

Frequent Dual Initiation in Human Immunodeficiency Virus-Based Vectors Containing Two Primer-Binding Sites: a Quantitative In Vivo Assay for Function of Initiation Complexes

Yegor A. Voronin^{1,2} and Vinay K. Pathak^{1*}

HIV Drug Resistance Program, National Cancer Institute at Frederick, Frederick, Maryland 21702,¹ and Department of Biochemistry, West Virginia University, Morgantown, West Virginia 25606²

Received 14 November 2003/Accepted 19 January 2004

We previously demonstrated that murine leukemia virus (MLV)-based vectors containing two primer-binding sites (PBSs) have the capacity to initiate reverse transcription more than once (Y. A. Voronin and V. K. Pathak, *Virology* 312:281–294, 2003). To determine whether human immunodeficiency virus (HIV)-based vectors also have the capacity to initiate reverse transcription twice, we constructed an HIV type 1 (HIV-1)-based vector containing the HIV-1 PBS, a green fluorescent protein reporter gene (*GFP*), and a second PBS derived from HIV-2 3' of *GFP*. Simultaneous initiation of reverse transcription at both the 5' HIV-1 PBS and 3' HIV-2 PBS was predicted to result in deletion of *GFP*. As in the MLV-based vectors, *GFP* was deleted in approximately 25% of all proviruses, indicating frequent dual initiation in HIV-based vectors containing two PBSs. Quantitative real-time PCR analysis of early reverse transcription products indicated that HIV-1 reverse transcriptase efficiently used the HIV-2 PBS. To investigate tRNA primer-RNA template interactions in vivo, we introduced several mutations in the HIV-2 U5 region. The effects of these mutations on the efficiency of reverse transcription initiation were measured by quantitative real-time PCR analysis of early reverse transcription products, with initiation at the HIV-1 PBS used as an internal control. Disruption of the lower and upper parts of the U5-inverted repeat stem reduced the efficiency of initiation 20- and 6-fold, respectively. In addition, disruption of the proposed interactions between viral RNA and tRNA^{Lys3} thymidine-pseudouridine-cytidine and anticodon loops decreased the efficiency of initiation seven- and sixfold, respectively. These results demonstrate the relative influence of various RNA-RNA interactions on the efficiency of initiation in vivo. Furthermore, the two-PBS vector system provides a sensitive and quantitative in vivo assay for analysis of RNA-RNA and protein-RNA interactions that can influence the efficiency of reverse transcription initiation.

Two copies of viral genomic RNA are packaged during assembly of a retroviral virion (8, 13, 15, 45). After entering a target cell, most retroviruses convert the viral RNA into double-stranded DNA through the process of reverse transcription (6, 66). Retroviruses use a tRNA as a primer to initiate DNA synthesis (14, 33, 38, 54–56, 74). Human immunodeficiency virus types 1 (HIV-1) and 2 (HIV-2) both use tRNA^{Lys3} as a primer for initiation of reverse transcription (14, 38, 56). The primer tRNA is specifically packaged into virions during assembly (21, 48, 49, 74). The tRNA placement requires unwinding of the primer tRNA and dissociation of the inverted repeat (IR)-PBS stem in the viral genome. The virally encoded nucleocapsid protein possesses nucleic acid chaperone activity and is known to facilitate nucleic acid hybridization, such as tRNA placement (28, 44, 57, 68).

Even though each HIV virion contains two RNAs and primer-binding sites (PBSs), it is currently unknown whether HIV-1 particles have the capacity to initiate reverse transcription on both genomic RNAs. Retroviral virions are estimated to contain 75 to 150 molecules of reverse transcriptase (7, 53, 69) and approximately eight primer tRNAs (35, 48, 49, 54), well in excess of the two available PBSs. Thus, the components necessary to initiate reverse transcription more than once are pres-

ent in each virion. It is possible that the number of initiation events per virion is limited by the number of PBSs that are functionally associated with a tRNA primer. In cell-free murine leukemia virus and avian leucosis virus virions, approximately 50% of genomic RNAs have a tRNA tightly bound to them (31, 54), but the distribution of the occupied PBSs among infectious virions is not known. In addition, it is possible that not all PBSs associated with tRNAs are used for initiation due to incorrect placement of the tRNA, improper association with reverse transcriptase or other components of the initiation complex, or potential inhibitory effects of RNA secondary structure near the PBS.

Initiation of reverse transcription on only one of the two copackaged genomic RNAs would leave the second RNA available as a template for recombination during minus-strand DNA synthesis. Consistent with this idea, analysis of proviruses generated by a recombining population in spleen necrosis virus showed that all infected cells contained a single provirus (34). This observation also supports the general belief that one provirus is formed during each infection event. However, it is possible that both nonrecombining and recombining virions are capable of initiating and completing synthesis of two full genomic DNAs. Inefficient minus-strand DNA transfer, intermolecular minus-strand DNA template switching (3, 34, 39), degradation of viral DNA, formation of one and two long terminal repeat (LTR) circles (41, 61, 65), autointegration (46, 62), and inefficient integration (19, 20, 29, 63) could reduce the efficiency of provirus formation.

* Corresponding author. Mailing address: HIV Drug Resistance Program, National Cancer Institute—Frederick, Building 535, Room 334, Frederick, MD 21702. Phone: (301) 846-1710. Fax: (301) 846-6013. E-mail: VPATKAK@ncifcrf.gov.

The regions surrounding the PBS are known to form a specific secondary structure, which plays a role in increasing the efficiency and specificity of initiation. Different elements of the structure have been implicated in the process (1, 2, 9–12, 22, 23, 47, 50, 59, 71). The U5-leader and the U5-IR stems are formed in the region; the stability of both of these stems has been shown to be important for initiation in murine leukemia virus, avian sarcoma and leucosis virus, and HIV-1 (9, 12, 22, 23, 50). Although the PBS forms the largest and most important site for primer tRNA binding, additional interactions between genomic RNA and primer tRNA have been proposed. In HIV-1, a region of the U5 sequence that is involved in forming a U5-leader stem, called the primer activation signal, has been proposed to interact with the thymidine-pseudouridine-cytidine (TΨC) arm of the tRNA (10, 11). These interactions were originally described in avian retroviruses and therefore seem to represent a common feature of the reverse transcription initiation complex (22, 23). In addition, studies of HIV-1 viruses modified to use tRNA^{His} as a primer indicated that an A-rich loop in the U5-IR stem interacts with the anticodon loop of the primer tRNA^{Lys3} (59, 71).

To date, the importance of various RNA-RNA interactions involved in initiation of reverse transcription has been determined by two experimental approaches. First, mutations that allow use of tRNAs other than tRNA^{Lys3} for initiation of DNA synthesis have been introduced into replication-competent virions, and the replication properties as well as the nature of reversion mutations that restore viral replication to wild-type levels have been studied (43, 71–73). Second, structural analysis of viral RNAs in complex with initiator tRNAs has been performed to discern the various RNA-RNA interactions (16–18, 30, 36). Both of these approaches have been highly successful in demonstrating the importance of RNA-RNA interactions. However, it has not been possible to determine the quantitative importance of specific RNA structures to the efficiency of initiation *in vivo*.

We recently demonstrated that when a murine leukemia virus-based vector contained an additional PBS (derived from spleen necrosis virus) in the middle of the genome, both the wild-type and the inserted PBS could be used for initiation of reverse transcription with similar efficiencies (70). Analysis of cells infected with the murine leukemia virus vector also showed that the two PBSs were used simultaneously in at least 25 to 30% of all infectious virions. This result indicates that in simple retroviruses, at least 25% of virions are capable of forming two functional reverse transcription complexes. However, it is not known whether lentiviruses also have the capacity for dual initiation in single infectious particles. Several factors, including the efficiency of tRNA packaging, tRNA placement, formation of functional initiation complexes, and the inhibitory effects of RNA secondary structures, could influence the ability to initiate reverse transcription twice.

To determine whether HIV-1 virions have the capacity to initiate reverse transcription more than once, we constructed an HIV-1-based vector that contained a second PBS derived from HIV-2. Analysis of infected cells indicated that a single HIV-1 infectious virion has the capacity to use the HIV-2 PBS efficiently and to initiate reverse transcription more than once frequently. We have also performed quantitative mutational analysis of the RNA-RNA interactions at the HIV-2 PBS.

Analysis of the efficiency of initiation in these vectors is quantitative because initiation at the HIV-1 PBS can provide an internal control. These studies indicate that vectors containing two PBSs provide a sensitive and quantitative *in vivo* assay to determine the relative effects of various RNA-RNA interactions on the efficiency of reverse transcription initiation.

MATERIALS AND METHODS

Construction of vectors. Plasmid names begin with p, while names of viruses derived from the plasmids do not. All HIV-1-based vectors were derived from pHR'CMVLacZ, which was a kind gift from Inder Verma (Salk Institute) (52). HIV-2 sequences were derived from the HIV-2 ROD10 variant (14). All vectors contained a hygromycin phosphotransferase B gene (*hygro*) (32), which was expressed from the internal ribosomal entry site (IRES) of encephalomyocarditis virus (37).

To construct pKD-GFP-IN, the ClaI-XhoI fragment of pHR'CMVLacZ was replaced with the green fluorescent protein gene (*GFP*), IRES, and neomycin phosphotransferase gene (*neo*) (40). The final structure of the GFP-IRES-*neo* cassette was similar to that described previously (77). IRES and *neo* were then removed by EcoRI-KpnI digestion, and the plasmid backbone was used to insert an IRES-*hygro* cassette obtained by digestion of pLW1 with PvuI and ClaI (42). The resulting plasmid was called pHGFP-Hy.

To construct vectors with two PBSs, the U5-PBS-leader sequence of HIV-2 ROD10 was PCR amplified in two steps. First, the U5-PBS region was amplified with forward primer ROD10-F (5'-GGAATTCAGTTAAGTGTGTGCTCCC ATCTCTCC-3') and reverse primer HIV2-PBS-R (5'-CTTCAAGTCCCTGTT CGGGCGCCAACC-3'). The PBS-leader was amplified with forward primer HIV2-PBS-F (5'-GCAGGTTGGCGCCGAACAGGGACTTG-3') and reverse primer ROD10-R (5'-GGGGTACCCGGGCACTCCGTCGTGGTTTGT TCC-3'). In the second step, the two PCR products were joined by mixing and amplifying them with primers ROD10-F and ROD10-R. Primers HIV2-PBS-F and HIV2-PBS-R introduced a C310T mutation in the HIV-2 PBS. Primers ROD10-F and ROD10-R were designed to contain an EcoRI and a KpnI site at the 5' and 3' end, respectively. The PCR product was digested with EcoRI and KpnI and inserted in the corresponding sites in pKD-GFP-IN, replacing the IRES-*neo* sequences. The resulting plasmid was called pHIV-PBS. The IRES-*hygro* cassette was inserted into the SmaI site downstream of the HIV-2 PBS region, resulting in ph2PBS.

