

RNA Degradation and Primer Selection by Moloney Murine Leukemia Virus Reverse Transcriptase Contribute to the Accuracy of Plus Strand Initiation*

(Received for publication, December 29, 1999, and in revised form, February 15, 2000)

Colleen D. Kelleher and James J. Champoux‡

From the Department of Microbiology, School of Medicine, University of Washington, Seattle, Washington 98195

During reverse transcription, plus strand DNA synthesis is initiated at a purine-rich RNA primer generated by the RNase H activity of reverse transcriptase (RT). Specific initiation of plus strand synthesis from this polypurine tract (PPT) RNA is essential for the subsequent integration of the linear viral DNA product. Based on current models, it is predicted that priming from sites upstream of the PPT may be tolerated by the virus, whereas efficient extension from RNA primers located downstream from the PPT is predicted to generate dead-end products. By using hybrid duplex substrates derived from the Moloney murine leukemia virus long terminal repeat, we investigated the extent to which RNase H degrades the viral RNA during time course cleavage assays, and we tested the capacity of the polymerase activity of RT to use the resulting cleavage products as primers. We find that the majority of the RNA fragments generated by RNase H are 2–25 nucleotides in length, and only following extensive degradation are most fragments reduced to 10 nucleotides or smaller. Although extensive RNA degradation by RNase H likely eliminates many potential RNA primers, based on thermostability predictions it appears that some RNA fragments remain stably annealed to the DNA template. RNA primers generated by RNase H within the long terminal repeat sequence are found to have the capacity to initiate DNA synthesis by RT; however, the priming efficiency is significantly less than that observed with the PPT primer. We find that Moloney murine leukemia virus nucleocapsid protein reduces RNase H degradation and slightly alters the cleavage specificity of RT; however, nucleocapsid protein does not appear to enhance PPT primer utilization or suppress extension from non-PPT RNA primers.

Retroviral replication involves a complex series of coordinated steps that are catalyzed by the viral reverse transcriptase (RT).¹ Reverse transcription of the viral RNA genome into double-stranded DNA is a necessary step during the viral

life cycle, generating a terminally redundant DNA product that subsequently becomes integrated into the host cell genome. During reverse transcription two template switches and displacement synthesis through a region of duplex DNA generate long terminal repeats (LTR) that both duplicate unique sequences (U5 and U3) lost during transcription of the proviral DNA and produce ends that can be recognized by the integration machinery of the virus. The final duplex DNA product has the structure U3-R-U5 . . . (coding sequence) . . . U3-R-U5 (reviewed in Refs. 1–3). Although RT alone appears to possess all the catalytic activities required to complete reverse transcription, several studies have provided evidence that the viral nucleocapsid protein (NC) may facilitate some steps (reviewed in Refs. 4 and 5).

The multifunctional reverse transcriptase carries a polymerase activity that utilizes both RNA and DNA strands as templates as well as an RNase H activity that cleaves RNA in hybrid duplex with DNA (reviewed in Ref. 3). RNase H is required to carry out the following three distinct steps during reverse transcription: generation of the plus strand RNA primer, removal of the plus and minus strand primers from the nascent DNA strands, and degradation of the RNA genome following minus strand DNA synthesis which is thought to make the DNA template accessible for plus strand synthesis (reviewed in Ref. 6).

The extent to which the RNase H activity of RT degrades RNA has been investigated using a number of model systems. During RNA-templated DNA synthesis some cleavage of the template RNA concomitant with nucleotide incorporation appears to occur, and although the frequency of cleavage seems to vary widely depending on the viral RT studied (7, 8), some longer RNA fragments probably remain available as substrates for subsequent synthesis-independent cleavage by RNase H (7–10). The extent to which synthesis-independent cleavage by RNase H contributes to RNA removal is less well characterized. *In vitro* studies using uniformly labeled RNA from the human immunodeficiency virus (HIV) *gag* region have shown that very extensive degradation by HIV RT results primarily in fragments ranging from ~6 to 14 nucleotides (nt) in length (11). Whether this fragment distribution is representative of RNase H cleavage activity on other regions of the viral sequence remains to be determined.

In contrast to the relatively nonspecific cleavage that removes genomic RNA from the nascent minus strand DNA, RNase H is also required to make a precise cleavage at the RNase H-resistant polypurine tract (PPT) sequence to generate the primer for plus strand synthesis. Accurate plus strand initiation from the PPT is required to generate an integration-competent final DNA product. The finding that some retroviruses possess a second copy of the PPT (cPPT) near the middle of the genome that efficiently primes synthesis, as well as evidence that plus strands are discontinuous in a number of

* This work was supported by Grant R37 CA51605 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Microbiology, Box 357242, University of Washington, Seattle, WA 98195-7242. Tel.: 206-543-8574; Fax: 206-543-8297; E-mail: champoux@u.washington.edu.

¹ The abbreviations used are: RT, reverse transcriptase; LTR, long terminal repeat; NC, nucleocapsid protein; HIV, human immunodeficiency virus; nt, nucleotide(s); bp, base pairs; PPT, polypurine tract; cPPT, central polypurine tract; Mo-MLV, Moloney murine leukemia virus; NCdd, zinc finger deleted mutant of Mo-MLV NC; DTT, dithiothreitol; ddTTP, dideoxy-TTP; ddCTP, dideoxy-CTP.

retroviral systems, indicates that priming from regions upstream of the PPT in addition to PPT priming may not be detrimental to the virus (12–19). By contrast, current models predict that plus strands initiated from primer RNAs downstream of the PPT will generate dead-end products (6). Since RNA fragments within the LTR are closer to the end of the minus strand template than the PPT, plus strands initiated from LTR primers are predicted to have a temporal advantage over PPT-primed products in completing plus strand synthesis and the second strand transfer. Thus, priming from within the LTR is predicted to result in the generation of products with one incorrect end, lacking the requisite integration signal normally present at the terminus of the LTR. These considerations raise the possibility that sequences with the potential to generate extendable primers following RNase H cleavage have been selectively excluded from the LTR region.

In this study, we sought to determine whether the RNase H activity of Moloney murine leukemia virus (Mo-MLV) RT degrades the LTR RNA to a sufficient extent that none of the non-PPT RNA fragments would be predicted to remain annealed. Additionally we investigated the effects of Mo-MLV NC on RT-catalyzed RNase H cleavage. Since we found that some longer RNA fragments persisted following extensive cleavage, we tested the capacity of the polymerase activity of RT to use these RNA fragments as primers. Our results indicate that some non-PPT RNA fragments generated by RNase H from within the LTR region or from elsewhere in the genome have the capacity to prime synthesis by the polymerase, although the efficiency of extension is significantly less than from the PPT primer.

EXPERIMENTAL PROCEDURES

Materials—Mo-MLV RT and calf intestinal alkaline phosphatase were purchased from Amersham Pharmacia Biotech. SuperScript II (RNase H⁻ RT) was from Life Technologies, Inc. Expression and purification of the RNase H domain of Mo-MLV RT have been described elsewhere (20). All other enzymes were obtained from New England Biolabs. Denaturing polyacrylamide gels (8.3 M urea) were prepared with Sequagel reagents from National Diagnostics. Synthesis and characterization of native Mo-MLV NC and the NCdd mutant (zinc finger replaced by a Gly-Gly linker) have been described (21, 22) and were generously provided by J. L. Darlix.

Nucleic Acids—RNA oligonucleotide “PPT Oligo” (5'-CCAGAAAA-GGGGG-3') was obtained from Integrated DNA Technologies, Inc. Salhind (5'-GAATCAGTCGACAAGCTTGTGCAC-3') and Salterm (5'-TCACTGGTCGACCGGGTTAACC-3') were used for amplification during M13INT construction. SP6S1 (5'-AGCTATTTAGGTGACACTA-TAGAATACGCATG-3') and SP6S2 (5'-CGTATTCTATAGTGTCACT-AAAT-3') were used to insert an SP6 promoter into M13INT. Amplification primers T7M13 and TermM13, used to generate M13LTR1, have been described (22). Construction of plasmids pGEMLTR2 and M13LTR2 has been described previously (22); pGEMLTR1 was similarly generated but contains a 738-bp insert from the Mo-MLV LTR (genome position 7758–8332 joined to 69–231 (23)) that, in addition to downstream sequences, includes the PPT and 68 nt of upstream sequence. M13LTR1 was constructed by cloning a polymerase chain reaction-amplified 782-bp region of pGEMLTR1, including the LTR insert, into M13mp7 DNA. M13INT was constructed by inserting an SP6 promoter oligonucleotide linker followed by a polymerase chain reaction-amplified region of the Mo-MLV integrase gene (genome position 5138–5824 (23)) into M13mp7 DNA so that the RNA transcripts are the same polarity as the viral RNA. Single-stranded phage and phagemid DNAs were isolated using established procedures (24). Single-stranded DNA inserts LTRi (807 nt), LTRΔPPTi (709 nt), and N-LTRi (738 nt) were excised and recovered from M13LTR1, M13LTR2, and M13INT, respectively, as described (22) except that the restriction enzyme digestion was with *Bam*HI.

