Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 9490–9494, September 1994
Genetics

High rate of mismatch extension during reverse transcription in a single round of retrovirus replication

GARY A. PULSINELLI AND HOWARD M. TEMIN†

McArdle Laboratory for Cancer Research, University of Wisconsin Medical School, 1400 University Avenue, Madison, WI 53706

Communicated by James Crow, May 23, 1994

ABSTRACT **We made spleen necrosis virus-based retroviral vectors with mutations at the 3′ end of the primer binding site region to observe the effects of terminal mismatches on retroviral replication. These vectors, when compared to a vector with the wild-type primer binding sequence, allowed us to assay the effects of the mutations on the viral titer during a single cycle of replication. The mutant vectors had titers that were comparable to the wild-type vector, indicating that reverse transcriptase has no trouble extending mismatches of as many as 3 bases under normal *in vivo* conditions. These results confirm and extend previous *in vitro* studies [Yu, H. & Goodman, M. (1992) *J. Biol. Chem.* 15, 10888–10896] that showed that such mismatch extension could occur in a cell-free system at high concentrations of incorrect nucleotides and in the absence of correct nucleotides. We now show that mismatch extension can occur during normal retroviral replication in cells and at normal physiological nucleotide concentrations.**

Retroviruses are RNA viruses that replicate through a DNA intermediate. The process by which viral RNA is copied into DNA is known as reverse transcription, and it is carried out by the viral-encoded enzyme reverse transcriptase (1, 2). This reverse transcription process is extremely error-prone, which gives retroviruses a very high mutation rate. This high mutation rate, in turn, leads to a high degree of genomic variability, which allows retroviruses to adapt quickly to changing environmental conditions.

Several features of the reverse transcription process can give rise to these mutational errors. First, the enzyme reverse transcriptase itself is highly error-prone, and it makes many mistakes in copying viral RNA into DNA; these mistakes include base-pair substitutions, frameshifts, deletions, and insertions, as well as combinations of these alterations (3, 4). Reverse transcription also involves two primer transfer steps (plus-strand primer transfer and minus-strand primer transfer; for further details of the reverse transcription process, see ref. 1), and these primer transfers also seem prone to errors, which can generate mutant progeny. In fact, the requirement for reverse transcriptase to facilitate these primer jumps may lead to the overall high error rate (5).

In a previous report, we (6) presented results from a system for studying deletions generated during a single round of retrovirus replication. This system uses a series of helper cells and nonhelper target cells to isolate proviruses that have undergone only one round of replication, thus allowing us to separate the events of a single round of replication from the cumulative effects of multiple rounds. The vector used in these experiments, JD216HyNeo, codes for two resistance genes, *hyg* and *neo* [*hyg* is the hygromycin B gene, conferring resistance to hygromycin B (Hyg^r); *neo* is the bacterial neomycin gene, conferring resistance to the neomycin analog G418 in mammalian cells].

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

In this report, we describe experiments with vectors derived from JD216HyNeo. These derivatives contain mutant bases at the 3′ end of the primer binding-site (PBS) region, and they enable us to assay the ability of the reverse transcriptase to extend mismatches during normal retroviral replication. After plus-strand primer transfer, these 3′ PBS bases normally pair with sequences copied from the tRNA that serves as the primer for minus-strand synthesis; the PBS mutants have a mismatch at this position. These mutants present a unique opportunity to study the effects of primer–template mismatches on normal retroviral replication.

The results from our studies with these vectors indicate that reverse transcriptase can extend mismatches ranging in size from 1 to 3 bases. Other investigators have studied mismatch extension by reverse transcriptase (7–10), but all of this work has been with cell-free systems containing high nucleotide concentrations and using purified reverse transcriptase [usually from human immunodeficiency virus type 1 (HIV-1)] to extend a single base mismatch on a nonviral template. Yu and Goodman (11) have shown that HIV-1 reverse transcriptase routinely extends mismatches of up to 3 bases, but again in a cell-free system containing unnaturally high concentrations of incorrect nucleotides and lacking correct nucleotides. Our results confirm and extend this previous work in a system close to the natural viral replication events that occur within cells at normal nucleotide concentrations. The mutant vectors had titers that were comparable to the wild-type vector, and sequencing of the resulting proviruses indicated that both wild-type and mutant sequences were found at the PBS region. The observed mixture of wild-type and mutant sequences is consistent with known mechanisms and nucleotide preferences of the mammalian DNA mismatch repair system (12).

