Identification of a preferred region for recombination and mutation in HIV-1 gag

Carrie Dykes, a Mini Balakrishnan, b Vicente Planelles, c Yonghong Zhu, d,1 Robert A. Bambara, b and Lisa M. Demeter a,d,*

a Department of Medicine, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA
b Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA
c Department of Pathology, University of Utah, School of Medicine, Salt Lake City, UT, USA
d Department of Microbiology and Immunology, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA

Received 19 September 2003; returned to author for revision 10 October 2003; accepted 24 February 2004

Abstract

We designed a cell culture-based system to test the hypothesis that recombination events during HIV-1 replication would be more frequent near the dimerization initiation sequence (DIS). A 459-bp region spanning the DIS through the 5'-end of gag was sequenced and analyzed to determine the frequency and distribution of crossover sites. We found a strong preference for recombination events occurring within a 112-nt-long region encompassing the gag AUG (64% of crossovers occurred in this region, compared to 10–14% in surrounding regions with similar lengths). Surprisingly, the region immediately surrounding the DIS was not a preferred site of recombination. Analysis of recombination events using RNA templates transcribed in vitro revealed a preference for crossover sites at the start of the gag coding region, similar to that observed in cell culture. This recombinogenic region was unusually G-rich and promoted extensive pausing by RT in vitro. Template features that induce RT pausing very likely contribute to the observed template switching events in gag during minus-strand synthesis. The region in gag that was a preferred site for recombination also had an approximately 2-fold higher mutation frequency compared to the rest of the region sequenced, but mutations were no more common in recombinant compared to non-recombinant clones, suggesting that recombination events were not mutagenic.

D © 2004 Elsevier Inc. All rights reserved.

Keywords: Human immunodeficiency virus-1; Reverse transcription; Recombination; Template switching; DIS

Introduction

The ability of HIV-1 to evolve rapidly in response to changes in selective pressure is a major barrier to the development of effective vaccines and treatments for this infection. Retroviral diversity is generated primarily during reverse transcription, a multi-step process in which the virally encoded reverse transcriptase (RT) generates double-stranded proviral DNA from the positive sense, single-stranded viral RNA genome (reviewed in Telesnitsky and Goff, 1997). Recombination and mutation events that occur during reverse transcription each contribute to retroviral diversity (Burke, 1997; Howell et al., 1991; Ji and Loeb, 1992; Preston et al., 1988; Quinones-Mateu et al., 2002; Roberts et al., 1988, 1995; Temin, 1993).

Retroviral recombination is possible because each virion contains two copies of the positive sense, single-stranded RNA genome (Hu and Temin, 1990a). Recombination occurs during reverse transcription when a newly synthesized DNA strand is transferred from the original template (or donor) to the second template (or acceptor), and can take place during minus-strand or plus-strand synthesis (Hu and Temin, 1990b; Stuhlmann and Berg, 1992; Yu et al., 1998). Strand transfer by RT is required for the completion of minus-strong stop and plus-strong stop DNA synthesis. At both these steps, strand transfer to the opposite end of the genome is required in order for DNA synthesis to proceed to completion. Strand transfers that are not essential for com-
pletion of reverse transcription can also occur throughout the length of the genome, during DNA polymerization by RT (Brincat et al., 2002; Delviks and Pathak, 1999; DeStefano et al., 1992; Peliska and Benkovic, 1992; Pfeiffer et al., 2000; Svarovskaia et al., 2000). When such transfer events occur between two nonidentical or genetically distinct RNA genomes, they result in a recombinant provirus.

Several tissue culture-based systems have been designed to look at retroviral recombination (Anderson et al., 1998; Bowman et al., 1998; Hu and Temin, 1990a, 1900b; Hu et al., 1997; Stuhlmann and Berg, 1992). These model systems produce heterozygous virions by stable transfection of a packaging cell line with two different parental viruses, each of which expresses a selectable marker. Recombination during reverse transcription of the two nonidentical RNA genomes results in a recombinant provirus that expresses both markers and is therefore easily selected. Studies using these model systems demonstrate that recombination occurs primarily during minus-strand synthesis, and exhibits high negative interference (Hu et al., 1997; Zhang et al., 2000). In addition, recombination events were rarely observed unless the two parental strains were co-packaged in the same virion.

Work by Dougherty et al. has emphasized the widespread occurrence of recombination throughout the HIV-1 genome (Jetzt et al., 2000; Yu et al., 1998; Zhuang et al., 2002). Using a heteroduplex tracking assay (HTA), these studies have evaluated the location and frequency of recombination events within the HIV-1 genome, during a single reverse transcription cycle. An average of three crossovers was detected per genome per replication cycle. In one study, a slightly higher frequency of crossovers was observed at the 5′-end of the pol coding region relative to the rest of the genome (Jetzt et al., 2000). In a more recent study, sequence analysis of a relatively small number of recombinants was suggestive of preferred sites of recombination in HIV (Zhuang et al., 2002).

The mechanisms leading to the promotion or inhibition of recombination at specific sites is poorly understood. Nonetheless, studies in vitro using purified components have provided insights into the strand transfer mechanism as well as template and protein factors that influence the process. RNA structures in the genome that cause pausing of the RT have been proposed to promote recombination (DeStefano et al., 1992, 1994; Diaz and DeStefano, 1996; Harrison et al., 1998; Kim et al., 1997; Palaniappan et al., 1996; Roda et al., 2002; Wu et al., 1995). Pausing was shown to enhance the RNase H activity of RT thereby facilitating template switching. We recently demonstrated in vitro, using RNA templates harboring the HIV-1 dimerization initiation sequence (DIS), that dimerization promotes recombination by a local proximity effect (Balakrishnan et al., 2001, 2003).

We wished to evaluate whether the region immediately surrounding the DIS was a preferred site for recombination in vivo. We designed a tissue culture model system to detect recombination events in the region near the DIS during reverse transcription, using env-deleted, pseudotyped HIV-1 vectors. We used two parental strains that differed by approximately 1/25 bases over a 5.1-kb region spanning the DIS through vpr, and analyzed recombination events with a 459-bp region of the provirus, spanning the DIS through the beginning of gag. The region includes hairpins and sequence features that affect RNA folding and pausing of the RT. The segment is large enough such that the effects of a wide variety of RNA sequence and structural features can be examined with respect to their relative contributions to recombination. Because of the unique features of our model system, we were able to map the location of recombination events at a high resolution. Using this model system, we have identified a preferred region for recombination near the beginning of gag, but found no preference for recombination occurring in regions closer to the DIS. Analysis of the region using a strand transfer assay system in vitro revealed transfer switching profiles within the 5′ gag region similar to what we have observed in cell culture.