Preparation of RNA—*In vitro* transcription was carried out using the RiboMAX kit (Promega) following the supplier's protocol. LTR and LTRΔPPT RNAs were produced using T7 RNA polymerase to transcribe *Bam*HI-linearized pGEMLTR1 and pGEMLTR2, respectively. *Eco*RI-linearized M13INT DNA was transcribed by SP6 RNA polymerase to produce N-LTR RNA. Full-length RNA transcripts of 753 nt for LTR,

655 nt for LTRΔPPT, and 705 nt for N-LTR were purified by electrophoresis in a 4% sequencing gel. Following UV shadowing, the RNA was extracted from gel slices by electroelution into 40 mM Tris acetate, pH 8.5, 2 mM EDTA. The RNA was ethanol-precipitated in 0.3 M sodium acetate, resuspended in 10 mM Tris-HCl, pH 8.0, 1 mM EDTA (TE), and stored at –80 °C. In some cases the 5'-triphosphate was removed by incubation with calf intestinal alkaline phosphatase in 200 mM Tris-HCl, pH 8.0, 10 mM ZnCl₂, 10 mM MgCl₂ for 30 min at 30 °C, followed by phenol/chloroform extraction and ethanol precipitation. The dephosphorylated RNA was 5' end-labeled in 20-μl reactions containing 0.2 μM RNA, 70 mM Tris-HCl, pH 7.6, 10 mM MgCl₂, 5 mM dithiothreitol (DTT), 0.3 μM [γ-³²P]ATP and 10 units of T4 polynucleotide kinase. Incubation was for 30 min at 37 °C; the reactions were stopped by the addition of 10 mM EDTA and incubated at 65 °C for 10 min to inactivate the kinase.

The nuclease P1 ladder was prepared by digesting 0.17 μg (77 nM) of end-labeled RNA with 0.01 ng of nuclease P1 in 20 mM sodium acetate, pH 5.2, for 30 min at 60 °C in a 10-μl reaction. The reaction was terminated by the addition of 9 volumes of Form-EDTA buffer (98% formamide, 10 mM EDTA), and aliquoted samples were stored at –20 °C. Under the conditions used, cleavage by nuclease P1 is enhanced at ribonucleotide A residues allowing the sequence identity of most bands to be clearly established. Separate RNA ladders generated using RNase ONE, RNase A, and RNase T1 were used to clarify the few remaining ambiguities.

Preparation of Hybrid Duplexes—Long RNA-DNA hybrid duplexes were annealed by incubating the indicated nucleic acids at 67 °C for 45 min in 200 mM KCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA. The PPT Oligo was annealed to single-stranded pGEMLTR1 or LTRi DNA under similar conditions except the reaction was heated to 90 °C and then slow-cooled to 25 °C.

Cleavage Assays—RNA-DNA hybrids were generated by annealing LTR RNA with complementary single-stranded pGEMLTR1 DNA or N-LTR RNA with complementary single-stranded M13INT DNA as described above. The cleavage reactions (30 μl) contained 1× reaction buffer (50 mM KCl, 50 mM Tris-HCl, pH 8.3, 6 mM MgCl₂, 5 mM DTT) plus 18 nM RNA annealed to 36 (pGEMLTR1) or 74 nM (M13INT) complementary DNA. Because the pGEMLTR1 and M13INT vector sizes differed by ~2-fold, the indicated DNA concentrations were used to equalize the amount of total DNA in the reactions. For assays using end-labeled RNA, cleavage was initiated by the addition of 2.5 (low enzyme) or 28 pmol (high enzyme) of Mo-MLV RT. The reactions were incubated at 37 °C, and aliquots of the reaction were stopped in 3 volumes of Form-EDTA buffer at the indicated times. For parallel control reactions, 0.08 units of *Escherichia coli* RNase H or an equivalent volume of RT dilution buffer (20% glycerol, 20 mM Tris-HCl, pH 8.3, 1 mg/ml bovine serum albumin, 2 mM DTT) was added in place of RT.

For RNA fragment analysis by exchange labeling, reactions were carried out as described above except the RNA was unlabeled during the cleavage reaction, and the reactions were stopped at the indicated time points by removing 8-μl aliquots into 2 μl of Tris/EDTA stop solution (255 mM Tris-HCl, pH 6.0, 24 mM EDTA) and heating for 10 min at 65 °C. The RNA fragments were exchange-labeled in 13-μl reactions containing 70 mM Tris-HCl, pH 7.6, 10 mM MgCl₂, 5 mM DTT, 400 μM ADP, 0.15 μM [γ-³²P]ATP, and 10 units of T4 polynucleotide kinase. The reactions were incubated for 30 min at 37 °C and terminated by the addition of 2.3 volumes of Form-EDTA buffer. For cleavage reactions carried out in the presence of NC protein, DNA templates LTRi and N-LTRi were used in place of the full-length phage DNA to avoid sequestration of NC by excess single-stranded DNA. Reactions containing 5 nM RNA annealed to 7.5 nM DNA in 1× reaction buffer plus 10 μM ZnCl₂ were preincubated with 30 μM NC protein or NCdd (3:1 NC:nt ratio) for 1 min at 37 °C. Cleavage was initiated by the addition of 0.7 pmol of RT, or for control reactions 0.014 units of *E. coli* RNase H or an equivalent volume of RT dilution buffer, and incubated at 37 °C for the indicated times. Reactions were terminated and exchange-labeled (where indicated) as described above. In all cases, samples were heated at 90 °C for 3 min prior to resolving the cleavage products on 20% sequencing gels.

Extension Assays—Substrates for the dideoxy-terminated extension assays were prepared by annealing unlabeled LTR or LTRΔPPT RNAs to the complementary M13-derived insert DNA templates, LTRi and LTRΔPPTi, respectively. For the oligonucleotide hybrid, PPT Oligo was annealed to LTRi DNA. RNA cleavage was initiated by the addition of 11.3 pmol of RT to 30-μl reactions containing hybrid substrates at 7.5 nM in 1× reaction buffer, and the reactions were incubated for 5 min at 37 °C. In control reactions, an equivalent number of units (156 units) of RNase H⁻ RT was added in place of RT. At 5 min, 12-μl aliquots of the cleavage reaction were added directly to prewarmed extension tubes

containing either 4 μl of ddTTP mix (1 \times reaction buffer, 200 μM each dATP, dGTP, and ddTTP and 0.3 μM [$\alpha\text{-}^{32}\text{P}$]dCTP) or 4 μl of ddCTP mix (1 \times reaction buffer, 200 μM each dATP, dGTP, and ddCTP and 0.3 μM [$\alpha\text{-}^{32}\text{P}$]dTTP), and incubation continued at 37 $^{\circ}\text{C}$ for 30 min. Following extension, stalled products were chased by the addition of 2 μl of 1.8 mM dCTP (ddTTP reactions) or 2 μl of 1.8 mM dTTP (ddCTP reactions). Incubation was continued at 37 $^{\circ}\text{C}$ for another 20 min, and the reactions were terminated by adding 2 μl of 100 mM EDTA. The samples were divided into two portions, and either an equal volume of H_2O ($-\text{NaOH}$) or an equal volume of 0.6 M sodium hydroxide ($+\text{NaOH}$) (to hydrolyze the RNA) was added, and the samples were heated at 60 $^{\circ}\text{C}$ for 30 min under a layer of mineral oil. Both the $-$ and $+$ NaOH samples were diluted in 3 volumes of Form-EDTA buffer and heated for 3 min at 90 $^{\circ}\text{C}$ prior to electrophoresis on a 20% sequencing gel. Dideoxy-termination assays that included NC were carried out as described above, except 10 μM ZnCl_2 was included in the 1 \times reaction buffer, and the hybrid substrates were incubated with 428 pmol of NC or an equivalent volume of NC dilution buffer (15 mM Tris-HCl, pH 6.0, 10 μM ZnCl_2 , 1 mM DTT) for 1 min at 37 $^{\circ}\text{C}$ prior to the initiation of cleavage with RT.

Run-off extension assays were carried out in 20- μl reactions. For the long hybrid substrates (LTR, LTR Δ PPT, and N-LTR), 10 nM template DNA (M13 derived inserts) and 20 nM RNA were present in the final reaction. For the oligonucleotide hybrid substrate, 10 nM LTRi DNA template and 7.5 nM PPT Oligo were used, and for RNA and template self-priming controls, the nucleic acid concentrations were 10 nM. The cleavage extension reactions contained the nucleic acid substrates in 1 \times reaction buffer plus 200 μM each dATP, dGTP, dCTP, 50 μM dTTP and 0.25 μM [$\alpha\text{-}^{32}\text{P}$]dTTP. Following the addition of 10.3 pmol of RT or for the control, an equal number of RNase H $^{-}$ RT units (200 units), the reactions were incubated at 37 $^{\circ}\text{C}$ for 30 min. Stalled products were chased by adding cold dTTP to 800 μM (6600:1 unlabeled:labeled nucleotide) and 400 units of RNase H $^{-}$ RT, and the incubation was continued at 37 $^{\circ}\text{C}$ for 30 min. Reactions were terminated by the addition of EDTA to 12 mM and heated for 10 min at 65 $^{\circ}\text{C}$ to inactivate the enzymes. The samples were divided, and from half of each reaction the RNA was hydrolyzed with NaOH as described for the dideoxy-terminated reactions. The samples were diluted in 3 volumes of Form-EDTA buffer, heated at 90 $^{\circ}\text{C}$ for 3 min, and separated on 6% sequencing gels. To quantify primer utilization, 5 discrete bands that did not co-migrate with RNA self-priming products were selected for each substrate. Using ImageQuant software, rectangles were drawn around the bands, and the volume was calculated with the volume quantitation feature. Because the background varied significantly along each lane, background volumes were determined for each band independently by quantifying the volume of an identical rectangle positioned in a clear area of the lane near the band of interest. Because the 3' ends of the extension products correspond to the end of the linear template, the approximate 5' end of each band could be determined from its size, based on a labeled 100-bp ladder and a sequencing reaction run in adjacent lanes. The pixel volumes for the individual bands were normalized based on the predicted number of labeled dT residues in the fragment.

