JOURNAL OF VIROLOGY, May 1992, p. 3093–3100
0022-538X/92/053093-08$02.00/0

Vol. 66, No. 5

5-Azacytidine and RNA Secondary Structure Increase the Retrovirus Mutation Rate

VINAY K. PATHAK† AND HOWARD M. TEMIN*

McArdle Laboratory, 1400 University Avenue, Madison, Wisconsin 53706

Received 5 December 1991/Accepted 14 February 1992

A broad spectrum of mutations occurs at a high rate during a single round of retrovirus replication (V. K. Pathak and H. M. Temin, Proc. Natl. Acad. Sci. USA 87:6019–6023, 1990). We have now determined that this high rate of spontaneous mutation can be further increased by 5-azacytidine (AZC) treatment or by regions of potential RNA secondary structure. We found a 13-fold increase in the mutation rate after AZC treatment of retrovirus-producing cells and target cells. The AZC-induced substitutions were located at the same target sites as previously identified spontaneous substitutions. The concordance of the AZC-induced and spontaneous substitutions indicates the presence of reverse transcription "pause sites," where the growing point is error prone. An analysis of nucleotides that neighbored substitutions revealed that transversions occur primarily by transient template misalignment, whereas transitions occur primarily by misincorporation. We also introduced a 34-bp potential stem-loop structure as an in-frame insertion within a *lacZα* gene that was inserted in the long terminal repeat (LTR) U3 region and determined whether this potential secondary structure increased the rate of retrovirus mutations. We found a threefold increase in the retrovirus mutation rate. Fifty-seven of 96 mutations were deletions associated with the potential stem-loop. We also determined that these deletion mutations occurred primarily during minus-strand DNA synthesis by comparing the frequencies of mutations in recovered provirus plasmids containing both LTRs and in provirus plasmids containing only one LTR.

The nucleoside analog 5-azacytidine (AZC) [4-amino-1-β-ribofuranosyl-5-triazine-2(1H)one] was first developed as a potential chemotherapeutic agent against acute myelogenous leukemia (14). AZC is incorporated into RNA and inhibits protein synthesis (19, 21). In addition, AZC decreases the stability of polyribosomes as well as the maturation of rRNA (28). AZC is also incorporated into DNA (29). Such incorporation can prevent methylation of DNA; the inhibition of methylation of DNA of some tissue culture fibroblast cells results in their differentiation into muscle or fat cells (32).

The mutagenic potential of AZC has been investigated in several organisms. AZC has been shown to be mutagenic in *Escherichia coli* (9), *Salmonella typhimurium* (5), and *Saccharomyces cerevisiae* (35) and at the thymidine kinase locus in mammalian cells (1). AZC increases sister chromatid exchanges in human lymphocytes (18). On the other hand, Landolph and Jones reported that AZC failed to induce ouabain or 6-thioguanine resistance in C3H10T1/2 and V79 cells, and they concluded that AZC was not mutagenic at these loci (17).

AZC also causes cancer in rats (6) and mice and transforms BHK cells (23). However, it was suggested that a nonmutational mechanism is involved in this oncogenesis.

AZC was used in early experiments to induce temperature-sensitive mutations in Rous sarcoma virus, murine sarcoma virus, and avian erythroblastosis virus (20). Many of the induced temperature-sensitive mutants have complex biological characteristics and contain two or three mutations.

Recently, AZC was shown to inhibit human immunodeficiency virus type 1 (HIV-1) replication to 60% of the normal levels in a human T-lymphocyte cell line (3). Under the conditions used, no significant cellular toxicity was observed, indicating that cell mortality was not the mechanism of inhibition of virus replication. The 2′-deoxy analog of AZC was an even more potent inhibitor of HIV replication.

Conditional lethal mutants of vesicular stomatitis virus were also induced after AZC treatment (26). Similarly, mutations were induced after AZC treatment of a togavirus strain at a frequency 200-fold higher than the spontaneous frequency (10). However, mutation frequencies at defined single codons in vesicular stomatitis virus were increased only slightly by AZC (11).

We previously described an assay to determine the forward mutation rate for spleen necrosis virus (SNV), an avian type C retrovirus, and to characterize all identified mutations at the molecular level (24) (Fig. 1). We used this system to study the forward mutations in retrovirus replication to determine whether or not AZC treatment increases the already high rate of retrovirus mutation. Since we and others (2) have suggested that the low processivity of reverse transcription plays a role in mutagenesis, we further tested whether or not potential secondary structure increased the mutation rate.

MATERIALS AND METHODS

Plasmid construction. VP212 (shown in Fig. 1) was described previously (24). The retrovirus shuttle vector pVP254, which contains the *lacZα* reporter fragment inserted into the *Sac*I restriction cleavage site in the 3′ long terminal repeat (LTR), 10 bp from the 5′ end of U3, was constructed with DNA fragments taken from pJD215 and pUC19 (7, 34). pUC19 was cleaved with *Eco*RI and *Hin*dIII in the polylinker, and a Klenow enzyme mediated fill-in reaction was performed to generate blunt ends (see Fig. 2). Four copies of an 8-bp *Aat*II linker (5′-GGACGTCC-3′) were inserted at the blunt ends. A total of 36 bp was inserted (32 bp from the *Aat*II linkers and 4 bp from the Klenow

* Corresponding author.
† Present address: Mary Babb Randolph Cancer Center, West Virginia University, Morgantown, WV 26506.

