Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 1197–1201, March 1987
Biochemistry

A promoterless retroviral vector indicates that there are sequences in U3 required for 3′ RNA processing

JOSEPH P. DOUGHERTY AND HOWARD M. TEMIN

McArdle Laboratory for Cancer Research, University of Wisconsin, Madison, WI 53706

Contributed by Howard M. Temin, November 5, 1986

ABSTRACT **We constructed a retrovirus vector in which all viral transcriptional signals are deleted during provirus formation. This vector would be safer and more useful in gene-transfer experiments. Construction of this vector involved the deletion of all of the U3 sequence except for 10 base pairs, required for integration, from the right-side long terminal repeat. We found that this deletion resulted in an inability to propagate virus efficiently. When we inserted sequences containing all of the signals for polyadenylylation of simian virus 40 late mRNA at the end of our vector, we were able to propagate virus efficiently. This result indicates that there are sequences in U3 upstream from AAUAAA that are essential for 3′ processing of viral RNA.**

There is considerable interest in using retrovirus vectors to introduce genes into the somatic cells of humans and into the germ line of animals. There are three inherent problems with using retrovirus vectors for gene therapy. The first is that there is potential for spread of vector virus in infected organisms and for spread to other organisms. The second is that there is the possibility of insertional activation of cellular oncogenes by viral transcriptional sequences in a target organism (1, 2). The third is that in cells in which stage-specific, tissue-specific, and/or inducible regulation is required, the viral transcriptional signals can interact with an inserted promoter that is being used to control a gene of choice (3–5).

Most retrovirus vectors are defective for retrovirus-encoded genes. They are packaged either by using helper virus or helper cells. Helper cells encode the viral gene products but the helper virus RNA is encapsidation-minus, so it is not packaged (6, 7). Therefore, vector virus stocks made from helper cells do not contain helper virus. However, there is still the possibility that endogenous retroviruses could supply helper functions and allow the spread of vector virus.

One way to overcome these problems is to design a retrovirus vector in which the viral transcriptional signals are lost during virus replication. With such a vector there could be no interaction between the viral transcriptional signals and those of an inserted promoter. Moreover, if an internal promoter were placed 3′ to the left-side R, the attachment site *att*R^+, the primer-binding site (PBS), and the encapsidation sequence (E) (see Fig. 1), spread of the vector even in the presence of endogenous helper virus would not be possible, since these sequences are essential for virus replication and would not be transcribed in the target cell.

We have constructed a spleen necrosis virus (SNV)-based retroviral vector in which the viral transcriptional signals are not present in the provirus in target cells. During the course of this work we found that there were sequences in U3 required for 3′ RNA processing. This result explains why 3′ RNA processing occurs at the 3′ end of the right-side R but not at the 3′ end of the left-side R.

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.

MATERIALS AND METHODS

Nomenclature. *hygro* refers to the hygromycin gene (8). Plasmids have a small p before their name (e.g., pJD214Hy), whereas virus made from those plasmids does not (e.g., JD214Hy).

Constructions. All constructions were made by standard techniques (9). pJD214Hy has already been described (10). pJD217Hy was made by deleting the *Sac* I–*Ava* I fragment (SNV map units 7.747–8.127) from the right-side long terminal repeat (LTR) and replacing it with an *Xho* I linker. These procedures resulted in a deletion of the entire U3 except for 10 base pairs (bp) at the 5′ end, which includes the left-side attachment site. It should be noted that in the left-side LTR, the 10 bp of U3 remaining in the right-side LTR were deleted. Thus, there is no homology between the two U3s of the left and right sides. pJD217SVHy was constructed by inserting the 565-bp *Nde* I–*Hin*dIII fragment from pSV2-neo (11) containing the simian virus 40 (SV40) early promoter into the *Xba* I site of pJD217Hy. pJD220Hy was made by inserting the 220-bp *Bam*HI–*Hin*dIII fragment containing the SV40 poly(A) site as described by Wickens and Stephenson (12) into the *Bam*HI site at the 3′ end of U5 in pJD217Hy. pJD220SVHy was constructed by inserting the 220-bp *Bam*HI–*Hin*dIII fragment into the *Bam*HI site of pJD217SVHy.

