

# Mutations in the RNase H Primer Grip Domain of Murine Leukemia Virus Reverse Transcriptase Decrease Efficiency and Accuracy of Plus-Strand DNA Transfer

Jean L. Mbisa, Galina N. Nikolenko, and Vinay K. Pathak\*

*HIV Drug Resistance Program, National Cancer Institute at Frederick, Frederick, Maryland*

Received 20 May 2004/Accepted 19 August 2004

The RNase H primer grip of human immunodeficiency virus type 1 (HIV-1) reverse transcriptase (RT) contacts the DNA primer strand and positions the template strand near the RNase H active site, influencing RNase H cleavage efficiency and specificity. Sequence alignments show that 6 of the 11 residues that constitute the RNase H primer grip have functional equivalents in murine leukemia virus (MLV) RT. We previously showed that a Y586F substitution in the MLV RNase H primer grip resulted in a 17-fold increase in substitutions within 18 nucleotides of adenine-thymine tracts, which are associated with a bent DNA conformation. To further determine the effects of the MLV RNase H primer grip on replication fidelity and viral replication, we performed additional mutational analysis. Using either  $\beta$ -galactosidase (*lacZ*) or green fluorescent protein (*GFP*) reporter genes, we found that S557A, A558V, and Q559L substitutions resulted in statistically significant increases in viral mutation rates, ranging from 2.1- to 3.8-fold. DNA sequencing analysis of nonfluorescent *GFP* clones indicated that the mutations in RNase H primer grip significantly increased the frequency of deletions between the primer-binding site (PBS) and sequences downstream of the PBS. In addition, quantitative real-time PCR analysis of reverse transcription products revealed that the mutant RTs were substantially inefficient in plus-strand DNA transfer relative to the wild-type control. These results indicate that the MLV RNase H primer grip is an important determinant of in vivo fidelity of DNA synthesis and suggest that the mutant RT was unable to copy through the DNA-RNA junction of the minus-strand DNA and the tRNA because of its bent conformation resulting in error-prone plus-strand DNA transfer.

Genetic diversity is a hallmark of retroviral populations resulting from a high rate of mutations during viral replication (4, 37, 39). This genetic variation is of clinical significance because it is the basis for antiviral drug resistance and escape from host immune responses readily exhibited by retroviruses such as human immunodeficiency virus type 1 (HIV-1) (12, 21, 25, 27, 35, 36). The rapid evolution of retroviruses is also an impediment to the design of broadly effective vaccines against HIV-1 (11, 40). Although host cell DNA polymerases and RNA polymerase II are involved in the replication of retrovirus genomes, error-prone replication by the virally encoded reverse transcriptase (RT) is most likely a major contributor to the high mutation rate of retroviruses (20).

The structure of RT and inherent nature of the reverse transcription process likely play an important role in the low fidelity of RT. Unlike most high-fidelity cellular DNA polymerases, RT lacks a classical 3'-5' exonuclease proofreading activity. Other structural features also known to affect RT fidelity include positioning of the template-primer complex at the polymerase active site that is dictated by contacts with RT residues and the local geometry of the polymerase active site (3, 15, 44). In addition, reverse transcription of the retroviral RNA genome requires two template-switching events, namely, minus-strand DNA transfer and plus-strand DNA transfer. It is hypothesized that in order to accommodate these two essential

events, RT has evolved to possess low template affinity and processivity (39), which can inadvertently result in template-switching mutations (30). Taken together, these factors contribute to the error-prone nature of DNA synthesis by RT and to the high mutation rate of retroviruses.

In spite of the potential role of RT structure in the accuracy of DNA synthesis, only a few studies have characterized the structural determinants of RT fidelity in vivo (13, 14, 22–24, 37, 44). These studies have identified the polymerase active site YXDD motif and other domains, such as the deoxynucleoside triphosphate (dNTP)-binding site, to be important in replication fidelity. Mutations in the YXDD motif of murine leukemia virus (MLV) and human T-cell leukemia virus type 1 were associated with changes in fidelity of DNA synthesis (13, 22). Mutational analysis of the dNTP-binding site and flanking residues of both MLV and HIV-1 has revealed the importance of the domain in replication fidelity. Mutating the residue that binds the base and ribose moiety of the incoming dNTP F155 to W in MLV and its HIV-1 homolog Y115 to A resulted in significant increases in the in vivo mutation rate, as did mutation of MLV dNTP-binding site flanking residue L151 to F (14, 24). In contrast, mutation of HIV-1 dNTP-binding site residue Q151 to N increased the accuracy of DNA synthesis (24). Intriguingly, mutations in the finger and primer grip subdomains of HIV-1, such as K65R, D76V, R78A, and W229A, affected the in vivo replication fidelity by increasing the accuracy of DNA synthesis (24). In addition, HIV-1 RTs resistant to 3'-azido-3'-deoxythymidine were shown to increase the in vivo mutation rate (23).

Structural determinants of RT fidelity are not restricted to

\* Corresponding author. Mailing address: HIV Drug Resistance Program, NCI-Frederick, P.O. Box B, Bldg. 535, Rm. 334, Frederick, MD 21702-1201. Phone: (301) 846-1710. Fax: (301) 846-6013. E-mail: vpathak@ncifcrf.gov.

residues close to the polymerase active site. We recently reported that the Y586F substitution in the MLV RNase H domain resulted in an approximately fivefold increase in the MLV mutation rate *in vivo*, which is the highest reported to date (44). Residue Y586 of MLV is equivalent to HIV-1 residue Y501, a constituent of the recently described RNase H primer grip domain, which contacts the DNA primer strand and positions the template strand near the RNase H active site, influencing RNase H cleavage efficiency and specificity (31, 34). Mansky et al. also recently reported that the HIV-1 mutant Y501W results in a 2.7-fold increase in the *in vivo* mutation rate (24). Sequence alignments indicate that six out of 11 HIV-1 RNase H primer grip residues have functional equivalents in the MLV RNase H domain (34). In the present study, mutational analysis of several residues in the MLV RNase H primer grip was carried out to further determine the role of this domain in MLV replication fidelity *in vivo*. The results show that mutation of certain residues in this domain result in an increased frequency of deletions between the primer-binding site (PBS), and sequences downstream, indicating an error-prone plus-strand DNA transfer.

#### MATERIALS AND METHODS

**Plasmids, retrovirus vectors, and mutagenesis.** Plasmids pLGPS and pRMBNB express MLV *gag* and *pol* genes from a truncated long terminal repeat ( $\Delta$ LTR) promoter (14, 26). Plasmid pSV $\alpha$ 3.6 encodes the  $\alpha$  subunit of the murine Na<sup>+</sup>,K<sup>+</sup>-ATPase gene and confers resistance to ouabain (19). Plasmid pGN-MLV-GFFP-IHy is a derivative of pES-GFFP (38) in which the NotI-NgoMIV fragment containing the GFFP-IRES-*neo* cassette was replaced by the PpuMI-BglII fragment from pKD-HIV-GFFP-IHy (29) containing the cytomegalovirus (CMV)-GFFP-IRES-*hygro* cassette. Plasmid pHCMV-G expresses the G envelope protein of vesicular stomatitis virus from a CMV promoter (43). The S557A, A558V, Q559L, Y586A, and T590A substitution mutations in the MLV RNase H primer grip domain of pRMBNB were generated by using a QuikChange site-directed mutagenesis kit (Stratagene). By using an alanine-scanning mutagenesis strategy, we substituted most of the residues with alanines, with the exception of A558 and Q559. The A558 residue, which was already an alanine was substituted with a valine, and Q559 was substituted with a leucine because of ease of primer design. The presence of the desired mutations and the absence of other mutations were verified by restriction enzyme digestions, followed by DNA sequencing.