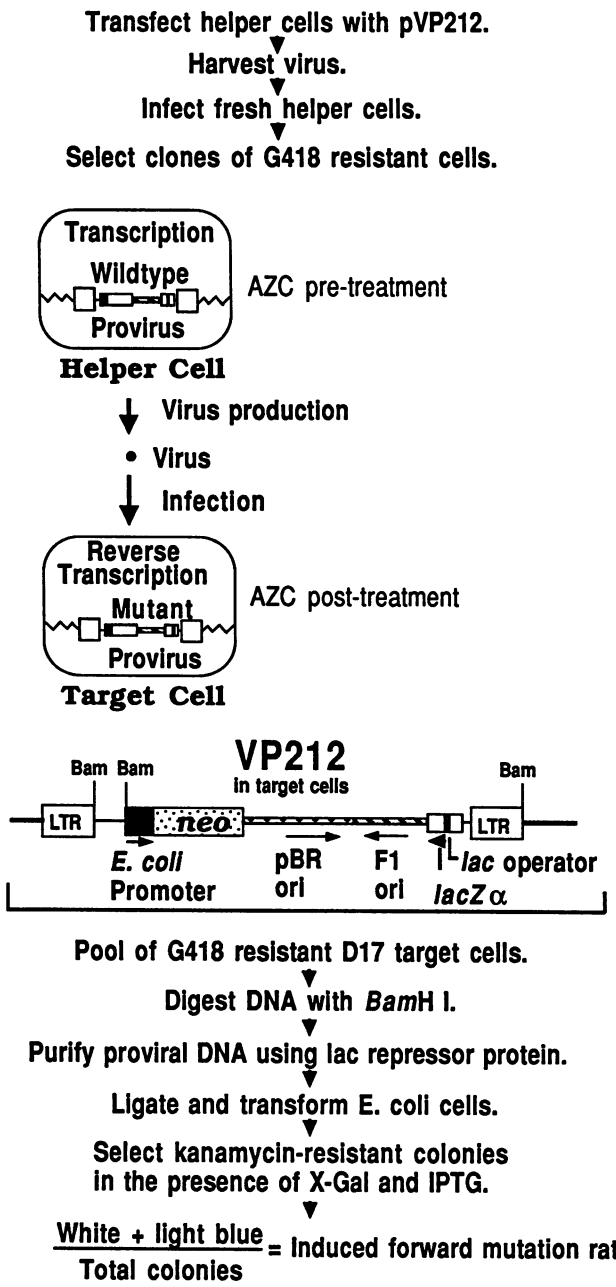

FIG. 1. Outline of experimental approach for determination of effects of AZC on forward mutation rate for a single cycle of retrovirus replication. A map of a provirus derived from an SNV-based retrovirus shuttle vector, pVP212, is shown in the center (see description in Materials and Methods). *neo* (stippled box), neomycin resistance gene; *E. coli* promoter (black box), bacterial promoter for expression of *neo*; pBR ori (hatched box), origin of replication from a pBR322-based plasmid; F1 ori (hatched box), origin of replication of M13 single-stranded phage; *lacZα*, the *lacZα* encoding gene from pUC 19; *lac* operator (narrow black box), *lac* gene operator sequence; Bam, recognition site for *Bam*HI. The experimental approach taken to generate pools of target cells infected with the vector viruses is also described in Materials and Methods. AZC pretreatment, treatment of virus-producing cells with AZC for 24 h before harvest; AZC posttreatment, treatment of target cells with AZC for 24 h after infection. The provirus DNAs were purified, ligated, and used to transform highly competent *E. coli* cells. Kanamycin-resistant bacterial colonies were selected in the presence of the 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) color indicator and the isopropyl-β-D-thiogalactopyranoside (IPTG) inducer. The ratio of white plus light-blue colonies to total colonies provided a forward mutation rate for a single retrovirus replication-cycle.

enzyme fill-in reaction). Standard DNA cloning procedures were used (31). A detailed description of the cloning steps used to generate pVP254 is available on request.

Mutagenesis protocol. Retrovirus shuttle vector plasmid pVP212 was transfected into SNV helper cells and produced virus VP212. As shown in the middle of Fig. 1, a provirus derived from VP212 contains *neo*, the bacterial aminoglycoside 3′-phosphotransferase gene from Tn*5* (13), which can be expressed in mammalian cells from the SNV LTR promoter and confers resistance to G418, a neomycin analog. The same *neo* gene can be expressed from a bacterial promoter in the vector and confers kanamycin resistance in bacterial cells. The pBR origin of replication allows replication of the vector in bacterial cells, and the F1 origin of replication from bacteriophage M13 can be used for isolation of single-stranded DNA for DNA sequencing. The vector also contains the *lacZα* gene. It is used as a reporter of mutant proviruses arising during a single round of retrovirus replication. In appropriate bacterial cells, the wild-type *lacZα* gene confers a blue color to bacterial colonies, whereas a mutated *lacZα* gene gives rise to white or light-blue colonies. In addition, the vectors contain the *lac* operator sequence, which binds tightly and specifically to the Lac repressor protein (22).

We transfected an SNV helper cell line (.2G) with pVP212 and selected G418-resistant cells (Fig. 1). Virus produced from these cells was used to infect fresh .2G helper cells. G418-resistant cells were selected, and individual cell clones were isolated and expanded (in the presence of anti-SNV antibodies to prevent the spread of virus during expansion) to generate helper cell clones infected with VP212. Next, virus produced from the helper cell clones was used to infect D17 cells, a dog cell line. Viral structural proteins are not produced in D17 cells, and therefore spread of the defective VP212 virus does not occur. Because the vector viral genome was introduced into helper cells by infection and single cell clones were isolated, mutant viruses that may have arisen during transfection could be recognized since they would give all mutant progeny when analyzed in bacteria. Mutations occurring in VP254 were studied by using the same protocol as that used for VP212, except that the DSDh helper cell line was used for transfection and infection (12).