Results

Experimental design

To measure the location of recombination events in the HIV-1 genome during reverse transcription, we designed two replication-defective, env-deleted HIV-1 vectors: pDHIV.3.B71 and pDHIV.3.Thy.1.JRCSF.KL9 (Fig. 1). To create a system that would allow the identification of recombinants by sequencing, we replaced the HIVNL4-3 BssHII/SalI fragment in pDHIV.3.Thy-1, encoding the gag, pol, vif, and vpr genes, with the corresponding fragment from pYK-JRCSF. Alignment of these regions from the two isolates shows an average of one substitution approximately every 25 bases, allowing fine mapping of crossover locations. In addition, the inclusion of a JRCSF marker upstream of the BssHII restriction site by mutagenesis (proviral positions 667–668) allowed for the identification of recombinants across the DIS hairpin. The human co-stimulatory gene, B7.1, and the murine Thy1.2 gene were substituted for nef in pDHIV.3.B7.1 and pDHIV.3.Thy1.JRCSF.KL9, respectively. The levels of B7.1 and Thy1 protein expressed on the cell surface were measured using flow cytometry. These design features allow one to monitor the efficiency and reproducibility of the co-infection step and obviate the need for the production of stably infected cell lines.

Fig. 2 details the experimental design used to create virus stocks from dually infected cells and to identify recombinants. This tissue culture model consists of three separate steps: (1) production of infectious homozygous parental virus stocks, (2) production of virus stocks from dually infected cells, and (3) screening for recombinant proviruses by sequence analysis of a PCR-amplified region of proviral DNA. More than a single round of infection is possible in Steps 1

and 2, because harvesting of infected cells occurs 3 days after transfection with the VSV-g expression vector, although the number of replication cycles is likely to be limited.

COS-1 cells were co-transfected with pHCMV-G (which encodes VSV-g) and either pDHIV.3.B71 or pDHIV.3.-Thy.1.JRCSF.KL9. The resulting VSV-g pseudotyped parental viruses were concentrated by centrifugation to produce high-titer virus stocks and used to co-infect fresh COS-1 cells at a multiplicity of infection (MOI) of 1.0 for each parental virus. The infectious units/ml, measured from five independent transfections, was $1.26 \times 10^5$ for DHIV.3.B7.1 and $2.66 \times 10^4$ for DHIV.3.Thy.1.JRCSF.KL9. After 24 h, the co-infected cells were transfected with pHCMV-G to produce a VSV-g pseudotyped virus stock that contained both homozygous parental virions, in which both copies of genomic RNA are either the NL4-3 or JR-CSF, as well as heterozygous virions, which contain both parental RNA genomes. The virus stocks generated from these co-infections were used to infect H9 cells at a MOI of 0.01. Genomic DNA was isolated after 48 h and the 5.1-kb region from the DIS hairpin through vpr was amplified using PCR. These PCR products were cloned into E. coli and a 459-bp region encompassing the DIS hairpin was sequenced to determine the frequency and distribution of crossovers. Clones from two separate infections using two independently generated virus stocks from dually infected cells were analyzed.

Estimating the efficiency with which heterozygous virions are produced is important because previous work had demonstrated that recombinant viruses are rarely detected after reverse transcription unless the two parental strains are co-packaged (Hu and Temin, 1990a). We analyzed the efficiency of infection at Step 2, using flow cytometry. After co-infection, approximately 40% of the cells were dually infected, and expressed both Thy1.2 and B7.1 cell surface proteins. Approximately 40% of cells were uninfected, expressing neither marker, and approximately 20% of the cells were singly infected, expressing either B7.1 or Thy.1. Experiments using isotype control antibodies demonstrated that the staining seen was specific for Thy1.2 and B7.1 (data not shown). Dual staining of singly infected cells also confirmed the specificity of antibody staining (data not shown). The relative proportion of dually infected cells was also approximately 40% in a second independent experiment. These studies demonstrate that dual infection occurred efficiently and reproducibly.
Identification of recombinant clones

Recombinants were identified by sequence analysis of PCR-amplified proviral DNA from cells infected by virus produced by dually infected cells. We screened 163 clones from the infection using the initial virus stock described above, and 75 clones from a second independently generated virus stock. There are 31 individual base pair differences between NL4-3 and JRCSF in the 459-bp region chosen for sequencing. We defined nucleotide changes that were less than three bases apart as a single marker, resulting in 25 markers for the purpose of identifying recombinants (Fig. 3).

Each clone was aligned against NL4-3 and JRCSF, and the position of each marker was categorized as being characteristic of either the NL4-3 or JRCSF parent. For the six markers that consisted of two nucleotides, only changes in both bases of the marker were counted as a crossover event. Five clones from the pooled experiments had only single nucleotide changes at one of these six markers; these nucleotide changes were therefore not counted as crossover events.

Using the first virus stock from dually infected cells, 33/163 (20%) of screened clones had a recombination event within the 459-bp region sequenced (Fig. 3A). Fifteen of these recombinant clones (45%) had a single crossover.