MATERIALS AND METHODS

Nomenclature. The "p" at the beginning of construct names (e.g., pJD216HyNeo) designates the plasmid vectors, whereas names without the "p" (e.g., JD216HyNeo) denote viruses (or proviruses) derived from the corresponding plasmid vector. All sequence coordinates are given relative to the 5′ end of the 5′ long terminal repeat (LTR) in the plasmid.

Plasmids and Vectors. The vector pJD216HyNeo is described in Dougherty and Temin (13). pGP1 contains the 1100-bp *Eco*RI fragment from pJD216HyNeo cloned into pTZ19R (Pharmacia), and pGP2 is the remaining pJD216HyNeo backbone ligated back together. The vectors pGP216mut1.3, pGP216mut2.1, pGP216mut2.2, pGP216mut2.3, and pGP216mut2.1Δ are PBS mutant derivatives of pJD216HyNeo (see Fig. 1).

Cells, Transfection, and Infection. D17 cells are derived from a dog osteosarcoma and are permissive for spleen necrosis virus (SNV) infection. C3A2 cells are a helper cell

Abbreviations: SNV, spleen necrosis virus; PBS, primer binding site; Hyg^r, hygromycin resistant; HIV-1, human immunodeficiency virus type 1; LTR, long terminal repeat.

†Deceased February 9, 1994.

line created by transfection of DNA for two defective reticuloendotheliosis virus, strain A, proviruses into D17 cells (14). DSN cells are an SNV-based helper cell line derived from D17 cells; they contain the coding sequences for viral proteins but virtually no other viral sequences (15). The DSDh cells are similar to DSN cells; they differ only in the selectable marker used in constructing them (16). DSN cells contain the *neo* gene for G418 resistance (17), while DSDh cells contain a mutant dihydrofolate reductase gene (18) for methotrexate resistance. Cell culture and selection were carried out as described (7).

Transfection was performed by the dimethyl sulfoxide/Polybrene method (19). Infections were done with 0.2 ml or 0.4 ml of virus (clarified by centrifugation at 1600 × *g* for 4 min) plus 0.2 ml of Polybrene (100 μg/ml) for 40 min at 37°C.

To obtain the Hyg^r titer, 10-fold serial dilutions of virus were used to infect D17 cells. These D17 cells were selected with hygromycin B, and the titer was determined at endpoint dilution (7).

PBS Mutagenesis. To facilitate mutagenesis, an *Eco*RI fragment from pJD216HyNeo containing the PBS was subcloned into pTZ18R, creating pGP1. Single-stranded phagemid DNA was isolated, and site-directed mutagenesis was performed by the method described in McClary *et al.* (20); the mutagenesis primers, obtained from Genosys (The Woodlands, TX) or synthesized on an Applied Biosystems model 391 DNA synthesizer, are (sequences listed 5′ → 3′; changed bases are underlined; mutant plasmids generated from each primer are listed in parentheses on the right)

pbs-mut1	CTCGTCCGGGACACCCTCCCC	(pGP216mut2.1)
pbs-mut1.3	CTCGTCCGGAGCACCCTCCCC	(pGP216mut1.3)
pbs-mut2	CTCGTCCGGGCCACCCTCCCC	(pGP216mut2.2, -mut2.1Δ)
pbs-mut2.3	CTCGTCCGGCCCACCCTCCCC	(pGP216mut2.3)

After verification of mutations 1.3, 2.1, 2.2, 2.3, and 2.1Δ by sequencing (Sequenase kit, United States Biochemical), the *Eco*RI fragment was cloned back into pGP2, regenerating an intact proviral vector. Restriction enzymes were obtained from Boehringer Mannheim unless otherwise noted. The final mutant plasmids were designated pGP216mut1.3, pGP216mut2.1, pGP216mut2.2, pGP216mut2.3, and pGP216mut2.1Δ.

PCR. To obtain material for PCR analysis of proviruses, confluent 60-mm dishes of cells were rinsed with Tris buffer and then lysed in 0.5 ml of PCR buffer with nonionic detergents and proteinase K [50 mM KCl/10 mM Tris, pH 8.3/1.8 mM $MgCl_2$/0.45% Nonidet P-40/0.45% Tween 20; 6 μl of proteinase K (20 mg/ml) was added per ml of lysis buffer just before use] (21). After transfer to a microcentrifuge tube, the lysates were incubated at 58°C for 1 hr and then stored at −20°C.