The steps in going from a parental provirus in helper cells to a provirus in the target cells are defined as a single replication cycle. These steps include transcription of proviral RNA by the cellular transcriptional machinery, packaging of viral RNA and release of viral particles, infection of target cells, reverse transcription of viral RNA, and integration of the newly synthesized viral DNA to generate a provirus.

The tight binding of the Lac repressor protein to *lac* operator sequences was used to purify proviral DNA fragments from complex mixtures of genomic DNA fragments and to clone the proviral DNA fragments with a high efficiency (24).

DNA isolation and Lac repressor protein-mediated purification of proviral DNA. Genomic DNA isolations were performed by using the phenol and chloroform extraction procedure previously described (31). Proviral DNAs were

TABLE 1. Recovery of proviruses after AZC treatment

AZC treatment and mutation type	No. of mutants/total recovered[a]	No. of different mutations/no. recovered[b]	Mutation rate
Pre- and posttreatment	36/1,392		0.026
Transversion		4/23	
Transition		3/3	
Deletion		1/9	
Frameshift		1/1	
Pretreatment	7/1,120		0.006
Transversion		1/6	
Deletion		1/1	
Posttreatment	7/540		0.013
Deletion		2/2	
Deletion with insertion		1/2	
Frameshift		1/2	
Insertional hypermutation		1/1	

[a] The first number indicates the total number of proviruses with a mutation, and the second number is the total number of proviruses recovered.
[b] The first number indicates the number of different mutations obtained, while the second number indicates the total number of proviruses recovered with the type of mutation indicated.

isolated by using Lac repressor protein as previously described, except that electroporation was used to transform DH5α cells (8).

DNA sequencing. All plasmid sequencing was performed by using the Sequenase kit and protocols (United States Biochemical), with either the dGTP or dITP reactions.

Cells, transfection, and infection. D17 cells, DSDh helper cells, and .2G helper cells were grown as previously described (24). Transfections were performed by the dimethyl sulfoxide-Polybrene procedure (15). Virus harvesting and virus infections were performed as previously described (12).

RESULTS

Recovery and characterization of proviruses after AZC treatment. To determine the effect of AZC treatment on retrovirus mutations, helper cells were either pretreated for 24 h with AZC (1 μM) before harvesting virus or the target D17 cells were treated (1 μM) for 24 h postinfection. In some experiments, both the pretreatment and the posttreatment were done. This concentration of AZC resulted in a reduction of viral titer to 5% of the control and minimal cellular toxicity.

AZC induced a broad spectrum of mutations (Table 1), which was similar to the spectrum of spontaneous mutations and included substitutions (transitions and transversions), frameshifts, and duplications (24). The rate of retrovirus mutation after AZC pre- plus posttreatment during one round of replication was 0.026. In comparison, the rate of spontaneous retrovirus mutation was 0.002 (24), indicating that AZC treatment led to a 13-fold increase in the retrovirus mutation rate. AZC pretreatment alone led to a threefold increase in the mutation rate, and AZC posttreatment alone led to a sixfold increase in the mutation rate, indicating that AZC affected retrovirus mutations during reverse transcription (posttreatment) and also during viral RNA transcription (pretreatment), perhaps by incorporation into viral genomic RNA.

Transitions (purine to purine or pyrimidine to pyrimidine) predominated over transversions (purine to pyrimidine or pyrimidine to purine) in our previous experiments with spontaneous mutations occurring in a single round of retrovirus replication (Table 2) (24). Fifteen of 19 (80%) of the independent spontaneous substitutions were transitions. In the presence of AZC, however, only three of eight independent substitutions were transitions, suggesting that AZC primarily induced transversion substitutions (Table 2). The probability of getting five or more transversions by chance is only about 5% (Fisher's exact test).

TABLE 2. Comparison of spontaneous and AZC-induced substitutions

Mutation type[a]	No. of mutations from following substitution origin:	
	Spontaneous[b]	AZC treated[c]
Transition	15	3
Transversion	4	5
Total	19	8

[a] Each mutant in this table is from a separate pool and therefore represents an independent event.
[b] From reference 24.
[c] From substitution data in Table 1.

Unexpectedly, five of eight independent AZC-induced substitutions occurred at nucleotide positions that were also sites of spontaneous mutation (Fig. 2).

Construction of shuttle vector VP254. To assess the effects of potential secondary structure, retrovirus shuttle vector