---

**Fig. 2.** Detection of HIV-1 recombination in cell culture, experimental design. Step 1. Generation of parental virus stocks. Schematic versions of the vectors from Fig. 1 are shown at the top. COS-1 cells were co-transfected with either pDHIV.3.B7.1 or pDHIV.3.Thy-1.JRCSF.KL9 and pHCMV-G. Parental viruses DHIV.3.B7.1 and DHIV.3.Thy-1.JRCSF.KL9 are shown as red or green hexagons, respectively. Step 2. Generation of virus from cells infected with both parental stocks. COS-1 cells were co-infected with the two parental stocks and then transfected 24 h later with pHCMV-G. B7.1 and Thy.1 protein expressed on the surface of the co-infected COS-1 cell are diagrammed in blue and purple, respectively. The virus stock harvested on day 5 contains a mixture of virions. Parental virions are shown in red or green and the heterozygous virions are shown in yellow. Step 3. Infection of H9 with virus produced from dually infected cells. Recombinant virus stocks were used to infect H9 cells at a low MOI (0.01). Total DNA was isolated from infected cells, PCR-amplified, cloned into *Escherichia coli*, and serial clones were sequenced across a 459-bp region encompassing the DIS and the first part of gag.
Fig. 3. Types of crossovers observed. The 459-bp region sequenced is represented as a horizontal line at the top, divided into four regions of equal length. The location of the DIS is marked with a black bar. NL4-3 sequence is shown in solid red; the JRCSF sequence in hatched green. Position of each sequence marker is shown as a dark oval. Panels A and B represent results from two independent experiments. Clones screened are grouped as to whether they exhibited one, two, or three crossovers. No clones with greater than three crossovers in this region were observed. Positions of crossovers are marked by changes in the color and pattern of the horizontal lines.
event each, 15 clones (39%) had two crossovers each, and 3 clones (15%) had three crossovers each. Using the second heterozygous stock, 12/75 (16%) of screened clones had a recombination event over the 459-bp region sequenced (Fig. 3B). Four of these recombinant clones (33%) had a single crossover event each, 6 clones (50%) had two crossovers each, and 2 clones (17%) had three crossovers each. The disproportionate occurrence of double and triple crossovers observed in each experiment likely is an example of the previously observed phenomenon of high negative interference (Hu et al., 1997).

The frequency of recombination seen using this model system was approximately 3-fold higher than frequency of recombination during PCR. When proviral DNA from cells infected separately with each parental stock were pooled before PCR amplification, 6/101 (6%) of clones contained one or more crossovers. Four clones had a single crossover each while only two clones had double crossovers. No clones with triple crossovers were seen.

**Distribution of crossover events**

The distribution of crossovers is shown in Fig. 4 for each heterozygous stock. Because the distances between markers are not uniform, the recombination frequency within each marker segment was corrected for distance. For this correction, the proportion of crossovers between two markers was divided by the number of nucleotides between the markers. Because the likelihood of recombination may be proportional to the length of the homologous region between the two parental strains, the distance-corrected recombination frequency should reflect the likelihood of recombination at each marker, independent of distance.

The most striking result evident from the data presented in Fig. 4 is that recombination frequency varied greatly with location in this 459-bp-long region. Specifically, recombination frequencies are very high in the region between markers 5–9 in both experiments. In comparison, recombination frequencies were lower between markers 1–4, 10–13, and 18–24. This difference was evident, irrespective of whether the percent of crossovers at each marker was corrected for distance. The significance of moderate to high crossover frequencies at certain markers (e.g., 14, 17, and 25) is unclear, because these events were observed in only one of the two experiments. The region spanning markers 5–9 accounted for 64% (35/54) of all crossovers in experiment #1 and 61% (14/23) of the crossovers in experiment #2. This 112-nt-long region was only one-quarter the length of the region sequenced. This preferred region for recombination begins 30 bases upstream of the gag AUG and extends 82 nt into the coding sequence. In contrast, the 112-nt-long regions upstream and downstream of this preferred region accounted for only 10% and 14% of the recombination events, respectively.

PCR-mediated recombination did not appear to be the basis for the preferred region of recombination that we observed. There were eight crossover events within the six clones that recombined in the PCR control experiment. One crossover occurred in the region encompassing markers 1–4, three occurred in the region encompassing markers 5–10, two occurred in the region including markers 11–16, and three occurred in the region from markers 17–25. Thus, the overall rate of PCR-mediated recombination was relatively low with a fairly even distribution of crossovers over the region sequenced.

---

**Fig. 4. Distribution of crossovers.** The x-axis lists the markers for recombination events, numbered 1 – 25. The y-axis is the number of crossovers at each marker divided by the total number of crossovers that occurred over the entire region, divided by the number of bases between each marker. Bars shown for each marker represent the distance-corrected crossover frequency between that marker and the marker immediately upstream. Results from the first heterozygous stock (solid black bars) and second heterozygous stock (striped bars) are shown separately.
**In vitro strand transfer system**

To better understand and analyze characteristics of the \textit{gag}-associated transfers, we designed a simple in vitro transfer system. The intention here was to determine whether the transfer profile observed in the virus could be reproduced in vitro. This approach would then allow us to identify factors that contribute to the creation of the preferred transfer site. The system was designed such that primer switching from the donor RNA onto the acceptor RNA followed by completion of synthesis on the acceptor would yield transfer products that are longer and therefore easily distinguished from the non-recombinant donor extension products (Fig. 5A). The donor and acceptor templates shared a homology of 320 nts, extending from the PBS to the beginning of \textit{gag}. Primer extensions performed in the presence of the acceptor template yielded a 374-nt transfer product, in addition to the 354-nt full-length donor extension product (Fig. 5B). Transfer efficiency is calculated as, $100 \times \frac{\text{transfer products}}{\text{transfer products} + \text{products of full-length extension on donor}}$. At 60 min, 11–13% transfer efficiency was measured, indicating that 11–13% of the full-

---

**Fig. 5.** An in vitro system to analyze template switching catalyzed by HIV-1 RT during minus-strand synthesis through the 5'-UTR-\textit{gag} RNA. (A) Schematic of the 5'-UTR-\textit{gag} RNA templates. A schematic representation of the HIV-1 genome highlighting some of the prominent regions is included at the top of the panel. Regions and nucleotide positions in the genomic RNA that correspond to the 354-nt donor and 368-nt acceptor RNA templates are indicated. Bold lines represent the NL4-3 donor (black) and JRCSF acceptor (gray) RNA templates. Hatched lines at the template ends denote plasmid-derived sequences in the templates. Regions corresponding to the PBS and DIS sequences are indicated. (B) Strand transfer assays catalyzed by HIV-1 RT. Minus-strand synthesis was initiated on the donor RNA template using the 5'-end labeled DNA primer (P) in the presence of the acceptor RNA. Reactions were terminated at 2.5, 5, 10, 15, 30, 45, and 60 min (lanes 1–7) and products were resolved on a 6% polyacrylamide gel under denaturing conditions. Full-length synthesis on the donor (F) resulted in a 354-nt extension product, while products resulting from template switching and synthesis on the acceptor generated a 374-nt transfer product (T). Lane 8 is a donor extension reaction in the absence of acceptor. Arrowheads on the right side of the gel indicate the major RT pause sites, while stars along with the numbers on the left indicate approximate marker positions in the cDNA. Pause sites corresponding to the SD hairpin (SD) and the kissing hairpin (DIS) are also indicated. Lane L—25-base DNA ladder.
length products derived from synthesis completed on the acceptor after transfer. We note here that the 374-nt product would consist of transfer products generated from single and odd numbers of crossovers, where synthesis was completed on the acceptor. Recombinants resulting from double and even numbers of crossovers would complete synthesis on the donor template and would therefore be indistinguishable from donor extension products.