Cells. D17 dog cells and D17-C3 dog helper cells were grown as described (13, 14). The D17-C3 dog cell line is a D17 Rev-A helper cell line that supplies trans functions for packaging of defective Rev without production of replication-competent helper virus (6). Selection for hygromycin-resistant cells was done in the presence of hygromycin at 100 μg/ml (Eli Lilly).

Transfections and Virus Assay. Transfections were done by the calcium phosphate precipitation method (15, 16). Virus titers were obtained as described (10, 13).

S1 Nuclease Mapping. RNA isolation and S1 nuclease mapping were performed as described (17–19). Probes were 3′ end-labeled with [α-^{32}P]dCTP (Amersham) and Klenow fragment of DNA polymerase I according to published protocols (9). For each sample, 10 μg of total cellular RNA was used.

RESULTS

Rationale of the Experiment. Fig. 1 diagrams the experimental design. The basic idea is to delete the enhancer and promoter sequences, which are present in U3, from the right-side LTR of the plasmid DNA. This vector DNA is then used to transfect helper cells. Viral transcription begins at R on the left side. Virus is harvested from the transfected helper cells, and target cells are infected. Since the right-side U3 from the plasmid will supply the sequences for both U3 copies in the resulting provirus, the transcriptional signals

Abbreviations: SV40, simian virus 40; LTR, long terminal repeat; SNV, spleen necrosis virus; PBS, primer-binding site; E, encapsidation sequence; TU, transforming units.

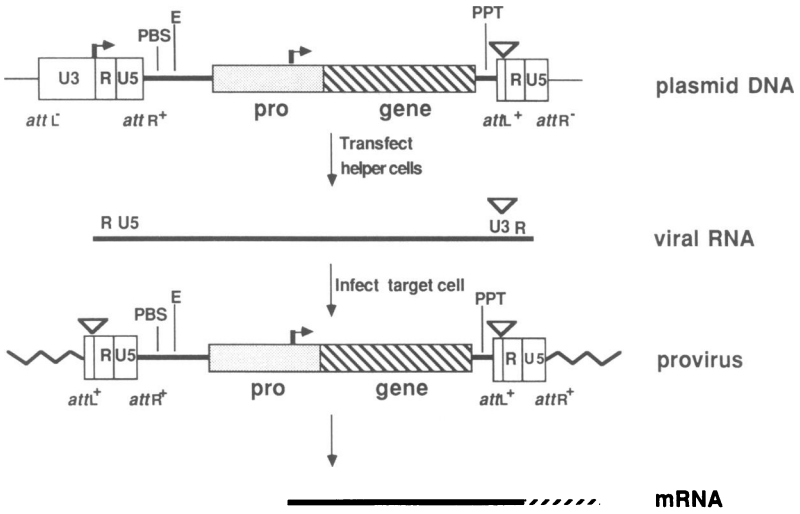

FIG. 1. Experimental design. Described are the steps involved in going from the plasmid DNA to the provirus in the target cell in which the inserted gene is transcribed from the internal promoter. Helper cells are transfected with the vector DNA. Viral RNA is transcribed in the helper cells; the viral RNA is packaged, and virions are released. The virus stocks are harvested and used to infect target cells. Proviruses are formed in the target cells, and the inserted gene is transcribed by the internal promoter. The boxes with the diagonal lines represent structural gene sequences. The stippled boxes represent inserted promoter sequences. Transcription initiation sites are indicated by arrows. The open boxes with U3, R, and U5 represent LTR sequences. The inverted triangles represent deleted U3 sequence. The thin lines represent pBR322 sequences. The jagged lines represent chromosomal sequences. pro, Promoter; PPT, polypurine tract; attR$^+$, the sequence that will form the right side of the attachment site; attL$^+$, the sequence that will form the left side of the attachment site; attL$^-$, the original provirus left-side attachment site that was deleted; and attR$^-$, the original right-side attachment site that was deleted.

originally deleted from the right side of the plasmid DNA will be deleted from both sides of the resulting provirus.