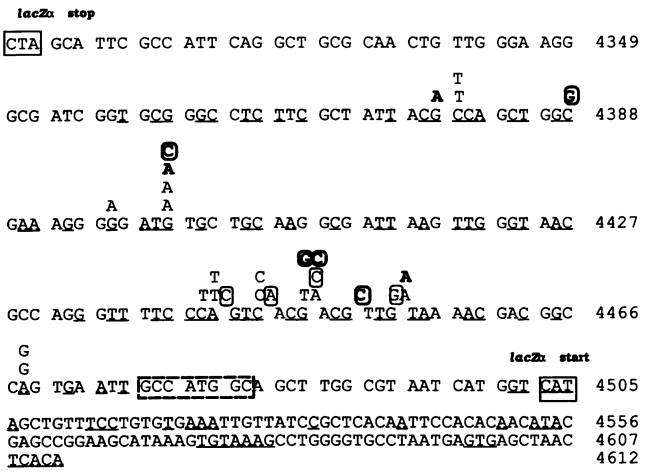

FIG. 2. Viral plus-strand nucleotide sequence of *lacZα* in pVP212, showing sites of spontaneous and AZC-induced substitution mutations. The *lacZα* stop and start codons are indicated with open boxes. Nucleotide numbers on each line are relative to the end of the *Sac*I restriction site near the 5′ end of the LTR. The box with dashed lines indicates the site of insertion of the potential stem-loop structure in VP254; an *Nco*I linker sequence (box) was replaced with four *Aat*II linkers. The sequence shown as triplets is the *lacZα* open reading frame, and the sequence shown as a continuous line is the *lacZα* promoter region. Underlined nucleotides are known mutational targets in *lacZα* (2, 30). Letters above the sequence indicate base pair substitutions found in this and previous work (24). Normal letters indicate spontaneous substitutions, including those found in VP254, and bold letters indicate substitution mutations induced by AZC treatment. Letters in rounded boxes indicate transversion substitutions. Each substitution shown was isolated from a different pool and therefore represents an independent mutation.

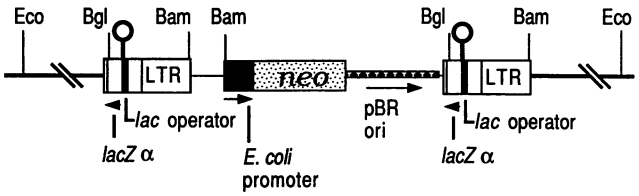

FIG. 3. Provirus derived from pVP254. *lacZα* (open boxes), *lacZα* gene from pUC19; *lac* operator (narrow black box), *lac* operator sequence; *neo* (stippled box), neomycin resistance gene; *E. coli* promoter (black box), bacterial promoter for expression of *neo*; pBR ori (hatched box), origin of replication from pBR322; thick black circle and vertical line, potential stem-loop structure formed by *Aat*II linkers; thick horizontal lines, cellular flanking DNA; Eco, *Eco*RI; Bgl, *Bgl*I; Bam, *Bam*HI cleavage sites.

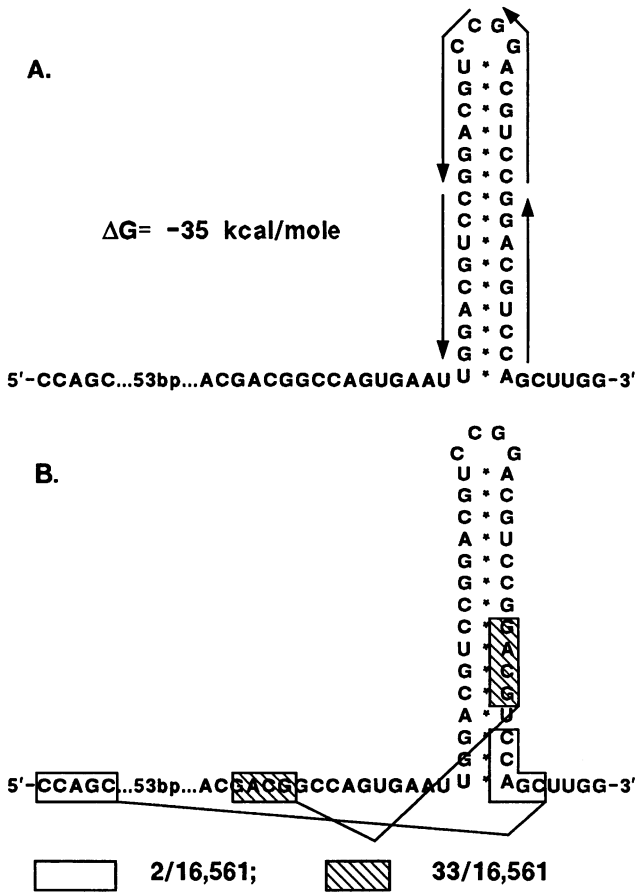

FIG. 4. (A) A predicted stem-loop structure formed by four *Aat*II linkers and its predicted ΔG. Black arrows indicate each *Aat*II linker. (B) Two deletions which initiated either at the base of the stem-loop or within the stem-loop are shown. Open boxes or hatched boxes indicate direct repeats at the deletion junctions. The frequency of these two deletions is also shown.

pVP254 (Fig. 3), which is similar to VP212, was constructed. The vector contains SNV LTRs and other *cis*-acting elements necessary for viral replication, a *neo* gene that can be expressed from the LTR promoter in mammalian cells and from a bacterial promoter in *E. coli* cells, and a pBR origin of replication for propagation in bacterial cells. The vector also contains a *lacZα* peptide gene fragment inserted within the U3 region of the LTR. The presence of the *lacZα* gene does not interfere with the promoter within U3. Because the *lacZα* gene is located within the LTR, the *lacZα* gene in the downstream U3 region is duplicated during each cycle of retrovirus replication and becomes a part of both LTRs in the provirus. Provirus DNA fragments containing only the downstream LTR can be recovered after *Bam*HI digestion of genomic DNA, and provirus DNA containing both LTRs can be recovered after *Eco*RI digestion of genomic DNA (as indicated in Fig. 3). Four copies of an 8-bp *Aat*II restriction site linker were inserted in-frame in the *lacZα* open reading frame resulting in a 36-bp insertion (four 8-bp linkers plus 4 bp from the fill-in reaction) (Fig. 2). The insertion potentially can form a 17-bp stem-loop structure (Fig. 4A).