The PCR was carried out on lysates from the resistant cell clones of interest. Primers were purchased from Genosys. The primers used were (sequences listed 5′ → 3′) U3-60, GCTTGCCCTGGCCACTAACCG, located in the U3 region of the 5′ LTR at position 60; and neo-rev, ACCCAAGCGGCCGGAGAACCT, located just past the initiation codon for the *neo* gene at position 2600 and in the opposite orientation to the vector. Amplifications were performed in a buffer consisting of 50 mM KCl, 10 mM Tris (pH 8.3), 1.8 mM $MgCl_2$, and 40 μM each dNTP. For amplification, 1–20 μl of lysate was added to this buffer with 5 pmol of each of the two primers in a total volume of 100 μl. After heating the reaction mixtures to 95°C for 12 min to denature the DNA and to destroy the proteinase K, 2 units of *Taq* polymerase (obtained from Perkin–Elmer/Cetus or Boehringer Mannheim) was added to each reaction mixture. Amplifications were performed for 35–40 cycles in a Coy Laboratory Products (Ann Arbor, MI) TempCycler (model 60) or on a Perkin–Elmer/Cetus Thermal Cycler (model 480). Products were analyzed by electrophoresis on 1% TBE agarose gels (1/10th of the total PCR reaction mixture per lane).

PCR Sequencing and Digestion. The primers used in sequencing were (sequences listed 5′ → 3′) U5-550, TTGGCTCGCCTACTGGGTGGG, located in the U5 region of the 5′ LTR at position 550; and E-seq, CCAAACCCTCGCGAGGTACCA, located at the beginning of the encapsidation sequence at position 720 and in the opposite orientation to the vector.

For PBS mutants 1.3, 2.1, 2.1Δ, and 2.3, the PCR products were analyzed by direct sequencing. Templates were sequenced using one of the following protocols:

(*i*) After removal of the mineral oil by chloroform extraction, primers and excess nucleotides were removed from the PCR reactions using a Linkers 6 Quick Spin column (Boehringer Mannheim) or using a Chroma Spin-400 column (Clontech), in both cases according to the manufacturer's directions. The PCR products were then denatured for sequencing by boiling (22). Sequencing was then performed using the Sequenase kit (United States Biochemical), essentially as described in the Sequenase manual, except that the annealing was done for 10 min at 37°C rather than at 65°C and the supplied Mn^{2+} buffer was added.

(*ii*) Alternatively, the PCR products were used directly for cycle sequencing with the AmpliTaq cycle sequencing kit (Perkin–Elmer/Cetus). This method was used to generate most of the sequence data in this report.

For analysis of PBS mutant 2.2, 1 μl of *Hae* III (New England Biolabs) was added to 9–18 μl of PCR product and incubated at 37°C for 1 hr. The resulting digests were analyzed by electrophoresis on a 1.5% agarose gel.

The computer program DNA STRIDER (23) was used for sequence organization and analysis.

RESULTS

To study the effects of template–primer mismatches on retroviral replication, we constructed vectors with mutations at the end of the PBS region (Fig. 1). The wild-type vector JD216HyNeo has the bases . . . cggGAT . . . at the 3′ end of the PBS. The mutant vector GP216mut1.3 contains the

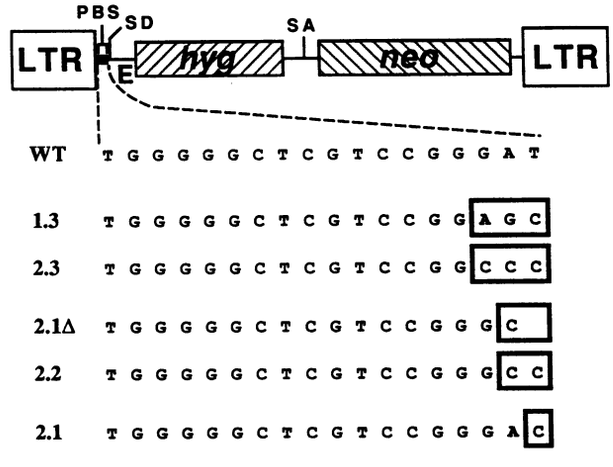

FIG. 1. PBS mutants. The vector JD216HyNeo is shown at the top. LTR, SNV LTR; E, SNV encapsidation sequence; SD, splice donor sequence; SA, splice acceptor sequence. The first sequence line shows the wild-type (WT) PBS sequence. The remaining lines show the PBS regions of the mutant constructs; the changed bases are boxed. Note that mutant 2.1Δ is 1 base shorter than the others.