To remove the cytomegalovirus promoter from pHGFP-Hy, pHMut, and ph2PBS, the vectors were digested with SalI and ClaI and ligated with the *GFP*-encoding NotI fragment derived from pGL1 (Gibco). The resulting plasmids were named pdhGFP-Hy, pdhMut, and pdh2PBS, respectively.

Mutations of the HIV-2 U5-PBS region were generated with a PCR mutagenesis protocol. Briefly, the region was amplified with *Taq* polymerase (Takara) with two sets of mutagenic primers and subcloned into the pPCR-Script vector by blunt-end ligation (Stratagene). After the presence of the desired mutation was confirmed by sequencing, the fragment was subcloned into the pdhGFP-Hy backbone to generate analogs of the pdh2PBS plasmid with the desired HIV-2 PBS mutations.

Cells, transfections, and virus propagation. 293T cells were obtained from the American Type Culture Collection. HeLa-tat-III cells were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, National Institute of Allergy and Infectious Disease, NIH, from William Haseltine and Ernest Terwilliger (60, 67). HeLa-tat-III cells express HIV-1 Tat protein, which allows transcription of Tat-deficient vectors from the HIV-1 LTR promoter. Cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum (HyClone Laboratories, Inc.). Penicillin (50 U/ml; Gibco) and streptomycin (50 µg/ml; Gibco) were also added to the medium. Cells were maintained in a 37°C incubator with 5% CO₂.

All transfections were performed by the calcium-phosphate precipitation method (CalPhos transfection kit; Clontech). For stable transfections of vectors into 293T cells, 4 µg of vector DNA was cotransfected with pCV1 (molar ratio, 3:1). pCV1 was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, National Institute of Allergy and Infectious Disease, NIH, from Flossie Wong-Staal (5). This plasmid expresses Tat and Rev proteins, which are necessary for HIV-1 expression. Cells were plated the day before transfection at 2×10^5 cells per 100-mm-diameter dish. Transfection was performed for 3 h, after which the cells were washed with 6 ml of medium and fresh medium was added. The day after transfection, the medium was replaced with medium containing 250 µg of hygromycin per ml to select stably transfected cells.

After 2 weeks, approximately 700 to 1,500 resistant colonies were pooled and expanded for each experiment.

Virus produced by transient transfections was used to measure viral titers and generate pools of infected cells for flow cytometry and Southern blotting analysis. 293T cells were plated at 2×10^6 cells per 100-mm-diameter dish and on the next day were transfected with either pdhGFP-Hy or pdh2PBS, pCMV Δ R8.2 (51), and pHCMV-G (75) (molar ratio, 1:1:1, with a total of 20 μ g of DNA). On the following day, the cells were washed with 6 ml of medium and fresh medium was added. The virus was collected 24 h later.

Tenfold serial dilutions of virus-containing supernatant were used to infect 2×10^5 HeLa-tat-III cells per 60-mm-diameter dish. Infections were performed in the presence of Polybrene (50 μ g/ml) for 2 h. The infected cells were subjected to hygromycin selection at 24 h postinfection (250 μ g of hygromycin per ml). Viral titers were determined by quantitation of hygromycin-resistant colonies approximately 2 weeks after infection.

Detection of GFP expression by flow cytometry. The HeLa-tat-III cells were infected at a multiplicity of infection of less than 0.1. Approximately 10,000 hygromycin-resistant infected colonies were pooled for analysis. The percentage of GFP-expressing cells was measured by flow cytometry (FACSCalibur; Becton Dickinson); at least 10,000 events were collected for each experiment, and the results were analyzed with CellQuest software (Becton Dickinson).

Southern blotting analyses. Genomic DNA was isolated from infected cells with the AquaPure isolation kit (Bio-Rad), and proviral structures were analyzed by Southern blot hybridization as previously described (27). A 1.2-kb DNA fragment containing *hygro* was used to generate a probe. Quantitation of bands was performed with a Phosphorimager and the Quantity One software program (Bio-Rad).

Quantitative real-time PCR analysis. The day before transfection, 2×10^6 293T cells stably expressing each vector were plated per 100-mm-diameter dish. The cells were transfected with pCMV Δ R8.2 and pHCMV-G (1:1 molar ratio, with a total of 8 μ g of DNA) overnight. After transfection, the cells were washed with 6 ml of medium and fresh medium was added. The virus was collected after 24 h, filtered (0.45- μ m pore size; Corning), and incubated with DNase I (Roche, 30 U/ml) and MgCl₂ (10 mM final concentration) at room temperature for 30 min to minimize carryover of transfected DNA. Approximately 3×10^6 293T cells were infected by incubation with virus-containing medium in the presence of 50 μ g of Polybrene per ml. Cells were harvested 0, 1, and 2 h after infection after one wash with 6 ml of phosphate-buffered saline. To allow the infection to proceed for 6 h, cells were washed with phosphate-buffered saline 2 h after infection and incubated with fresh medium for an additional 4 h.

The QIAmp DNA Blood mini kit (Qiagen) was used to extract total cellular DNA from infected cells. DNA from approximately 10^5 cells was used for each real-time PCR assay with an ABI Prism 7700 sequence detector (Applied Biosystems). The R-U5 region was detected with forward primer 5'-AGCTTGCC TTGAGTGTCTCAA-3', reverse primer 5'-TGACTAAAAGGGTCTGAGGG ATCT-3', and probe 5'-6-carboxyfluorescein (FAM)-AGAGTCACACAACAC ACGGGCACACTA-6-carboxytetramethylrhodamine (TAMRA)-3'. The U5- Ψ region was detected with forward primer 5'-TCTGTTGTGTGACTCTGGTAA CTAGAGA-3', reverse primer 5'-CCGTGCGCGCTTCAG-3', and probe 5'-F AM-CCCGAACAGGGACTTGAAGCGCAAAG-TAMRA-3'. The GFP region was detected with forward primer 5'-TCAGACACAACATTGAGGATG GA-3', reverse primer 5'-CGCCGATTGGAGTGTCTGT-3', and probe 5'-F AM-CCGTGCAGCTGGCCGACCAT-TAMRA-3'. The *hygro* region was detected with forward primer 5'-ACGAGGTCGCCAACATCTTC-3', reverse primer 5'-CGCGTCTGCTGCTCCAT-3', and probe 5'-FAM-CAAGCCAACC ACGGCCTCCAGA-TAMRA-3'. The final concentrations of primers and probes were 600 and 75 nM, respectively.

Threefold serial dilutions of pdhGFP-Hy were used to generate a standard curve ranging from 17 to 1,000,000 copies of DNA per PCR. The same dilutions were used to generate a standard curve for each primer-probe set, which allowed accurate measurement of relative amounts of DNA products detected by different sets. The correlation coefficient for all standard curves was >0.99 . The amount of each PCR product in the sample was determined from a standard curve generated with the particular primer set.

To normalize the amount of DNA analyzed in the real-time PCR experiments, each sample was also analyzed with a primer set designed to detect the human porphobilinogen deaminase gene (*PBGD*) (GenBank accession number M95623) (kindly provided by Michael Piatak, AIDS Vaccine Program, Science Applications International Corporation, Frederick, Md.) (76). The *PBGD* primer set included the forward primer 5'-AGGGATCACTCAGGCTCTTCT-3', reverse primer 5'-GCATGTTCAAGCTCCTTGGTAA-3', and probe 5'-FAM-CAGGCTTTCTCTCCAATCTGCCGGA-TAMRA-3'.

RESULTS

Structure of HIV-based vectors containing two PBSs and protocol to determine dual initiation. We constructed two HIV-based retroviral vectors to determine whether two initiation events can occur in the same virion during reverse transcription (Fig. 1A). Vector dhGFP-Hy contains the 5' and 3' LTRs, PBS, packaging signal (Ψ), polypurine tract, and other *cis*-acting elements required for vector propagation. dhGFP-Hy also expresses *hygro*, which confers resistance to hygromycin, and *GFP* from a single bicistronic transcript initiated from the 5' HIV LTR. Efficient transcription that is driven from the HIV-1 LTR requires Tat, which is not expressed from dhGFP-Hy or dh2PBS. Therefore, efficient transcription of these vectors requires expression of Tat in both producer and target cells. Translation of *hygro* is facilitated by the presence of the IRES.

We constructed pdh2PBS by insertion of a 224-bp sequence containing the PBS derived from the ROD10 isolate of HIV-2 between *GFP* and the IRES (Fig. 1A). The PBS that is located just downstream of the 5' LTR is referred to as the 5' PBS, and the HIV-2 PBS is referred to as the 3' PBS throughout the text. The predicted structures of the HIV-1 and HIV-2 U5-leader regions including the PBS in complex with the tRNA primer are shown in Fig. 1B. HIV-2, like HIV-1, uses tRNA^{Lys3} as a primer, and 17 of 18 nucleotides of the HIV-2 PBS are identical to those of the HIV-1 PBS (Fig. 1C). However, there is very little homology between the sequences surrounding the PBSs in HIV-1 and HIV-2. The inserted sequence contains the complete HIV-2 U5 (129 nucleotides), including the HIV-2 attachment site (*att*), the PBS (18 nucleotides), and a portion of the 5' untranslated leader sequence (77 nucleotides). The PBS of the ROD10 isolate of HIV-2 differs slightly from that of many other HIV-1 and HIV-2 isolates (26). In most isolates, position 310 is occupied by C, whereas ROD10 has T, which has been proposed to occur due to infrequent use of a variant of tRNA^{Lys}, tRNA^{Lys5} (26, 64). Both nucleotides can base pair with the eighth nucleotide (G) of the primer tRNA^{Lys3}. In order to eliminate any effects of different PBS sequences or use of different tRNAs on the efficiency of initiation at the HIV-1 and HIV-2 PBSs, nucleotide 310 of the HIV-2 PBS was mutated from T to C (Fig. 1B).