**Cells, transfections, and infections.** D17 dog osteosarcoma cells and 293T cells were obtained from the American Type Culture Collection. The D17-based cell lines ANGIE P or A3GFP11, which contain a single GA-1 or MP-1 provirus, respectively, also express an amphotropic MLV envelope (13, 44). The GN-MLV-GFFP cell line contains a single MLV-based provirus derived from the vector pGN-MLV-GFFP-IHy. To construct the cell line, 293T cells were cotransfected with vector pGN-MLV-GFFP-IHy, helper construct pLGPS, and envelope construct pHCMV-G. The resulting pseudotyped virus was used to infect 293T cells. After hygromycin selection, several nonfluorescent cell clones (verified by flow cytometry) were isolated, expanded, and characterized by Southern blot analysis and infection assays. A cell clone containing one full-length provirus and producing the best virus titer was named GN-MLV-GFFP and used as a virus producer cell line. Cells were grown in Dulbecco modified Eagle medium (HyClone Laboratories, Inc.) supplemented with 6% calf serum (D17 cells) or 10% fetal calf serum (293T cells), 50 U of penicillin (Gibco)/ml, and 50  $\mu$ g of streptomycin (Gibco)/ml. Transfection of ANGIE P and A3GFP11 cells was performed by using the dimethyl sulfoxide-Polybrene method (18). Transfection of 293T cells was carried out by using calcium-phosphate precipitation (33) (CalPhos Transfection Kit; Clontech). Infections were performed in the presence of Polybrene (50  $\mu$ g/ml).

**In vivo single-replication-cycle fidelity assay.** ANGIE P or A3GFP11 cells were plated at a density of  $2 \times 10^5$  cells per 60-mm-diameter dish and 24 h later were cotransfected with either wild-type or mutant pRMBNB and pSV $\alpha$ 3.6. The transfected cells were selected for resistance to ouabain ( $10^{-7}$  M), and the resistant colonies were pooled and expanded. The culture medium from pooled cells containing either GA-1 or MP-1 virus was used to infect D17 target cells plated at a density of  $2 \times 10^5$  cells per 60-mm-diameter dish as previously

described (13). After selection for resistance to G418, GA-1-infected D17 cells were stained with X-Gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside), and *lacZ* inactivation was determined by counting blue and white colonies, whereas MP-1-infected cells were observed under a fluorescence microscope to determine the frequency of *GFP* inactivation (8, 13, 44).

**Isolation of single nonfluorescent cell clones by FACS.** G418-resistant colonies from MP-1 virus-infected cells were pooled, passaged, and subjected to fluorescence-activated cell sorting (FACS) (CloneCyt Plus System, FACS Vantage SE; BD Biosciences) to isolate individual nonfluorescent clones that did not express functional *GFP*. The clones were sorted into 96-well plates, and the nonfluorescent phenotype was subsequently verified by fluorescence microscopy.

**Isolation of genomic DNA, PCR, and sequence analysis.** After FACS, single nonfluorescent cell clones grown in 96-well plates were expanded into 24-well plates and then into 60-mm-diameter dishes. The cell clones were harvested and lysed to isolate genomic DNA by using an Bio-Rad AquaPure Genomic DNA Isolation Kit (Bio-Rad). Provirus-specific DNA fragments were amplified by PCR by using Takara Hot-Start *Taq* DNA polymerase (Takara Mirus Bio., Inc.) and the following sets of forward and reverse primers encompassing the region between the 5'LTR and internal ribosome entry site (IRES) of MP-1 provirus: MP-1623F (5'-T CACTCCTTCTCTAGGCGCCGGAATTGG-3') and MP-2390R (5'-GGAATTG GCCGCTCACTGTACAGCTCG-3'), U3-5538F (5'-CCAATCAGTTCGCTTC TCGCT-3') and MP-2390R, or U3-5538F and MP-2938R (5'-GTTCAATCATGC GAAACGATCC-3').

**Quantitative real-time PCR analysis.** For real-time PCR analysis, we used virus produced by transfection of the GN-MLV-GFFP cell line because it produced high-titer virus. GN-MLV-GFFP cells were cotransfected with either wild-type or mutant pRMBNB and pHCMV-G. Medium was changed 24 h posttransfection, and virus was collected 24 h later. The virus was used to infect 293T target cells for 1, 3, and 6 h after which cells were washed once with PBS. For harvesting cells at 3 and 6 h postinfection, fresh medium was added to the cells at 1 h postinfection. Total cellular DNA was extracted from infected cells by using the QIAmp DNA Blood Minikit (Qiagen). DNA from  $\sim 10^5$  cells was used for each real-time PCR assay with an ABI Prism 7700 sequence detector (Applied Biosystems). The primer and probe sets for the RU5, *hygro*, and U5- $\Psi$  regions and the PCR conditions used were as previously described (10, 42). We used a primer and probe set designed to detect the human porphobilinogen deaminase gene to normalize for the amount of DNA analyzed in the real-time PCR experiments (41). Threefold dilutions of the MLV-based vector pMMQD3 (6) were used to generate a standard curve ranging from 17 to 1,000,000 copies of DNA per PCR assay. The same dilutions were used to generate a standard curve for each primer and probe set, which allowed accurate measurement of relative amounts of DNA products detected by different sets. The amount of each PCR product was determined from a standard curve generated with that particular primer and probe set. The correlation coefficient for all standard curves was  $>0.99$ .

To normalize for the efficiency of initiation of viral DNA synthesis by the different viruses, we quantified the amount of viral RNA from the viral preparations that were used for infection by RT-PCR, as previously described (10).

#### RESULTS

**Effect of mutating MLV RNase H primer grip residues on RT replication fidelity.** We previously showed that mutating the MLV RNase H primer grip domain residue Y586 to F significantly decreased replication fidelity (44). To further determine the role of this domain on replication fidelity, we generated the following mutations: S557A, A558V, Q559L, Y586A, and T590A, encompassing five of the six residues that constitute the domain. The effect of these mutations on the *in vivo* MLV mutation rate was individually assessed and compared to that of wild-type RT by using a previously described single-round replication assay in which the *lacZ* gene served as a mutation reporter gene (Fig. 1). Briefly, GA-1 is an MLV-based vector that expresses *lacZ* and *neo* from the viral LTR promoter; the *neo* open reading frame is translated from an IRES. ANGIE P cells express the amphotropic MLV envelope and contain a single GA-1 provirus. Plasmid pSV $\alpha$ 3.6 confers resistance to ouabain. The ANGIE P cells were cotransfected with either wild-type or mutant *gag-pol* constructs and pSV $\alpha$ 3.6.

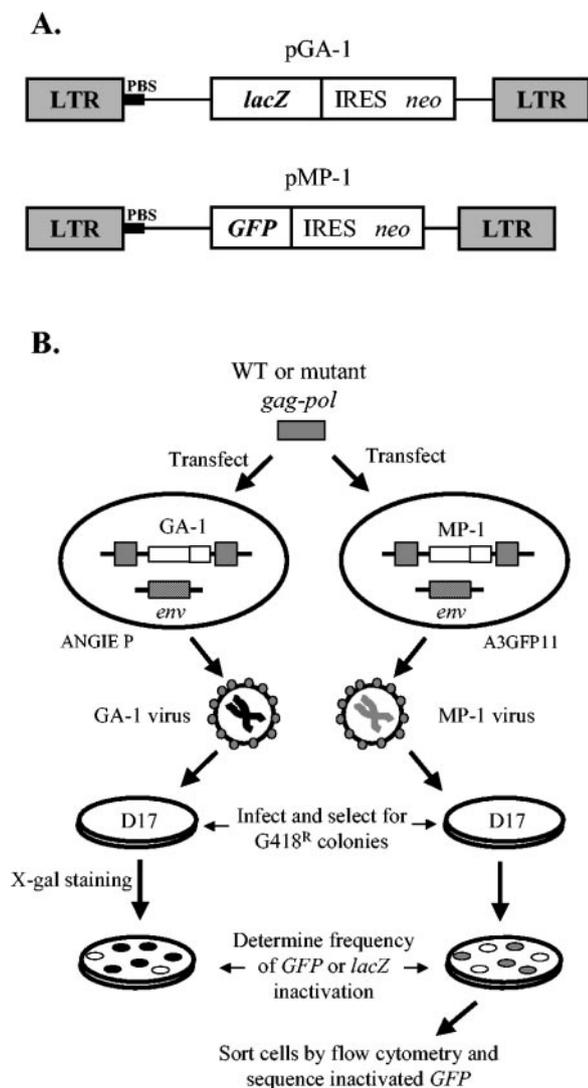


FIG. 1. Single-replication-cycle in vivo fidelity assay. (A) Structures of MLV-based vectors pGA-1 and pMP-1. The schematic shows the relative positions of the LTRs, PBS, and reporter genes *lacZ* and *GFP*. The neomycin resistance gene (*neo*) is expressed from an IRES. (B) Experimental protocols. Wild-type (WT) or mutant MLV *gag-pol* constructs were cotransfected separately with pSV $\alpha$ 3.6 into ANGIE P or A3GFP11 cell lines. Both cell lines stably express an amphotropic MLV envelope (*env*) and an integrated MLV-based provirus expressing either *lacZ* (GA-1 provirus) or *GFP* (MP-1 provirus), respectively. The GA-1 or MP-1 viruses produced were used to infect D17 target cells. G418-resistant (G418<sup>R</sup>) infected cell clones were quantified to determine the frequency of inactivation of the *lacZ* or *GFP* genes. In addition, *GFP* virus-infected cell clones were pooled, individual non-fluorescent cells were isolated by flow cytometry, and the nature of the *GFP*-inactivating mutations was characterized by PCR amplification, followed by DNA sequencing.