**Distribution of strand transfer events in vitro**

The donor and acceptor RNA templates were generated from the NL4-3 and JRCSF molecular clones of HIV-1, respectively. As with the analysis in vivo, the naturally occurring sequence variations were used as markers to map the crossover sites within the recombinant transfer products.

To determine whether RT template switching within the 5’-untranslated region (UTR) and early part of *gag* showed features similar to those observed during reverse transcription in the virus, the transfer products were isolated, cloned, and sequenced. Fig. 6A presents the distribution of crossover sites as determined from three independent experiments. Note that unlike the case with the cell culture recombinants, only crossovers through marker 14 can be analyzed. The bulk of transfer products resulted from a single template-switching event per template. An average of 1 in 30 clones was the result of triple crossovers within a template. Only products of single crossover events were included in the analysis in Fig. 6A.

Crossovers were observed across the length of the template, indicating that most regions within the in vitro template had the potential to support some homologous

---

**Fig. 6. Distribution of transfers in vitro.** Template description is same as in Fig. 5. Open circles on the templates correspond to relative positions of natural, single, or double nucleotide base differences between the donor and acceptor templates, and are designated as markers 1 through 14. These were used in determining the region of primer terminus transfer. (A) Transfer products were isolated, amplified by PCR, cloned, and sequenced to determine the distribution of crossover sites within the template. Bars between two markers represent transfer frequency within that marker segment and were calculated as (100 × transfers within marker segment/total transfers analyzed). Values from three independent experiments, each comprising of about 45–50 recombinant clones, are presented. Only products resulting from single crossover events were included in the analysis. (B) Comparing crossover distribution profiles of PCR-generated recombinant products (solid bars) to those generated by RT-mediated template switching (hatched bars). For examining PCR-generated recombinants, cDNAs were separately generated from donor and acceptor templates. Equal amounts of each were mixed together and amplified by PCR (see Materials and methods for details). Values were averaged from a minimum of three independent experiments, each consisting of 45–50 recombinant clones. The frequency of PCR recombinants was normalized to the observed average transfer frequency of 20%. The estimated frequency of RT recombinants was obtained by subtracting this estimated PCR transfer frequency at each site from the observed average transfer frequency shown in A. Marker segments are highlighted by parenthesis, with the segment length in nucleotides indicated in italics.
recombination (Fig. 6A). The distribution profile did not reveal any striking preference or high frequency of crossovers at the dimerization site. In fact only 20–30% of the total crossovers occurred within the 84-nt segment at the end of the template, following marker 2. About 30–40% of transfers mapped to the 88-nt segment between markers 2 and 5. This region spans from the DIS hairpin to 18 nt into gag. Finally, 45–50% of crossovers were localized within the 110-nt segment between markers 5 and 11, comprising the 5′-end of gag coding region.

To address the contribution of PCR-generated recombinants, cDNA extension products independently generated using the donor and the acceptor RNAs were mixed together in equal quantities, and subjected to identical PCR amplification conditions as the transfer products. The amplified products were cloned and sequenced. Faithful amplification should yield products having the pure donor or pure acceptor sequences. Results from four independent experiments revealed that 15–25% of the amplified products contained markers from both donor and acceptor. This meant that 15–25% of the recombinants analyzed were very likely generated by PCR. We therefore analyzed the crossover sites within the PCR recombinants. Fig. 6B compares the distribution of sites for the PCR-generated crossovers (corrected for the observed 20% transfer frequency) and the crossovers estimated to be due to RT alone (obtained by subtracting the PCR-generated crossover frequency from the observed crossover frequency at each marker). Interestingly, crossover sites of the PCR recombinants showed a distinctly different distribution profile than the RT-generated recombinants, with the bulk of the crossovers occurring in the template segment between markers 1 and 5. This suggested that the preferred region of crossovers in gag was genuine and not influenced by the PCR-generated recombinants.

Comparison of recombinants from cell culture and in vitro

To compare the crossover profiles of the cell culture and in vitro generated recombinants, the distance-corrected distributions were plotted together (Fig. 7). Regions upstream of marker 1 and downstream of marker 14 were excluded, as these regions were not common to both systems. Distance correction for results in vitro was done in the same way as for the cell culture data, by dividing the frequency of crossovers measured between two given markers by the distance between those markers. Significantly, the two profiles showed similarities in that a higher frequency of crossovers occurred in the start of the gag region (markers 5–9). In both systems, relatively fewer crossovers occurred in the template region downstream of marker 9. A slightly higher frequency of crossovers were observed in the PBS-DIS region (markers 1–3) in the in vitro system as compared to the cell culture-based system. A proportion of these likely represent PCR recombinants, as suggested from Fig. 6B.

Sequence features of the preferred site for recombination in gag

The sequence content of the preferred site for recombination at the 5′-end of gag (markers 5–9) differed significantly from another region incorporating markers 17–25 in which recombination was much less frequent. The gag region had 35 crossovers in the first experiment, was 112 nt long, had a GC content of 51% (57/112) and a purine content of 72%. There were 46 G residues throughout the
entire sequence and there were 4 independent stretches of 3 or more G residues. In contrast, the region encompassed by markers 17–25 had only two crossovers in the first experiment, had a lower GC content of 37% (41/111), and a similar purine content of 66% (73/111). There were half as many G residues (25) in this region, and no stretches of 3 or more G residues.

**Distribution of recombination events and base substitutions**

In the course of analyzing the sequences of clones obtained in the cell culture system, we noted the presence of a number of mutations (these are marked in blue in Fig. 8). The occurrence of these mutations appeared to cluster in specific regions (see, e.g., the region between markers 2 and 3 in Fig. 8). We therefore characterized the frequency, nature, and distribution of mutations in all sequenced clones. This was done to determine the correlation between crossover events and mutations, and to establish whether certain segments of the analyzed sequence were more mutagenic than others.

Only 3 of 369 mutational events were deletions; no nucleotide insertions were observed. G to A substitutions were the most frequent and comprised 63.1% and 50.6% of all substitutions observed in the first and second experiments, respectively. A to G substitutions were second in order of frequency. The remaining types of base substitutions occurred at a frequency of less than 5%. The overall mutation frequency and type of base substitution were similar between recombinant and non-recombinant clones (data not shown).