Previously, we observed that insertion of the homologous PBS in the middle of the retroviral vector resulted in a high frequency of deletions through reverse transcriptase switching templates during the process of reverse transcription due to the formation of direct-repeat sequences in the vector (our unpublished results). Therefore, the HIV-2-derived PBS region, which lacks significant homology to the HIV-1 PBS, was chosen for insertion in the middle of the HIV-1-based vector. In this vector, direct-repeat deletion through reverse transcriptase template switching is not expected to occur because the inserted HIV-2 sequence has only 49% homology to the HIV-1 sequence (Fig. 1C).

Reverse transcription of vectors with two PBSs. Each viral RNA containing two PBS regions can initiate reverse transcription at the 5' PBS (Fig. 2A), the 3' PBS (not shown), or both PBSs (Fig. 2B). Each of these initiation scenarios results in a provirus with a different structure. A GFP-positive provirus, the structure of which was identical to that of the original

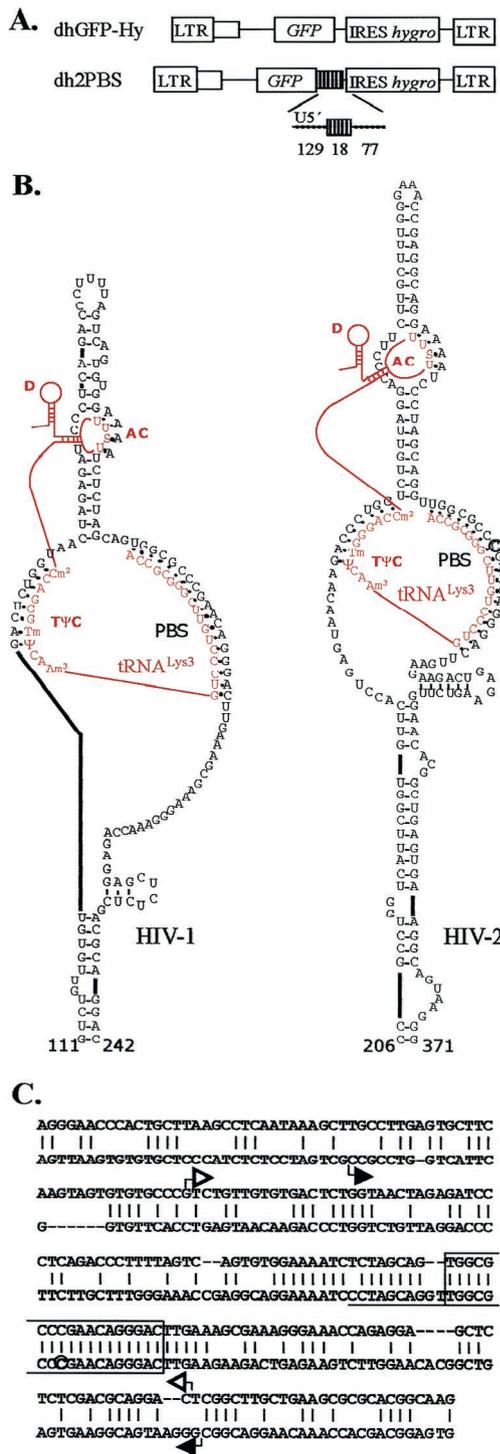


FIG. 1. Structures of vectors, predicted secondary structures of HIV-1 and HIV-2 PBS regions, and sequence homology comparison of HIV-1 and HIV-2 PBS regions. (A) Structures of HIV-1-based vectors. *GFP* is expressed from the HIV-1 LTR promoter, and translation of *hygro* is facilitated by the IRES. Vector dhGFP-Hy has only the 5' PBS derived from HIV-1 (open box between 5' LTR and *GFP*). Vector dh2PBS has an additional 3' PBS fragment derived from HIV-2 (striped box between *GFP* and IRES *hygro*). The 3' PBS fragment consists of 129 nucleotides of the HIV-2 U5 (dotted line labeled U5'), 18 nucleotides of HIV-2 PBS (striped box), and 77 nucleotides of the 5' untranslated leader from HIV-2 (dotted line). (B) Comparison of predicted secondary structures of the HIV-1 (adapted from Beerens

vector, is generated when initiation occurs only at the 5' PBS and reverse transcription is completed through strand transfer and DNA synthesis steps expected for replication of a normal retrovirus (24) (Fig. 2A). If initiation occurs only at the 3' PBS, an unusually long minus-strand strong-stop DNA containing R, U5, PBS, Ψ , and *GFP* is formed. After completion of reverse transcription, a proviral structure that contains a second PBS, Ψ signal, and *GFP* downstream of the 3' LTR is formed (not shown) (70). Initiation of reverse transcription at both the 5' and 3' PBSs leads to the formation of a provirus with a *GFP*-negative structure, which does not contain *GFP* and Ψ sequences (Fig. 2B). Initiation at the 5' PBS results in the formation of normal minus-strand strong-stop DNA, which is transferred to the 3' end of the viral RNA through homology between the 3' and 5' R regions. Initiation at the 3' PBS results in the formation of a minus-strand DNA that encodes *GFP*, Ψ , and PBS.

DNA synthesis initiating at the 3' PBS does not extend past the PBS because initiation of reverse transcription at the 5' PBS results in degradation of the R and U5 regions by the RNase H activity of reverse transcriptase (4). Due to the lack of homology to the R region, this product cannot be used for minus-strand DNA transfer; as a result, it is not involved in subsequent steps during reverse transcription. The RNA sequences corresponding to *GFP*, Ψ , and 5' PBS are degraded during minus-strand DNA synthesis initiating from the 3' PBS. Therefore, DNA synthesis that is initiated at the 5' PBS and transferred to the 3' end of the viral RNA is extended to the 3' PBS but not farther. Plus-strand DNA synthesis is initiated at the polypurine tract and extended through U3, R, U5, and the first 18 nucleotides of the primer tRNA. The DNA is then transferred to the 3' PBS and reverse transcription is completed, resulting in the *GFP*-negative structure.

To summarize, one of the consequences of inserting a second functional PBS in the HIV-based vector is that initiation of reverse transcription at both the 5' and 3' PBS regions can result in the formation of the *GFP*-negative provirus structure,

and Berkhout [10]) and HIV-2 (adapted from Freund et al. [30]) PBS regions in complex with tRNA^{Lys3}. Nucleotides of tRNA that interact with the viral genomic RNA are shown in red; the rest of the sequence is indicated by the thin line. The thick line connects different parts of the genomic RNA. The numbers at the base of each predicted structure refer to nucleotide positions in the HIV-1 and HIV-2 genomes, starting at the beginning of the 5' R regions. The circled nucleotide within the HIV-2 PBS indicates a mutation introduced into the HIV-2 ROD10 sequence to make it identical to PBSs from other HIV-1 and HIV-2 isolates. The tRNA^{Lys3} T^ψC loop (T^ψC), anticodon loop (AC), and D loop (D) are labeled. Modified bases in tRNA^{Lys3} are defined as follows: S, 5-methoxycarbonylmethyl-2-thiouridine; Cm, methylcytidine; Tm, 2'-O-methyl-5-methyluridine; Ψ , pseudouridine; Am³, 1-methyladenosine. (C) Homology comparison of HIV-1 (top line) and HIV-2 (bottom line) nucleotide sequences in the vicinity of the PBS regions. The HIV-2 sequence shown is one of the fragments used to generate the dh2PBS vector and consists of the U5 region, PBS, and 77 nucleotides of the 5' untranslated leader. The homologous HIV-1 sequence shown indicates that 49% of the nucleotides are identical (vertical lines). The boxed sequence indicates 18 nucleotides of the PBS, and the underlined sequence represents the *att* site. The white and black arrowheads indicate the sequences that form the predicted secondary structures of HIV-1 and HIV-2, respectively, shown in B. The mutated nucleotide 310 is circled.

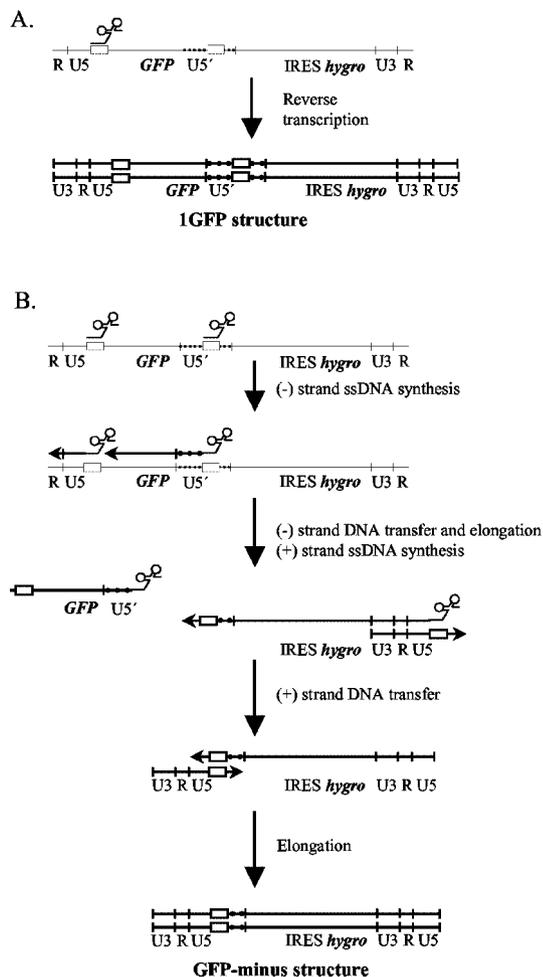


FIG. 2. Reverse transcription of vectors containing two PBSs. Thin lines indicate genomic RNA, thick lines indicate newly synthesized DNA, and dotted lines indicate HIV-2-derived sequences. The partial cloverleaf symbol represents tRNA used to initiate DNA synthesis. White boxes represent the 5' and 3' PBSs. Horizontal arrows show the direction of DNA synthesis. The U5, U3, and R regions form the HIV-1 LTR. The U5 region of the HIV-2 LTR is indicated as U5'. (A) Initiation of reverse transcription at the 5' PBS, followed by completion of normal reverse transcription, results in the formation of a GFP-positive proviral structure. (B) Initiation of reverse transcription at both the 5' and 3' PBS regions can result in the formation of a GFP-negative proviral structure (see Results for a detailed description).

which cannot express *GFP*. It is important to note that multiple initiations in a single virion do not always lead to the GFP-negative provirus structure, and the observed frequency of deletions provides the lowest estimation for the frequency of multiple initiations (see Discussion) (70).