Recovery and characterization of VP254 proviruses. Genomic DNAs from pools of target cells derived from three helper cell clones were digested with *Bam*HI, and 16,561 proviruses were recovered (Table 3). For a single cycle of retrovirus replication of VP254, the mutation rate (0.006) was threefold higher than the mutation rate of VP212 (24). DNA sequence analysis of the mutant proviruses indicated that 57 of 96 (59%) mutant proviruses contained a mutation associated with the putative stem-loop structure. The remaining 39 mutant proviruses contained mutations within the *lacZα* gene but not within the stem-loop structure. The rate of mutations not within the stem-loop structure (39 of 16,561) was identical to the rate of VP212 mutations (37 of 16,867), indicating that the location of the *lacZα* gene within the LTR, as opposed to its location between the LTRs in VP212, did not affect its mutation rate. (This identity also indicates the reproducibility of the system.) The threefold increase in the VP254 mutation rate was thus the result of the presence of the stem-loop structure. The rate of deletions was previously determined as 2×10^{-6}/bp per cycle (25). In contrast, the rate of deletions within the stem-loop structure was 9.6×10^{-5}/bp per cycle, nearly 50-fold higher than the average rate of deletions, indicating that potential nucleic acid secondary structures are deletion hotspots.

Analysis of deletions within the stem-loop. A deletion that initiated in the potential stem-loop structure constituted one-third of all mutations (Fig. 4B, hatched boxes). A total of 40 bp was deleted. A direct repeat of four nucleotides was present at the deletion junctions, indicating that it occurred by the previously discussed misalignment mechanism (24). A similar deletion that initiated in the potential stem-loop structure occurred only twice, having a direct repeat of five nucleotides and deleting 120 bp.

All deletions that occurred entirely within the stem-loop structure altered the number of copies of the 8-bp *Aat*II linkers and thus altered the *lacZα* reading frame. In a few proviruses, the 8-bp sequences were duplicated to form five or six *Aat*II linkers, also altering the *lacZα* reading frame.

Deletions involving the potential stem-loop primarily occurred during minus-strand DNA synthesis. We determined whether the deletions involving the potential stem-loop structure occurred primarily during minus-strand DNA synthesis. We recovered proviruses after digestion of the genomic DNA with *Eco*RI, an enzyme that does not cut within the provirus (Fig. 3). The resulting provirus plasmids contained both LTRs and two copies of the *lacZα* gene. If the mutations occurred during minus-strand DNA synthesis, they would be present in both copies of the *lacZα* genes, and these provirus plasmids would have a mutant colony phenotype. On the other hand, mutations occurring during the elongation step of plus-strand DNA synthesis would be

TABLE 3. Recovery of VP254 proviruses after *Bam*HI digestion

Mutation in downstream LTR and mutation type	No. of mutants/no. recovered[a]	No. of different mutations/no. recovered[b]	Mutation rate
Mutations in stem-loop	57/16,561		0.003
−1 linker		6	
−2 linkers		16	
+1 linker		2	
Other deletion		33	
Mutations not in stem-loop	39/16,561		0.002
Transition		10/18	
Transversion		3/8	
Deletion		6/6	
Deletion with insertion		3/4	
Frameshift		2/2	
Insertion		1/1	
Total	96/16,561		0.006

[a] The first number indicates the total number of proviruses with a mutation, and the second number is the total number of proviruses recovered.

[b] The first number indicates the number of different mutants obtained, while the second number indicates the total number of proviruses recovered with the type of mutation indicated. Mutations in the stem-loop were identical in each category.

present in only one copy of the *lacZα* gene, and the provirus plasmids would contain one mutant *lacZα* gene and one wild-type *lacZα* gene; these proviruses would result in a blue colony.

We recovered 5,260 proviruses after *Eco*RI digestion of the genomic DNA (Table 4). The mutation rate (0.008) was similar to the mutation rate (0.006) after *Bam*HI digestion. If a significant proportion of the mutations had occurred during the elongation step of plus-strand DNA synthesis, a lower mutation rate would be expected. The results indicated that a majority of the deletion mutations recovered involving the potential stem-loop structure occurred during minus-strand DNA synthesis.

Deletions without direct repeats, deletions with insertions, and insertional hypermutation. During analysis of AZC-induced mutations, we observed two deletions without short direct repeats or insertions at the deletion junctions (Fig. 5A). This class of deletions, named deletions without direct repeats, occurs less frequently than the deletion or deletion with insertion mutations. A deletion without direct repeats after a mistake in plus-strand strong-stop DNA synthesis was previously reported (27).

TABLE 4. Recovery of VP254 proviruses after *Eco*RI digestion

Mutation in both LTRs and mutation type	No. of mutants/no. recovered[a]	No. of different mutations/no. recovered[b]	Mutation rate
Mutations in stem-loop	36/5,260		0.007
−1 linker		1/4	
−2 linkers		1/10	
+1 linker		1/6	
+2 linkers		1/1	
Other deletion		2/15	
Mutations not in stem-loop	4/5,260		0.001
Transition		1/1	
Transversion		1/1	
Deletion		1/1	
Frameshift		1/1	
Total	40/5,260		0.008

[a] The first number indicates the total number of proviruses with a mutation, and the second number is the total number of proviruses recovered.

[b] The first number indicates the number of different mutants obtained, while the second number indicates the total number of proviruses recovered with the type of mutation indicated. Mutations in the stem-loop that changed the number of linkers were identical in each category except for other deletion.