Thermostability Calculations—Melting temperatures (T_m) were calculated using the nearest-neighbor model, incorporating thermodynamic parameters for RNA-DNA hybrid duplexes (Ref. 25 and references therein). Thermostability plots were generated by calculating the T_m values for all possible RNA oligonucleotides that could be generated during cleavage of substrate sequences, using a fixed oligonucleotide length for each independent plot. An Excel macro program was written to iteratively repeat the calculation after shifting the frame of reference down the sequence by 1 base. The calculated T_m values were plotted as a function of the position of the 3' end of the hypothetical RNA oligonucleotide along the linear sequence. T_m predictions were calculated for hybrid substrates LTR, N-LTR (IN), and ENV corresponding to the Mo-MLV genome positions 7758–8332 joined to 69–231, 5138–5824, and 6285–6985, respectively (23). Calculations were based on the concentrations of the components of the cleavage reactions (10 nM hybrid, 50 mM monovalent cation and 6 mM Mg^{2+}).

RESULTS

Analysis of RNA Fragment Thermostabilities—Although sequence recognition has been demonstrated to play a role in facilitating the selection of the correct RNA fragment for plus strand initiation by RT (26–30), removal of all potentially competing RNA fragments by the RNase H activity of RT, particularly from the LTR region, could provide a mechanism to help ensure accurate initiation. To begin to investigate

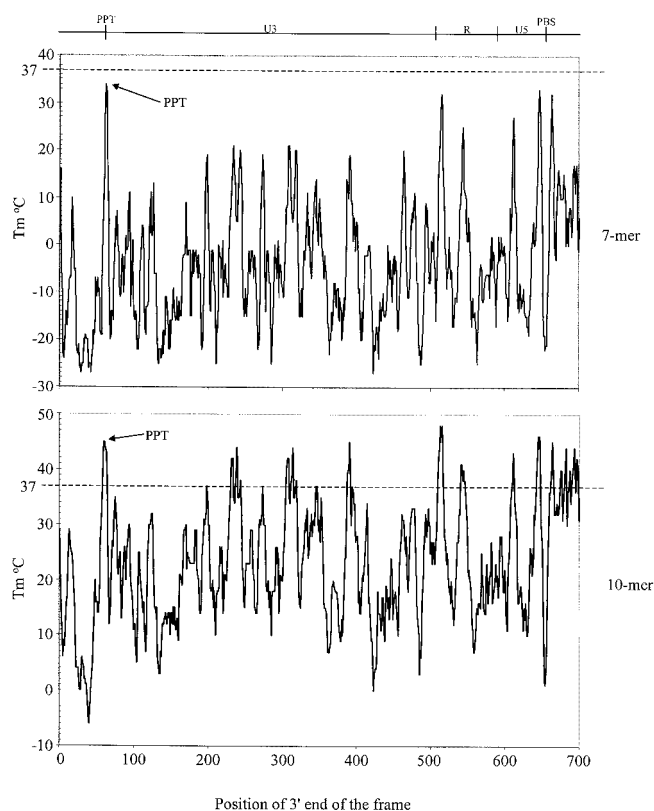


FIG. 1. Hybrid thermostability predictions for short RNA oligonucleotides within the viral LTR sequence. The predicted hybrid melting temperatures for all possible RNA fragments of 7 bases (upper plot) or 10 bases (lower plot) were determined. Melting temperatures are shown plotted as a function of the position of the 3' end of RNA fragments with the indicated frame length that could be generated in the LTR and surrounding sequence. Horizontal dotted lines indicate the position of a T_m of 37 $^{\circ}\text{C}$. Arrows show the position of the PPT sequence. Line drawing (top) shows features of the LTR and surrounding sequence for reference (PPT, U3, R, U5, and primer-binding site (PBS)).

whether genomic RNA could be effectively removed from the minus strand DNA through RT-catalyzed RNase H cleavage, we determined the predicted T_m for RNA fragments derived from viral sequences using thermodynamic parameters based on the nearest-neighbor model for thermostability prediction of RNA/DNA hybrid duplexes (25). Initially, the relationship between thermostability and the size of the RNA fragments was assessed for potential products within the LTR and surrounding sequences. For each of a series of fixed frame lengths, the T_m of the sequence within the frame was calculated and then the frame was repositioned down the sequence by one nucleotide, and the T_m calculation was repeated. Iterative application of this procedure yielded the T_m values shown in Fig. 1 for frames lengths of 7 and 10 nt, plotted as a function of the position of the 3' end of the frame. Based on this analysis we found that whereas none of the possible 7-mers have T_m values above 37 $^{\circ}\text{C}$, 12% of the 10-mers have T_m values at or above 37 $^{\circ}\text{C}$ and would thus be predicted to remain stably annealed. Interestingly, we note that the fragment corresponding to the PPT primer has the highest T_m among regions either immediately upstream of the PPT site or downstream within U3.

To assess whether the thermostabilities for RNA fragments that could be generated in the LTR region are representative of viral sequences elsewhere in the genome and to gain an understanding of the proportion of the fragments at various lengths that may remain annealed to the minus strand DNA, similar thermostability calculations were carried out using two addi-

TABLE I

Percentage of RNA fragments predicted to remain annealed at 37 °C

Thermostability calculations were based on the nearest neighbor model using fixed frame lengths (see "Experimental Procedures"). The values in the table indicate the percentage of RNA oligonucleotides of indicated length with a calculated $T_m \geq 37$ °C.

Length ^a	LTR ^b	IN ^c	ENV ^d
<i>nt</i>			
7	0	0	0
8	1	1	0
9	4	4	2
10	12	11	6
11	26	27	22
12	41	45	42
13	56	68	64
14	71	81	82
15	82	88	89
16	92	94	95
17	95	96	98
18	97	98	99

^a RNA oligonucleotide frame length.

^b Mo-MLV LTR and surrounding sequences as defined in text (750 nt).

^c Segment of Mo-MLV IN gene corresponding to N-LTR as defined in text (705 nt).

^d Segment of Mo-MLV envelope gene (700 nt, positions 6285–6985).

tional, randomly chosen, non-LTR regions of viral sequence. When the three regions were compared, little difference in the proportion of the RNA fragments, from 7 to 18 nt in length, that would remain annealed at 37 °C was observed (Table I). In all cases, it appears that degradation of the genomic RNA by RT to fragments 7 nt and smaller would be required for effective removal of the RNA from the minus strand DNA template.

Analysis of RNA Fragments Generated by RT Using 5' End-labeled RNA—As a first step toward analyzing the RNA fragments generated by the RNase H activity of Mo-MLV RT, we followed the kinetics of appearance of RNA cleavage products during time course assays, *in vitro*. RT-catalyzed RNase H cleavage was carried out using a 705-nt hybrid substrate derived from an arbitrarily selected region of the viral genome. Hybrid duplexes were generated by annealing 5'-³²P-labeled RNA of plus strand sequence from the viral integrase gene (N-LTR) to single-stranded phage DNA containing a complementary minus strand insert. Although the use of 5' end-labeled RNA limited the analysis to the 5'-terminal cleavage products generated by RT, the experimental design allowed the identity of the RNA fragments to be determined precisely. Cleavage by RNase H was assayed at high RT concentrations (50:1 RT to hybrid ratio) that approximate the ratio estimated to be found in the virion (3) and allowed analysis of the products that are generated following very extensive degradation, and at lower RT concentrations (5:1 RT to hybrid ratio) to characterize early cleavage products. Under both conditions, the predominant 5' end-labeled cleavage product generated early by RNase H was 16 nt in length, although other products ranging from 6 to 26 nt in length were also observed (Fig. 2A, lane 5; Fig. 2B, lanes 3–5). Upon continued incubation with RT, particularly at the high enzyme concentration, the initial cleavage products were further degraded to produce 7–15-nt-long fragments (Fig. 2B, lane 7). These results are consistent with earlier reports in which preferred cleavages by HIV RT were shown to occur at a distance of 15–20 nt from an RNA 5' end (5' end-directed cleavage) (32–34). Interestingly, significant cleavage at sites more distal from the labeled 5' end only constituted a significant portion of the total product at very early time points and only when high enzyme concentrations were used (Fig. 2B, lane 4 and data not shown). By contrast, cleavage by *E. coli* RNase H was not 5' end-directed, as evidenced by the relatively uniform distribution of products generated during