GFP expression in vectors containing two PBSs. We determined the replication ability of vectors dhGFP-Hy and dh2PBS with the protocol outlined in Fig. 3. Each vector, HIV-1 helper construct pCMV Δ R8.2, and the pHCMV-G envelope expression construct were cotransfected into 293T cells, and the resulting virus was collected 24 h later. The virus was used to infect HeLa-tat-III cells, which express Tat protein and therefore allow vector expression in the absence of the helper construct. Infected cells were selected for resistance to hygromycin,

and pools of infected cells were analyzed by flow cytometry to determine the frequency of infected cells that expressed *GFP*. In addition, DNAs isolated from pools of infected cells and from single infected-cell clones were analyzed by Southern blotting to determine proviral structures.

The titers of each vector were determined in three independent experiments by selecting infected cells for resistance to hygromycin and quantifying the number of resistant colonies. The average dhGFP-Hy titer was $104 \times 10^4 \pm 20 \times 10^4$ CFU/ml. The titer of vector dh2PBS was similar to that of the control vector dhGFP-Hy ($56 \times 10^4 \pm 12 \times 10^4$ CFU/ml), indicating that the presence of the 3' PBS did not have a deleterious effect on viral replication.

To determine whether the presence of the functional 3' PBS led to the deletion of *GFP*, we performed flow cytometry analysis of pools of cells infected with the dhGFP-Hy and dh2PBS viruses. Five independent experiments were performed to determine the average frequencies of GFP-positive and GFP-negative cells; a representative flow cytometry analysis is shown in Fig. 4A. Analysis of cells infected with dhGFP-Hy virus and selected for hygromycin resistance indicated that <1% of the infected cells were GFP-negative. In contrast, flow cytometry analysis of cells infected with dh2PBS virus and selected for hygromycin resistance indicated that the frequency of GFP-negative cells was $24 \pm 3.6\%$. This frequency of GFP-negative cells is very similar to the 26% frequency of GFP-negative cells observed after infection with a murine leukemia virus-based vector containing a second PBS derived from spleen necrosis virus (70). Because dh2PBS virus is expected to have two functional PBSs and control vector dhGFP-Hy has only one functional PBS, the higher proportion of GFP-negative cells was correlated with the presence of two PBSs. This result strongly suggests that in 24% of the cells infected with

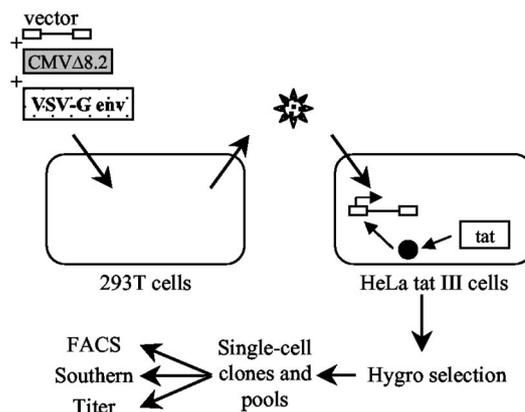


FIG. 3. Experimental design. To analyze replication of HIV-1-based vectors, plasmids encoding the vectors (pdhGFP-Hy or pdh2PBS), *gag-pol* and other accessory genes (pCMV Δ R8.2), and the vesicular stomatitis virus G envelope (pHCMV-G) were cotransfected into 293T cells. Virus was collected and used to infect HeLa-tat-III cells, which express Tat protein and facilitate transcription from the HIV-1 promoter. Cells were subjected to hygromycin drug selection until resistant colonies formed. Single-cell clones as well as pools of hygromycin-resistant colonies were analyzed by flow cytometry and Southern analysis. Quantitation of the hygromycin-resistant colonies was used to determine viral titers.

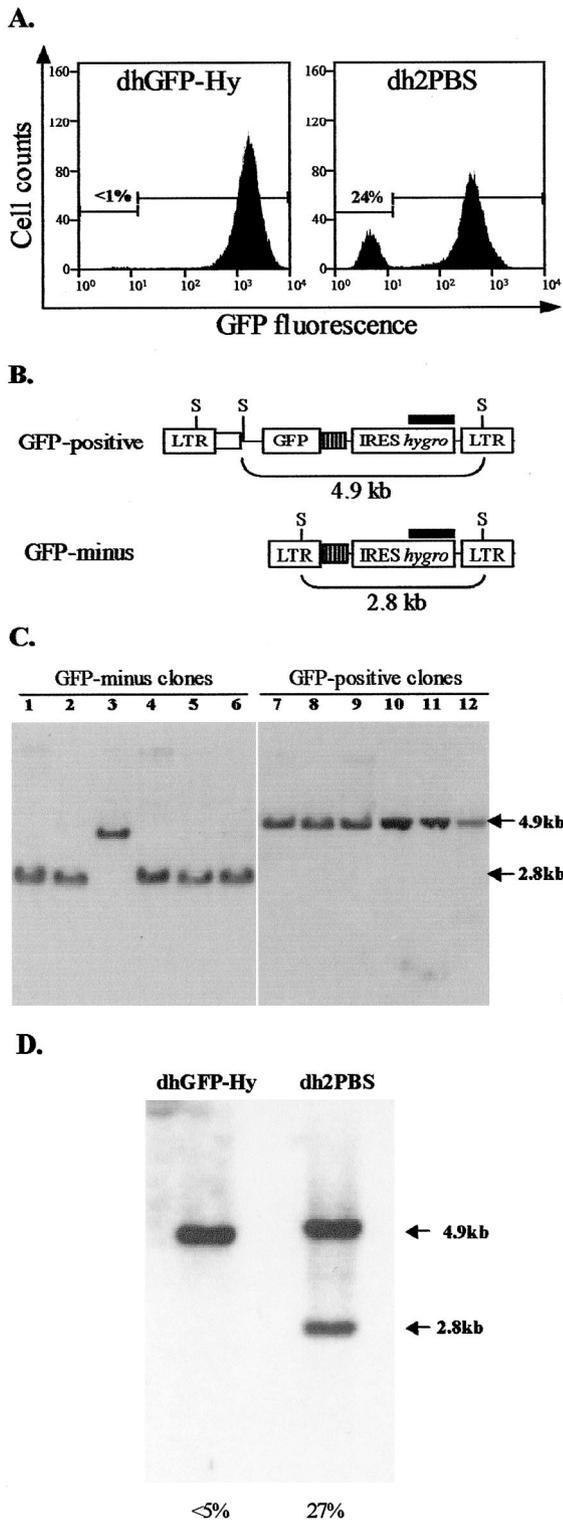


FIG. 4. Flow cytometry and Southern blotting analyses. (A) Representative flow cytometry analyses of HeLa-tat-III cells infected with dhGFP-Hy (left panel) or dh2PBS (right panel) virus. Infected cells were selected for resistance to hygromycin, and pools of resistant cells were analyzed for expression of GFP. Average proportions of GFP-negative cells determined in five independent experiments are shown. At least 5,000 events were analyzed in each experiment. (B) Structures of proviruses and sizes of corresponding bands. The *SacI* restriction sites (S) and the location of the Hygro probe (black bar) are shown. Digestion of GFP-positive proviral structures is expected to generate a

dh2PBS virus, GFP was deleted through initiation of reverse transcription at both the 5' and 3' PBSs.

Analysis of proviral structures to determine the frequency of GFP deletion. The structures of proviruses generated after infection with dh2PBS vectors are outlined in Fig. 4B. When genomic DNA of infected cells was digested with *SacI*, the GFP-positive proviral structure was expected to generate a 4.9-kb band; in contrast, the GFP-negative proviral structure was expected to generate a smaller 2.8-kb band. Both bands can be visualized by hybridization with the Hygro probe, and the relative proportion of each band can be quantified by Phosphorimager analysis.

We first analyzed genomic DNAs isolated from clones of single HeLa-tat-III cells infected with dh2PBS virus (Fig. 4C). Fluorescence-activated cell sorting was performed to identify and isolate 10 GFP-positive and 10 GFP-negative clones. Genomic DNAs from each of the cell clones were digested with *SacI* and analyzed by Southern blotting; representative results obtained from six GFP-positive and six GFP-negative cell clones are shown in Fig. 4C. Five of the six GFP-negative clones generated the 2.8-kb band, confirming the presence of a GFP-negative provirus structure. The presence of the predicted 2.8-kb band in the GFP-negative cell clones provided strong evidence that reverse transcription was initiated at both the 5' and 3' PBSs in the same virion and resulted in the formation of the GFP-negative proviral structure, as shown in Fig. 3B. One GFP-negative clone (Fig. 3C, lane 3) generated a band that was larger than the 2.8-kb band expected from a GFP-negative structure and smaller than the 4.9-kb band expected from a GFP-positive proviral structure (Fig. 3C). We hypothesize that this clone contained an abnormal provirus that was generated through an error occurring during reverse transcription. All six of the GFP-positive cell clones generated the expected 4.9-kb band. Of the 10 GFP-negative clones analyzed, seven generated the expected 2.8-kb band, and of the 10 GFP-positive clones, nine generated the expected 4.9-kb band (data not shown).

Next, we analyzed genomic DNAs from five independent pools of infected cells to determine the relative proportions of GFP-positive and GFP-negative proviral structures. Representative results from one set of pools are shown in Fig. 4D. The frequency of GFP-negative proviral structures was estimated by quantitative Phosphorimager analysis of the 4.9- and 2.8-kb bands. Genomic DNA from cells infected with dhGFP-Hy was used as a control. As expected, analysis of DNA from dhGFP-Hy-infected cells revealed a single 4.8-kb band after *SacI* digestion; this band is slightly smaller than the 4.9-kb band generated by dh2PBS because it lacks the 3' PBS fragment. Because only the 5' PBS was present in this vector, deletion of

4.9-kb band, and digestion of the GFP-negative proviral structure is expected to generate a 2.8-kb band. (C) Southern blotting analysis of single-cell clones infected with dh2PBS virus. Analysis of six single-cell clones that were nonfluorescent (GFP-negative) and six clones that were fluorescent (GFP-positive) is shown. (D) Southern blotting analysis of pools of cells infected with dhGFP-Hy and dh2PBS viruses. The proportions of the GFP-negative proviruses among all proviruses were quantified for five independent infections by Phosphorimager analysis and are shown as a percentage below each lane.

GFP through initiation of reverse transcription at two PBS sites was not expected to occur. Quantitation of the 2.8- and 4.9-kb bands observed from five independent infections with dh2PBS virus indicated that, on average, approximately $27 \pm 2\%$ of the cells contained the GFP-negative proviral structure. This *GFP* deletion frequency is in agreement with the 24% frequency of GFP-negative cells observed after flow cytometry analysis of infected cells (Fig. 4A).