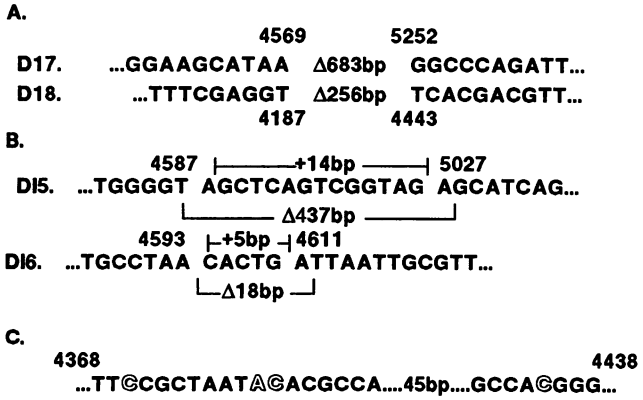

FIG. 5. (A) Deletions without direct repeats. Two independent deletion mutants without direct repeats or insertions at the deletion junctions are shown. Sequences of 683 and 256 bp were deleted. The numbers above and below the lines refer to the sequence in Fig. 2. (B) Deletions with insertions. Final products of deletions with insertions DI5 and DI6 are shown. The numbers in brackets above the line indicate the number of base pairs inserted, and the numbers in brackets below the line indicate the number of base pairs deleted. (C) Insertional hypermutation. The nucleotide sequence surrounding the insertional events is shown. The outlined nucleotides were inserted in a single provirus. The numbers above the line refer to the sequence in Fig. 2.

We also characterized two deletions with insertions which were similar to those previously described (Fig. 5B) (24). DI5 deleted 371 bp and inserted 14 bp at the deletion junction. DI6 deleted 18 bp and inserted 5 bp at the deletion junction. The origin of the inserted sequences is unknown. The 14 bp inserted in D15 are not in VP212.

In addition to these deletion events, a provirus with three independent insertions in a stretch of 68 bp was identified (Fig. 5C). At the first position, a C nucleotide was inserted adjacent to another C; at the second position, an AC dinucleotide was inserted adjacent to another AC dinucleotide; and at the third position, a C nucleotide was inserted between an A and a G. The first and second insertions could have resulted from a duplication of the adjacent nucleotides. We hypothesize that these insertional events resulted from the action of an error-prone reverse transcriptase that formed as a result of an error occurring during the transcription or translation of the *pol* gene. We and others previously described hypermutations that resulted in G to A substitutions (24, 33).

DISCUSSION

Our observation that AZC treatment during retrovirus replication increased the retrovirus mutation rate indicates that the high retrovirus spontaneous mutation rate can be further increased by treatment with mutagenic agents. This is in contrast to results with vesicular stomatitis virus and poliovirus when AZC and other mutagens were used (11) and suggests that some chemical treatment of retrovirus infection might increase variation in the retrovirus population.

TABLE 5. Mechanisms of spontaneous and AZC-induced substitutions

Substitution type and source	Nucleotide change	Sequence change at mutant site[a]	Potential dislocation mutagenesis	Predicted strand of AZC incorporation[b]
Transitions				
Spontaneous[c]	G to A	ATGT to ATAT[1]	No	NA
	G to A	GGGG to GGAG	No	NA
	G to A	ACGA to ACAA[2]	Yes	NA
	G to A	TTGT to TTAT[3]	No	NA
	C to T	CGCC to CGTC	No	NA
	C to T	TCCC to TCTC	No	NA
	C to T	CCCA to CCTA	No	NA
	C to T	CACG to CATG[4]	No	NA
	T to C	AGTC to AGCC	Yes	NA
	A to G	CCAG to CCGG	Yes	NA
AZC	G to A	ACGC to ACAC	No	Minus-strand DNA
	G to A	ATGT to ATAT[1]	No	Minus-strand DNA
	G to A	TTGT to TTAT[3]	No	Minus-strand DNA
Transversions				
Spontaneous[c]	A to C	CCAG to CCCG	Yes	NA
	C to A	GTCA to GTAA	Yes	NA
	G to C	ACGA to ACCA[2]	Yes	NA
	T to G	GTTG to GTGG	Yes	NA
AZC	C to G	GGCG to GGGG	Yes	RNA
	G to C	ACGA to ACCA[2]	Yes	Minus-strand DNA
	G to C	ACGT to ACCT	Yes	Minus-strand DNA
	G to C	ATGT to ATCT[1]	No	Minus-strand DNA
	C to G	CACG to CAGG[4]	Yes	RNA

[a] Superscripts 1 to 4 designate substitutions occurring at the same nucleotide positions. Boldface letters indicate mutation target and substituted bases. Underlined letters indicate CpG dinucleotides in which AZC-induced substitutions occurred.
[b] NA, not applicable.
[c] From reference 24.

From these data, it can be calculated that similar treatment of a population of a wild-type retrovirus with a 10-kb genome would result in an average of two mutations per genome after one round of replication. This prediction is consistent with the results of early experiments to generate temperature-sensitive mutations by using AZC treatment (20).

Six of eight AZC-induced substitutions occurred at G nucleotides, and the remaining substitutions occurred at a C nucleotide (Fig. 2 and Table 5). Substitutions at G nucleotides could have occurred after incorporation of AZC opposite the G nucleotides during minus-strand DNA synthesis; subsequently, dislocation mutagenesis or misincorporation could have occurred during plus-strand DNA synthesis. The two substitutions at C nucleotides could have occurred by incorporation of AZC into the viral genomic RNA during RNA transcription and then by dislocation mutagenesis or misincorporation during minus-strand DNA synthesis.