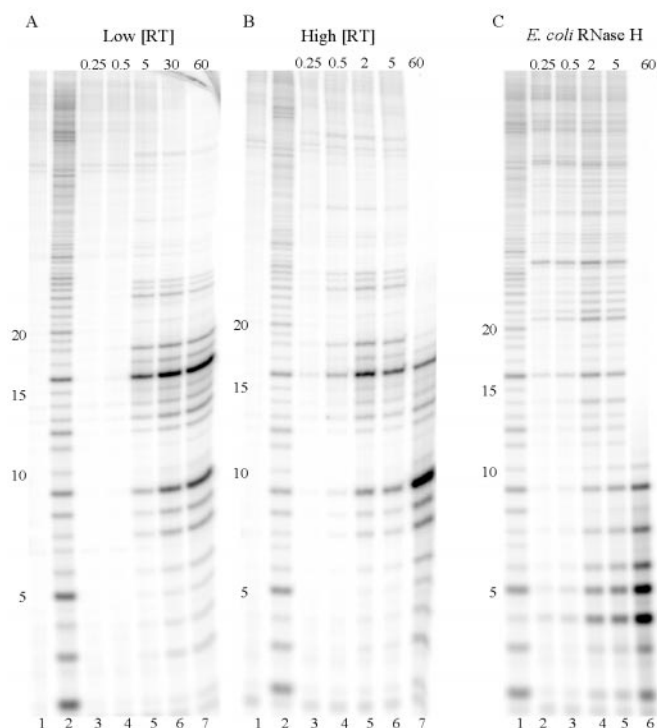
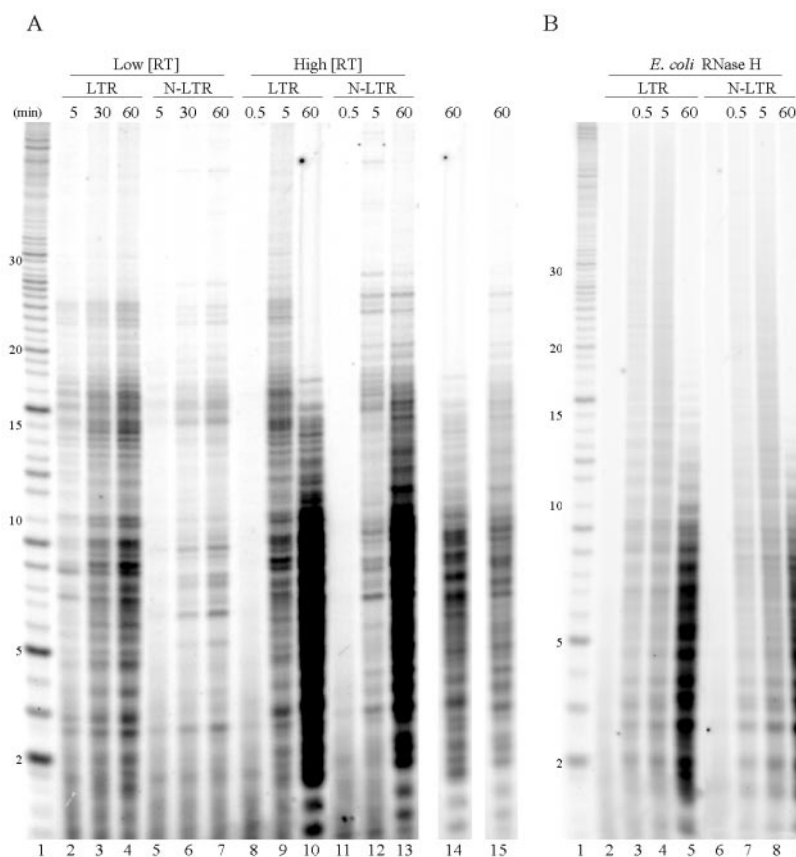


FIG. 2. RNase H cleavage of 5' end-labeled N-LTR hybrid substrate. 5' end-labeled N-LTR RNA was annealed to complementary single-stranded M13INT DNA and incubated with RT at a 5:1 RT to hybrid molar ratio (A), a 50:1 RT to hybrid ratio (B), or with *E. coli* RNase H (C). Aliquots of the RNase H cleavage reactions were stopped at the time points indicated above each lane (min), and the products were separated on 20% sequencing gels. A and B, lane 1, control reactions using RT dilution buffer in place of RT; lane 2, P1 ladder generated using 5' end-labeled N-LTR RNA (lengths indicated in bases to the left, extra band between 6 and 7 nt bands is background and is not seen during cleavage with RNase H); lanes 3–7, time course cleavage of N-LTR hybrid duplex. For C, lane 1, nuclease P1 ladder; lanes 2–6, time course cleavage of N-LTR hybrid substrate.

the initial 5 min of the cleavage reaction (Fig. 2C, lanes 2–5).

Analysis of RT-generated RNA Fragments by Exchange Labeling—The fast rate of 5' end-directed cleavage by RT (30) has generally made cleavage of the remainder of long hybrid substrates difficult to study using end-labeled RNA. To investigate more fully the kinetics and the size distribution of RNA fragments generated by RT along the length of hybrid duplexes, while avoiding the complication of disproportionate representation inherent with the use of internally labeled RNA, RT-catalyzed RNase H products were labeled using the exchange reaction of polynucleotide kinase. Thus, each RNA fragment generated should carry equivalent ³²P label at the 5' end. To assess whether cleavage of the LTR sequence differs from elsewhere on the genome, cleavage was compared using hybrid substrates derived from LTR and non-LTR (N-LTR) regions of the viral sequence. As in the previous assay, product formation was examined at both low and high RT to substrate ratios during time course cleavage reactions. By using the low enzyme concentration, RNase H products from both hybrids ranged from ~2 to 25 nt in length (Fig. 3A, lanes 2–7). Under these conditions, the size distribution of the cleavage products did not change during the course of the reaction, but rather, the total amount of product generated increased with longer incubation times (Fig. 3A, compare lane 2 with 4 and lane 5 with 7). By the latest time point at the high enzyme concentration (enzyme:hybrid ratio of 50:1), 92% of the labeled RNA fragments for the LTR hybrid and 87% for the non-LTR hybrid were 10 nt or smaller (Fig. 3A, lanes 10 and 13–15); the re-

FIG. 3. RNase H cleavage of LTR and N-LTR hybrid substrates, analysis by exchange labeling of the cleavage products. LTR hybrid was prepared by annealing unlabeled LTR RNA to complementary single-stranded pGEMLTR1 DNA, and N-LTR hybrid was prepared by annealing unlabeled N-LTR RNA to complementary single-stranded M13INT DNA. Hybrid duplexes were incubated with RT for the period indicated *above* each lane (min). Following termination of the reaction, the RNA fragments were 32 P-labeled using the exchange reaction and separated on 20% sequencing gels. **A**, RT-catalyzed RNase H cleavage at RT: hybrid ratios of 5:1 (*lanes 2–7*) or 50:1 (*lanes 8–15*). *Lane 1*, N-LTR P1 ladder generated using 5' end-labeled N-LTR RNA (lengths indicated in bases to the left). Lighter exposures of *lanes 10* and *13* are shown in *lanes 14* and *15*, respectively. **B**, *E. coli* RNase H-catalyzed cleavage. *Lane 1*, N-LTR nuclease P1 ladder; *lanes 2* and *6*, control reactions using RT dilution buffer in place of RNase H; *lanes 3–5*, cleavage products using LTR substrate; *lanes 7–9*, cleavage products using N-LTR substrate.



remainder of the RNA consisted primarily of fragments ranging from 11 to 25 nt in length. Only following very extensive degradation (60 min at the high RT concentration) was the median fragment length reduced to 6 nt. Fig. 3A, *lanes 14* and *15*, shows the high enzyme, 60-min cleavage products at lower exposure for comparison. The disappearance of some products migrating at ≥ 15 nt with long incubation at the high RT concentration confirms that longer fragments (15–25 nt) generated at the early time points remained stably annealed and available for subsequent cleavage by RT (Table I; Fig. 3A, compare *lane 9* with *14* and *lane 12* with *15*).

When cleavage products generated with the two different substrates were compared, a pattern emerged in which most products fell roughly into three general size classes as follows: class I consisting of fragments 24–26 nt in length, class II consisting of 14–18-mers, and class III consisting of 7–10-mers (Fig. 3A). All three classes were also seen with the 5' end-labeled substrate (Fig. 2). By contrast, cleavage catalyzed by *E. coli* RNase H (Fig. 3B) or by the RNase H domain alone of Mo-MLV RT (20) (not shown) resulted in products with a uniform size distribution until the latest time points (see "Discussion").