The fact that the 2.8-kb band was not observed for dhGFP-Hy vectors indicates that the GFP-negative proviral structure was formed only when the vector contained two PBSs. This result also strongly supports the view that during infection with dh2PBS, reverse transcription was initiated at both the 5' and 3' PBSs in the same virion, which resulted in the formation of the GFP-negative proviral structures.

Quantitative PCR analysis of reverse transcription initiation in vectors containing two PBSs. To determine the kinetics and relative efficiencies of reverse transcription initiation at the 5' and 3' PBSs, we analyzed the products of reverse transcription 0.5, 1, 2, and 6 h after infection. The locations of the real-time PCR primer and probe sets are shown in Fig. 5B. The primer sets R-U5 and GFP were used to detect R-U5 and GFP products, which provided a measure of the initiation of reverse transcription at the 5' and 3' PBSs, respectively. The Hygro primer set was used to detect reverse transcription products after minus-strand DNA transfer. The U5- Ψ primer set was used to detect late reverse transcription products after plus-strand DNA transfer. Four independent infection experiments were performed, and the copy number of each target sequence in each sample was determined three times with DNA from approximately 10^4 to 10^5 infected cells. The amounts of cellular DNAs analyzed were normalized by determining the copy number of the cellular *PBGD* gene.

To determine the kinetics of reverse transcription initiation, 293T cells were infected with dhGFP-Hy virus, and the amounts of the R-U5 product at 0.5, 1, and 2 h after infection were compared with the amount of the R-U5 product detected at 6 h. Previous studies have indicated that under the experimental conditions used in this study, the reverse transcription products reach maximal levels 6 h after infection (our unpublished results). The total amount of R-U5 products 0.5, 1, and 2 h after infection were 2 to 4%, 12 to 20%, and 45 to 75% of the products detected 6 h after infection, respectively (data not shown).

Because only a small percentage of reverse transcription initiation events occurred at 0.5 h after infection (2 to 4%) and most of the initiation events occurred 2 h after infection (45 to 75%), we compared the relative amounts of reverse transcription products for the dhGFP-Hy and dh2PBS vectors 1 and 2 h after infection (Fig. 5C). To compare the relative efficiencies of initiation at the 3' PBS, the amount of the R-U5 product was set to 100%. The primary reverse transcription product detected in cells infected with dhGFP-Hy 1 h after infection was R-U5. The reverse transcription products detected by the Hygro, GFP, and U5- Ψ primer sets were less than 20% of the R-U5 product value, indicating that most of the minus-strand strong-stop DNAs were not transferred to the 3' end of the genome and extended to the Hygro primer set. In contrast to dhGFP-Hy reverse transcription products, the profile of reverse transcription products detected 1 h after infection with

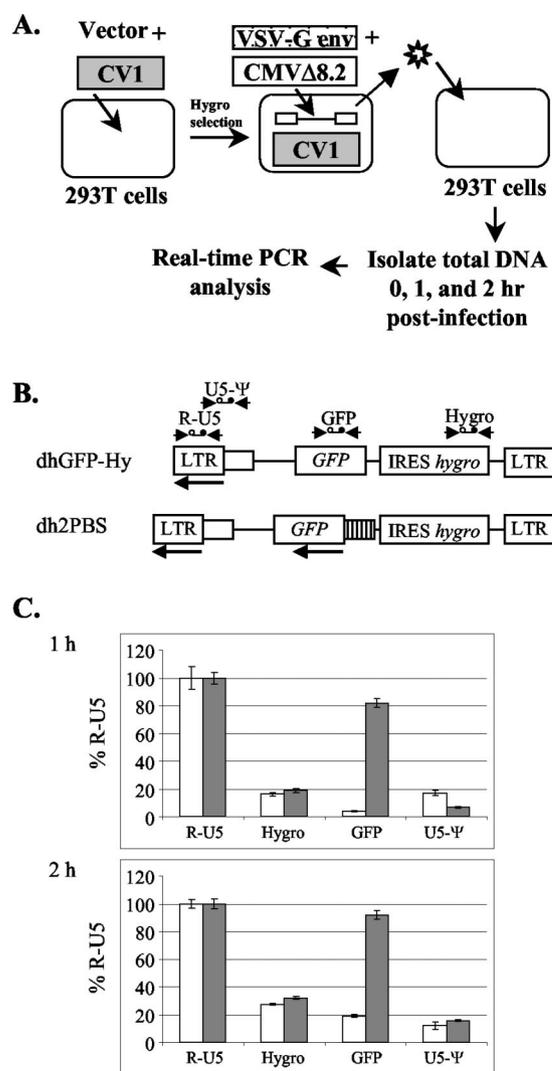


FIG. 5. Quantitative real-time PCR analysis. (A) Experimental protocol. Stable cell lines were generated by cotransfection of 293T cells with the *tat* and *rev* expression plasmid (pCV1) and vector DNA (pdhGFP-Hy or pdh2PBS), followed by hygromycin selection. After transfection with vesicular stomatitis virus G envelope-encoding plasmid (pCMV-G) and a helper construct (pCMVΔ8.2), virus was collected and treated with DNase I to reduce the possibility of DNA contamination from producer cells. Infections were performed for 0, 1, and 2 h, and total cellular DNA was isolated for real-time PCR analysis. (B) The locations of real-time PCR primer-probe sets are shown above the structures of dhGFP-Hy and dh2PBS. The early products of reverse transcription that could form before minus-strand DNA transfer are represented as arrows below the structures of dhGFP-Hy and dh2PBS vectors. (C) Comparison of relative amounts of reverse transcription products determined 1 and 2 h after infection of 293T cells. The amounts of DNA products determined after dhGFP-Hy infection (white bars) and dh2PBS infection (gray bars) are shown. Four independent experiments were analyzed, and each DNA sample was analyzed by real-time PCR three times. The average copy numbers of various DNA products determined are expressed as a percentage of the DNA products determined with the R-U5 primer-probe set. The error bars represent the standard error of the mean.

dh2PBS indicated that the levels of GFP products (82%) were similar to the levels of the R-U5 products. The high levels of GFP products indicated that reverse transcription was initiated at the 5' and 3' PBS sites with similar efficiencies.

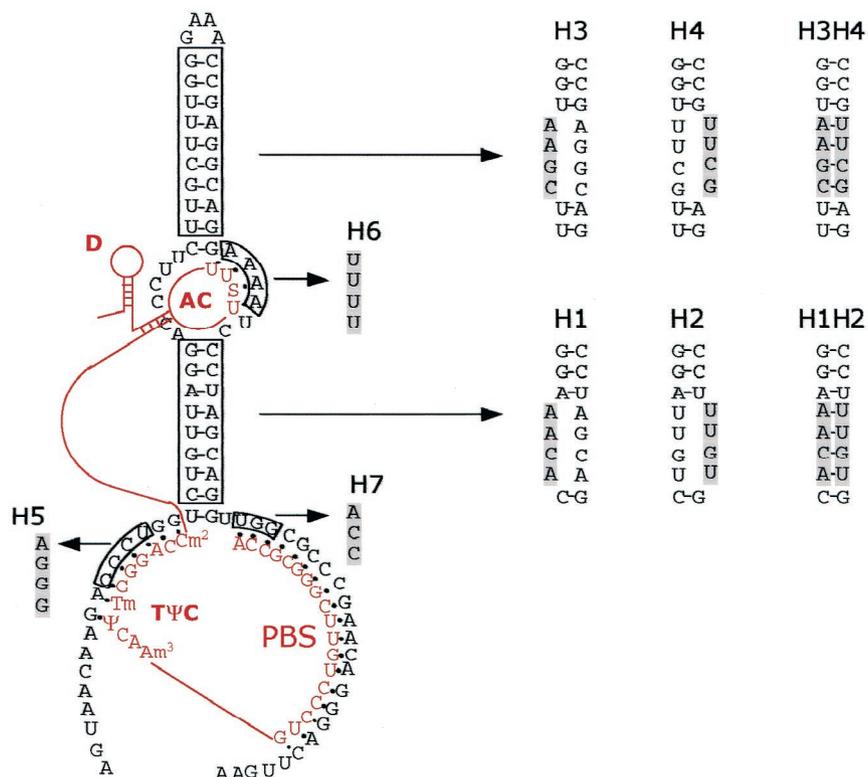


FIG. 6. Predicted structure of tRNA-HIV-2 RNA complex and mutations introduced to disrupt or restore specific features of the predicted structure. The portions of the tRNA^{Lys3} sequences that interact with the genomic RNA are shown in red; the tRNA regions not involved in interactions with genomic RNA are shown as lines and schematic hairpin structures. TΨC, TΨC loop; AC, anticodon loop; D, D-loop. The structures of the modified nucleotides are described in the Fig. 1 legend. Replaced nucleotides are shown with a gray background.

Analyses of the reverse transcription products 2 h after infection with the dh2PBS virus indicated that the relative amount of GFP products was approximately 90% of that of the R-U5 products, which was significantly higher than that of the GFP products detected for the dhGFP-Hy virus (20%). The larger amounts of GFP products detected for the dh2PBS virus confirmed that initiation of reverse transcription at the 3' PBS occurred with high efficiency.

Mutagenesis of HIV-2 PBS region. We used the vectors containing two PBSs to investigate the impact of different RNA-RNA interactions within the tRNA-template complex on the efficiency of initiation at the 3' PBS derived from HIV-2. Several mutations were introduced into the HIV-2 PBS and U5 sequences in dh2PBS (Fig. 6). Mutations H1 and H2 disrupted the complementarity of the lower portion of the U5-IR stem by mutating the sequences distal and proximal to the PBS, respectively; the double mutation H1H2 restored the complementarity and the stem structure. Similarly, mutations H3 and H4 disrupted the complementarity of the upper part of the U5-IR stem by mutating the regions distal and proximal to the PBS, respectively; the double mutation H3H4 restored the complementarity and the stem structure. Mutation H5 targeted a portion of the U5 region that was previously proposed to interact with the TΨC arm of tRNA^{Lys3} (30). Mutation H6 was designed to block the proposed interaction between the A-rich loop of viral RNA with the AC loop of the primer tRNA (16, 30). Finally, mutation H7 disrupted annealing of

the 3'-terminal CCA of the tRNA with the 5' region of the HIV-2 PBS.