Characterization of Constructions. Fig. 2 diagrams the vectors we used for this study. In all cases we used the *hygro* gene as the selectable marker. In the experiment described in Fig. 2, we transfected the vector DNA into D17-C3 dog helper cells, harvested virus 5 days posttransfection, infected D17 cells, selected for hygromycin resistance, and obtained a virus titer in TU/ml. pJD214Hy served as our positive control vector in which the viral enhancer and promoter

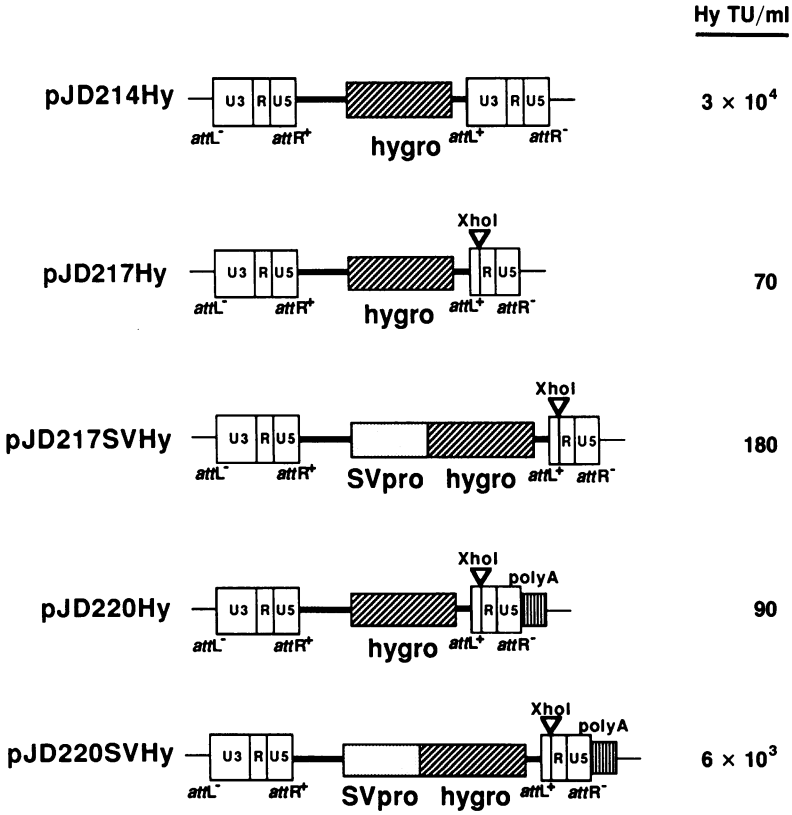

FIG. 2. Constructs and titers. The boxes with the diagonal lines represent *hygro* gene sequences. SVpro refers to the SV40 early gene promoter. The open boxes with U3, R, and U5 represent LTR sequences. The inverted triangles indicate U3 deletions. *Xho* I indicates that an *Xho* I linker replaced deleted U3 sequences. The boxes with vertical lines and polyA above them refer to the SV40 late mRNA poly(A) sequences. attR$^+$, attL$^+$, attR$^-$, and attL$^-$ are the same as in the legend to Fig. 1. Titers obtained with virus generated from the different vectors are indicated in transforming units (TU)/ml on the right.

sequences were not deleted. JD214Hy virus gave a titer of 3 × 10^4 TU/ml (Fig. 2).

pJD217Hy has the viral transcriptional sequences deleted from the right-side LTR, but it does not contain an internal promoter to drive expression of the *hygro* gene in the target cells. With JD217Hy virus we obtained a low titer (Fig. 2). The titer obtained with JD217Hy virus represents background expression since there is no internal promoter to drive the expression of *hygro* (see *Discussion*).

pJD217SVHy has the right-side LTR enhancer and promoter deleted and has the SV40 promoter inserted to transcribe the *hygro* gene in the D17 target cells (Fig. 2). With JD217SVHy virus we obtained a very low titer, 180 TU/ml, compared to JD214Hy, 3 × 10^4 (Fig. 2). We had expected a titer close to that of JD214Hy, but the titers obtained with JD217SVHy were the same as the background levels obtained with pJD217Hy (Fig. 2). This result led us to believe we were deleting U3 sequences important for propagation of the vector. It had been hypothesized that 3′ RNA processing occurs at the end of R on the right-side LTR but not on the left-side LTR because sequences in U3 are required for 3′ RNA processing (20, 21). Since U3 sequences from the right-side LTR, but not the left-side LTR, are present in the viral RNA (Fig. 1), 3′ RNA processing only occurs at the right side. We believed that with JD217SVHy, 3′ RNA processing was not occurring at the end of R so the viral RNA was either too big to be efficiently packaged (22) or was unstable, leading to degradation (23).