transition mutation for each of the last three bases (. . . cg-gAGC . . . at the 3′ end of the PBS); transitions were chosen since they are the most common type of base substitution error made by reverse transcriptase (3), and we wanted to generate a situation that was as representative of real reverse transcriptase errors as possible. The mutant vectors GP216mut2.1, 2.2, and 2.3 change the last one, two, or three bases, respectively, to C.

The PBS region is involved in both primer transfer steps of retroviral replication. First, this region is the site at which the tRNA primer binds and initiates minus-strand synthesis. We predicted that these mutations would not interfere with this initiation, since the tRNA primer would still have at least 15 normal PBS bases with which to pair and the small disruption would be at the 5′ end of the primer. Second, the PBS region is the transfer point for the plus-strand primer. At the completion of minus-strand DNA synthesis, the minus-strand DNA contains the complement of the PBS sequence from the viral vector RNA. The 3′ end of the plus-strand primer contains the PBS sequence copied from the tRNA primer. Normally, plus-strand primer transfer results in the pairing of these two complementary PBS sequences, followed by extension to complete the double-stranded viral genome. However, with our mutants, plus-strand primer transfer should result in mismatches between the mutant bases copied from the vector PBS (now present in the minus-strand DNA) and the wild-type bases copied from the normal tRNA primer (now present at the 3′ end of the plus-strand primer DNA) (Fig. 2). Such mismatches could either be easily extended as observed *in vitro* (12), with little or no effect on viral titer, or they might not be extended, reducing viral titer. Further, proviruses with wild-type sequences in the PBS can only arise if the reverse transcriptase extends the mismatch between the wild-type bases from the tRNA and the mutant bases from the vector PBS, since the tRNA is the only source of wild-type bases present during reverse transcription.

The mutant vectors were introduced into DSDh helper cells by transfection, followed by selection for Hyg^r clones. These clones were designated Step 1 producer clones. In some cases, virus harvested from these clones was used to infect a further round of DSDh helper cells; clones derived from these infections were designated Step 2 producer clones. The proviruses in these Step 1 and Step 2 producer clones were verified by PCR amplification followed by direct sequencing of the resulting PCR products. Only virus obtained from clones containing proviruses with the correct mutant sequence was used for subsequent infections. Virus harvested from either Step 1 or Step 2 producer clones was then serially diluted and used to infect D17 target cells, which were selected for Hyg^r clones, and the Hyg^r titer was obtained.

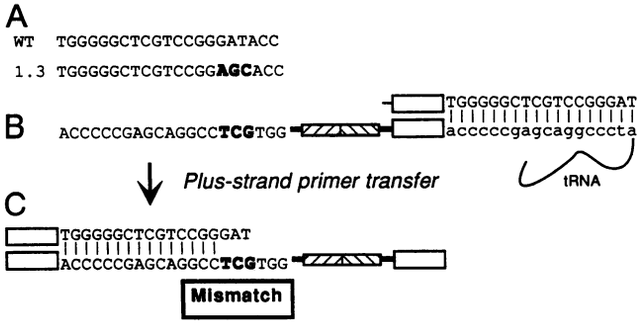

FIG. 2. Predicted effects of PBS mutants. (*A*) Sequences of the wild-type (WT) PBS and mutant 1.3 proviral PBS. The mutant bases are shown in outlined letters. (*B*) Viral DNA after completion of plus-strand DNA primer synthesis (top strand) and minus-strand DNA synthesis (bottom strand). White boxes, LTR sequences; hatched boxes, *hyg* and *neo* coding regions; thick lines, viral sequences. (*C*) Viral DNA after plus-strand primer transfer. The last three bases of the wild-type plus-strand primer cannot pair with the mutant PBS in the minus-strand DNA, creating a primer–template mismatch. Correct base pairing is indicated by vertical lines.

