. & Goff, S. P. (1988) J. Mol. Biol. 199, 47-59.<br>27. Whitcomb. J. M., Kumar, R. & Hughes, S. H. (1990) J. V.
- 27. Whitcomb, J. M., Kumar, R. & Hughes, S. H. (1990) J. Virol. 64, 4903-4906.
- 28. Smith, J. S., Kim, S. & Roth, M. J. (1990) J. Virol. 64, 6286-6290.
- 
- 29. Flavell, A. J. & Ish-Horowicz, D. (1983) Cell 34, 415-419.<br>30. Roth, M. J., Schwartzberg, P. L. & Goff, S. P. (1989) Cell 30. Roth, M. J., Schwartzberg, P. L. & Goff, S. P. (1989) Cell 58, 47-54.
- 31. Pauza, C. D. (1990) Virology 179, 886-889.
- Varela-Echavarría, A., Garvey, N., Preston, B. D. & Dougherty, J. P. (1992) J. Biol. Chem. 267, 24681-24688.
- 33. Pathak, V. K. & Temin, H. M. (1990) Proc. Natl. Acad. Sci. USA 87, 6024-6028.
- 34. Varela-Echavarrfa, A., Prorock, C. M., Ron, Y. & Dougherty, J. P. (1993) J. Virol. 67, 6357-6364.
- 35. Peliska, J. A. & Benkovic, S. J. (1992) Science 258, 1112-1118.