To test this hypothesis, we inserted a 220-bp fragment containing all of the sequences required for polyadenylylation of SV40 late mRNA at the end of our U3-deleted vectors (Fig. 2). pJD220SVHy, which uses the SV40 early promoter to drive *hygro* expression in infected target cells, gave titers close to that of pJD214Hy (Fig. 2). pJD220Hy gave a low background equivalent to that of pJD217Hy (Fig. 2).

S1 Nuclease Mapping Indicated That 3′ RNA Processing Does Not Occur at the End of R with Our U3-Deleted Vectors. We transfected D17 cells with the vectors described in Fig. 2, isolated total cellular RNA 48 hr posttransfection, and analyzed the RNA with the probes indicated in Fig. 3. pJD216NeoHy is a vector previously described (10) that contains the same 3′ end as pJD214Hy. For another study, we established a D17 cell clone containing a single provirus copy of JD216NeoHy (10). The RNA used in Fig. 3, lane 1, was purified from this cell line. If 3′ RNA processing occurred at the end of R for RNA derived from pJD216NeoHy and pJD214Hy, then a major 580-base protected fragment is expected. This was the case as can be seen in Fig. 3, lanes 1 and 2. With an intact U3, 3′ RNA processing occurred at the expected place. If 3′ RNA processing occurred at the end of R for RNA derived from U3-deleted vectors pJD217Hy and pJD217SVHy, then protection of a 140-base fragment would be expected using probe B (Fig. 3). Protection of a 140-base protected fragment was not observed (Fig. 3, lanes 3 and 4). S1 nuclease analysis of RNA derived from U3-deleted vectors pJD220Hy and pJD220SVHy, containing the inserted SV40 poly(A) signal, showed that 3′ RNA processing did not occur at the end of R but occurred in the SV40 poly(A)