The range of Hyg^r titers observed was similar to that seen with the wild-type vector (24), indicating that the mutations did not interfere with either minus-strand or plus-strand priming or with any other part of retroviral replication. We used PCR to amplify the Hyg^r D17 clones and determined the sequence at the PBS by one of three methods: (*i*) direct sequencing after column purification, (*ii*) direct cycle sequencing, or (*iii*) restriction analysis for mutant 2.2. This sequence analysis revealed that 50–80% of the Hyg^r clones maintained the mutant PBS sequence, depending on the specific mutant; most of the remainder reverted exactly to the wild-type sequence; and in some cases both types of provirus were present (Table 1) (these clones with both proviruses presumably arose when a cell containing a provirus with an unrepaired mismatch in the PBS region divided into two cells, each of which then contained a different proviral sequence; see *Discussion*). The presence of the wild-type and mixed sequences indicated that, at least some portion of the time, the reverse transcriptase was able to extend mismatches of 1–3 bases. As further support for this hypothesis, we found one Hyg^r provirus derived from mutant 1.3 (3 changed PBS bases) in which only the last base remained mutant; the first two reverted to wild type, suggesting that they were copied from the tRNA, whereas the last base was not, and then reverse transcriptase extended the 2-base mismatch to create the combined PBS provirus (see footnote *, Table 1; see *Discussion*).

Because we were concerned about the possibility that contaminating wild-type viruses were responsible for the observed wild-type Hyg^r clones, we constructed the GP216mut2.1Δ mutant. The mutant vector 2.1Δ changes the last two bases to a single C, which is the equivalent of a 2-base mismatch (see Figs. 1 and 3). This vector allows us to discriminate between a contaminating wild-type vector and a true mismatch extension, since the two types of clones will have different sequences (Fig. 3); the pseudo-wild-type vector resulting from mismatch extension should be one base shorter than the true wild-type vector. We observed no Hyg^r proviruses with completely wild-type PBS regions but found many mismatch-extended proviruses or both mismatch-extended and mutant 2.1Δ proviruses (Table 1). Thus, the wild-type clones we observed are, in fact, the result of mismatch extension and not contamination.

Table 1. Proviral sequences of infected Hyg^r D17 clones from PBS mutant vectors

PBS mutant		Provirus sequence distribution			
Name	No. of changes	Mutant	Wild type	Both	Other
mut1.3	3	15	2	3	1*
mut2.3	3	10	2	0	0
mut2.2	2	9	10	1	0
mut2.1Δ	2	4	12†	10	0
mut2.1	1	20	7	1	1‡

*Clone 1.3-2(D)-6H has portions of both wild-type and mutant sequence:

Wild-type vector	AACATT TGGGGGCTCGTCCG GAT ACCTC
Clone 1.3-2(D)-6H	AACATT TGGGGGCTCGTCCG GAC ACCTC
Mutant 1.3 vector	AACATT TGGGGGCTCGTCCG AGC ACCTC

†For mut2.1Δ, "wild type" is defined as the read-through revertant, which has a deletion of 1 base 3′ to the PBS.

‡Clone 2.1-7(D)-3H has a PBS deletion with an insertion; the last base of the PBS is changed to G, matching neither the wild-type T nor the mut2.1 C.