Mutation frequency differed widely among the segments defined by the recombination markers. For example, mutation frequencies (corrected for length) for regions defined by markers 3, 5, and 8, are much higher than rates in regions defined by markers 4, 10, 12, 16, and 23 (data not shown). These differences were seen in both recombinant and non-recombinant clones. The preferred sites for mutations appeared to contain primarily G to A substitutions, and were observed almost exclusively in regions where there were runs of three or more G residues (see Fig. 8, in the regions between markers 2–3, 4–5, 5–6, 7–8, 14–15). These preferred sites for mutations were observed in recombinant as well as non-recombinant clones (data not shown).

Mutations at these sites did not occur more commonly in clones that recombined at the adjacent markers compared to clones that did not recombine at the adjacent markers (data not shown).

There was some overlap between areas that were more likely to undergo recombination events, and those at which mutations were more likely to occur. However, there were exceptions to this finding. For example, region 3 underwent frequent mutation but had low levels of recombination. Thus, the two events were not strictly linked.

To assess the contribution of PCR to the types and distribution of mutations observed in the cell culture system, we cloned and sequenced amplified products from PCR reactions, using cloned pNL4-3 DNA as substrate and similar reaction conditions as the proviral DNA amplifications. The three most frequent types of mutations that occurred during PCR in the absence of reverse transcription were A to G (40%), T to C (26%), and G to A (23%) substitutions. G to A mutations were therefore less common during PCR than during HIV infection (23% vs. 51–63%, respectively) and there were no observed hotspots for G to A mutations in runs of G residues. While PCR may contribute to the overall rate of mutation we observed, it does not appear to be responsible for the high percentage of G to A mutations, or the preference for G to A substitutions in stretches of G residues seen after reverse transcription.

**Discussion**

It is important to determine whether recombination occurs randomly throughout the HIV genome, or whether there are preferred sites for recombination. If such sites exist, they could influence the location of recombination events, particularly under circumstances in which the effective replicating pool of HIV is limited (see Brown, 1997 for a more detailed discussion of this concept). Proposed mechanisms for recombination during retrovirus replication suggest that structure in the RNA templates and properties of the proteins involved in replication would result in preferred regions for crossovers (Balakrishnan et al., 2003; Coffin, 1979; DeStefano et al., 1992, 1994; Harrison et al., 1998; Peliska and Benkovic, 1992; Roda et al., 2002; Wu et al., 1995). Specifically, regions that promote pausing of the RT or interactions between the two genomic RNAs have shown a propensity for primer transfer during synthesis in vitro (Balakrishnan et al., 2001, 2003; DeStefano et al., 1992, 1994; Diaz and DeStefano, 1996; Harrison et al., 1998; Kim et al., 1997; Mikkelsen et al., 1998, 2000; Palaniappan et al., 1996; Roda et al., 2002; Wu et al., 1995). Analyses to date of recombination of HIV-1 in cell culture demonstrate crossovers throughout the genome with a suggestion of the existence of preferred locations (Jetzt et al., 2000; Zhuang et al., 2002). We have identified the presence of a preferred site of recombination near the 5’-end of the gag gene, using a tissue culture system that allows a fine analysis of crossover locations in the 5’-UTR and gag regions of the HIV-1 genome. A parallel analysis of crossovers within the same sequences in vitro revealed the consistent finding that this region of gag is a preferred site for transfer events, supporting the hypothesis that this region of the genome promotes recombination by HIV RT. The power of this dual approach is that the cell culture system shows recombination events in a biological setting, and the reconstituted system allows for the identification of important elements leading to recombination specificity within the RNA and protein molecules employed. This relationship
provides the key to the ultimate mechanistic interpretation of natural recombination events.

Recent studies by Zhuang et al. (2002) suggested the presence of local hotspots for recombination throughout the genome and also identified a preferred location in the 5′-UTR of gag. Our tissue culture system also has the potential to identify recombination events over a larger region spanning the dimerization site in the 5′-UTR, through the vpr coding sequence. Our system has some novel features that facilitate a high-resolution analysis of

![Fig. 8](image-url)
recombination events. All of the progeny proviruses are sequenced and no selection markers are used to enrich for recombinants. This enables us to identify all recombinants generated with no possible loss through selection. Because recombination has been shown to occur in vitro within short stretches of homology (DeStefano et al., 1997; Kim et al., 1997; Palaniappan et al., 1996; Wu et al., 1995), recombination events measured over large regions of the genome using lower resolution strategies such as HTA may not detect some preferred sites of recombination.
Previous analyses of strand transfer recombination in vitro lead to proposals that RT pausing during synthesis results in a concentration of RNase H-directed cleavages of the template genomic RNA near the pause site (DeStefano et al., 1992, 1994, 1997; Roda et al., 2002). This localized cleavage of genomic RNA reveals a single-stranded region of the newly made primer that can interact with the other genomic RNA to initiate the recombination event. Results from the transfer assays in vitro presented here indicated a possible correlation between the location of RT pausing and the preferred region of crossovers in gag. Three prominent pause sites were observed between markers 5 and 9, and correlated with the region were the bulk of transfers occurred. Interestingly, locations of these pause sites also correlated with homopolymeric nucleotide runs in the template RNA. Homopolymeric nucleotide runs have been shown to promote pausing by RT during DNA synthesis (Ji et al., 1994; Klarmann et al., 1993).

Pausing promotes recombination, presumably because the template is degraded more efficiently by the RNase H activity of the enzyme when polymerization is stalled. This increased degradation can then provide a location for the second template to invade the enzyme/primer-template complex and promote transfer of the RT and DNA primer terminus (Balakrishnan et al., 2003; DeStefano et al., 1992, 1994; Kim et al., 1997; Roda et al., 2002; Wu et al., 1995).