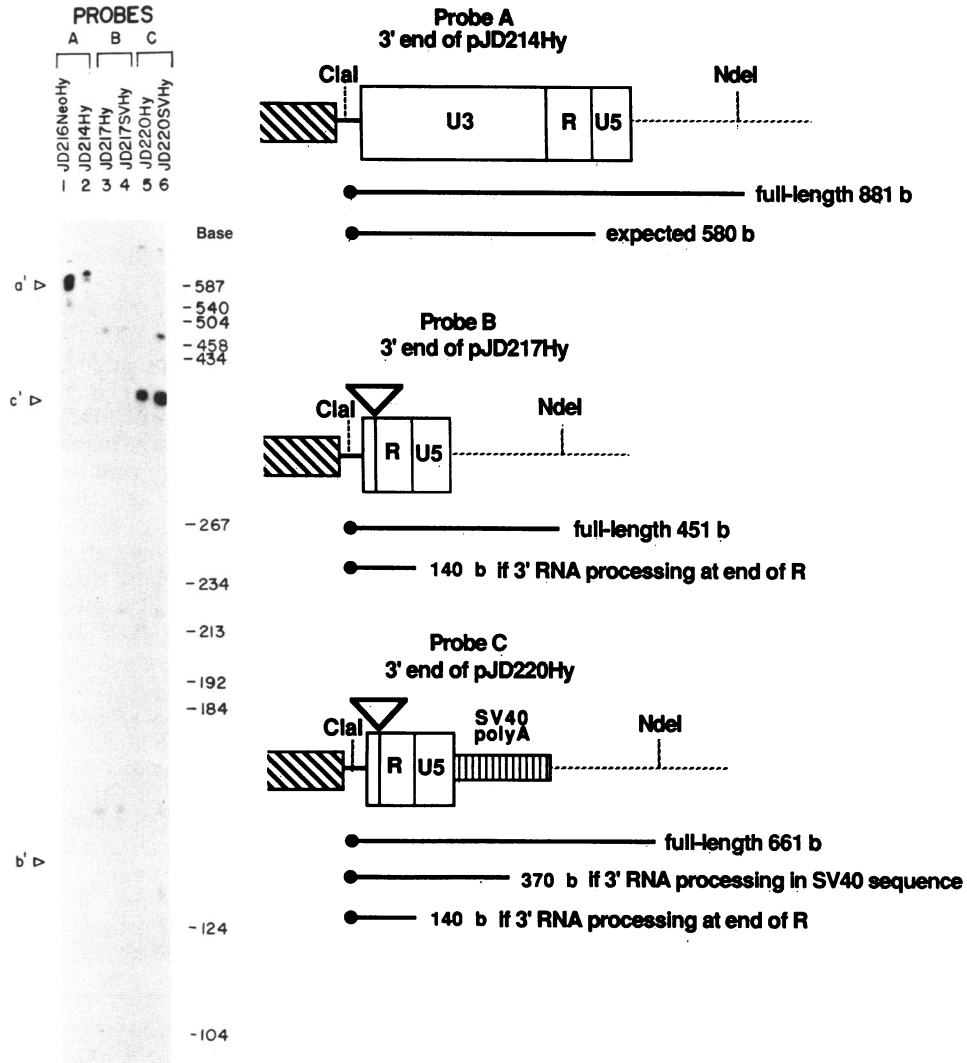

FIG. 3. S1 nuclease mapping. (*Right*) The probes used are designated A, B, and C. Probe A was prepared from the 3′ end of pJD214Hy. Probe B was prepared from the 3′ end of pJD217Hy. Probe C was prepared from the 3′ end of pJD220Hy. All probes were labeled at the *Cla* I site as indicated. They were then cut with *Nde* I as indicated. The large boxes represent the right-side LTR from each vector. The boxes with the diagonal lines represent the 3′ end of the *hygro* gene. The dashed horizontal lines represent pBR322 sequences. The inverted triangles indicate the U3 deletions. The box with the vertical lines and SV40 polyA above it represents the sequences required for polyadenylylation of SV40 late mRNA. The sizes of the full-length probes and the expected sizes of protected fragments are indicated in bases (b). (*Left*) The origin of the RNA used is indicated at the top of each lane. The probe used in each case is also indicated at the top of the gel. The sizes of the markers used are indicated in bases. a′, b′, and c′ refer to the positions where protected fragments of 580, 140, and 370 bases would be expected. Bowing of the gel occurred during electrophoresis, so the 580-base band in lane 2 appears to be running higher than expected.

sequence, as evidenced by the lack of a 140-base protected fragment and the presence of a 370-base protected fragment (Fig. 3, lanes 5 and 6).

The Proviral Structure of JD220SVHy in Infected D17 Cells Had the Expected Structure. As can be seen in Fig. 2, we inserted an *Xho* I linker in place of the deleted U3 sequence. If the hypothesis we described in Fig. 1 was correct, then an *Xho* I site should be present on both sides of the proviruses in JD220SVHy-infected D17 cells. Therefore, *Xho* I-digested genomic DNA from D17 cells infected with JD220SVHy should give a 2.2-kilobase-pair (kbp) fragment if hybridized to *hygro*-specific sequences. To test this hypothesis, we expanded individual D17 cell clones infected with JD220SVHy, isolated genomic DNA from the cell clones, digested the genomic DNA with *Xho* I, electrophoresed the DNA on a 1% agarose gel, blotted to nitrocellulose, and hybridized the blot with a *hygro*-specific probe. Fig. 4 shows the analysis of genomic DNA from 5 different JD220SVHy-infected D17 cell lines. Each had a 2.2-kbp fragment when probed with *hygro*-specific sequence (Fig. 4). We tested another 10 cell lines, and they gave the same results (data not shown).

DISCUSSION

We constructed a retroviral vector such that the viral transcriptional signals are not present in an infected target cell. This vector ensures that viral transcriptional signals cannot interact with an inserted promoter that is being utilized for stage-specific, tissue-specific, and/or inducible regulation of a gene of interest. Moreover, this vector ensures that there will be no spread of vector virus in an infected host even in the presence of endogenous helper virus. Essential cis-acting sequences (such as the left-side R, the PBS, the *att*R^+, and the E) will not be present in the mRNA transcribed from the provirus because there will be no viral transcriptional signals, and the internal promoter will be 3′ to essential cis-acting sequences (Fig. 1).