The Effects of NC on RNA Degradation by RNase H—Since NC protein has nucleic acid annealing as well as helix destabilizing properties and has been proposed to interact directly with RT (35–38), the effect of Mo-MLV NC on the rate and sequence specificity of RNase H cleavage by RT was investigated. Assays similar to those described above were carried out, except the LTR and non-LTR hybrid substrates were incubated with NC at a 3:1 NC:nt ratio prior to the initiation of cleavage with RT. The NC concentration used was based on the ratio that we had previously found to promote strand annealing and to facilitate displacement synthesis (22); some protein in our preparation is likely inactive since others have reported the promotion of strand annealing at lower protein to nucleic acid

ratios. By using 5' end-labeled RNA (not shown) or unlabeled RNA that was subsequently labeled by the exchange method (Fig. 4), the presence of NC resulted in a reproducible reduction in the extent of cleavage by RT. In the presence of NC, some RNA persisted as fragments up to ~ 27 nt in length following 60 min of incubation with RT (Fig. 4, *lanes 9* and *17*), whereas very little RNA longer than ~ 18 nt remained in parallel reactions in the absence of NC (Fig. 4, *lanes 5* and *13*). We observed that the differences in cleavage were only apparent at later time points; the extent of the cleavage and the product distribution were nearly identical in the presence or absence of NC at 5 min (Fig. 4, compare *lane 3* with *7* and *lane 11* with *15*). Thus it appears that the effect of NC may be to decrease the further degradation of primary cleavage products rather than to decrease the rate of RNase H cleavage, as has been previously suggested (39). Notably, a zinc finger-deleted mutant of NC (21) did not affect cleavage by RT, and cleavage by *E. coli* RNase H was unaffected by either form of NC (data not shown). A control reaction in which NC was added following cleavage but prior to exchange labeling demonstrated that the reduction in the amount of product generated by RT in the presence of NC was not attributable to an inhibitory effect of NC on the exchange labeling reaction (data not shown).

Priming Capacity of RNA Fragments Produced by RT: Dideoxy-terminated Extensions—Since some RNA fragments derived from RNase H cleavage of the viral LTR region were found to be long enough to remain stably annealed to the DNA template, we investigated whether these fragments have the capacity to prime synthesis by RT. For purposes of comparison, RNA-DNA hybrids either included the PPT sequence or were constructed from the region just downstream of the PPT (Fig. 5A, LTR and LTR Δ PPT). In this assay, unlabeled hybrid duplexes were incubated with RT for 5 min to initiate RNase H cleavage, following which three of the four dNTPs plus one dideoxynucleotide were added to allow limited primer exten-

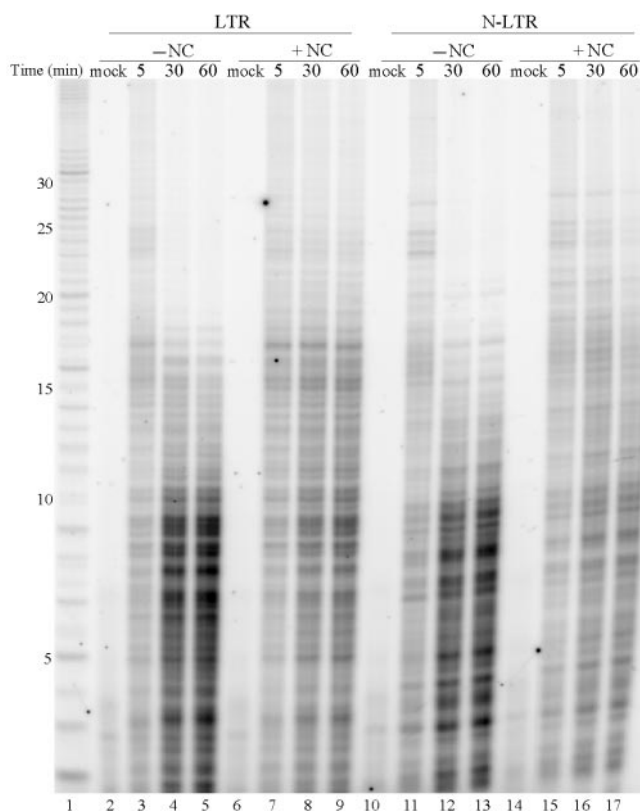


FIG. 4. The effect of NC protein on RNase H cleavage. LTR or N-LTR hybrid duplexes were generated by annealing unlabeled LTR or N-LTR RNA with LTRi or N-LTRi single-stranded DNA inserts, respectively. Hybrid substrates were preincubated with NC (lanes 6–9 and 14–17) or NC dilution buffer (lanes 2–5 and 10–13) for 1 min prior to the initiation of cleavage with RT, or addition of RT dilution buffer in place of RT for control reactions (lanes 2, 6, 10, and 14). Following cleavage for the period indicated above each lane, the reactions were terminated, and the RNA fragments were ^{32}P -labeled using the exchange reaction. Lane 1, N-LTR nuclease P1 ladder generated using 5' end-labeled N-LTR RNA (lengths indicated in bases to the left). The products were separated in a 20% sequencing gel.

sion by the polymerase activity of RT. To ensure that all possible priming events were visualized, two complementary reactions were carried out as follows: in the first, [α - ^{32}P]dTTP and dideoxy-CTP along with dATP and dGTP were included in the nucleotide mix; the complementary reaction contained [α - ^{32}P]dCTP and dideoxy-TTP along with dATP and dGTP. Thus, if dideoxy termination from a given primer occurred before incorporation of the labeled nucleotide in one reaction, incorporation of the label during extension from the same primer would be ensured in the complementary reaction. Moreover, extension products from any one primer would only be visualized in one of the two reactions but never in both. Since RNA primers with the same 3' ends need not have the same 5' ends, some additional bands were expected among the extension products due solely to heterogeneity of the RNA 5' ends. To eliminate this variable, a fraction of each extension sample was treated with NaOH to remove the RNA primer. Following the alkaline hydrolysis treatment, the discrete number of RNA primers (RNA 3' ends supporting extensions by RT) could be estimated by summing the bands present in the two complementary reactions.

To determine accurately where the extension products from the PPT primer would migrate, control reactions were carried out using a 15-mer RNA oligonucleotide corresponding to the plus strand primer RNA (Fig. 5A, PPT Oligo). Extension from the PPT Oligo by an RNase H⁻ mutant of RT in the ddCTP

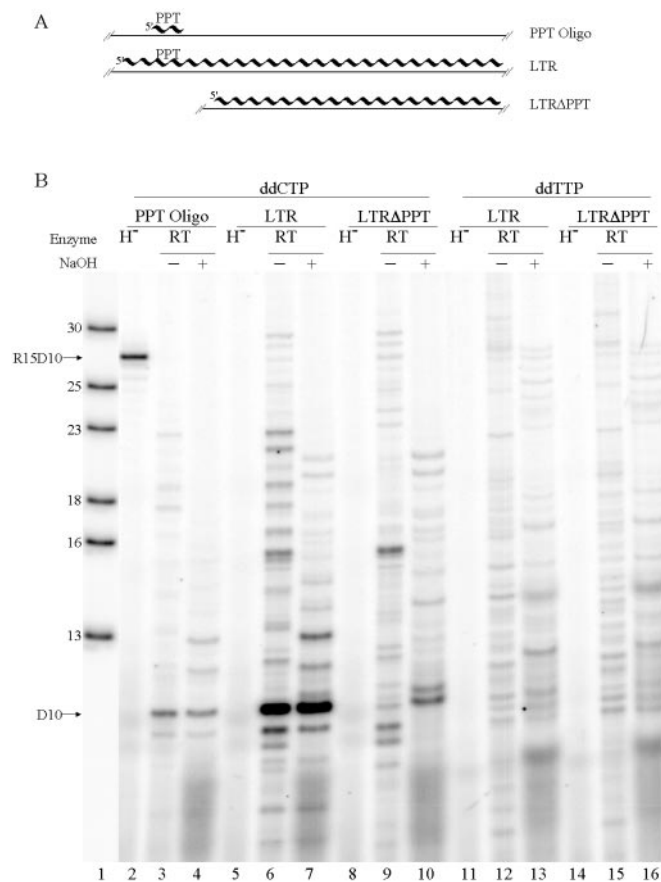


FIG. 5. Dideoxy-terminated extensions from RNase H-generated RNA primers on the LTR and LTRAPPT hybrid substrates. A, shown schematically are the different hybrid substrates used in the dideoxy termination assay. Straight lines indicate DNA strands, and wavy lines indicate RNA strands. The template strand for PPT Oligo and LTR hybrids was single-stranded LTRi DNA and for the LTRAPPT hybrid the template was single-stranded LTRAPPTi DNA. The 15-mer PPT Oligo anneals to the complement of the PPT sequence. The PPT cleavage site on the LTR substrate is located 68 nt downstream from the LTR RNA 5' end. B, shown are the products of RT-catalyzed DNA synthesis on the substrates shown in A: lanes 2–4, PPT Oligo; lanes 5–7 and 11–13, LTR; lanes 8–10 and 14–16, LTRAPPT. Following a 5-min incubation with RT (lanes 3, 4, 6, 7, 9, 10, 12, 13, 15, and 16) to allow cleavage or RNase H⁻ RT as a control (lanes 2, 5, 8, 11, and 14), either ddCTP-labeling mix (see “Experimental Procedures”), lanes 2–10, or ddTTP-labeling mix, lanes 11–16, was added. Following a 30-min extension reaction, stalled products were chased by the addition of excess cold dTTP (lanes 2–10) or dCTP (lanes 11–16). RNA was removed by alkaline hydrolysis in lanes marked +. The samples were run on a 20% sequencing gel, and only the portion of the gel containing visible bands is shown. Deoxynucleotide size markers are shown in lane 1. Arrow R15D10 shows the position of ddCTP-terminated 10-mer DNA extension product from the PPT in which the RNA primer remains attached to the DNA 5' end. Arrow D10 shows the position of ddCTP terminated 10-mer DNA extension products from the PPT primer in which the RNA has been removed by RNase H.

reaction resulted in a product consisting of the 15-mer RNA primer plus a ^{32}P -labeled 10-nt-long DNA extension (Fig. 5B, lane 2, R15D10). Extension catalyzed by wild-type RT in an otherwise identical reaction resulted predominantly in a product corresponding to the 10-nt-long labeled DNA lacking the RNA primer (Fig. 5B, lane 3, D10). No change in the position of this band was detected when the sample was treated with NaOH (Fig. 5B, lane 4), confirming that the RNA primer had been removed at the RNA-DNA junction by RNase H (26, 40–42).