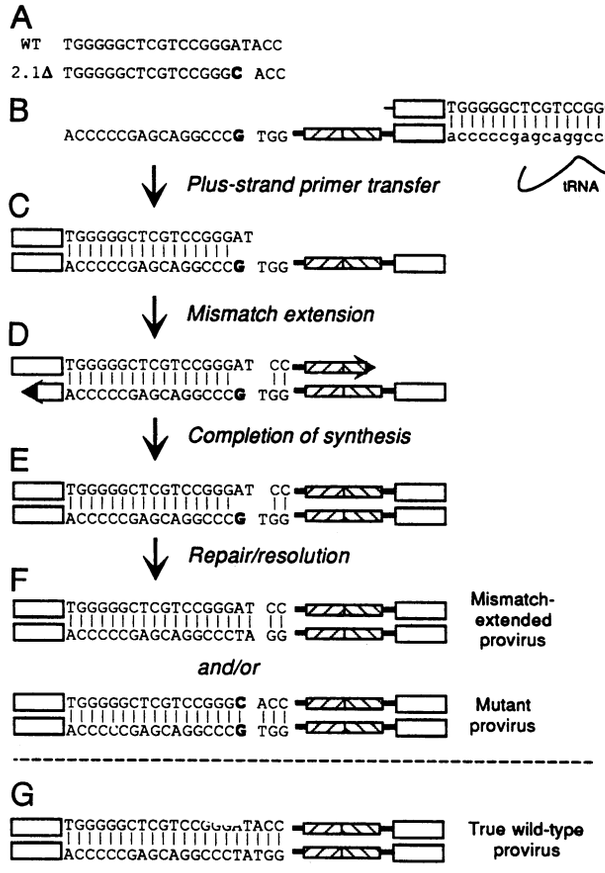

FIG. 3. Predicted outcome of mismatch extension of mutant 2.1Δ vector. The spaces in the sequences are included only to emphasize that these sequences are 1 base shorter than the wild-type sequence in these regions; no actual gaps exist in the DNA. (*A*) Sequences of the wild-type (WT) PBS and mutant 2.1Δ proviral PBS. The mutant base is shown in boldfaced type. (*B*) Viral DNA after completion of plus-strand primer DNA synthesis (top strand) and minus-strand DNA synthesis (bottom strand). Symbols are as in Fig. 2. (*C*) Viral DNA after plus-strand primer transfer. The last 2 bases of the wild-type primer cannot pair with the mutant PBS in the minus-strand DNA, creating a mismatch in the DNA. (*D*) Viral DNA after extension of the plus-strand DNA past the mismatch and extension of the minus-strand DNA. (*E*) Viral DNA after completion of DNA synthesis, containing a mismatch at the end of the PBS. (*F*) Possible outcomes of mismatched PBS: repair to mismatch-extended provirus, repair to mutant provirus, or no repair followed by cell division to yield a clone with two different proviruses (not shown). Note that the mismatch-extended provirus is 1 base shorter than the true wild-type provirus shown in *G*.

DISCUSSION

To study the effects of primer–template mismatches on retroviral replication, we mutated the bases at the 3′ end of the PBS in the vector JD216HyNeo (Fig. 1). We made several mutations, ranging in length from 1 to 3 bases. These mutations left at least 15 bases to which the tRNA could bind and prime minus-strand DNA synthesis. After completion of minus-strand synthesis and priming of plus-strand DNA synthesis, the plus-strand primer DNA molecule should contain the wild-type bases at the end of the PBS (copied from the normal tRNA), whereas the minus-strand DNA should contain mutant bases at the end of the PBS (copied from the mutant vector) (Fig. 2). After plus-strand primer transfer, the two sequences should be mismatched, presenting us a unique opportunity to study the effects of primer–template mismatches on normal retroviral replication.

When we tested these vectors in cell culture, we saw no apparent reduction in the Hyg^r titer (24). Many of the proviruses from the Hyg^r cells that we amplified and sequenced retained the mutant PBS sequence, but many others had an exact reversion to wild type (Table 1). The presence of the wild-type PBS in a significant portion of the Hyg^r proviruses indicates that the reverse transcriptase is able to extend mismatches as long as 3 bases (and possibly longer, as this is the longest mismatch we have studied). Yu and Goodman (11) reported a similar result in an *in vitro* system containing a high concentration of incorrect nucleotides and lacking correct nucleotides; we now show that such mismatch extension can occur *in vivo* and under normal reverse transcription conditions (including normal, physiological nucleotide concentrations). Further support for this conclusion is provided by the structure of Hyg^r clone 1.3-2(D)-6H (see footnote *, Table 1). This clone has a mixed structure at the end of the PBS; that is, the next-to-last 2 bases are derived from the tRNA (and are hence wild type), while the last base is derived from the vector PBS (and is hence mutant). This partial reversion indicates that the plus-strand primer transfer occurred just before the copying of the final base of the PBS in the tRNA, and the reverse transcriptase was then able to extend past the 2-base mismatch.

This finding of a high rate of mismatch extension during normal retroviral replication is particularly important in reference to the work of Perrino *et al.* (7). They report that the high HIV-1 misincorporation rate is not due to a higher frequency of incorporating incorrect nucleotides into the growing strand but rather to a higher frequency of extending these incorrect nucleotides after they are incorporated. The high rate of mismatch extension now observed by us in an SNV vector system thus suggests that this mechanism is also operative during normal retroviral replication and not solely in a cell-free system.