On the basis of in vitro studies of RNA dimerization, the DIS sequence in HIV-1 has been postulated to be responsible for initiating dimerization between the viral RNA genomes (Laughrea and Jette, 1994; Paillart et al., 1994, 1996, 1997). Earlier analyses of recombination had suggested that the MLV DIS promotes crossovers, possibly based on proximity of the RNA templates (Mikkelsen et al., 1998, 2000). However, recent analyses of HIV-1 have shown that regions of the viral RNA containing the DIS are not uniquely more recombinogenic than others (Zhuang et al., 2002), an observation that is supported by our data obtained both in vivo and in vitro. Our previously published analysis of recombination events promoted by the HIV-1 DIS in vitro, using shorter RNA templates, revealed a proximity effect facilitated by DIS-induced template-template dimerization (Balakrishnan et al., 2001, 2003). This effect of the DIS on recombination could be translated to other templates by simply inserting the 35-nt DIS hairpin within an ectopic sequence. The crossovers observed near the DIS in both the MLV cell culture system and in vitro are likely promoted by such a mechanism. On the basis of our current analysis with longer templates containing the gag region, we conclude that such proximity effects, though existent, are very likely local and not dominant in the context of large regions or the entire virus. It is also possible that there are numerous other sites of low-affinity interaction between long RNA templates that provide a general alignment of sequences between the two templates. In such a case, the proximity effect at the DIS would not particularly stand out in the context of the entire viral genome.

We believe the crossovers we have described primarily reflect recombination events occurring during reverse transcription, rather than an artifact of PCR, based on the differences in both the frequency and relative distribution of crossovers observed in cell culture compared to PCR. Moreover, the transfer distribution data obtained in vitro matched the results obtained in culture, supporting their validity. Although the number of replication cycles occurring in our cell culture system is limited, we cannot completely rule out the possibility that more than one round of replication is occurring at the time of infection with heterozygous virus. Therefore, we cannot define with certainty an absolute rate of recombination under these experimental conditions.

Another novel feature of the studies presented here is that the frequency and location of mutations can be compared to recombination events. We have observed that G to A mutations occurred at a high frequency within homopolymeric runs of G residues. We also identified some clones with G to A hypermutation, in which serial G to A mutations occur (Vartanian et al., 1991). This type of base substitution preferably occurs at GpG or GpA dinucleotides (Janini et al., 2001; Vartanian et al., 1991, 1994), but can also occur in homopolymeric stretches of G residues (Vartanian et al., 1994). It has been suggested that DNA synthesis by RT from a template with homopolymeric stretches of Cs can deplete dCTP pools, increasing the chance of misincorporation at the end of the homopolymeric region. More recently, the cellular protein APOBEC3G, a cytidine deaminase that is expressed in human T lymphocytes (including the H9 cell line), has been implicated in the production of retroviral G to A hypermutation (reviewed in KewalRamani and Coffin, 2003). Although an intact vif gene (which was present in our HIV vectors) counteracts the mutagenic effects of APOBEC3G in cell culture, overexpression of APOBEC3G in a cell line that normally does not express this cellular protein can induce G to A hypermutation in HIV with an intact vif gene (Zhang et al., 2003). Thus, the relative extent to which HIV RT and APOBEC3G contribute to the G to A mutations observed in our model system is not clear at present.

In summary, we have defined a region of the HIV-1 genome near the beginning of gag that is much more recombinogenic than surrounding regions of the 5'-UTR and gag. The presence of this preferred site for recombination may influence the likelihood of progeny viruses having a recombination event at this site, under conditions in which the effective replicating pool of HIV-1 is limited. The preferred site is flanked by similar sized areas that are virtually inert for recombination. Analysis of recombination over the same sequences in vitro recapitulated the preferred site for recombination. The
profile and location of pause sites within this region implies a role for RT pausing in promoting the transfers. These results present a basis for future work to analyze the mechanistic basis of actual recombination events in vivo. In addition, we have identified that this region in gag also undergoes frequent G to A substitutions, which cluster in regions of homopolymeric stretches of G residues. Highly G-rich regions may promote both recombination and G to A substitutions, but these most likely occur through independent mechanisms. The experimental cell culture model system described here should prove useful in defining whether other regions of the HIV-1 genome are also highly prone to undergo recombination or mutation, and to correlate sequence features with the likelihood of these events.

**Materials and methods**

**Reagents**

The following reagents were obtained through the AIDS Research and Reference Reagent program, Division of AIDS, National Institute of Allergy and Infectious Diseases: the infectious molecular clone pNL4-3 was obtained from Malcolm Martin, pYK-JRCSF was obtained from Irvin SY Chen and Yoshio Koyanagi, and the HeLa-CD4-LTR-β-gal cell line was obtained from Michael Emerman. pHCMV-G is a vector which expresses vesicular stomatitis virus (VSV)-g protein under the control of the human cytomegalovirus immediate early promoter (Yee et al., 1994). pBJ-hB7.1 contains the leader sequence and extracellular domain of the human co-stimulatory gene, B7.1 (Challita-Eid et al., 1998). The vector HIV-1NL4-3-luc-env(--vprX is a derivative of pNL4-3 which has a deletion in env, the reporter luciferase gene in place of nef, and a frameshift mutation at codon 64 of the vpr open reading frame (Planelles et al., 1995a). pHIV-thy-1 is a derivative of pNL4-3 in which nef has been replaced with the mouse gene, Thy1.2 (Planelles et al., 1995b). Recombinant HIV-1 RT (specific activity 20,000 units/mg) was supplied by the Genetics Institute (Cambridge, MA). T7 RNA polymerase, T4 polynucleotide kinase, DNase I, restriction endonucleases, dNTPs, rNTPs, 20,000 units/mg) was supplied by the Genetics Institute (Cambridge, MA). T7 RNA polymerase, T4 polynucleotide kinase, DNase I, restriction endonucleases, dNTPs, rNTPs, and RNase inhibitor were purchased from Roche (Indianapolis, IN) and RNase H from Invitrogen (Carlsbad, CA).

**Cell culture**

COS-1, SV40-transformed African green monkey kidney cells [American Type Culture Collection (ATCC), Manassas, VA], were grown in the presence of Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal bovine serum (FBS), l-glutamine (2 mM), penicillin (100 units/ml), and streptomycin (100 units/ml). HeLa-CD4-LTR-β-gal cells were grown in DMEM supplemented with 10% FBS, l-glutamine (2 mM), penicillin (100 units/ml), streptomycin (100 units/ml), hygromycin (100 μg/ml; Sigma, St. Louis, MO), and geneticin (200 μg/ml; Sigma). H9 human lymphoma cells (ATCC) were grown in RPMI supplemented with 20% FBS, penicillin (100 units/ml), and streptomycin (100 units/ml).