Cleavage and extension by RT using the LTR substrate in ddCTP-terminated reactions generated products that co-migrated with the D10 band from the oligonucleotide control

experiment (Fig. 5B, compare lanes 3 and 4 with 6 and 7). In addition, following removal of the RNA primers by alkaline hydrolysis, 6–8 major bands not present in the oligonucleotide control reactions were observed (Fig. 5B, lane 7), presumably due to extensions from non-PPT RNA fragments. Supporting this conclusion is the finding that products migrating at many of the same positions were generated when parallel reactions were carried out using the LTR Δ PPT substrate (Fig. 5B, compare lane 7 to 10). Consistent with the fact that PPT-primed extensions would not be expected to be labeled in the ddTTP-terminated reactions (because termination precedes the labeled nucleotide incorporation), very little difference was found between the products generated with the LTR and LTR Δ PPT substrates (Fig. 5B compare lanes 12 and 13 with 15 and 16). Based on the number of discrete bands identified following alkaline hydrolysis of the products for both the ddCTP and ddTTP-terminated reactions, we estimate that within the LTR and surrounding sequence, approximately 15–20 RNA fragments capable of priming synthesis by RT are generated by the RNase H activity. Control reactions for each assay in which the RNase H⁻ mutant of RT was used in place of the wild-type enzyme demonstrated that the extension products observed were due solely to priming from RNA fragments generated by RT-associated RNase H (Fig. 5B lanes 5, 8, 11, and 14).

The resolution of the RT extension products by the high percentage sequencing gels revealed that most of the bands shifted following NaOH treatment, indicating that the products retained one or more 5' ribonucleotides. However, the major band (corresponding to the D10 extension product from the PPT) was unaffected by NaOH (Fig. 5B, lanes 6 and 7), consistent with complete removal of this RNA primer by RNase H. Thus, although RNase H appeared to remove efficiently the PPT RNA primer from its extension product, the enzyme failed to cleave the RNA from the DNA extensions for most of the non-PPT primers (Fig. 5B, compare lane 6 with 7, lane 9 with 10, lane 12 with 13, and lane 15 with 16).

A comparison of the intensity of the bands when resolved on sequencing gels consistently suggested that priming from the PPT (Fig. 5B, lanes 6 and 7, D10) was much more efficient than from all non-PPT RNA primers generated by RT within the LTR. Since the 10-nt-long DNA extension product from the PPT contains only one dT, the amount of PPT product may actually be an underestimation with respect to extensions from other primers that could incorporate multiple labeled dT residues. We considered the possibility that efficient removal of the PPT RNA by RNase H (see above) may allow "recycling" or reutilization of the PPT RNA primer by RT. By excising and counting the radioactivity in bands corresponding to the 10-nt DNA extension product from time course extension assays using either the PPT Oligo or LTR substrates (data not shown), it was determined that limited recycling likely does occur under these experimental conditions. Since similar recycling was not observed in experiments designed to generate longer extension products from the PPT (see below), recycling appears to be dependent on the small size of the DNA extension product generated by incorporation of the ddCTP; the T_m of the 10 nt PPT extension product is 14 °C and thus is not predicted to remain annealed. Similar recycling was observed by Gotte *et al.* (40) in dideoxynucleotide-terminated extensions from the PPT in the HIV system and thus appears to be a property unique to this particular experimental design.

Since extension from RNA fragments located downstream of the PPT by RT would be predicted to be detrimental to the virus, we investigated whether NC protein could suppress the utilization of non-PPT RNA fragments as primers. Although a slight (~1.7-fold) enhancement in overall priming was ob-

served in the presence of NC, quantitation of the amount of PPT extension product showed an insignificant increase relative to non-PPT-directed products (data not shown).

Priming Capacity of RNA Fragments Produced by RT: Run-off Extensions—To extend the analysis of the capacity of RT to utilize RNase H-generated products within the LTR region as primers and to estimate quantitatively the efficiency of non-PPT RNA primer extension by RT relative to extension from the PPT, we carried out a series of cleavage/extension assays in which the end of the linear DNA template was used to create a defined end point to extension. As before, the substrates used for the experiment consisted of two hybrid duplexes from the viral LTR that either contained (LTR) or lacked (LTR Δ PPT) the PPT sequence (see Fig. 5A). A third substrate of similar length was derived from a non-LTR region of the genome as described above (N-LTR). The hybrids were incubated with RT and dNTPs, including one α -³²P-labeled nucleotide, to allow RNase H-mediated cleavage of the RNA and generation of a labeled DNA extension product in a single step experiment. Following the reaction, the samples were divided, and from half the RNA portion of the product was removed by alkaline hydrolysis. In a parallel assay, a PPT RNA oligonucleotide primer was used in place of the full-length RNA to determine where PPT-primed extension products would migrate on a sequencing gel. Incubation of the PPT Oligo hybrid with RNase H⁻ RT resulted in a discrete extension product which, when treated with NaOH to remove the 15-nt RNA primer, generated a slightly faster migrating, 697-nt-long DNA product (Fig. 6A, lanes 1 and 2). The small shift in mobility was confirmed in other experiments (not shown). As before, the PPT Oligo hybrid reactions containing wild-type RT resulted almost exclusively in the formation of the 697-nt product, either with or without NaOH treatment, due to the efficient removal of the PPT RNA primer by the RNase H activity of RT (Fig. 6A, lanes 3 and 4).

When full-length LTR hybrid substrates containing the PPT sequence were incubated with RT, RNase H cleavage and polymerase extension resulted in numerous discrete labeled products. The predominant product co-migrated precisely with the 697-nt control product (Fig. 6A, compare lane 3 with 5) and, like the PPT Oligo control product, was unaffected by NaOH (Fig. 6A, compare lane 5 with 6). This product is very likely the result of extension from the PPT plus strand primer generated by RNase H cleavage. In otherwise identical reactions carried out with the LTR Δ PPT or N-LTR substrates, an array of discrete extension products (Fig. 6, B and C, lanes 1) was formed. Most of these products as well as most of the minor LTR products underwent a shift in mobility following NaOH treatment (compare Fig. 6, A–C, compare lanes + with lanes -). Parallel control reactions designed to test whether any of the observed products were due to RNA self-priming (Fig. 6A, lanes 7 and 8, Fig. 6, B and C, lanes 3 and 4), DNA template self-priming, or extension of contaminating non-RNase H-generated RNA fragments (data not shown) indicated that only RNA self-priming generated a few minor species that co-migrated with products from the cleavage/extension reactions.