We determined whether the substitutions observed here or earlier (24) could have occurred by the dislocation mutagenesis model proposed by Kunkel (2, 16) by looking for a next nucleotide effect. In this model, a transient misalignment of the template strand and the nascent strand occurs, so that the nucleotide adjacent to the mutation site is copied. A second misalignment that mispairs the copied nucleotide with the nucleotide at the mutation site then occurs, and polymerization continues, resulting in a mismatched nucleotide in the DNA. This model requires that the substituted nucleotide is present in the template adjacent to the mutation site. We found that only 3 of 18 transitions that occurred spontaneously or after AZC treatment contained the substituted nucleotide adjacent to the template site (Fig. 2 and Table 5). In contrast, eight of nine transversions (four AZC induced and four spontaneous) contained the substituted nucleotide adjacent to the template site. These results suggest that transversions occur primarily by dislocation mutagenesis, while most transitions occur by misincorporation.

This hypothesis is in contrast to the results of studies based on in vitro experiments with reverse transcriptase from avian myeloblastosis virus and HIV-1 (2, 30). When HIV-1 reverse transcriptase mutation rates were measured in vitro, it was found that a majority of the mutational hotspots involved transitions, and six of seven transition hotspots could be explained by a dislocation mechanism. It is possible that SNV reverse transcriptase and HIV reverse transcriptase induce transitions by different mechanisms.

Alternatively, transitions may occur by different mechanisms in vivo and in vitro or both.

The observation that transversions are less likely than transitions to occur by misincorporation suggests that the polymerase active site discriminates less efficiently between two purines or two pyrimidines but can more easily discriminate between a purine and a pyrimidine.

Previous work by others showed that there are 114 sites in the *lacZα* gene and its promoter sequence that when mutated result in a detectable phenotypic change in colony color (2, 30). Only 13 of these sites were found spontaneously mutated in our experiments. If all target sites mutate at equal frequencies, then five of eight AZC-induced mutations are not expected to occur at the same sites as the spontaneous mutants. The observation that five of eight AZC-induced mutations occurred at 4 of these 13 sites of spontaneous mutation (Fig. 2) suggested that a special feature of these target sites led to an increased frequency of reverse transcription errors; that is, these sites are hotspots. AZC induced transversion substitutions at some hotspots where only spontaneous transitions were observed. The concordance of AZC-induced transversions and spontaneous transitions suggested that AZC increased the tendency of the reverse transcription growing point to make errors at these hotspots and altered the nature of the mutations induced. Analysis of the *lacZα* sequence for potential stem-loop structures with the UWGCG programs indicated no potential stem-loop structures at the hotspots.

The potential nucleic acid secondary structure that we tested was also a mutational hotspot. The potential secondary structure was composed of short direct repeats. Therefore, we do not know whether the sequence was a deletion hotspot because it was composed of direct repeats or because it could potentially form a stem-loop structure. To distinguish between these possibilities, it will be necessary to test a potential stem-loop structure that does not have direct repeats or to test short direct repeats that cannot form secondary structure. We have previously shown that a direct repeat of 110 nucleotides is a deletion hotspot (24). Since some of the deletions involved with the stem-loop involve an alteration of the number of direct repeats and others do not, it is likely that both the direct repeats and the potential stem-loop structure contribute to the 50-fold-increased mutation rate.

We also showed that the mutations occurred in the potential stem-loop structure primarily during minus-strand DNA synthesis. Experiments are now under way to determine whether or not spontaneous mutations occur primarily during minus-strand DNA synthesis.

ACKNOWLEDGMENTS

We thank N. Drinkwater, W.-S. Hu, J. Jones, L. Mansky, G. Pulsinelli, and S. Yang for useful comments on the manuscript and T. Jacoby, B. Pietz, and J. Schoening for technical assistance.

The research in our laboratory is supported by Public Health Service grants CA-22443 and CA-07175 from the National Cancer Institute. H.M.T. is an American Cancer Society Research Professor.