After selection with ouabain, virus-producing cells were pooled and the virus produced from the pooled cells was used to infect target D17 cells. The infected cells were selected for resistance to G418, a neomycin analog, and the resulting colonies were stained with X-Gal. The frequency of *lacZ* inactivation was then determined by counting white and blue colonies.

Viruses containing wild-type RT inactivated *lacZ* at a frequency of  $5.5\% \pm 0.5\%$  (Table 1), which is comparable to

results obtained previously (13, 14, 44). In contrast, three of the five mutants (S557A, A558V, and Q559L) significantly increased the mutation rate by 2.1- to 3.8-fold ( $P < 0.003$ ; Student's *t* test) relative to wild-type RT. The two mutants with the largest effect on fidelity, S557A and A558V, also dramatically reduced virus titers by ca. 100- and 500-fold, respectively, compared to wild-type RT ( $6.5 \times 10^3$  CFU/ml). In contrast, Q559L only reduced virus titers by twofold. Of the remaining two mutants, T590A did not significantly change the frequency of *lacZ* inactivation or the virus titer ( $P > 0.1$ ; Student *t* test), whereas the effect on replication fidelity by Y586A, a nonconservative mutation compared to the previously described Y586F, could not be determined because it reduced virus titers by >10,000-fold compared to wild-type RT. Thus, the majority of residues within the MLV RNase H primer grip reduced replication fidelity. Using a previously described assay (13), we also determined the RT activities of the S557A, Q559L, and T590A mutant RTs compared to wild-type RT. We found that the S557A mutation had the largest effect on RT activity (21% of wild-type), whereas Q559L and T590A had RT activities of 69 and 56%, respectively, compared to wild-type RT. The RT activities were normalized to the amount of capsid protein quantified by Western blotting (data not shown). That the RT activities were within fivefold of the wild type and that wild-type levels of processed capsid proteins were present on Western blots indicated that the mutations did not significantly influence proteolytic processing. Thus, virus titers of mutants were reduced to a greater extent than RT activities, suggesting that other steps in reverse transcription were also affected by these mutations.

**Analysis of mutations generated by the S557A mutant with *GFP* as a reporter gene.** To analyze the nature of replication errors made by mutant RTs, we characterized the mutations generated by S557A because it had one of the highest mutation

TABLE 1. Effect of mutating MLV RNase H primer grip residues on RT replication fidelity

MLV RT genotype	Virus titer (CFU/ml) <sup>a</sup>	No. of white colonies/total no. of colonies <sup>b</sup>	Frequency (% of <i>lacZ</i> inactivation $\pm$ SEM) <sup>c</sup>	Relative change in inactivation of <i>lacZ</i> <sup>d</sup>
Wild type	$(6.5 \pm 3.8) \times 10^3$	59/1,109	$5.5 \pm 0.5$	1.0
S557A	$(6.4 \pm 2.6) \times 10^1$	635/3,643	$18.4 \pm 2.4^e$	3.3
A558V	$(1.4 \pm 0.1) \times 10^1$	76/337	$21.1 \pm 2.0^e$	3.8
Q559L	$(3.5 \pm 1.1) \times 10^3$	281/2,394	$11.7 \pm 0.2^e$	2.1
Y586A	<1	NC <sup>f</sup>	NA <sup>g</sup>	NA
Y586F <sup>h</sup>	$(3.7 \pm 0.6) \times 10^2$	582/2,203	$26.4 \pm 3$	5.4
T590A	$(9.0 \pm 5.9) \times 10^3$	50/719	$6.7 \pm 0.5$	1.2

<sup>a</sup> The average virus titers  $\pm$  the standard errors of the mean were determined by serial dilutions and infections, followed by counting the G418<sup>R</sup> colonies.

<sup>b</sup> That is, the number of mutant colonies that displayed a white-colony phenotype and the total number of colonies counted in two to three independent experiments.

<sup>c</sup> The frequency of *lacZ* inactivation was calculated as follows: (number of white colonies in each experiment/total number of colonies)  $\times$  100.

<sup>d</sup> The relative change in inactivation of *lacZ* gene was calculated as follows: the frequency of *lacZ* inactivation observed with mutant MLV RT/the frequency of *lacZ* inactivation observed with wild-type MLV RT.

<sup>e</sup> The frequency of *lacZ* inactivation was significantly higher than for the wild type ( $P < 0.003$ ; Student *t* test).

<sup>f</sup> No colonies were present after G418 selection, signifying a >10<sup>4</sup>-fold reduction in titer.

<sup>g</sup> NA, not applicable.

<sup>h</sup> Data were from reference 44 [the wild-type virus titer was  $(9.9 \pm 1.6) \times 10^3$ , and the mutant frequency was  $4.9\% \pm 0.2\%$ ].

TABLE 2. Spectrum of *GFP*-inactivating mutations associated with wild-type, S557A, and Y586F RTs

Mutation type	Wild-type RT		S557A RT			Y586F RT		
	No. of mutants sequenced <sup>a</sup> (proportion of total) <sup>b</sup>	Mutant frequency <sup>c</sup> (%)	No. of mutants sequenced (proportion of total)	Mutant frequency (%)	Relative increase in mutant frequency <sup>d</sup>	No. of mutants sequenced (proportion of total)	Mutant frequency (%)	Relative increase in mutant frequency
PBS deletions <sup>e</sup>	9 (0.10)	0.09	17 (0.27)	0.79	8.8	23 (0.36)	1.83	20.3
Other deletions <sup>f</sup>	29 (0.32)	0.28	16 (0.26)	0.76	2.7	13 (0.20)	1.01	3.6
Duplications <sup>g</sup>	7 (0.08)	0.07	7 (0.11)	0.32	4.6	4 (0.06)	0.30	4.3
All temp.-switch <sup>h</sup>	45 (0.50)	0.44	40 (0.65)	1.90	4.3	40 (0.63)	3.19	7.3
Substitutions	32 (0.36)	0.31	20 (0.32)	0.94	3.0	17 (0.27)	1.39	4.5
Frameshifts	13 (0.14)	0.12	2 (0.03)	0.09	0.8	7 (0.11)	0.56	4.7
Total	90 (1.0)	0.87	62 (1.0)	2.93	3.4	64 (1.0)	5.07	5.8

<sup>a</sup> That is, the number of mutants containing *GFP*-inactivating mutations identified by DNA sequencing.

<sup>b</sup> That is, the proportion of mutants identified by DNA sequencing containing a specific type of *GFP*-inactivating mutation (e.g., the proportion of mutants with PBS deletions for wild-type RT is  $9 \div 90 = 0.10$ ).

<sup>c</sup> Mutant frequencies were determined by multiplying the proportion of mutants sequenced by the overall mutant frequency (e.g., the mutant frequency for PBS deletions for wild-type RT is  $0.10 \times 0.87 = 0.09$ ).

<sup>d</sup> That is, the fold increase in mutant frequency for mutant RTs relative to the wild-type RT (e.g., the relative increase in the frequency of PBS deletions for S557A is  $0.82 \div 0.09 = 9.1$ ).

<sup>e</sup> Nonfluorescent clones containing deletion mutations between the PBS and *GFP* or IRES.

<sup>f</sup> Nonfluorescent clones containing deletion mutations other than PBS deletions.

<sup>g</sup> Nonfluorescent clones containing duplicated sequences.