To test the effects of these mutations on the efficiency of initiation at the 3' PBS derived from HIV-2, four independent infections were performed; 1 and 2 h after infection, total cellular DNA was isolated. The amounts of early reverse transcription products generated by initiation at the HIV-2 PBS (GFP products) were measured and compared with the amounts of early products generated by initiation at the 5' PBS (R-U5 products), which served as an internal control. Most of the copy numbers detected for the GFP and R-U5 primer sets were at least fivefold greater than those for the negative control (<50 copies). The absolute copy numbers of the R-U5 products varied between different infections from 2×10^3 to 5×10^5 due to variation in transfection, infection, and DNA recovery efficiencies.

The results obtained are summarized in Table 1. The amounts of the GFP products detected 1 and 2 h after infection with dhGFP-Hy were 3.3 and 19.7% of the amounts of the R-U5 products, respectively. These percentages were calculated as $(\text{GFP/R-U5}) \times 100$. Because only the 5' PBS is present in dhGFP-Hy, the result indicates that 3.3 and 19.7% of the reverse transcription initiation events underwent minus-strand DNA transfer and were extended to form the GFP products 1 and 2 h after infection, respectively. The GFP/R-U5 percentages detected after infection with dhGFP-Hy represented the background signal due to initiation at the 5' PBS; therefore,

TABLE 1. Quantitative real-time PCR analysis of effects of mutations in the HIV-2 PBS on reverse transcription initiation efficiency

Cct ^a	Time after infection ^b (h)	GFP/R-U5 ^c (%) ± SE	Adjusted GFP/R-U5 ^d (%) ± SE	Initiation efficiency (%) ± SE ^e	Avg decrease in initiation efficiency ^f (fold)
dhGFP-Hy	1	3.3 ± 0.4	0	0	NA
	2	19.7 ± 0.8	0	0	
dh2PBS	1	85.2 ± 3.3	81.9 ± 3.3	100	NA
	2	92.5 ± 2.2	72.8 ± 2.2	100	
H1	1	7.0 ± 0.5	3.7 ± 0.5	4.5 ± 0.7	20.9
	2	22.8 ± 2.2	3.1 ± 2.2	5.1 ± 2.4	
H2	1	7.4 ± 1.3	4.1 ± 1.3	5.0 ± 1.6	14.3
	2	28.1 ± 2.2	8.4 ± 2.3	11.6 ± 3.1	
H1H2	1	93.1 ± 5.0	89.9 ± 5.0	109.7 ± 6.0	1.0
	2	88.4 ± 10.5	68.7 ± 10.5	94.5 ± 14.3	
H3	1	15.7 ± 1.0	12.4 ± 1.0	15.1 ± 1.2	6.1
	2	32.9 ± 1.6	13.2 ± 1.5	18.0 ± 2.2	
H4	1	20.3 ± 1.1	17.0 ± 1.1	20.7 ± 1.4	4.6
	2	36.3 ± 3.1	16.6 ± 3.1	22.8 ± 4.2	
H3H4	1	81.9 ± 3.5	78.6 ± 3.5	95.9 ± 4.3	1.1
	2	80.6 ± 5.5	60.9 ± 5.5	83.6 ± 7.6	
H5	1	13.2 ± 1.7	9.9 ± 1.7	12.0 ± 2.0	7.2
	2	31.5 ± 3.7	11.8 ± 3.7	16.3 ± 5.1	
H6	1	16.3 ± 1.3	13.0 ± 1.3	15.8 ± 1.6	6.2
	2	31.6 ± 1.9	11.9 ± 1.9	16.3 ± 2.6	
H7	1	6.9 ± 1.1	3.6 ± 1.1	4.3 ± 1.4	21.6
	2	23.3 ± 1.0	3.6 ± 1.0	5.0 ± 1.4	

^a The genotypes of the vectors and the mutants are shown in Fig. 1 and 6, respectively.

^b Real-time PCR analysis was performed 1 or 2 h after infection of target cells.

^c The GFP/R-U5 percentages were calculated as follows: four independent infections were performed, and the copy numbers of the R-U5 and GFP products were determined by quantitative real-time PCR analysis as described in Materials and Methods. The ratio of the GFP and R-U5 copy numbers provides a measure of the reverse transcription initiation efficiency at the 3' PBS relative to the 5' PBS.

^d The GFP/R-U5 percentages detected for dhGFP-Hy 1 and 2 h postinfection were subtracted from the corresponding percentages for all other vectors to adjust for the background signal generated by initiation at the 5' PBS.

^e Initiation efficiency was measured relative to that of the wild-type HIV-2 PBS present in dh2PBS as follows: $100 \times (\text{adjusted GFP/R-U5 percentage for mutant} / \text{adjusted GFP/R-U5 percentage for dh2PBS})$.

^f Decrease in initiation efficiency is defined as $100 / \text{adjusted GFP/R-U5 percentage}$. NA, not applicable.

these GFP/R-U5 percentages were subtracted from the GFP/R-U5 percentages obtained for all vectors containing a wild-type or mutated 3' PBS (Table 1, adjusted GFP/R-U5 column).

In contrast to dhGFP-Hy, the GFP/R-U5 percentages for dh2PBS 1 and 2 h after infection were 85.2 and 92.5%, respectively. Because dh2PBS contained a wild-type HIV-2 PBS region, the adjusted GFP/R-U5 percentage for this vector was set to 100%. Thus, the adjusted GFP/R-U5 percentages for the mutants represent the efficiency of initiation at the 3' PBS relative to the efficiency of initiation at the 5' PBS of dh2PBS. With the exception of H2, the initiation efficiencies obtained 1 and 2 h after infection were in agreement for all mutants.

As shown in Table 1, the H1 mutant, which had a substitution of the distal portion of the lower stem, exhibited initiation efficiencies of approximately 4.5 and 5.1% relative to the wild-type HIV-2 PBS. The result indicates an average decrease in

initiation efficiency of 20.9-fold. Similarly, the H2 mutant, which had a substitution of the proximal portion of the lower stem, exhibited initiation efficiencies of 5 and 11.6% relative to the wild-type HIV-2 PBS (average 14.3-fold decrease). However, the H1H2 mutant, in which the complementarity and the lower stem were restored, exhibited initiation efficiencies of 109.7 and 94.5% relative to the wild-type HIV-2 PBS. This result indicates that the lower part of the U5-IR stem is important for the efficiency of initiation.

Similarly, mutants H3 and H4, in which the distal and proximal portions, respectively, of the upper portion of the U5-IR stem were replaced, exhibited a 6.1- and 4.6-fold reduction in initiation efficiency, respectively, relative to the wild-type HIV-2 PBS. By contrast, the H3H4 mutant, in which complementarity and the upper part of the U5-IR stem were restored, exhibited initiation efficiencies of 95.9 and 83.6% relative to the wild-type HIV-2 PBS. These results demonstrate that the upper portion of the U5-IR stem is also important for maintaining the efficiency of initiation.

Mutant H5, in which the proposed interaction between the primer activation signal in U5 and tRNA^{Lys3} was disrupted, exhibited a 7.2-fold decrease in the efficiency of initiation. Similarly, mutant H6, in which the proposed interaction between the anticodon loop of tRNA^{Lys3} and the A loop in U5 was disrupted, exhibited a 6.2-fold decrease in the efficiency of initiation, indicating that the A-loop sequence influences the efficiency of initiation.

Finally, mutant H7, in which the interaction between the 3' CCA of tRNA^{Lys3} and the PBS was disrupted, exhibited the largest decrease in the efficiency of initiation (22-fold). The severe reduction in initiation efficiency was consistent with the previous observations indicating that primer-template mismatches in the 5' region of the PBS are highly detrimental to viral replication (25, 58, 70).

DISCUSSION

The results of this study show that an HIV-1-based vector containing two PBSs has the capacity to initiate reverse transcription more than once. The results of flow cytometry and Southern blotting analyses are in close agreement with each other. The data indicate that 27% of the infectious HIV-1 particles derived from vectors containing two PBSs initiated reverse transcription twice, which resulted in GFP deletion. The 27% frequency of GFP deletion is very similar to that previously observed in murine leukemia virus-based vectors containing a second PBS derived from spleen necrosis virus. These results suggest that lentiviruses and gammaretroviruses have similar capacities to initiate reverse transcription more than once. It is important to note that multiple initiations in a single virion do not always result in the formation of a GFP-negative provirus. For example, if two initiations occur on one of the copackaged RNAs but minus-strand DNA transfer occurs intermolecularly, a GFP-positive proviral structure is formed. A more thorough consideration of various reverse transcription scenarios for vectors containing two PBSs was described previously (70). Briefly, it was assumed that initiations of DNA synthesis occur randomly on the four PBSs and that intramolecular and intermolecular minus-strand DNA transfers occur with similar efficiencies; based on these as-

sumptions, the predicted frequency of the GFP-negative proviral structures is 25% if all of the virions initiate reverse transcription twice. Thus, the observation in the current study that 27% of the HIV-1 vectors containing two PBSs generated a GFP-negative proviral structure represents an underestimate of virions that underwent dual initiation of reverse transcription.

The results of our study show that HIV-1 proteins can efficiently use the HIV-2 PBS region to initiate reverse transcription *in vivo*. These results confirm and extend previous observations from *in vitro* studies indicating that HIV-1 reverse transcriptase can initiate reverse transcription with the tRNA^{Lys3} primer annealed to the HIV-2 PBS (17). Real-time PCR analysis of early reverse transcription products indicated that the HIV-2 PBS was used at least as efficiently as the HIV-1 PBS.

Our study demonstrates that vectors containing two PBSs provide a quantitative *in vivo* assay for analysis of reverse transcription initiation. Several features of the assay described here contribute to the quantitative analysis of initiation. First, viral replication is limited to a single cycle, and early products of reverse transcription are analyzed so that the potential effects of the mutations on other aspects of viral replication (for example, transcription and integration) are eliminated from the analysis. Second, quantitative PCR analysis was used to enhance the sensitivity of the assay so that the effects of mutations on the initiation efficiency could be determined over at least a 20-fold range. Third, and perhaps most important, the 5' PBS was used as an internal control to eliminate the possible influence of differences in virus production and infection between wild-type and mutant vectors. The use of the 5' PBS as an internal control is based on the reasonable assumption that initiations at the 5' and 3' PBSs occur as independent events; however, it is possible that changes in initiation efficiency of the 3' PBS may affect the initiation efficiency of the 5' PBS and vice versa.