Quantitative analysis of the samples run on sequencing gels was carried out to determine the relative efficiencies of PPT to non-PPT RNA primer utilization by RT. Analysis of five selected bands from the LTR and N-LTR substrates that did not co-migrate with RNA self-priming products suggested that no single RNA fragment was extended as efficiently as the PPT primer (Fig. 6D and Table II). The same bands were reproducibly observed in each independent experiment (6 for the LTR substrate and 3 for the N-LTR substrate). Although there was some variability between experiments, it appeared that approximately 5-fold less product was generated from the most

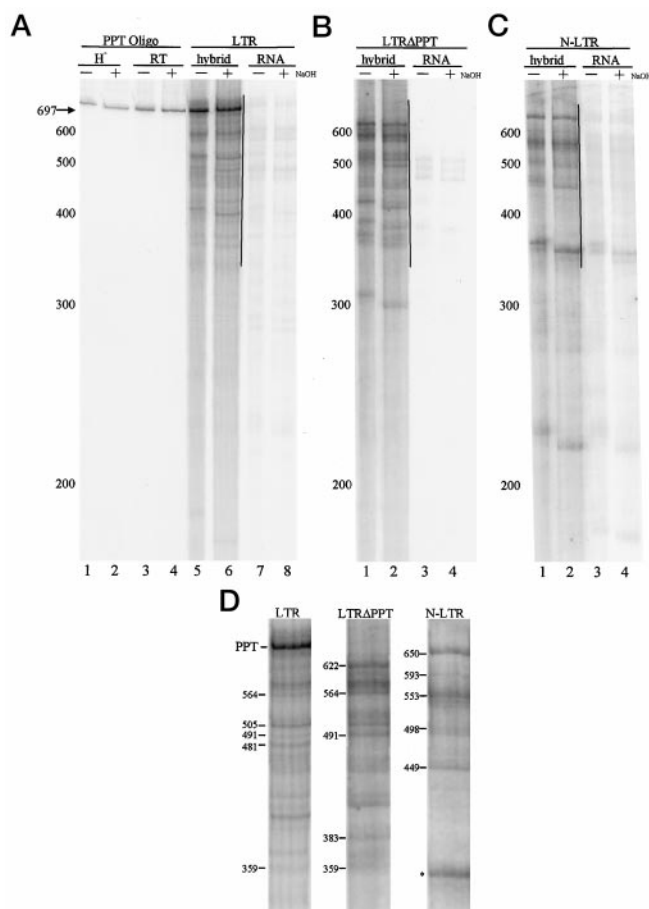


FIG. 6. Run-off DNA extension products primed by RNase H-generated RNA primers. RNA cleavage and extension were carried out in a single-step reaction by adding RT and nucleotides, including [α - 32 P]dTTP, to preformed RNA-DNA hybrid substrates. **A**, for lanes 1–4, the hybrid substrate was the PPT Oligo annealed to LTRi. For lanes 5 and 6, the substrate was LTR RNA annealed to LTRi DNA. Lanes 7 and 8 show a control for self-priming where the LTR RNA was “annealed” in the absence of DNA. Reactions were carried out with RT (lanes 3–8) or RNase H⁻ RT (lanes 1 and 2), and the RNA was removed by treatment with NaOH from an aliquot of each sample (lanes 2, 4, 6, and 8). **B**, for lanes 1 and 2, the substrate was LTRΔPPT RNA annealed to LTRΔPPTi DNA. Lanes 3 and 4 are RNA self-priming controls where LTRΔPPT RNA is annealed in the absence of DNA. Reactions were carried out with RT, and the RNA was removed by alkaline hydrolysis from an aliquot of each sample (lanes 2 and 4). **C**, substrates are as described for **B** except the DNA is N-LTRi and the RNA is N-LTR RNA. Reactions were as described above for **B**. The positions of the DNA size markers are indicated to the left of **A–C**. The vertical lines mark the portions shown enlarged in **D**. **D**, enlarged regions of lane 6 from **A**, lane 2 from **B**, and lane 2 from **C**. PPT indicates the band corresponding to the 697-nt DNA product initiated at the PPT RNA primer. Approximate DNA lengths are indicated to the left for the bands chosen for quantitative analysis. The band marked with an asterisk co-migrates with an RNA self-priming product.

efficiently used RNA primers compared with priming from the PPT; the majority generated substantially less product. Consistently, less non-PPT primer extension from the LTR substrate was observed as compared with the LTRΔPPT and N-LTR substrates (see “Discussion”).

DISCUSSION

Because successful generation of a provirus depends on accurate plus strand initiation at the PPT, we have been interested in understanding factors that may influence correct primer selection by RT. We have focused here on the viral LTR region because of the critical position it occupies in the genome with respect to plus strand synthesis. Due to the proximity to

the end of the template, plus strands incorrectly initiated from within the LTR are predicted to complete synthesis of plus strand strong stop DNA and thus complete the second template switch before PPT-primed extensions. Therefore, in contrast to regions upstream of the PPT where initiation from secondary sites by RT is tolerated by the virus, use of RNA fragments from the LTR region as plus strand primers by RT is predicted to generate dead-end products. In the present study, we have characterized the RNase H cleavage products generated by RT within the LTR to determine whether efficient and complete removal of the RNA in this region may provide a mechanism to ensure that plus strands are not initiated downstream of the PPT. Non-PPT RNA fragments that persisted following RNA degradation were tested for their capacity to prime DNA synthesis by RT.

To determine the extent to which selection of the PPT primer is influenced by the availability of potential RNA priming fragments, limited RNA degradation by the RT-associated RNase H was compared on LTR- and non-LTR-containing substrates. In both cases the resulting products primarily ranged from ~2 to 25 nt in length when analyzed by exchange labeling the RNase H cleavage products. Following extended incubation with high concentrations of RT, most of the RNA was further degraded, although ~10% of the RNA persisted as species longer than 10 nt. A similar product distribution was observed when cleavage reactions were carried out using 5' end-labeled RNA. Thus, RNA degradation by RNase H to fragments sufficiently small to dissociate from the DNA almost certainly limits the potential for incorrect priming events in the LTR. However, these results also suggest that even following considerable RNase H degradation, some RNA may remain stably annealed to the DNA template (see Table I). Since it has been found that plus strand synthesis begins as early as 10–30 min after the start of reverse transcription (17, 43), it appears possible that a substantial amount of RNA may remain hybridized to the DNA template when PPT primer selection occurs.

At early times, the RNase H cleavage products clustered in 3 size groups centered on ~24, ~16, and ~9 nt. Similar product distributions have been observed by others; the larger cleavage products are thought to result from a coordination between the polymerase domain of RT interacting with the 5' RNA end and the RNase H domain which appears to be spatially separated from the polymerase catalytic site by ~15–21 nt (9, 33, 44–46). Subsequent directional processing by RNase H during multi-cycle reactions appears to generate products in the smallest size range (9, 10, 32–34, 44, 47–49). That the majority of the cleavage products appeared to be coordinated by the RNA 5' end is consistent with recent findings suggesting that the kinetics of 5' end-directed cleavage are favorable over cleavage at internal sites (30).

Interestingly, when the RNA products from both the LTR and non-LTR regions of the genome were visualized by the exchange labeling method (and therefore carried an equivalent signal regardless of size or proximity to the 5' end), the proportion of the fragments in the 3 size classes was maintained over the time course of kinetic assays carried out at lower enzyme concentrations. Additionally, the amount of labeled product appeared to accumulate during the reaction, probably reflecting the progressive degradation of very long or full-length RNA substrates. That the concentration of the 3 classes of RNA fragments generated over time increased while the size distribution of the fragments remained unchanged during limited cleavage suggests that the initial size classes reflect stable products that are less favorable substrates for further degradation by RT than longer or full-length RNA molecules. By contrast, at late time points in the presence of very high en-

TABLE II
Ratio of non-PPT to PPT primer utilization by RT

LTR			N-LTR		
Approximate length ^a	Normalized utilization ^b	S.D.	Approximate length	Normalized utilization	S.D.
697 (PPT)	1.00	0.00	650	0.21	0.06
564	0.04	0.02	593	0.05	0.00
505	0.10	0.03	553	0.17	0.09
491	0.04	0.01	498	0.06	0.02
481	0.06	0.01	449	0.09	0.01

^a Approximate lengths of primer extension products (see Fig. 6D).

^b The amount of extension from non-PPT primers was normalized to the extension from the PPT primer and corrected for the relative number of labeled dT residues in the fragment. All calculations were based on results from two independent experiments.

zyme concentrations, the two larger size classes were reduced relative to the smallest fragments. Importantly, complete degradation of the RNA to very small products (*i.e.* 10 nt and shorter) appears to occur only after the uncleaved substrate has been exhausted. The classes of RNA products generated by RT contrasts with the very uniform distribution of products generated by cleavage with *E. coli* RNase H or by the RNase H domain alone of RT (20), both of which appear to lack sequence specificity and do not appear to be coordinated by the RNA 5' end. Thus, the polymerase domain of RT appears to play a key role in defining the RNA products generated by the RNase H activity on long hybrid duplexes.

Although the extent to which the RNA was degraded by RT appeared to be diminished when Mo-MLV NC was included in the reaction, the effect did not correlate with a decrease in the RNase H catalytic activity; the initial rates of product generation in the presence or absence of NC appeared to be identical. Our results contrast with those of others (50) who found that HIV NCp7 enhanced cleavage by HIV RNase H; however, the extent to which the enhancement was due to the facilitated formation of additional hybrid duplex substrates by NC remains unclear. Of particular note, we found that NC did not affect cleavage by *E. coli* RNase H which efficiently degraded the available RNA to fragments of ~9 nt or smaller. This result indicates that the reduced cleavage by RT in the presence of NC was not due to a reduction in the thermostability of the longer RNA fragments and, in accord with the conclusions of others (38), is suggestive of a possible specific interaction between RT and its cognate NC.

The finding that some RNA persisted downstream from the PPT led us to investigate mechanisms that may prevent use of non-PPT RNA fragments in this region for plus strand priming, since it is predicted that plus strands initiated within the LTR would generate dead-end products, unable to integrate. Although several *in vitro* studies carried out using synthetic RNA oligonucleotides have found that RT has an extremely limited capacity to extend RNA primers lacking the PPT sequence (30, 51–53), *in vivo* studies, as well as cell-free assays using longer substrates, have provided evidence for the utilization of some non-PPT RNA primers by RT (16–19, 41, 54–56). Since such studies have not identified non-PPT extensions arising from LTR-derived RNAs, we tested the possibility that cleavage within the LTR by RT-associated RNase H may selectively yield RNA products that are not extendable by the polymerase activity of RT.