<sup>h</sup> Total of all nonfluorescent clones containing template-switching mutations, CPBS deletions, other deletions, and duplications.

rates and a sufficiently high virus titer to permit the experiment to be performed. We used *GFP* as a reporter gene because it is smaller than the *lacZ* gene (and therefore amenable to PCR amplification) and because it affords easier isolation of cells expressing inactivated *GFP* genes by FACS. MP-1 is an MLV-based vector that expresses *GFP* and *neo* from the LTR promoter; the *neo* open reading frame is translated from an IRES. A3GFP11 is a cell line that expresses the amphotropic MLV envelope and contains a single MP-1 provirus (44). The wild-type or S557A *gag-pol*-expressing construct was cotransfected with pSV $\alpha$ 3.6 into the A3GFP11 cell line (Fig. 1). The virus produced was then used to infect D17 target cells, and G418-resistant colonies were selected. The frequency of *GFP* inactivation was determined by examining the G418-resistant colonies by fluorescence microscopy. The frequency of *GFP* inactivation by wild-type RT was  $0.87\% \pm 0.31\%$ , which is comparable to the results obtained previously (44). In contrast, S557A mutant RT exhibited a 3.4-fold higher frequency of *GFP* inactivation, which was not statistically different from that obtained by using *lacZ* as a reporter gene (3.3-fold higher; Table 1).

To determine the nature of the mutations introduced into *GFP* by the wild-type or S557A mutant RT, individual nonfluorescent clones were isolated by FACS, the mutated *GFP* sequences were amplified by PCR, and their DNA sequences were determined as previously described (44). However, in the present study we amplified not only the *GFP* gene but also the region between the 5'LTR and *neo* gene of the MP-1 provirus, which includes the *GFP* open reading frame, with different sets of primers. This strategy enabled us to amplify provirus-specific DNA from the majority of nonfluorescent clones because the minimal neomycin-resistant integrated provirus would require both the promoter in the 5'LTR and *neo* expression. To rule out disproportionate clonal expansion, clones from the same infection possessing the same inactivating mutations were counted only once even though this strategy could have underestimated potential mutational hotspots. Taken together, this approach provided an accurate representation of the spec-

trum of *GFP*-inactivating mutations introduced by wild-type and mutant RTs.

**S557A mutant is associated with an increase in deletion mutations between PBS and *GFP* or IRES.** DNA sequencing of amplified proviral DNA from nonfluorescent clones showed that both S557A and wild-type RTs introduced a wide range of mutations including deletions, duplications, substitutions, and frameshifts (Table 2). The relative changes in mutant frequencies for frameshift and substitution mutations introduced by the S557A RT compared to wild-type RT were modest (0.8- and 3.0-fold, respectively). In contrast, mutations associated with RT template-switching events (deletions and duplications) accounted for the largest increase in mutant frequency by the S557A mutant RT relative to wild-type RT (1.90% versus 0.44%, representing a 4.3-fold increase). Of the template-switching mutations introduced by the S557A mutant, the highest increase in relative mutant frequency was for deletions, which ranged from 963 to 2,222 bp in length, between the PBS and sequences in *GFP* or IRES (an 8.8-fold increase relative to wild-type RT), which was a much higher increase than for the other classes of mutations that had only 0.8- to 4.6-fold changes in the relative mutant frequency.

Sequencing data showed that, although the 5'-end junctions of the PBS deletions contained various sequences, they comprised sequences within a 24-nt region ranging from a few nucleotides upstream of the start of the PBS up to and including the last nucleotide of the 18-nucleotide (nt) PBS (Fig. 2). This result indicated that the PBS region is a hotspot for template-switching mutations associated with the S557A mutant RT and that incomplete copying of the PBS followed by an error-prone plus-strand transfer could be involved in generating the mutations. In contrast, even though the 3'-end junctions of the deletions also varied, they contained sequences from a region more than 1,300 bp in length, ranging from the start of the *GFP* open reading frame to the end of the IRES, which suggested that no specific acceptor sequences were required for the template switch. Interestingly, 35% of the dele-

Clone	Junction in PBS	Junction in GFP or IRES	Length of deletion (bp)
WT PBS	<sup>562</sup> TCTTTCATT <b>TGGGGGCTCGTCCGGGAT</b> <sup>568</sup>		
WT	17-4	<sup>2781</sup> TGGGCG	2,195
	13-6	<sup>1968</sup> TCAGAC	1,390
	22-1	<sup>2426</sup> TCTTGA	1,848
	14-16	<sup>2235</sup> GCCGCT	1,659
	5-14*	<sup>2762</sup> GATGAT	2,186
	10-5	<sup>2507</sup> AAGCCA	1,936
	4-21	<sup>2639</sup> GTACCC	2,070
	10-9	<sup>1655</sup> ACCTTC	1,087
	2-13	<sup>2207</sup> TCCTCC	1,645
	S557A	6-24*	<sup>1700</sup> ATGAAG
7-11*		<sup>1876</sup> TCTCGG	1,287
10-16		<sup>2795</sup> GAGGCG	2,209
7-22		<sup>1648</sup> GGTCAC	1,063
2-15		<sup>2795</sup> CGGATC	2,216
10-10		<sup>2118</sup> ACATGG	1,534
1-22		<sup>2596</sup> ATGGCT	2,007
10-9		<sup>2803</sup> TCGTAT	2,222
2-13		<sup>2040</sup> GCCTCG	1,460
10-8		<sup>1906</sup> CTCCCA	1,327
5-10		<sup>2206</sup> CTCCCC	1,628
4-14		<sup>2585</sup> GTCAAA	2,007
7-6		<sup>2533</sup> AAAGGC	1,956
14-23	<sup>2776</sup> TATGGA	2,200	
1-4	<sup>1533</sup> TGAATG	963	
5-5	<sup>2609</sup> TTCAAC	2,043	
9-1*	<sup>1984</sup> GGATGG	1,418	
Y586F	33-14	<sup>1772</sup> GATGAC	1,183
	31-10	<sup>2307</sup> TGAGGG	1,719
	34-6	<sup>2270</sup> ATGTTA	1,683
	8-25*	<sup>2319</sup> AACCTG	1,730
	6-7	<sup>1566</sup> GAAGGT	982
	31-15	<sup>1783</sup> aTACAA	1,199
	8-23	<sup>1474</sup> CAAGGG	891
	34-16	<sup>2677</sup> CACATG	2,094
	4-9	<sup>2635</sup> CGACGG	1,453
	33-20*	<sup>2778</sup> AGCTTG	2,195
	6-1	<sup>1920</sup> TGGCGG	1,345
	4-8	<sup>2796</sup> CGGGAT	2,216
	5-32	<sup>2625</sup> AGGATG	2,043
	5-4	<sup>1990</sup> ATCCGT	1,410
	32-15	<sup>2720</sup> GCCCCC	2,141
	32-3	<sup>1656</sup> CCGTCA	1,078
	32-2	<sup>2258</sup> GCCTTT	1,682
	8-15*	<sup>2792</sup> GTCGAG	2,218
	7-36	<sup>1810</sup> TGACAC	1,445
	31-8	<sup>2083</sup> CGCCCT	1,512
6-31	<sup>2007</sup> ACCATT	1,460	
33-5	<sup>2592</sup> GGCTCT	2,023	
33-9	<sup>1593</sup> TCACCC	1,031	

FIG. 2. Sequence analysis of deletion mutations between the PBS and GFP or IRES. The complete wild-type (WT) PBS sequence (shaded nucleotides) with flanking sequences is shown compared to the deletion junctions of the nine wild-type, 17 S557A, and 23 Y586F nonfluorescent clones with deletions between the PBS and GFP or IRES. The number of deleted nucleotides for each clone is indicated in the last column. Short direct repeats at the deletion junctions are shown in boldface. Nucleotide numbers above the sequences refer to the number beginning at the start of the 5'LTR. Seven clones with an insertion from a different part of the provirus between the junctions are shown with an asterisk after the clone number. Clone numbers were designated by the number of the infection plate followed by the number of the clone.

tion junctions contained no homology, whereas the rest had a homology of <4 nt. Taken together, the sequencing data indicated that the increase in the mutation rate associated with the S557A mutant RT was largely due to a unique class of tem-

plate-switching mutations involving deletions between the PBS and sequences downstream that were a result of error-prone plus-strand DNA transfer.