Analysis of the H1, H2, and H1H2 mutants indicates that the lower part of the U5-IR stem is important for efficient initiation. Similarly, analysis of the H3, H4, and H3H4 mutants indicated that the upper part of the U5-IR stem is also important for efficient initiation. The fact that the compensatory mutations (H1H2 and H3H4) restored initiation efficiency to the wild-type level provides strong evidence that these regions interact with each other to form stems that are important for efficient initiation of reverse transcription. The observation that the H1 and H2 mutations had a greater effect on initiation efficiency (15- to 20-fold reduction) than the H3 and H4 mutations (5- to 6-fold reduction) suggests that the lower part of the stem is quantitatively more important for initiation efficiency than the upper part of the stem. However, it is possible that these mutations resulted in the formation of other RNA secondary structures that were detrimental to reverse transcription initiation.

Mutations H5 and H6 disrupted the proposed interactions with the TΨC and the AC loops, respectively (14, 16). The results suggest that these U5 sequences are important for efficiency of initiation; however, it will be necessary to restore these interactions by using mutant tRNAs to verify that the U5 sequences indeed interact with elements of the primer tRNA. Mutation H7 reduced the initiation efficiency approximately

22-fold. A similar mutation at the HIV-1 PBS reduced the viral titer by 1,700-fold (data not shown). The reduction in viral titer is much more severe than the reduction in initiation efficiency, suggesting that the initiation events that occur at the mutated PBS are aberrant. It is possible that the lack of homology between the PBS and the primer tRNA results in synthesis of DNA products that lack an *att* site, which further reduces the viral titer.

These results are in good agreement with previous studies of HIV-2 PBS structure. The structures of the U5-IR and U5-Ψ stems as well as the existence of an unpaired A-loop have been confirmed by chemical probing of viral genomic RNA (14). The base pairing of the A-loop of viral RNA with the AC loop of tRNA^{Lys3} was proposed based on studies with antisense oligonucleotides targeted to these regions (16). Interestingly, *in vitro* studies of HIV-2 reverse transcription showed significant amounts of elongated products in the absence of AC-loop interactions (17). This is consistent with the relatively mild reduction in initiation observed in our study for the H6 mutant (sixfold reduction). Finally, interactions of viral genomic RNA with TΨC and AC loops of the primer tRNA were demonstrated previously by nuclease digestion of initiation complexes (30).

In this study, we used vectors containing two functional PBSs that use the same primer tRNA. These vectors were designed to ensure that the HIV-1 proteins could package the specific primer tRNA and carry out tRNA placement. In future studies, vectors containing two PBSs could be modified to analyze the ability of HIV-1 proteins to package heterologous tRNAs and carry out proper tRNA placement. Additionally, mutational analysis of the HIV-1 PBS could be performed by exchanging the positions of the HIV-1 and HIV-2 PBSs. Finally, mutational analysis of virally encoded reverse transcriptase, nucleocapsid, and integrase could be performed to identify determinants of viral proteins that play an important role in initiation of reverse transcription.

ACKNOWLEDGMENTS

We especially thank Michael Piatak and the late Kalachar Suryanarayana for sharing their expertise in using the *PBGD* gene in real-time PCR experiments. We also especially thank Wei-Shau Hu and Alan Rein for critical reading of the manuscript and Anne Arthur for expert editorial help. We also thank Jean Mbisa, Olga Nikolaitchik, Terence Rhodes, and David Thomas for critical reading of the manuscript and Sook-Kyung Lee for isolation of pCV-1 plasmid DNA.

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

REFERENCES

1. Aiyar, A., D. Cobrinik, Z. Ge, H. J. Kung, and J. Leis. 1992. Interaction between retroviral U5 RNA and the T psi C loop of the tRNA^{Tyr} primer is required for efficient initiation of reverse transcription. *J. Virol.* **66**:2464-2472.
2. Aiyar, A., Z. Ge, and J. Leis. 1994. A specific orientation of RNA secondary structures is required for initiation of reverse transcription. *J. Virol.* **68**:611-618.
3. Anderson, J. A., R. J. Teufel, 2nd, P. D. Yin, and W. S. Hu. 1998. Correlated template-switching events during minus-strand DNA synthesis: a mechanism for high negative interference during retroviral recombination. *J. Virol.* **72**:1186-1194.
4. Artzi, H. B., J. Shemesh, E. Zeelon, B. Amit, L. Kleiman, M. Gorecki, and A. Panet. 1996. Ribonuclease H activity during initiation of reverse transcription using tRNA^{Lys}/RNA primer/template of human immunodeficiency virus. *Arch. Biochem. Biophys.* **325**:209-216.
5. Arya, S. K., C. Guo, S. F. Josephs, and F. Wong-Staal. 1985. Trans-activator gene of human T-lymphotropic virus type III (HTLV-III). *Science* **229**:69-73.