REFERENCES

1. **Amacher, D. E., and G. N. Turner.** 1987. The mutagenicity of 5-azacytidine and other inhibitors of replicative DNA synthesis in the L5178Y mouse lymphoma cell. Mutat. Res. **176:**123–131.
2. **Bebenek, K., J. Abbotts, J. D. Roberts, S. H. Wilson, and T. A. Kunkel.** 1989. Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. J. Biol. Chem. **264:**16948–16956.
3. **Bouchard, J., M. C. Walker, J.-M. Leclerc, N. Lapointe, R. Beaulieu, and L. Thibodeau.** 1990. 5-Azacytidine and 5-azadeoxycytidine inhibit human immunodeficiency virus type 1 replication in vitro. Antimicrob. Agents Chemother. **34:**206–209.
4. **Bouck, N., D. Kokkinakis, and J. Ostrowsky.** 1984. Induction of a step in carcinogenesis that is normally associated with mutagenesis by nonmutagenic concentrations of 5-azacytidine. Mol. Cell. Biol. **4:**1231–1237.
5. **Call, K., J. Jensen, H. Liber, and W. Thilly.** 1986. Studies of mutagenicity and clastogenicity of 5-azacytidine in human lymphoblasts and salmonella typhimurium. Mutat. Res. **160:**249–257.
6. **Carr, B. I., J. G. Reilly, S. S. Smith, C. Winberg, and A. Riggs.** 1984. The tumorigenicity of 5-azacytidine in the male Fischer rat. Carcinogenesis **5:**1538–1590.
7. **Dougherty, J. P., and H. M. Temin.** 1986. High mutation rate of a spleen necrosis virus-based retrovirus vector. Mol. Cell. Biol. **6:**4387–4395.
8. **Dower, W. J., J. F. Miller, and C. Ragsdale.** 1988. High efficiency transformation of E. coli by high voltage electroporation. Nucleic Acids Res. **16:**6127–6145.
9. **Fucik, V., S. Zadrazil, Z. Sormova, and F. Sorm.** 1965. Mutagenic effects of 5-azacytidine in bacteria. Collect. Czech. Chem. Commun. **30:**2883–2885.
10. **Halle, S.** 1968. 5-Azacytidine as a mutagen for arboviruses. J. Virol. **2:**1228–1229.
11. **Holland, J. J., E. Domingo, J. Carlos de la Torre, and D. A. Steinhauer.** 1990. Mutation frequencies at defined single codon sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. J. Virol. **64:** 3960–3962.
12. **Hu, W.-S., and H. M. Temin.** 1990. Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. Proc. Natl. Acad. Sci. USA **87:**1556–1560.
13. **Jorgensen, R. A., S. J. Rothstein, and W. J. Reznikoff.** 1979. A restriction enzyme cleavage map of Tn5 and location of a region encoding neomycin resistance. Mol. Gen. Genet. **117:**65–72.
14. **Karon, M., L. Sieger, S. Leimbrock, J. Z. Finkelstein, M. E. Nesbit, and J. J. Swaney.** 1963. 5-Azacytidine, a new agent for the treatment of acute leukemia. Blood **42:**359–365.
15. **Kawai, S., and M. Nishizawa.** 1984. New procedure for DNA transfection with polycation and dimethyl sulfoxide. Mol. Cell. Biol. **4:**1172–1174.
16. **Kunkel, T. A., and P. S. Alexander.** 1986. The base substitution fidelity of eucaryotic DNA polymerases—mispairing frequencies, site preferences, insertion preferences, and base substitution by dislocation. J. Biol. Chem. **261:**160–166.
17. **Landolph, J. R., and P. A. Jones.** 1982. Mutagenicity of 5-azacytidine and related nucleosides in C3H10T1/2 clone 8 and V79 cells. Cancer Res. **42:**817–832.
18. **Lavia, P., M. Ferraro, A. Micheli, and G. Olivieri.** 1985. Effect of 5-azacytidine (5-azaC) on the induction of chromatid aberrations (CA) and sister-chromatid exchanges (SCE). Mutat. Res. **149:**463–467.
19. **Li, L. H., E. J. Olin, H. H. Buskirk, and L. M. Reineke.** 1970. Cytotoxicity and mode of action of 5-azacytidine on L1210 leukemia. Cancer Res. **30:**2760–2769.
20. **Linial, M., and D. Blair.** 1982. Genetics of retroviruses, p. 649–783. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses: molecular biology of tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Lu, L. J. W., and K. Randerath.** 1979. Effects of 5-azacytidine on transfer RNA methyltransferase. Cancer Res. **39:**940–948.
22. **Mossing, M. C., and M. T. Record, Jr.** 1985. Thermodynamic origins of specificity in the lac repressor-operator interaction adaptability in the recognition to mutant operator sites. J. Mol. Biol. **186:**295–305.
23. **Olsson, L., and J. Forchhammer.** 1984. Induction of the metastatic phenotype in a mouse tumor model by 5-azacytidine. Proc. Natl. Acad. Sci. USA **81:**3389–3393.
24. **Pathak, V. K., and H. M. Temin.** 1990. Broad spectrum of *in vivo* forward mutations, hypermutations, and mutational hot-

spots in a retroviral shuttle vector after a single replication cycle: substitutions, frameshifts, and hypermutations. Proc. Natl. Acad. Sci. USA **87**:6019–6023.

25. **Pathak, V. K., and H. M. Temin.** 1990. Broad spectrum of *in vivo* forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: deletions and deletions with insertions. Proc. Natl. Acad. Sci. USA **87**:6024–6028.
26. **Pringle, C. R.** 1970. Genetic characteristics of conditional lethal mutants of vesicular stomatitis virus induced by 5-fluorouracil, 5-azacytidine, and ethyl methane sulfonate. J. Virol. **5**:559–567.
27. **Pulsinelli, G., and H. M. Temin.** 1991. Characterization of large deletions occurring during a single round of retrovirus replication: a novel deletion mechanism involving errors in strand transfer. J. Virol. **65**:4786–4797.
28. **Reichman, M., H. Karlan, and S. Panmon.** 1973. Destructive processing of the 45S ribosomal precursor in the presence of 5-azacytidine. Biochim. Biophys. Acta **299**:173–175.
29. **Riggs, A. D., and P. A. Jones.** 1983. 5-methylcytosine, gene regulation and cancer. Adv. Cancer Res. **40**:1–30.
30. **Roberts, J. D., B. D. Preston, L. A. Johnston, A. Soni, L. A. Loeb, and T. A. Kunkel.** 1989. Fidelity of two retroviral reverse transcriptases during DNA-dependent DNA synthesis in vitro. Mol. Cell. Biol. **9**:469–476.
31. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
32. **Taylor, S. M., and P. A. Jones.** 1979. Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine. Cell **17**:771–779.
33. **Vartanian, J.-P., A. Meyerhans, B. Åsjö, and S. Wain-Hobson.** 1991. Selection, recombination, and G-A hypermutation of human immunodeficiency virus type 1 genomes. J. Virol. **65**: 1779–1788.
34. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
35. **Zimmermann, F. K., and I. Scheel.** 1984. Genetic effects of 5-azacytidine in Saccharomyces cerevisiae. Mutat. Res. **139**:21–24.