**The Y586F mutant is also associated with deletion mutations between PBS and GFP or IRES.** Our previous analysis of GFP-inactivating mutations associated with the Y586F mutant concentrated on mutations introduced within the GFP gene by using a single set of primers at the 5' and 3' ends of the gene for PCR amplification. This approach could have missed the GFP-inactivating mutations involving the PBS. To determine whether deletion mutations involving the PBS or other types of mutations are associated with the Y586F mutant, we used the PCR primers in the 5'LTR and neo to amplify provirus-specific DNA from Y586F mutant-infected cells. Similar to S557A mutant analysis, this approach resulted in amplification of provirus-specific DNA from a majority of Y586F-infected non-fluorescent clones. DNA sequencing of the PCR products also showed that, similar to the S557A mutant, mutations associated with the RT template-switching events accounted for the largest increase in relative mutant frequency relative to wild-type RT (3.19% versus 0.44%, representing a 7.3-fold increase). Furthermore, deletions between the PBS and GFP or IRES accounted for the largest increase in mutant frequency among all of the template-switch mutations, representing a 20.3-fold increase in the relative mutant frequency. In contrast, the other classes of mutations had modest increases in relative mutant frequency of 3.6- to 4.7-fold. Examination of the deletion junctions showed that they were similar to those associated with the S557A mutant, suggesting that they were derived by a similar mechanism (Fig. 2).

**Both Y586F and S557A increase the frequency of substitution mutations near adenine-thymine tracts (A tracts).** We previously showed that the Y586F mutation increases the frequency of substitution mutations in regions associated with nucleotide sequences AAAA, TTTT, or AATT, known as A tracts (44). A tracts are known to induce bends in DNA (5). Analysis of substitution mutations introduced by the Y586F mutant in the present study showed that 22 of 32 substitution mutations (69%) were within 18 nt of A tracts (the total includes 17 mutants from Table 2 and 15 mutants from a second separate experiment [data not shown]). The distance between the polymerase active site and the RNase H cleavage site in an HIV-1 crystal structure in complex with a template-primer and dNTP substrate is 18 nt (16). If the substitution mutations were distributed randomly, we would expect 39% to be located within 18 nt of an A tract in GFP (278 of 717 nt); thus, the 69% frequency is significantly higher than that expected by random distribution ( $P < 0.0003$ ;  $\chi^2$  test) and is similar to the 81% observed previously (44). Analysis of substitution mutations induced by the S557A mutant showed that 60% of the mutations were within 18 nt of an A tract (12 of 20 mutations), which is also significantly higher than that expected by random distribution ( $P = 0.05$ ;  $\chi^2$  test). In contrast, the proportion of substitution mutations associated with A tracts with wild-type RT was 50% (16 of 32 mutations), which is not statistically different from that expected by random distribution ( $P = 0.2$ ;  $\chi^2$  test). Therefore, these results indicate that the Y586F and S557A mutants increased the frequency of substitution mutations near A tracts.

**Quantitative real-time PCR analysis of viral DNA synthesis by S557A and Y586F mutant viruses.** We hypothesized that the incomplete copying of the PBS by both S557A and Y586F mutant RTs was due to an inability of the mutant RTs to copy through the DNA-RNA junction of the minus-strand DNA and the tRNA because of its bent conformation (32). We postulated that the incomplete copying was followed by an error-prone plus-strand DNA transfer, resulting in deletions between the PBS and *GFP* or IRES. To test this hypothesis, we analyzed the effects of RNase H primer grip mutations on initiation of viral DNA synthesis. This step involves the RNase H primer grip passing through a RNA-DNA junction. In contrast to the junction encountered during plus-strand strong-stop DNA synthesis, this junction is part of the primer strand and is formed during the initiation process by the extension of the tRNA primer. However, the RNase H primer grip has to pass through it because it lags behind the polymerase active site. In addition, we also investigated the effects of RNase H primer grip mutations on plus- and minus-strand DNA transfer steps. Wild-type or mutant *gag-pol*-expressing constructs were cotransfected with the envelope plasmid pHCMV-G into GN-MLV-GFP cells, which contain an integrated copy of an MLV-based provirus. Virus was collected 48 h after transfection and used to infect target 293T cells. We monitored initiation of viral DNA synthesis and DNA strand-transfer events by analyzing products of reverse transcription at 1, 3, and 6 h postinfection by quantitative real-time PCR assay. Initiation of DNA synthesis was observed by analyzing early reverse transcription products with an RU5 primer and probe set (Fig. 3A). RNA was isolated from a fraction of the viral preparation used for infection, quantified by RT-PCR, and used to normalize the amount of RU5 product. The amount of RU5 product from cells infected with mutant virus was then compared to the amount from cells infected with wild-type virus (set at 100%) (Fig. 3B). This analysis showed that initiation of DNA synthesis was not significantly affected by the S557A and Y586F mutations.

Primer and probe sets in the *hygro* and the U5-Ψ region were used to detect products after minus- and plus-strand DNA transfer, respectively (Fig. 3A). To measure minus-strand DNA transfer efficiency, the amount of *hygro* DNA synthesized was compared to the amount of RU5 DNA (Fig. 3C). To measure plus-strand DNA transfer efficiency, the amount of U5-Ψ DNA synthesized was compared to the amount of *hygro* DNA (Fig. 3D). We measured strand transfer efficiency up to 6 h postinfection because our previous results have indicated that the majority of reverse transcription events occur within the first 6 h (42). The analysis showed that minus-strand DNA transfer was less efficient at 1 and 3 h postinfection for both S557A and Y586F mutant viruses by as much as three- and ninefold, respectively, relative to wild-type virus, but was similar for all three viruses at 6 h postinfection. This result indicates that minus-strand DNA transfer occurred slowly by the mutant viruses but was as efficient as that by wild-type virus later during infection. In contrast, plus-strand DNA transfer was significantly less efficient for both S557A and Y586F by 22- and 18-fold at 3 h postinfection and by five- and threefold at 6 h postinfection, respectively. The significant deficiency in plus-strand DNA transfer by the mutant viruses by using the quantitative real-time PCR assay was consistent with the mu-

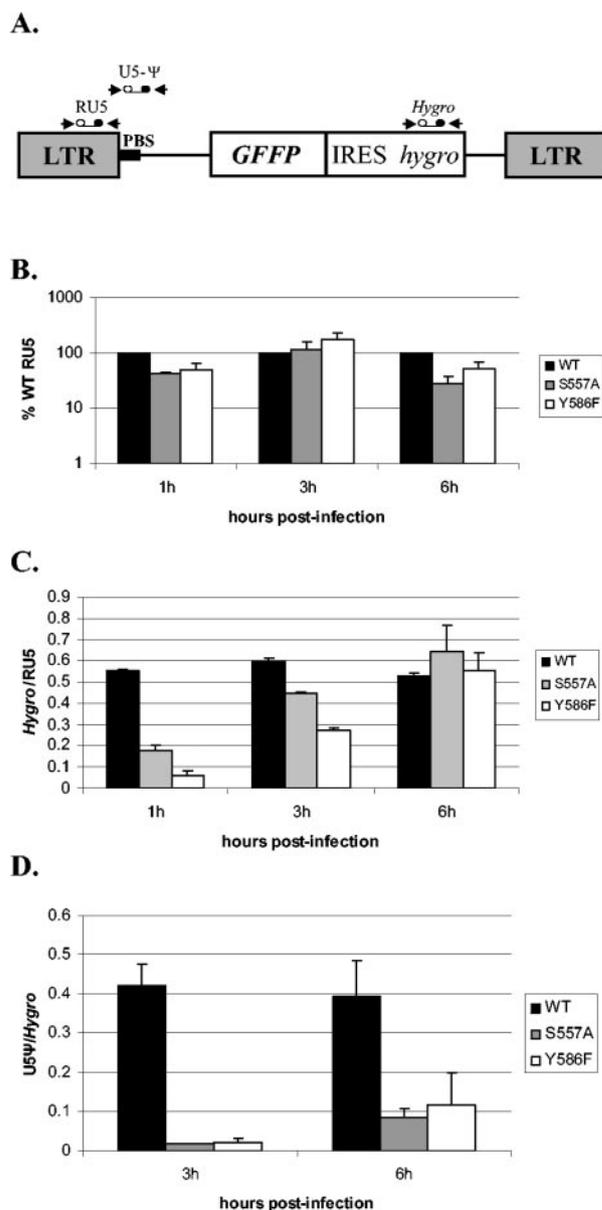


FIG. 3. Effect of S557A and Y586F mutations on viral DNA synthesis. (A) Locations of real-time PCR primer-probe sets used are shown above the structure of the MLV-GFP vector. (B) Initiation of viral DNA synthesis by mutant viruses was determined by analyzing the amount of RU5 product at 1, 3, and 6 h postinfection relative to wild-type virus (set at 100%). 293T target cells were infected by virus produced by transfection of the GN-MLV-GFP helper cell line. The amount of RU5 product was normalized to the amount of viral RNA in each viral preparation used for infection, which was determined by RT-PCR. Two independent experiments were analyzed, and each DNA sample was analyzed by real-time PCR twice. (C) Efficiency of minus-strand DNA transfer by mutant viruses was determined by analyzing the amount of *hygro* gene product relative to RU5 product at 1, 3, and 6 h postinfection for each virus relative to wild-type virus. The experiments were carried out as described in panel B. (D) The efficiency of plus-strand DNA transfer by mutant viruses was determined by analyzing the amount of U5-Ψ product relative to *hygro* gene product at 3 and 6 h postinfection for each virus relative to wild-type virus. The experiments were carried out as described in panel B. The error bars in all of the graphs represent standard error of the mean.

tation analysis data showing a defect in plus-strand DNA transfer by both S557A and Y586F mutants.