We must point out a possible difficulty with these data. The PBS mutant vectors showed a remarkable tendency to revert to wild type upon transfection or infection of certain types of helper cells (namely, the C3A2 and DSN lines). Reversion frequencies in these cells ranged from 50% to 100% (24). This tendency was not observed with the helper cell line DSDh or target cell line D17. For this reason, most of the work described in this paper was performed using DSDh helper cells to provide virus for infecting the D17 target cells, and all Step 1 and Step 2 producers used to generate the described results were screened (by sequencing or digestion) to be sure that they maintained the PBS mutation; however, we could not exclude the possibility that some residual contaminating wild-type proviruses were present at some low level and contributed to the wild-type proviruses in the D17 target cells or that the unknown reversion tendency was active in the target cells at a much lower frequency than in the helper cells.

To address this point, we used the vector GP216mut2.1Δ. This vector changes the last 2 bases of the PBS to a single C (AT to C). This mutant thus allows us to distinguish a true reversion to wild-type from a mismatch extension, since the mismatch extension will be 1 base shorter than the true revertant (Fig. 3). All of the progeny of this vector contained either a mutant PBS, a mismatch-extended PBS, or both PBS sequences; we saw no instances of true reversion to complete wild type. Thus, the phenomenon we are studying is, in fact, true mismatch extension and not an artifact of contamination with wild-type vectors.

This high rate of mismatch extension strongly contradicts our original hypothesis about the origin of the PBS deletions. This hypothesis was dependent on the inability of the reverse transcriptase to extend such mismatches and instead to use them to create deletions. This contradiction, along with other data and a new hypothesis, will be discussed in another report.

Because the mismatch extension creates an unpaired region in the integrated provirus, we attempted to interpret the

distribution of mutant and wild-type proviruses in terms of what is known about mammalian DNA mismatch repair. Not much is known about the details of this process, but some data from *in vitro* systems is available (refs. 25 and 26; reviewed in ref. 13). Results from these systems indicate that the repair process is initiated at a nick in the DNA, with the nicked strand being the strand that is repaired. The retroviral integration process proceeds in a way that leaves a gap at the 5′ end of the plus strand [and minus strand, but this gap is too far away to be used by the system, which seems limited to about 1 kb (13)], which is later repaired (27). This gap could facilitate the repair process and guide it to the plus strand; the result would be a tendency to repair the wild-type strand to match the mutant strand. This tendency would account for the observed preference for the 1.3, 2.3, and 2.1 mutants to maintain the mutant sequence at such high frequency. Alternatively, the repair of the mismatched PBS sequences might simply be part of the normal gap-repair process required for retroviral integration.

The repair process also has pronounced tendency to repair certain mismatches more efficiently than others (13). The 1.3 and 2.1 mutants are almost ideal substrates for repair; they have the mismatches that are most easily repaired (indeed, the 2.1 G·T mismatch is the most easily repaired, and the cell may even have special mechanisms for doing so). Mutants 2.2, 2.3, and 3 are mixed—they contain some mismatches that are easily repaired and others whose repair is dependent on sequence context. The 2.3 mutant appears to be in a good context, because it is repaired as easily as the 1.3 mutant. The 2.2 and 3 mutants may be in a less good context, judging from the more even distribution between mutant and wild-type proviruses that we observed. The 2.1Δ mutant contains two mismatches, which appear to be intractable to repair in the current sequence context. Since the mismatch is therefore not repaired, cell division results in two cells with different sequences, which would account for the observed tendency to see both sequences in the 2.1Δ clones. [Berwin and Barklis (28) clearly show that a retroviral vector with a mismatched PBS can integrate and give rise to two populations of cells with identical integration positions but different PBS sequences; Thomas *et al.* (26) also showed that the repair process is not 100% effective and depends on the sequence being repaired.] A system based on the ability of the reverse transcription process to create these unpaired regions might be valuable in providing further insight into the mammalian DNA mismatch repair process.