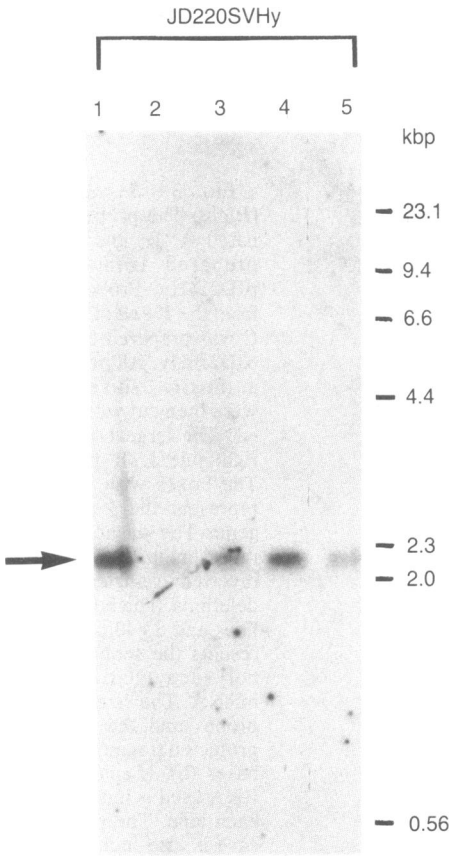

FIG. 4. Analysis of genomic DNA from JD220SVHy-infected D17 clones. Five individual D17 cell clones infected with JD220SVHy were expanded. Genomic DNA was isolated and digested with *Xho* I; this was followed by electrophoresis on a 1% agarose gel and blotting of the gel to nitrocellulose. The blot was then hybridized with a ^{32}P-labeled *hygro*-specific probe followed by autoradiography. Markers are indicated in kbp. The arrow indicates the bands seen. They are 2.2 kbp in size.

With JD220Hy and JD217Hy virus stocks, a small number of background colonies was seen. No colonies were expected, since there was no internal promoter present in these vectors to drive expression of the *hygro* gene. We have thought of two possible explanations for this result. One is that the proviruses integrated next to active cellular promoters that drove *hygro* gene expression. Another explanation is that recombination occurred between the helper cell sequences, which contain intact copies of the SNV LTR, and the transfected vector. If the first explanation were correct, then we would expect the *Xho* I site, inserted in place of the deleted U3 sequence, to be regenerated on both sides of the provirus, whereas if the second explanation were correct, then *Xho* I sites would not be found on both sides of the provirus. The second possibility seems more likely because in 9 of 10 proviruses from JD220Hy-infected cells, the *Xho* I site was not regenerated on both sides (data not shown).

While this work was in progress, Yu *et al.* (24) reported construction of a murine leukemia virus-based retroviral vector that would delete the viral enhancer and part of the viral promoter. Their construction contains overlapping homology between the left-side U3 and the partially deleted right-side U3, so that during growth of a helper cell line into which their deleted vector was transfected, homologous recombination could occur to reconstitute the right-side U3, so that the vector would no longer be self-inactivating. Our vector contains no overlapping homologies, so homologous recombination between vector molecules to regenerate a wild-type U3 on the right side is not possible. Furthermore, insertional activation of a protooncogene by the viral LTR will not occur with our vector.

Deletion of most of U3 appears to have resulted in a loss of correct 3′ end processing of viral RNA even though AAUAAA, a sequence that is known to be essential for polyadenylylation (23), was still present 17 bp from the end of R (Figs. 2 and 3). Because normal 3′ end processing was not occurring at the end of R, we were not able to obtain virus stocks with titers close to the titer of our control, JD214Hy. The lack of normal 3′ end processing might have resulted in lower titers either because 3′ end processing of viral RNA in the transfected helper cells occurred in cellular sequences far downstream from R, so the viral RNA was too large to be packaged efficiently (22), or because without normal 3′ end processing viral RNA was unstable and rapidly degraded (23). We feel that the second of these two possibilities is more likely, since we detected very little JD217Hy- or JD217SVHy-specific RNA during our S1 nuclease analysis (Fig. 3, lanes 3 and 4). When we added a poly(A) site to the vector at the end of U5, we were able to obtain higher titer virus stocks (Fig. 2). However, the titer obtained with JD220SVHy was still lower by a factor of 5 than that obtained with JD214Hy (Fig. 2). One explanation is that the SV40 promoter is weaker by a factor of 5 than the SNV promoter in D17 cells (25), and the difference in titer is a reflection of the difference in promoter strength. Another explanation is that in the target cells the SV40 poly(A) site is not present at the end of the provirus, so that normal 3′ end processing of mRNA transcribed from the provirus in the target cell is not occurring, resulting in mRNA instability.