## DISCUSSION

The data presented in this study confirm and extend our previous finding that the MLV RNase H primer grip domain plays a significant role in viral replication fidelity (44). Using *lacZ* or *GFP* as the reporter gene, we found that mutation of three of the six residues that constitute the domain resulted in 2.1- to 3.8-fold increases in the *in vivo* mutation rates. Although most of the mutations in the RNase H primer grip resulting in an increase in the mutation rate also showed significant decreases in virus titers, we do not believe that the former is a consequence of the latter. We have previously reported mutations in RT that significantly reduced virus titers but did not affect mutation rates (14). DNA sequencing analysis of *GFP*-inactivating mutations introduced by the mutant RTs showed that the increases in the mutation rate were primarily due to an increase in the frequency of deletions between the PBS and *GFP* or IRES. In the present study, we amplified provirus-specific DNA from nonfluorescent clones by using primers between the 5'LTR and *neo*. Because expression of genes from the provirus would require the 5'LTR promoter and expression of *neo* to overcome G418 selection, this strategy enabled DNA amplification from the smallest possible neomycin-resistant integrated provirus and resulted in amplification of DNA from >90% of nonfluorescent clones. Thus, this approach provided an accurate representation of the spectrum of *GFP*-inactivating mutations introduced by wild-type and mutant RTs.

We observed that the increase in the frequency of deletions between the PBS and *GFP* or IRES in mutant RT viruses was the result of an error-prone plus-strand DNA transfer that is often preceded by an incomplete copying of the tRNA sequences that anneal to the PBS (see proposed model, Fig. 4). The incomplete copying of the tRNA primer sequences could be due to the inability of the RT to copy through the DNA-RNA junction of the minus-strand DNA and the tRNA because of its bent conformation (32). The premature termination may occur after copying through the DNA-RNA junction because the RNase H primer grip lags behind the polymerase active site and may be unable to progress through the bend in the duplex at the DNA-RNA junction. The global structure of the duplex formed by the chimeric DNA-tRNA strand and the cDNA strand of both MLV and HIV-1 has been shown to be significantly distorted (9, 32). The duplex assumes an H-form structure at the DNA:RNA hybrid portion and a B-form structure at the DNA:DNA end, resulting in structural discontinuity at the junction. In the MLV duplex, this discontinuity causes a change in the direction of the helix with a bend of  $18 \pm 3^\circ$ , a large negative buckle at the junction base-pair step T5 · A14-T6 · a13, and a gradual increase in the minor groove width from the DNA section toward the hybrid section (32). Plus-strand DNA transfer is then error-prone, possibly due to a decrease in the length of homology between the plus-strand strong-stop DNA and the complementary minus-strand DNA region or due to weaker interactions between the mutant RT and the template-primer. Although our proposed model depicts minus-strand DNA transfer to occur to the copackaged

RNA that has not initiated DNA synthesis, it is possible that minus-strand DNA transfer could occur intramolecularly or to a copackaged RNA that has initiated DNA synthesis. In this case, minus-strand DNA synthesis could continue not by pairing RU5 sequences but by pairing the complete or incomplete PBS sequence at the deletion junction with the PBS sequence in the minus-strand DNA. Alternatively, the deletions could form as a result of incomplete minus-strand DNA synthesis (for example, due to slow processivity of mutant RTs), forcing the plus-strand strong-stop DNA to anneal to the 3' end of the prematurely terminated minus-strand DNA resulting in deletion of the region between the PBS and the termination point of minus-strand DNA synthesis (Fig. 4). Both mechanisms would lead to errors occurring during plus-strand DNA transfer.

Our previous finding showed an increased frequency of substitution mutations by the Y586F mutant near A tracts (44). Similar to DNA-RNA junctions, A tracts are associated with a bent DNA conformation at their junction with a G/C base pair, which results in a narrowed minor groove (5). Analysis of substitution mutations associated with the Y586F and S557A mutants in the present study also showed an increased frequency of substitution mutations within 18 nt of A tracts relative to wild-type RT. This result indicates that MLV RNase H primer grip mutants are unable to induce a proper conformation of the template-primer duplex at the polymerase active site when a bent conformation is present, resulting in error-prone replication.

Our hypothesis that error-prone plus-strand DNA transfer is responsible for the increase in deletion mutations between the PBS and *GFP* or IRES was tested further by comparing the efficiency of plus-strand DNA transfer by the mutant and wild-type RTs, by using a quantitative real-time PCR assay. This analysis showed that the efficiency of accurate plus-strand DNA transfer was significantly decreased in mutant RT viruses relative to wild-type virus, which is consistent with the observed increase in deletions. In contrast, the efficiency of minus-strand DNA transfer, although diminished in mutant viruses at early time points, was similar to that of wild-type virus by 6 h post-infection. It is possible that the longer region of homology in minus-strand strong-stop DNA compared to plus-strand strong-stop DNA makes minus-strand DNA transfer more efficient than plus-strand DNA transfer in mutant viruses despite the weaker RT and template-primer interactions.

RT encounters another RNA-DNA junction during initiation of viral DNA synthesis, which is formed by the tRNA primer and nascent DNA strand. In this case, the hybrid duplex consists of the chimeric tRNA-DNA strand and the template RNA strand. In contrast to the significant effect of RNase H primer grip mutations on plus-strand DNA transfer, their effect on initiation of DNA synthesis was insignificant relative to wild-type virus. One explanation for this difference could be that the RNA-DNA junction is formed during initiation of minus-strand DNA synthesis and becomes part of the primer strand, whereas the DNA-RNA junction during plus-strand strong-stop DNA synthesis is part of the template strand. Alternatively, the MLV RNase H primer grip mutants could be better at copying through the DNA-RNA junction when it is in a duplex with an RNA strand rather than a DNA strand. The crystal structure of the HIV-1 hybrid duplex formed by a chimeric DNA-tRNA strand and template RNA strand has been