6. Baltimore, D. 1970. RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* **226**:1209–1211.
7. Bauer, G., and H. M. Temin. 1980. Radioimmunological comparison of the DNA polymerases of avian retroviruses. *J. Virol.* **33**:1046–1057.
8. Beemon, K., P. Duesberg, and P. Vogt. 1974. Evidence for crossing-over between avian tumor viruses based on analysis of viral RNAs. *Proc. Natl. Acad. Sci. USA* **71**:4254–4258.
9. Beerens, N., and B. Berkhout. 2000. In vitro studies on tRNA annealing and reverse transcription with mutant HIV-1 RNA templates. *J. Biol. Chem.* **275**:15474–15481.
10. Beerens, N., and B. Berkhout. 2002. The tRNA primer activation signal in the human immunodeficiency virus type 1 genome is important for initiation and processive elongation of reverse transcription. *J. Virol.* **76**:2329–2339.
11. Beerens, N., F. Groot, and B. Berkhout. 2001. Initiation of HIV-1 reverse transcription is regulated by a primer activation signal. *J. Biol. Chem.* **276**:31247–31256.
12. Beerens, N., F. Groot, and B. Berkhout. 2000. Stabilization of the U5-leader stem in the HIV-1 RNA genome affects initiation and elongation of reverse transcription. *Nucleic Acids Res.* **28**:4130–4137.
13. Bender, W., and N. Davidson. 1976. Mapping of poly(A) sequences in the electron microscope reveals unusual structure of type C oncornavirus RNA molecules. *Cell* **7**:595–607.
14. Berkhout, B., and I. Schoneveld. 1993. Secondary structure of the HIV-2 leader RNA comprising the tRNA-primer binding site. *Nucleic Acids Res.* **21**:1171–1178.
15. Billeter, M. A., J. T. Parsons, and J. M. Coffin. 1974. The nucleotide sequence complexity of avian tumor virus RNA. *Proc. Natl. Acad. Sci. USA* **71**:3560–3564.
16. Boulme, F., F. Freund, S. Gryaznov, P. E. Nielsen, L. Tarrago-Litvak, and S. Litvak. 2000. Study of HIV-2 primer-template initiation complex using antisense oligonucleotides. *Eur. J. Biochem.* **267**:2803–2811.
17. Boulme, F., F. Freund, and S. Litvak. 1998. Initiation of in vitro reverse transcription from tRNA^{Lys3} on HIV-1 or HIV-2 RNAs by both type 1 and 2 reverse transcriptases. *FEBS Lett.* **430**:165–170.
18. Brule, F., R. Marquet, L. Rong, M. A. Weinberg, B. P. Roques, S. F. Le Grice, B. Ehresmann, and C. Ehresmann. 2002. Structural and functional properties of the HIV-1 RNA-tRNA(Lys)₃ primer complex annealed by the nucleocapsid protein: comparison with the heat-annealed complex. *RNA* **8**:8–15.
19. Butler, S. L., M. S. Hansen, and F. D. Bushman. 2001. A quantitative assay for HIV DNA integration in vivo. *Nat. Med.* **7**:631–634.
20. Butler, S. L., E. P. Johnson, and F. D. Bushman. 2002. Hum. immunodeficiency virus cDNA metabolism: notable stability of two-long terminal repeat circles. *J. Virol.* **76**:3739–3747.
21. Cen, S., H. Javanbakht, S. Kim, K. Shiba, R. Craven, A. Rein, K. Ewalt, P. Schimmel, K. Musier-Forsyth, and L. Kleiman. 2002. Retrovirus-specific packaging of aminoacyl-tRNA synthetases with cognate primer tRNAs. *J. Virol.* **76**:13111–13115.
22. Cobrinik, D., A. Aiyar, Z. Ge, M. Katzman, H. Huang, and J. Leis. 1991. Overlapping retrovirus U5 sequence elements are required for efficient integration and initiation of reverse transcription. *J. Virol.* **65**:3864–3872.
23. Cobrinik, D., L. Soskey, and J. Leis. 1988. A retroviral RNA secondary structure required for efficient initiation of reverse transcription. *J. Virol.* **62**:3622–3630.
24. Coffin, J. M., S. H. Hughes, and H. E. Varmus. 1997. *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
25. Das, A. T., and B. Berkhout. 1995. Efficient extension of a misaligned tRNA-primer during replication of the HIV-1 retrovirus. *Nucleic Acids Res.* **23**:1319–1326.
26. Das, A. T., B. Klaver, and B. Berkhout. 1997. Sequence variation of the human immunodeficiency virus primer-binding site suggests the use of an alternative tRNA(Lys) molecule in reverse transcription. *J. Gen. Virol.* **78**:837–840.
27. 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.
28. De Rocquigny, H., C. Gabus, A. Vincent, M. C. Fournie-Zaluski, B. Roques, and J. L. Darlix. 1992. Viral RNA annealing activities of human immunodeficiency virus type 1 nucleocapsid protein require only peptide domains outside the zinc fingers. *Proc. Natl. Acad. Sci. USA* **89**:6472–6476.
29. Follenzi, A., L. E. Ailles, S. Bakovic, M. Geuna, and L. Naldini. 2000. Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nat. Genet.* **25**:217–222.
30. Freund, F., F. Boulme, S. Litvak, and L. Tarrago-Litvak. 2001. Initiation of HIV-2 reverse transcription: a secondary structure model of the RNA-tRNA(Lys₃) duplex. *Nucleic Acids Res.* **29**:2757–2765.
31. Fu, W., B. A. Ortiz-Conde, R. J. Gorelick, S. H. Hughes, and A. Rein. 1997. Placement of tRNA primer on the primer-binding site requires pol gene expression in avian but not murine retroviruses. *J. Virol.* **71**:6940–6946.
32. Gritz, L., and J. Davies. 1983. Plasmid-encoded hygromycin B resistance: the sequence of hygromycin B phosphotransferase gene and its expression in *Escherichia coli* and *Saccharomyces cerevisiae*. *Gene* **25**:179–188.
33. Harada, F., R. C. Sawyer, and J. E. Dahlberg. 1975. A primer ribonucleic acid for initiation of in vitro Rous sarcoma virus deoxyribonucleic acid synthesis. *J. Biol. Chem.* **250**:3487–3497.
34. Hu, W. S., and H. M. Temin. 1990. Retroviral recombination and reverse transcription. *Science* **250**:1227–1233.
35. Huang, Y., J. Mak, Q. Cao, Z. Li, M. A. Wainberg, and L. Kleiman. 1994. Incorporation of excess wild-type and mutant tRNA^{Lys3} into human immunodeficiency virus type 1. *J. Virol.* **68**:7676–7683.
36. Isel, C., G. Keith, B. Ehresmann, C. Ehresmann, and R. Marquet. 1998. Mutational analysis of the tRNA^{Lys3}/HIV-1 RNA (primer/template) complex. *Nucleic Acids Res.* **26**:1198–1204.
37. Jang, S. K., H. G. Krausslich, M. J. Nicklin, G. M. Duke, A. C. Palmenberg, and E. Wimmer. 1988. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J. Virol.* **62**:2636–2643.
38. Jiang, M., J. Mak, A. Ladha, E. Cohen, M. Klein, B. Rovinski, and L. Kleiman. 1993. Identification of tRNAs incorporated into wild-type and mutant human immunodeficiency virus type 1. *J. Virol.* **67**:3246–3253.
39. Jones, J. S., R. W. Allan, and H. M. Temin. 1994. One retroviral RNA is sufficient for synthesis of viral DNA. *J. Virol.* **68**:207–216.
40. Jorgensen, R. A., S. J. Rothstein, and W. S. Reznikoff. 1979. A restriction enzyme cleavage map of Tn5 and location of a region encoding neomycin resistance. *Mol. Gen. Genet.* **177**:65–72.
41. Ju, G., and A. M. Skalka. 1980. Nucleotide sequence analysis of the long terminal repeat (LTR) of avian retroviruses: structural similarities with transposable elements. *Cell* **22**:379–386.
42. Julias, J. G., T. Kim, G. Arnold, and V. K. Pathak. 1997. The antiretrovirus drug 3'-azido-3'-deoxythymidine increases the retrovirus mutation rate. *J. Virol.* **71**:4254–4263.
43. Kang, S. M., Z. Zhang, and C. D. Morrow. 1999. Identification of a human immunodeficiency virus type 1 that stably uses tRNA^{Lys1,2} rather than tRNA^{Lys,3} for initiation of reverse transcription. *Virology* **257**:95–105.
44. Khan, R., and D. P. Giedroc. 1992. Recombinant human immunodeficiency virus type 1 nucleocapsid (NCp7) protein unwinds tRNA. *J. Biol. Chem.* **267**:6689–6695.
45. Kung, H. J., J. M. Bailey, N. Davidson, P. K. Vogt, M. O. Nicolson, and R. M. McAllister. 1975. Electron microscope studies of tumor virus RNA. *Cold Spring Harb. Symp. Quant. Biol.* **39**:827–834.
46. Lee, M. S., and R. Craigie. 1994. Protection of retroviral DNA from auto-integration: involvement of a cellular factor. *Proc. Natl. Acad. Sci. USA* **91**:9823–9827.
47. Leis, J., A. Aiyar, and D. Cobrinik. 1993. Regulation of initiation of reverse transcription of retroviruses, p. 33–47. *In* A. M. Skalka and S. Goff (ed.), *Reverse transcriptase*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
48. Levin, J. G., and J. G. Seidman. 1979. Selective packaging of host tRNAs by murine leukemia virus particles does not require genomic RNA. *J. Virol.* **29**:328–335.
49. Mak, J., M. Jiang, M. A. Wainberg, M. L. Hammarskjold, D. Rekosh, and L. Kleiman. 1994. Role of Pr160^{gag-pol} in mediating the selective incorporation of tRNA^{Lys} into human immunodeficiency virus type 1 particles. *J. Virol.* **68**:2065–2072.
50. Murphy, J. E., and S. P. Goff. 1989. Construction and analysis of deletion mutations in the U5 region of Moloney murine leukemia virus: effects on RNA packaging and reverse transcription. *J. Virol.* **63**:319–327.
51. Naldini, L., U. Blomer, F. H. Gage, D. Trono, and I. M. Verma. 1996. Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc. Natl. Acad. Sci. USA* **93**:11382–11388.
52. Naldini, L., U. Blomer, P. Gally, D. Ory, R. Mulligan, F. H. Gage, I. M. Verma, and D. Trono. 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**:263–267.
53. Panet, A., and Z. Kra-Oz. 1978. A competition immunoassay for characterizing the reverse transcriptase of mammalian RNA tumor viruses. *Virology* **89**:95–101.
54. Peters, G., F. Harada, J. E. Dahlberg, A. Panet, W. A. Haseltine, and D. Baltimore. 1977. Low-molecular-weight RNAs of Moloney murine leukemia virus: identification of the primer for RNA-directed DNA synthesis. *J. Virol.* **21**:1031–1041.
55. Peters, G. G., and J. Hu. 1980. Reverse transcriptase as the major determinant for selective packaging of tRNAs into avian sarcoma virus particles. *J. Virol.* **36**:692–700.
56. Ratner, L., W. Haseltine, R. Patarca, K. J. Livak, B. Starcich, S. F. Josephs, E. R. Doran, J. A. Rafalski, E. A. Whitehorn, K. Baumeister, et al. 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature* **313**:277–284.
57. Rein, A., L. E. Henderson, and J. G. Levin. 1998. Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication. *Trends Biochem. Sci.* **23**:297–301.
58. Rhim, H., J. Park, and C. D. Morrow. 1991. Deletions in the tRNA(Lys) primer-binding site of human immunodeficiency virus type 1 identify essential regions for reverse transcription. *J. Virol.* **65**:4555–4564.

59. **Rigour, M., V. Goldschmidt, F. Brule, C. D. Morrow, B. Ehresmann, C. Ehresmann, and R. Marquet.** 2003. Structure-function relationships of the initiation complex of HIV-1 reverse transcription: the case of mutant viruses using tRNA(His) as primer. *Nucleic Acids Res.* **31**:5764–5775.
60. **Rosen, C. A., J. G. Sodroski, K. Campbell, and W. A. Haseltine.** 1986. Construction of recombinant murine retroviruses that express the human T-cell leukemia virus type II and human T-cell lymphotropic virus type III trans activator genes. *J. Virol.* **57**:379–384.
61. **Shank, P. R., and H. E. Varmus.** 1978. Virus-specific DNA in the cytoplasm of avian sarcoma virus-infected cells is a precursor to covalently closed circular viral DNA in the nucleus. *J. Virol.* **25**:104–114.
62. **Shoemaker, C., J. Hoffman, S. P. Goff, and D. Baltimore.** 1981. Intramolecular integration within Moloney murine leukemia virus DNA. *J. Virol.* **40**:164–172.
63. **Sirven, A., F. Pflumio, V. Zennou, M. Titeux, W. Vainchenker, L. Coulombel, A. Dubart-Kupperschmitt, and P. Charneau.** 2000. The human immunodeficiency virus type-1 central DNA flap is a crucial determinant for lentiviral vector nuclear import and gene transduction of human hematopoietic stem cells. *Blood* **96**:4103–4110.
64. **Soderberg, K., L. Denekamp, S. Nikiforow, K. Sautter, R. C. Desrosiers, and L. Alexander.** 2002. A nucleotide substitution in the tRNA^{Lys} primer binding site dramatically increases replication of recombinant simian immunodeficiency virus containing a human immunodeficiency virus type 1 reverse transcriptase. *J. Virol.* **76**:5803–5806.
65. **Swanstrom, R., W. J. DeLorbe, J. M. Bishop, and H. E. Varmus.** 1981. Nucleotide sequence of cloned unintegrated avian sarcoma virus DNA: viral DNA contains direct and inverted repeats similar to those in transposable elements. *Proc. Natl. Acad. Sci. USA* **78**:124–128.
66. **Temin, H. M., and S. Mizutani.** 1970. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* **226**:1211–1213.
67. **Terwilliger, E., J. Proulx, J. Sodroski, and W. A. Haseltine.** 1988. Cell lines that express stably env gene products from three strains of HIV-1. *J. Acquir. Immune Defic. Syndr.* **1**:317–323.
68. **Tsuchihashi, Z., and P. O. Brown.** 1994. DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein. *J. Virol.* **68**:5863–5870.
69. **Vogt, V. M., and M. N. Simon.** 1999. Mass determination of rous sarcoma virus virions by scanning transmission electron microscopy. *J. Virol.* **73**:7050–7055.
70. **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.
71. **Wakefield, J. K., S. M. Kang, and C. D. Morrow.** 1996. Construction of a type 1 human immunodeficiency virus that maintains a primer binding site complementary to tRNA^{His}. *J. Virol.* **70**:966–975.
72. **Wakefield, J. K., H. Rhim, and C. D. Morrow.** 1994. Minimal sequence requirements of a functional human immunodeficiency virus type 1 primer binding site. *J. Virol.* **68**:1605–1614.
73. **Wakefield, J. K., A. G. Wolf, and C. D. Morrow.** 1995. Hum. immunodeficiency virus type 1 can use different tRNAs as primers for reverse transcription but selectively maintains a primer binding site complementary to tRNA^{3Lys}. *J. Virol.* **69**:6021–6029.
74. **Waters, L. C., and B. C. Mullin.** 1977. Transfer RNA in RNA tumor viruses. *Prog. Nucleic Acid Res. Mol. Biol.* **20**:131–160.
75. **Yee, J. K., A. Miyahara, 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.
76. **Yoo, H. W., C. A. Warner, C. H. Chen, and R. J. Desnick.** 1993. Hydroxymethylbilane synthase: complete genomic sequence and amplifiable polymorphisms in the human gene. *Genomics* **15**:21–29.
77. **Zhang, W. H., C. K. Hwang, W. S. Hu, R. J. Gorelick, and V. K. Pathak.** 2002. Zinc finger domain of murine leukemia virus nucleocapsid protein enhances the rate of viral DNA synthesis in vivo. *J. Virol.* **76**:7473–7484.