The first parental vector, pDHIV.3.B71, derived from pNL4-3, was designed to have a deletion of env and to express the human co-stimulatory gene, B7.1, on the cell surface of infected cells early after infection. B7.1 was excised from the vector, pBJ-hB7.1 using HindIII and XhoI. The HindIII end was filled in using the Klenow fragment of E. coli DNA polymerase I. HIV-1NL4-3-luc-env(--vprX was digested with MluI and XhoI. The MluI end was filled in using the Klenow fragment of E. coli DNA polymerase I. The B7.1 fragment was then ligated to the fragment from the HIV vector to produce pDHIV.4.B7, which now had the B7.1 gene in place of nef. Because pDHIV.4.B7 had an inactive vpr gene, a functional vpr gene from pNL4-3 was substituted, by subcloning the 2.3-kb AgeI–SalI fragment from pNL4-3 into pDHIV.4.B71, to create pDHIV.3.B71 (Fig. 1). The integrity of the vpr gene was confirmed by sequencing. Expression of B7.1 was confirmed by transfecting COS-1 cells with pDHIV.3.B71 (Superfect; Qiagen, Santa Clarita, CA) and staining with a B7.1 monoclonal antibody [α-B7.1-phycocerythrin (PE); Becton Dickinson, San Jose, CA], followed by flow cytometry.

The second parental HIV vector, pDHIV.3.Thy1.JRCSF.KL9, was designed to contain 5.1 kb from pYK-JRCSF, spanning the DIS through vpr; have a deletion in env, and to express the mouse gene product, Thy1.2, on the cell surface of infected cells early after infection. A 0.7-kb fragment of the envelope gene was deleted from the construct HIV-Thy-1 using two BglII sites within the gp120-coding sequence, to generate pDHIV.3.Thy1. The 5.1 kb BssHII–SalI fragment from pYK-JRCSF, spanning the DIS through vpr, was substituted for the corresponding NL4-3 sequence in pDHIV.3.Thy1 to create pDHIV.3.Thy1.JRCSF. To identify recombinants across the DIS, a subclone of pDHIV.3.Thy1.JRCSF was constructed and used to introduce mutations corresponding to the JRCSF sequence upstream of the DIS hairpin (NL4-3 proviral positions 667–668, T to A to G, respectively). The 2.9-kb AarII–SphI fragment from pHIV.3.Thy1.JRCSF was subcloned into pQE-31 (Qiagen) to generate pQE-31-JRCSF. The two nucleotide substitutions were introduced using PCR-mediated site-directed mutagenesis (Quikchange, Stratagene, La Jolla, CA) and the following primers: QE-31/JRCSF-1 (5’-CGCAGAAAGAGCAAGCCAGAGG-3’) and QE-31/JRCSF-2 (5’-CCTCTGCTTCTCTTTCGC-3’). The resulting clone, pQE-31-JRCSF.KL9, was sequenced in both directions to
verify the presence of the mutations and the absence of spurious mutations (Perkin-Elmer/Applied Biosystems, Foster City, CA). The AarII–Spfi fragment of pQE-31. JRCSF.KL9 was subcloned into pDHIV.Thyl.JRCSF to create pDHIV.Thyl.JRCSF.KL9 (Fig. 1). This vector was sequenced from proviral bases 667–5801 to verify the presence of the DIS mutation and confirm the published sequence of JRCSF. The expression of Thyl.2 was verified by transfection of COS-1 cells with pDHIV.Thyl.1. JRCSF.KL9 (Superfect; Qiagen), followed by flow cytometric analysis using a FITC-conjugated anti-thy 1.2 antibody (Caltag, Burlingame, CA).

Production of parental DHIV.3.Thyl.JRCSF and DHIV.3.B71 viral stocks

COS-1 cells were transiently transfected with pHCMV-G (Akkina et al., 1996; Burns et al., 1993), and either pDHIV.3.Thyl.JRCSF.KL9 or pDHIV.3.B71 (Superfect; Qiagen). Supernatants were harvested 72 h later and clarified by centrifugation at 500 × g. Virus stocks were concentrated by ultracentrifugation and stored at −80 °C. Relative infectious titers of each stock were determined using HeLa-CD4-LTR-β-gal indicator cells, as previously described (Kimpton and Emerman, 1992).

Production of virus from dually infected cells stocks

Twenty-four hours after seeding in 10-cm plates, 1 × 10^6 COS-1 cells were treated with 10 μg/ml of polybrene (Sigma) in PBS, and then co-infected with DHIV.3. Thyl.JRCSF.KL9 and DHIV.3.B71 at a MOI of 1.0 for each virus. Twenty-four hours after infection, cells were transiently transfected with pHCMV-G (Superfect; Qiagen). Supernatant was harvested 72 h later and clarified by centrifugation at 500 × g. The relative infectious titer was determined using HeLa-CD4-LTR-β-gal indicator cells as previously described (Kimpton and Emerman, 1992).

Flow cytometry

Flow cytometric analysis was performed in a FacsScan (Becton-Dickinson) using the manufacturer’s recommended protocols. PE-conjugated-anti B71 antibody (Becton Dickinson), anti-Thy1-FITC (Caltag), and their respective isotype control antibodies [mouse IgG1-PE (Pharmingen, San Diego, CA) and rat IgG_{2a}-FITC (Caltag)] were diluted 1:10 with phosphate-buffered saline (PBS) containing 2% FBS, 0.5 mM EDTA, and 0.02% sodium azide before use.

Production and sequencing of proviral clones

H9 cells were infected with virus stock from dually infected cells, at an MOI of 0.01 in the presence of 10 μg/ml of polybrene. After 72 h, aliquots of infected cell pellets were harvested and frozen at −80 °C. Genomic DNA was extracted from 1 × 10^6 H9 cells using the QiAmp DNA Blood Kit (Qiagen), and PCR amplified using the Expand Long Template PCR system (Roche Molecular Biochemical, Indianapolis, IN), and the following primers: LTR-3 (5′-GCTCAAAATGTTGTGTGCC-3′) and SAL-2 (5′-ACAAGCAGTTTAGCT-GACTTCC-3′). The resulting 5 kb product, which included the DIS hairpin, gag, pol, vif, and vpr, was cloned into pCR-XL-TOPO (TOPO XL PCR cloning kit; Invitrogen) and cloned DNA was purified from bacterial cultures (Wizard SV 96 Plasmid DNA purification kit; Promega, Madison, WI). Clones were digested with BamHI (MBI Fermentas, Amherst, NY) before purifying the DNA for sequencing (Strataprep 96 PCR purification kit; Stratagene). Recombinant clones were identified by sequencing a 459-base pair region spanning the DIS through the first 112 codons of gag (NL4-3 proviral nt 666–1125).