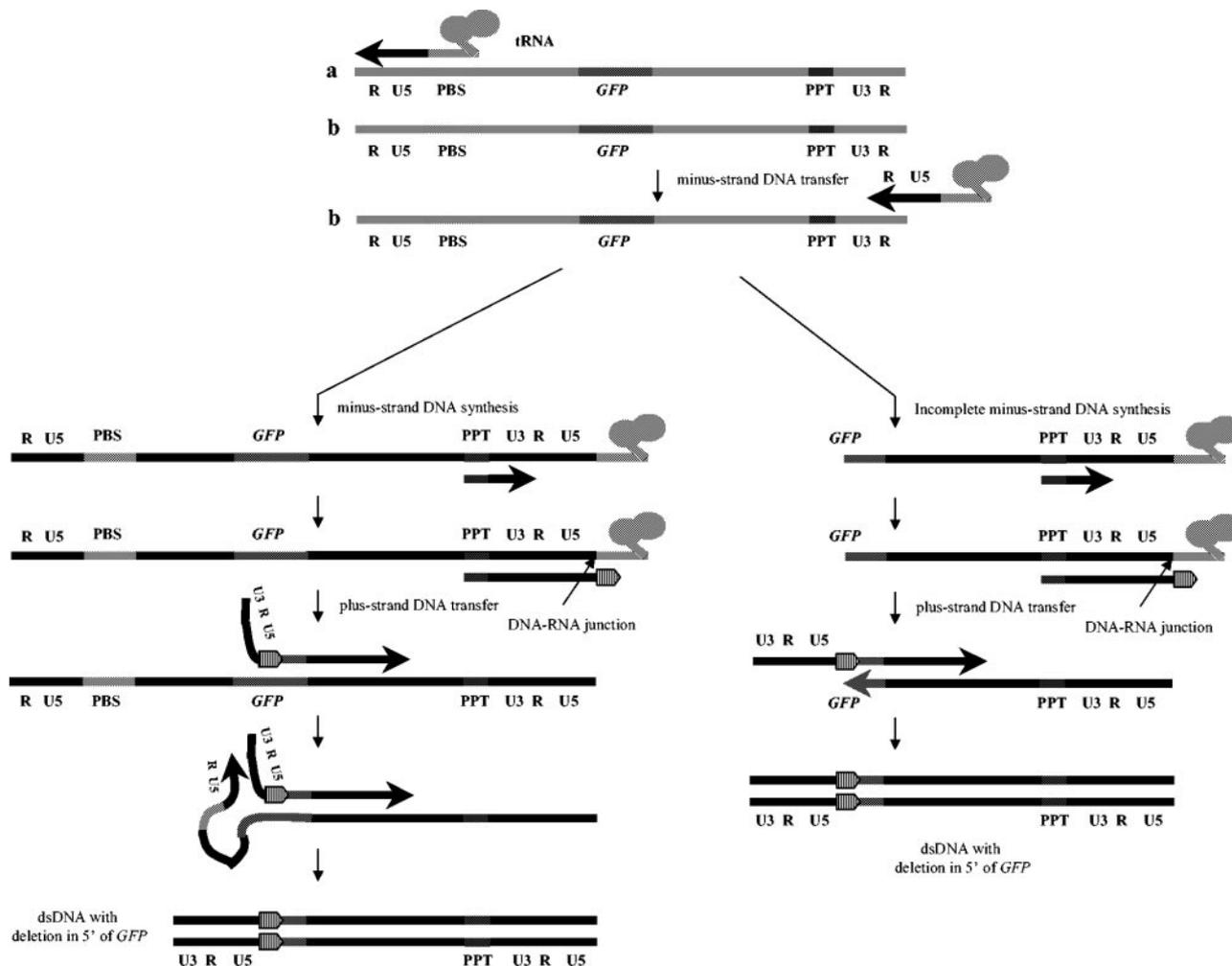


FIG. 4. Proposed mechanism for the formation of deletions between the PBS and *GFP* or IRES. Initiation of DNA synthesis is proposed to start from only one of the two copackaged RNA genomes labeled a. Minus-strand DNA transfer then occurs onto the copackaged RNA labeled b, which did not initiate reverse transcription, resulting in minus-strand DNA that has R, U5, and PBS at its 3' end. Plus-strand DNA synthesis initiates from the PPT as expected but fails to copy through the DNA-RNA junction of the minus-strand DNA and the tRNA because of its bent conformation, resulting in incomplete copying of the PBS (stripped arrowhead) and error-prone plus-strand DNA transfer into the *GFP* gene or IRES because of decreased homology and weaker RT-template interactions. The R and U5 regions from the minus-strand DNA then hybridize with the complementary R and U5 regions from the plus-strand strong-stop DNA, forming a loop from the PBS into *GFP* or IRES that is excised and repaired by the host DNA repair machinery. The completion of plus-strand DNA synthesis downstream of *GFP* or IRES and minus-strand DNA synthesis upstream of the R region results in the formation of a double-stranded DNA with a deletion between the PBS and *GFP* or IRES that is subsequently integrated into the host genome to form a deleted provirus. Alternatively, error-prone plus-strand DNA transfer could occur because of incomplete minus-strand DNA synthesis, which would also result in deletions between the PBS and the sequences downstream of the premature minus-strand DNA synthesis termination point. RNA molecules are shown in gray, whereas DNA molecules are shown in black.

reported to have an A-form structure with minor structural perturbations at the r(cpa)·d(TpG) base pair step (28). Intriguingly, HIV-1 RNase H primer grip mutants have been shown to significantly affect initiation of viral DNA synthesis but not plus-strand DNA transfer (17). The fact that HIV-1 RT is a heterodimer, whereas MLV RT is a monomer and that the residues constituting the HIV-1 RNase H primer grip reside in both the p66 and p51 subunits (7, 34), could influence the nature of RT and template-primer interactions by the two viruses. The MLV RT structure has recently been reported (7), but unfortunately the RNase H domain is not sufficiently resolved to directly examine whether the contacts between the RNase H primer grip residues and the DNA primer strand are similar to

those observed in HIV-1. The mutations in the RNase H primer grip could also indirectly affect viral DNA synthesis by affecting RNase H activity. However, in vitro studies have shown that mutations at A558 and Q559 of MLV RT do not affect its RNase H activity (1, 2). In contrast, the RNase H activity of the Y586F mutant is reduced by 20-fold. The effect on RNase H activity of mutations at S557 has not been determined.

In summary, our findings indicate that the MLV RNase H primer grip domain plays a significant role in dealing with structural constraints in the template-primer complex introduced by A tracts or DNA-RNA junctions. This role not only ensures proper positioning of the template-primer at the polymerase active site, resulting in incorporation of the correct

nucleotides, but also helps in preventing premature dissociation of the RT from the template-primer duplex, thereby facilitating efficient and accurate DNA synthesis.

#### ACKNOWLEDGMENTS

We especially thank Wei-Shau Hu for intellectual input throughout the project and Anne Arthur for expert editorial help. We also thank Patricia Henry, Sook-Kyung Lee, and Evgenia Svarovskaia for critical reading of the manuscript.

This study was supported by the HIV Drug Resistance Program, National Cancer Institute.