The indication that sequences in U3 are important for normal 3′ RNA processing is important because it suggests that retroviruses contain sequences upstream of the

AAUAAA that are essential for normal 3′ RNA processing. It further suggests that 3′ RNA processing occurs at the end of the right-side R and not at the end of the left-side R because U3 sequences from the right-side LTR, but not the left-side LTR, are present in the viral RNA, so 3′ RNA processing can only occur at the right side.

We thank S. Coe, J. Couch, S. Hellenbrand, S. Hinz, and J. Rein for technical assistance; S. Peltz for help with the S1 nuclease mapping; M. Wickens for a gift of pSVL−141/+79; C. Gelinas, T. Gilmore, and B. Sylla for many useful discussions and suggestions during the course of this work; and R. Dornberg, J. Embretson, C. Gelinas, T. Gilmore, C. Miller, A. Panganiban, S. Peltz, R. Risser, and J. Ross for helpful comments on the manuscript. This research was supported by Public Health Service Research Grants CA-22443 and CA-07175. J.P.D. was supported by a fellowship from the American Cancer Society and by Public Health Research Service Award CA-09075 from the National Institutes of Health. H.M.T. is an American Cancer Society Research Professor.

1. Neel, B. J., Hayward, W. S., Robinson, H. L., Fang, J. & Astrin, S. M. (1981) *Cell* **23,** 323–334.
2. Payne, G. S., Courtneidge, S. A., Crittenden, L. B., Fadly, A. M., Bishop, J. M. & Varmus, H. E. (1981) *Cell* **23,** 311–322.
3. Moreau, P., Hen, R., Wasylyk, B., Everett, R. D., Gaub, M. P. & Chambon, P. (1981) *Nucleic Acids Res.* **9,** 6047–6068.
4. Banerji, J., Rusconi, S. & Schaffner, W. (1981) *Cell* **27,** 299–308.
5. Gruss, P., Dhar, R. & Khoury, G. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 943–947.
6. Watanabe, S. & Temin, H. M. (1983) *Mol. Cell. Biol.* **3,** 2241–2249.
7. Mann, R., Mulligan, R. C. & Baltimore, D. (1983) *Cell* **33,** 153–159.
8. Gritz, L. & Davies, J. (1983) *Gene* **25,** 79–188.
9. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
10. Dougherty, J. P. & Temin, H. M. (1986) *Mol. Cell. Biol.* **6,** 4387–4395.
11. Southern, P. J. & Berg, P. (1982) *J. Mol. Appl. Genet.* **1,** 327–341.
12. Wickens, M. & Stephenson, P. (1984) *Science* **226,** 1045–1051.
13. Emerman, M. & Temin, H. M. (1984) *Cell* **39,** 459–467.
14. Watanabe, S. & Temin, H. M. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 5986–5990.
15. Graham, F. L. & van der Eb, A. J. (1973) *Virology* **52,** 456–467.
16. Wigler, M., Sweet, R., Sim, G. K., Wold, B., Lacy, E., Maniatis, T., Silverstein, S. & Axel, R. (1979) *Cell* **16,** 777–785.
17. Ross, J. (1976) *J. Mol. Biol.* **106,** 403–420.
18. Berk, A. J. & Sharp, P. A. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 1274–1278.
19. Weaver, R. F. & Weissmann, C. (1979) *Nucleic Acids Res.* **7,** 1175–1193.
20. Benz, E. W., Wydro, R. M., Nadal-Ginard, B. & Dino, D. (1980) *Nature (London)* **288,** 665–669.
21. Temin, H. M. (1982) *Cell* **28,** 3–5.
22. Gelinas, C. & Temin, H. M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 9211–9215.
23. Nevins, J. R. (1983) *Annu. Rev. Biochem.* **52,** 441–466.
24. Yu, S.-F., von Ruden, T., Kantoff, P. W., Garber, C., Seiberg, M., Ruther, U., Anderson, W. F., Wagner, E. F. & Gilboa, E. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 3194–3198.
25. Embretson, J. & Temin, H. M. (1986) *J. Virol.* **60,** 662–668.