Determination of PCR recombination rate

To determine the PCR recombination rate, H9 cells were infected at an MOI = 0.01 with either DHIV.3.B7.1 or DHIV.3.Thyl.JRCSF.KL9. Forty-eight hours later, cells were harvested and washed once with PBS, and harvested cell pellets were frozen at −80 °C. Genomic DNA from each separate infection was then pooled in a 1:1 ratio, and amplified as described above. PCR products were cloned and sequenced as described above.

Determination of the types and distribution of PCR mutations

To assess the contribution of the PCR Expand enzyme to the types and distribution of mutations in our system, we PCR-amplified 10 ng of pNL4-3 using the same PCR conditions that were used to amplify proviral DNA from cells infected with heterozygous virus. This PCR product was purified and cloned into pCR-XL-TOPO, and clones were prepared for sequencing as described above (“Production and Sequencing of Provir al Clones”).

Generation of RNA templates for in vitro strand transfer assays

Genomic sequences from the NL4-3 and JRCSF strains of HIV-1 were amplified by PCR and cloned into pBlueScript II KS(+) (Stratagene), to create the pNL182-520 donor and pJRC150-502 acceptor constructs for generation of RNA templates using an approach described previously (Balakrishnan et al., 2001). PCR primers Sac1/PBS-187 (5′-
CTGGAGCTCTTGGCGCCCGAAGAGGG) and BamH1/520 (5′-GTCGAGATCTCCCTCTCTTCTAAGCC) were used to amplify the genomic RNA region from the PBS through gag position 520 to create the donor construct, while primers Sac1/H5150 (5′-CTCGAGCTCCCTTCTGTTCTAAGCC) and JRC/BamH1-520 (5′-GTCGAGATCTCCCTCTCTGATATTCTAACAGG) were used to amplify the genomic region from the U5 position 150 through gag position 502 to create the acceptor construct. Restriction enzyme sites within the primers are underlined. PCR fragments were digested with HindIII end labeled primer 24 (5′-CCCAGTATTGTCCTACAGCC). Briefly, the 5′-end labeled primer (50 fmol) was mixed with donor (25 fmol) RNA in 50 mM Tris–HCl, pH 8.0, 1 mM EDTA, and 50 mM KCl. Mixtures were incubated at 80 °C for 1 min and slow-cooled to room temperature to facilitate primer annealing. Transfer assays were performed in 12 μl final reaction volume. Two units (50 ng) of HIV-1 RT were preincubated for 2 min at room temperature with 2 nM template termini of primer-template cDNA were used to amplify the genomic RNA region from the PBS through gag position 520 to create the donor construct, while primers Sac1 and BamH1 and cloned into pBluescript II KS (+) (Stratagene). RNA templates were prepared by in vitro run-off transcription from BamH1 linearized plasmids as previously described (Balakrishnan et al., 2001). Full-length transcripts were purified on 6% denaturing polyacrylamide gel and analyzed for integrity.

In vitro transfer assays

Assays were set up as previously described (Balakrishnan et al., 2001). Synthesis was initiated on the donor template using the DNA primer MB24 (5′-CCCAGTATTGTCCTACAGCC). Briefly, the 5′-end labeled primer (50 fmol) was mixed with donor (25 fmol) RNA in 50 mM Tris–HCl, pH 8.0, 1 mM EDTA, and 50 mM KCl. Mixtures were incubated at 80 °C for 1 min and slow-cooled to room temperature to facilitate primer annealing. Transfer assays were performed in 12 μl final reaction volume. Two units (50 ng) of HIV-1 RT were preincubated for 2 min at room temperature with 2 nM template termini of primer-template and 8 nM acceptor, in 50 mM Tris–HCl, pH 8.0, 80 mM KCl, and 1 mM DTT. Acceptor RNA was excluded in donor extension assays. Reactions were initiated by the addition of MgCl2 and dNTPs at a final concentration of 6 mM and 50 μM, respectively. Following incubation at 37 °C, reactions were terminated at the indicated times by addition of 12 μl of termination buffer (90% formamide, 10 mM EDTA, pH 8.0 and 0.1% each of xylene cyanole and bromophenol blue). Reaction products were resolved on 6% polyacrylamide–urea gels and visualized by phosphorimaging using ImageQuant software (Molecular Dynamics).

Analysis of transfer products

Transfer assays were performed for 1 h, the cDNA transfer products were then PAGE purified, amplified by PCR using primers Sac1/150 and BamH1/520, and cloned into pBlue-script II KS (+) as previously described (Balakrishnan et al., 2001). Individual clones, representing recombinant transfer products, were sequenced using M13 (−20) primers by automated sequencing to identify crossover sites. To address the contribution of PCR to the generation of false recombinant products, cDNAs separately generated from primer extensions on donor and acceptor RNA templates were mixed together and subject to identical PCR amplification as the transfer cDNAs. The amplified products were cloned and sequenced to determine the percentage of recombinants generated and the distribution of crossover sites in these products.

Acknowledgments

This work was supported in part by National Institutes of Health grants R01-AI-041387 and R01-GM-49573. Sequences reported here were submitted to GenBank (accession numbers AY508126 through AY508158). We thank Roslyn Chang for technical assistance with analyses done in vitro.

References

regions of heteropolymer templates by human immunodeficiency virus reverse transcriptase. J. Virol. 66 (11), 6370–6378.


Janini, M., Rogers, M., Bixr, D.R., McCutchan, F.E., 2001. Human immunodeficiency virus type 1 DNA sequences genetically damaged by hypermutation are often abundant in patient peripheral blood mononuclear cells and may be generated during near-simultaneous infection and activation of CD4+ T cells. J. Virol. 75 (17), 7973–7986.


Laughrue, M., Jette, L., 1994. A 19-nucleotide sequence upstream of the 5′ major splice donor is part of the dimerization domain of human immunodeficiency virus 1 genomic RNA. Biochemistry 33 (45), 13464–13474.


mediated by human immunodeficiency virus reverse transcriptase in vitro is promoted by pausing and results in misincorporation. J. Biol. Chem. 270 (1), 325–332.