#### REFERENCES

- Blain, S. W., and S. P. Goff. 1995. Effects on DNA-synthesis and translocation caused by mutations in the RNase-H domain of Moloney murine leukemia virus reverse transcriptase. *J. Virol.* **69**:4440–4452.
- Blain, S. W., and S. P. Goff. 1993. Nuclease activities of Moloney murine leukemia virus reverse transcriptase: mutants with altered substrate specificities. *J. Biol. Chem.* **268**:23585–23592.
- Chowdhury, K., N. Kaushik, V. N. Pandey, and M. J. Modak. 1996. Elucidation of the role of Arg 110 of murine leukemia virus reverse transcriptase in the catalytic mechanism: biochemical characterization of its mutant enzymes. *Biochemistry* **35**:16610–16620.
- Coffin, J. M. 1995. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science* **267**:483–489.
- Crothers, D. M., T. E. Haran, and J. G. Nadeau. 1990. Intrinsically bent DNA. *J. Biol. Chem.* **265**:7093–7096.
- Dang, Q., and W. S. Hu. 2001. Effect of homology length in the repeat region on minus-strand DNA transfer and retroviral replication. *J. Virol.* **75**:809–820.
- Das, D., and M. M. Georgiadis. 2004. The crystal structure of the monomeric reverse transcriptase from Moloney murine leukemia virus. *Structure* **12**:819–829.
- Delviks, K. A., W. S. Hu, and V. K. Pathak. 1997. Psi(−) vectors: murine leukemia virus-based self-inactivating and self-activating retroviral vectors. *J. Virol.* **71**:6218–6224.
- Fedoroff, O. Y., M. Salazar, and B. R. Reid. 1996. Structural variation among retroviral primer-DNA junctions: solution structure of the HIV-1 negative-strand Okazaki fragment r(ccca)d(ATGA)·d(TCATTGGG). *Biochemistry* **35**:11070–11080.
- Fu, W., and W. S. Hu. 2003. Functional replacement of nucleocapsid flanking regions by heterologous counterparts with divergent primary sequences: effects of chimeric nucleocapsid on the retroviral replication cycle. *J. Virol.* **77**:754–761.
- Gaschen, B., J. Taylor, K. Yusim, B. Foley, F. Gao, D. Lang, V. Novitsky, B. Haynes, B. H. Hahn, T. Bhattacharya, and B. Korber. 2002. AIDS: diversity considerations in HIV-1 vaccine selection. *Science* **296**:2354–2360.
- Goulder, P. J. R., C. Brander, Y. H. Tang, C. Tremblay, R. A. Colbert, M. M. Addo, E. S. Rosenberg, T. Nguyen, R. Allen, A. Trocha, M. Altfeld, S. Q. He, M. Bunce, R. Funkhouser, S. I. Pelton, S. K. Burchett, K. McIntosh, B. T. M. Korber, and B. D. Walker. 2001. Evolution and transmission of stable CTL escape mutations in HIV infection. *Nature* **412**:334–338.
- Halvas, E. K., E. S. Svarovskaia, and V. K. Pathak. 2000. Development of an in vivo assay to identify structural determinants in murine leukemia virus reverse transcriptase important for fidelity. *J. Virol.* **74**:312–319.
- Halvas, E. K., E. S. Svarovskaia, and V. K. Pathak. 2000. Role of murine leukemia virus reverse transcriptase deoxyribonucleoside triphosphate-binding site in retroviral replication and in vivo fidelity. *J. Virol.* **74**:10349–10358.
- Harris, D., P. N. S. Yadav, and V. N. Pandey. 1998. Loss of polymerase activity due to Tyr to Phe substitution in the YMDD motif of human immunodeficiency virus type-1 reverse transcriptase is compensated by Met to Val substitution within the same motif. *Biochemistry* **37**:9630–9640.
- Huang, H. F., R. Chopra, G. L. Verdine, and S. C. Harrison. 1998. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **282**:1669–1675.
- Julias, J. G., M. J. McWilliams, S. G. Sarafianos, E. Arnold, and S. H. Hughes. 2002. Mutations in the RNase H domain of HIV-1 reverse transcriptase affect the initiation of DNA synthesis and the specificity of RNase H cleavage in vivo. *Proc. Natl. Acad. Sci. USA* **99**:9515–9520.
- Kawai, S., and M. Nishizawa. 1984. New procedure for DNA transfection with polycation and dimethyl sulfoxide. *Mol. Cell. Biol.* **4**:1172–1174.
- Kent, R. B., J. R. Emanuel, Y. Benneriah, R. Levenson, and D. E. Housman. 1987. Ouabain resistance conferred by expression of the cDNA for a murine Na<sup>+</sup>, K<sup>+</sup>-ATPase alpha-subunit. *Science* **237**:901–903.
- Kim, T., R. A. Mudry, Jr., C. A. Rexrode II, and V. K. Pathak. 1996. Retroviral mutation rates and A-to-G hypermutations during different stages of retroviral replication. *J. Virol.* **70**:7594–7602.
- Malim, M. H., and M. Emerman. 2001. HIV-1 sequence variation: drift, shift, and attenuation. *Cell* **104**:469–472.
- Mansky, L. M. 2000. In vivo analysis of human T-cell leukemia virus type 1 reverse transcription accuracy. *J. Virol.* **74**:9525–9531.
- Mansky, L. M., and L. C. Bernard. 2000. 3'-Azido-3'-deoxythymidine (AZT) and AZT-resistant reverse transcriptase can increase the in vivo mutation rate of human immunodeficiency virus type 1. *J. Virol.* **74**:9532–9539.
- Mansky, L. M., E. Le Rouzic, S. Benichou, and L. C. Gajary. 2003. Influence of reverse transcriptase variants, drugs, and Vpr on human immunodeficiency virus type 1 mutant frequencies. *J. Virol.* **77**:2071–2080.
- Mellors, J. W., H. Z. Bazmi, R. F. Schinazi, B. M. Roy, Y. Hsiou, E. Arnold, J. Weir, and D. L. Mayers. 1995. Novel mutations in reverse transcriptase of human immunodeficiency virus type 1 reduce susceptibility to foscarnet in laboratory and clinical isolates. *Antimicrob. Agents Chemother.* **39**:1087–1092.
- Miller, A. D., and C. Buttimore. 1986. Redesign of retrovirus packaging cell-lines to avoid recombination leading to helper virus production. *Mol. Cell. Biol.* **6**:2895–2902.
- Moore, C. B., M. John, I. R. James, F. T. Christiansen, C. S. Witt, and S. A. Mallal. 2002. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science* **296**:1439–1443.
- Mueller, U., G. Maier, A. M. Onori, L. Cellai, H. Heumann, and U. Heinemann. 1998. Crystal structure of an eight-base pair duplex containing the 3'-DNA-RNA-5' junction formed during initiation of minus-strand synthesis of HIV replication. *Biochemistry* **37**:12005–12011.
- Nikolenko, G. N., E. S. Svarovskaia, K. A. Delviks, and V. K. Pathak. 2004. Antiretroviral drug resistance mutations in human immunodeficiency virus type 1 reverse transcriptase increase template-switching frequency. *J. Virol.* **78**:8761–8770.
- Pathak, V. K., and H. M. Temin. 1990. Broad-spectrum of in vivo forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: deletions and deletions with insertions. *Proc. Natl. Acad. Sci. USA* **87**:6024–6028.
- Rausch, J. W., D. Lener, J. T. Miller, J. G. Julias, S. H. Hughes, and S. F. J. Le Grice. 2002. Altering the RNase H primer grip of human immunodeficiency virus reverse transcriptase modifies cleavage specificity. *Biochemistry* **41**:4856–4865.
- Salazar, M., O. Y. Fedoroff, and B. R. Reid. 1996. Structure of chimeric duplex junctions: solution conformation of the retroviral Okazaki-like fragment r(ccca)d(AATGA)·d(TCATTGGG) from Moloney murine leukemia virus. *Biochemistry* **35**:8126–8135.
- Sambrook, J., and D. W. Russell. 2001. *Molecular cloning: a laboratory manual*, 3rd ed. vol. 3. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Sarafianos, S. G., K. Das, C. Tantillo, A. D. Clark, J. Ding, J. M. Whitcomb, P. L. Boyer, S. H. Hughes, and E. Arnold. 2001. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J.* **20**:1449–1461.
- Shankar, P., M. Russo, B. Harnisch, M. Patterson, P. Skolnik, and J. Lieberman. 2000. Impaired function of circulating HIV-specific CD8<sup>+</sup> T cells in chronic human immunodeficiency virus infection. *Blood* **96**:3094–3101.
- Stebbing, J., S. Patterson, and F. Gotch. 2003. New insights into the immunology and evolution of HIV. *Cell Res.* **13**:1–7.
- Svarovskaia, E. S., S. R. Cheslock, W. H. Zhang, W. S. Hu, and V. K. Pathak. 2003. Retroviral mutation rates and reverse transcriptase fidelity. *Front. Biosci.* **8**:D117–D134.
- Svarovskaia, E. S., K. A. Delviks, C. K. Hwang, and V. K. Pathak. 2000. Structural determinants of murine leukemia virus reverse transcriptase that affect the frequency of template switching. *J. Virol.* **74**:7171–7178.
- Temin, H. M. 1993. Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. *Proc. Natl. Acad. Sci. USA* **90**:6900–6903.
- van der Groen, G., P. N. Nyambi, E. Beirnaert, D. Davis, K. Franssen, L. Heyndrickx, P. Ondoa, G. Van der Auwera, and W. Janssens. 1998. Genetic variation of HIV type 1: relevance of interclade variation to vaccine development. *AIDS Res. Hum. Retrovir.* **14**(Suppl. 3):S211–S221.
- Voronin, Y. A., and V. K. Pathak. 2004. Frequent dual initiation in human immunodeficiency virus-based vectors containing two primer-binding sites: a quantitative in vivo assay for function of initiation complexes. *J. Virol.* **78**:5402–5413.
- Voronin, Y. A., and V. K. Pathak. 2003. Frequent dual initiation of reverse transcription in murine leukemia virus-based vectors containing two primer-binding sites. *Virology* **312**:281–294.
- Yee, J. K., A. Miyanojara, P. Laporte, K. Bouic, J. C. Burns, and T. Friedmann. 1994. A general method for the generation of high-titer, pantropic retroviral vectors: highly efficient infection of primary hepatocytes. *Proc. Natl. Acad. Sci. USA* **91**:9564–9568.
- Zhang, W. H., E. S. Svarovskaia, R. Barr, and V. K. Pathak. 2002. Y586F mutation in murine leukemia virus reverse transcriptase decreases fidelity of DNA synthesis in regions associated with adenine-thymine tracts. *Proc. Natl. Acad. Sci. USA* **99**:10090–10095.