An alternative interpretation for the wild-type/mutant PBS distribution is that the synthesis of the plus-strand primer failed to reach the end of the tRNA and instead the primer transferred prematurely. Our data cannot rule out this possibility; indeed, we believe that this is exactly what occurred with clone 1.3-2(D)-6H (see footnote *, Table 1). However, if incomplete synthesis were the primary source of the proviruses with mutant sequences, then the mutants of similar length should have similar distributions of mutant and wild-type sequences. This is true for mutants 1.3 and 2.3 but not for 2.2 and 2.1Δ. Further, premature transfer would also be expected to lead to more recombined PBS sequences similar to clone 1.3-2(D)-6H, which we did not observe.

G.A.P. dedicates this manuscript to the memory of H.M.T. We thank T. Jacoby, B. Pietz, J. Schoening, and B. Seufzer for their technical assistance; W. Sugden, A. Panganiban, and D. Loeb for critical reading of the manuscript; and K. Boris-Lawrie, D. Burns, L. Mansky, and S. Yang for discussion and helpful comments. This work was supported by Grants CA-07175 and CA-22443 from the U.S. Public Health Service. H.M.T. was an American Cancer Society research professor; G.A.P. was supported by Training Grant CA-09135 from the National Institutes of Health.

1. Coffin, J. M. (1990) in *Virology*, eds. Fields, B. N. & Knipe, D. M. (Raven, New York), 2nd Ed., Vol. 1, pp. 1437–1500.
2. Varmus, H. & Brown, P. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (Am. Soc. Microbiol., Washington, DC), pp. 53–108.
3. Pathak, V. & Temin, H. M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6019–6023.
4. Pathak, V. K. & Temin, H. M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6024–6028.
5. Temin, H. M. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 6900–6903.
6. Pulsinelli, G. & Temin, H. M. (1991) *J. Virol.* **65**, 4786–4797.
7. Perrino, F. W., Preston, B. D., Sandell, L. L. & Loeb, L. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8343–8347.
8. Mendelman, L. V., Petruska, J. & Goodman, M. F. (1990) *J. Biol. Chem.* **265**, 2338–2346.
9. Creighton, S., Huang, M.-M., Cai, H., Arnheim, N. & Goodman, M. F. (1992) *J. Biol. Chem.* **265**, 2633–2639.
10. Bakhanashvili, M. & Hizi, A. (1993) *FEBS Lett.* **319**, 201–205.
11. Yu, H. & Goodman, M. F. (1992) *J. Biol. Chem.* **267**, 10888–10896.
12. Modrich, P. (1991) *Annu. Rev. Genet.* **25**, 229–253.
13. Dougherty, J. P. & Temin, H. M. (1986) *Mol. Cell. Biol.* **6**, 4387–4395.
14. Watanabe, S. & Temin, H. M. (1983) *Mol. Cell. Biol.* **3**, 2241–2249.
15. Dougherty, J. P., Wisniewski, R., Yang, S., Rhode, B. W. & Temin, H. M. (1989) *J. Virol.* **63**, 3209–3212.
16. Hu, W.-S. & Temin, H. M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1556–1560.
17. Southern, P. J. & Berg, P. J. (1982) *J. Mol. Appl. Genet.* **1**, 327–341.
18. Simonsen, C. C. & Levinson, A. D. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2495–2499.
19. Kawai, S. & Nishizawa, M. (1984) *Mol. Cell. Biol.* **4**, 1172–1174.
20. McClary, J. A., Witney, F. & Geisselsoder, J. (1989) *BioTechniques* **7**, 282–289.
21. Higuchi, R. (1989) in *PCR Technology*, ed. Erlich, H. A. (Stockton, New York), pp. 31–38.
22. Kusukawa, N., Uemori, T., Asada, K. & Kato, I. (1990) *BioTechniques* **9**, 66–72.
23. Marck, C. (1988) *Nucleic Acids Res.* **16**, 1829–1836.
24. Pulsinelli, G. A. (1994) Ph.D. thesis (Univ. of Wisconsin, Madison).
25. Holmes, J. J., Clark, S. & Modrich, P. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 5837–5841.
26. Thomas, D. C., Roberts, J. D. & Kunkel, T. A. (1991) *J. Biol. Chem.* **266**, 3744–3751.
27. Whitcomb, J. M. & Hughes, S. H. (1992) *Annu. Rev. Cell Biol.* **8**, 275–306.
28. Berwin, B. & Barklis, E. (1993) *Nucleic Acids Res.* **21**, 2399–2407.