We analyzed the capacity of RT to utilize the RNA fragments it generates as primers, employing dideoxynucleotides to terminate extension at discrete positions. Consistent with the results of others (26, 41, 57), we found that the PPT primer was accurately generated and extended by RT when the sequence was maintained in its native context embedded in a much longer hybrid. In addition however, we also observed extension by RT from 15 to 20 non-PPT RNA primers. The identity and the rate of utilization of the non-PPT priming RNAs were

nearly identical between the LTR and LTRΔPPT substrates when termination was with ddTTP, suggesting that RNase H cleavage and primer utilization within the LTR region were not influenced by the presence of the PPT sequence in the LTR substrate under these conditions. However, when analyzed using the run-off extension assay, the amount of product generated on the LTRΔPPT substrate was consistently greater than the amount generated by non-PPT primers on the LTR hybrid, despite identity between the sequences over the final 655 bp. This effect is likely the result of displacement of potential downstream non-PPT RNA primers by extensions initiated from the much more efficient PPT primer. We speculate that efficient utilization of the PPT primer *in vivo*, followed by RNA displacement (22, 58), may be one mechanism to suppress incorrect priming in the LTR region. These conclusions are in accord with those from studies looking at primer utilization in regions upstream of the PPT in the spleen necrosis virus system (59), but contrast with those of Powell and Levin (26) who found no sites of initiation used by HIV RT from the region surrounding the HIV PPT, perhaps reflecting intrinsic differences between the efficiency or specificity of primer selection by HIV and Mo-MLV RTs.

Although we found that RT accurately cleaved the RNA to generate the PPT primer, a number of other RNA fragments generated within the LTR by RNase H remained following cleavage, some of which were capable of priming synthesis by RT. The finding that most non-PPT RNA primers were extended at much less than 20% the efficiency of the PPT primer is consistent with work in HIV where the efficiency of non-PPT RNA primer utilization from internal sites on the genome is generally less than 10% that of priming from the cPPT (55). However, to the extent that the non-LTR sequence used in our study is representative of the remaining viral sequence, the similarities between both RNA degradation and non-PPT primer utilization on the LTR and non-LTR substrates suggest an unanticipated lack of selection within the LTR for sequences that are either completely degraded by RNase H or incapable of priming synthesis by RT.

Examination of reverse transcription also suggests that some dead-end products may be generated *in vivo* and thus at some level can be tolerated by the virus. LTR-LTR circle junctions, once thought to be the substrates for integration, are now known to be dead-end products of reverse transcription (60–62). Sequence analysis of circle junctions has revealed that many contain deletions within U3, some of which could be explained by plus strand priming downstream from the PPT (63–65).

It remains possible that the relative importance of RNA degradation *versus* efficient PPT extension may vary between viral systems. For example, while AMV RT appears to degrade RNA much less efficiently than Mo-MLV RT (7, 8), from *in vivo* studies it appears that avian C-type retroviruses may be relatively more promiscuous in plus strand RNA primer selection (17, 19, 54, 66). Therefore, we speculate that in the case of the

avian retroviruses, inefficient RNA degradation may provide a mechanism to limit competition with the PPT RNA for plus strand initiation by RT. Alternatively, for Mo-MLV and perhaps others, more extensive degradation of the RNA and very efficient extension of the PPT primer may provide the key to generating the correct product for integration.

Acknowledgments—We thank Sharon Schultz for helpful discussions and comments on the manuscript and J. L. Darlix for providing the NC proteins. We are grateful to Jamie Winshell for superb technical work, for assistance with the preparation of the manuscript, and for developing the T_m macro.

REFERENCES

- Coffin, J. M. (1996) in *Virology* (Fields, B. N., Knipe, D. M., Howley, P. M., Chanock, R. M., Monath, T. P., Melnick, J. L., Roizman, B., and Straus, S. E., eds) pp. 1767–1847, Raven Press, Ltd., New York
- Telesnitsky, A., and Goff, S. P. (1993) in *Reverse Transcriptase* (Skalka, A. M., and Goff, S. P., eds) pp. 49–83, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Katz, R. A., and Skalka, A. M. (1994) *Annu. Rev. Biochem.* **63**, 133–173
- Rein, A., Henderson, L. E., and Levin, J. G. (1998) *Trends Biochem. Sci.* **23**, 297–301
- Darlix, J. L., Lapadat Tapolsky, M., de Rocquigny, H., and Roques, B. P. (1995) *J. Mol. Biol.* **254**, 523–537
- Champoux, J. J. (1993) in *Reverse Transcriptase* (Skalka, A. M., and Goff, S. P., eds) pp. 103–118, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- DeStefano, J. J., Mallaber, L. M., Fay, P. J., and Bambara, R. A. (1994) *Nucleic Acids Res.* **22**, 3793–3800
- DeStefano, J. J., Buiser, R. G., Mallaber, L. M., Myers, T. W., Bambara, R. A., and Fay, P. J. (1991) *J. Biol. Chem.* **266**, 7423–7431
- Gopalakrishnan, V., Peliska, J. A., and Benkovic, S. J. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10763–10767
- Oyama, F., Kikuchi, R., Crouch, R. J., and Uchida, T. (1989) *J. Biol. Chem.* **264**, 18808–18817
- Mizrahi, V. (1989) *Biochemistry* **28**, 9088–9094
- Tobaly Tapiero, J., Kupiec, J. J., Santillana Hayat, M., Canivet, M., Peries, J., and Emanoil Ravier, R. (1991) *J. Gen. Virol.* **72**, 605–608
- Stetor, S. R., Rausch, J. W., Guo, M. J., Burnham, J. P., Boone, L. R., Waring, M. J., and Le Grice, S. F. (1999) *Biochemistry* **38**, 3656–3667
- Charneau, P., and Clavel, F. (1991) *J. Virol.* **65**, 2415–2421
- Charneau, P., Alizon, M., and Clavel, F. (1992) *J. Virol.* **66**, 2814–2820
- Miller, M. D., Wang, B., and Bushman, F. D. (1995) *J. Virol.* **69**, 3938–3944
- Boone, L. R., and Skalka, A. M. (1981) *J. Virol.* **37**, 109–116
- Hsu, T. W., and Taylor, J. M. (1982) *J. Virol.* **44**, 47–53
- Taylor, J. M., Cywinski, A., and Smith, J. K. (1983) *J. Virol.* **48**, 654–659
- Schultz, S. J., and Champoux, J. J. (1996) *J. Virol.* **70**, 8630–8638
- De Rocquigny, H., Ficheux, D., Gabus, C., Allain, B., Fournie Zaluski, M. C., Darlix, J. L., and Roques, B. P. (1993) *Nucleic Acids Res.* **21**, 823–829
- Kelleher, C. D., and Champoux, J. J. (1998) *J. Biol. Chem.* **273**, 9976–9986
- Shinnick, T. M., Lerner, R. A., and Sutcliffe, J. G. (1981) *Nature* **293**, 543–548
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (eds) (1999) *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc., New York
- Sugimoto, N., Nakano, S., Katoh, M., Matsumura, A., Nakamuta, H., Ohmichi, T., Yoneyama, M., and Sasaki, M. (1995) *Biochemistry* **34**, 11211–11216
- Powell, M. D., and Levin, J. G. (1996) *J. Virol.* **70**, 5288–5296
- Pullen, K. A., Rattray, A. J., and Champoux, J. J. (1993) *J. Biol. Chem.* **268**, 6221–6227
- Rattray, A. J., and Champoux, J. J. (1989) *J. Mol. Biol.* **208**, 445–456
- Robson, N. D., and Telesnitsky, A. (1999) *J. Virol.* **73**, 948–957
- Schultz, S. J., Zhang, M., Kelleher, C. D., and Champoux, J. J. (1999) *J. Biol. Chem.* **274**, 34547–34555
- Deleted in *prima*
- DeStefano, J. J., Mallaber, L. M., Fay, P. J., and Bambara, R. A. (1993) *Nucleic Acids Res.* **21**, 4330–4338
- DeStefano, J. J. (1995) *Nucleic Acids Res.* **23**, 3901–3908
- Palaniappan, C., Fuentes, G. M., Rodriguez Rodriguez, L., Fay, P. J., and Bambara, R. A. (1996) *J. Biol. Chem.* **271**, 2063–2070
- Tsuchihashi, Z., Khosla, M., and Herschlag, D. (1993) *Science* **262**, 99–102
- Tsuchihashi, Z., and Brown, P. O. (1994) *J. Virol.* **68**, 5863–5870
- Herschlag, D., Khosla, M., Tsuchihashi, Z., and Karpel, R. L. (1994) *EMBO J.* **13**, 2913–2924
- Cameron, C. E., Ghosh, M., LeGrice, S. F. J., and Benkovic, S. J. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 6700–6705
- DeStefano, J. J. (1996) *J. Biol. Chem.* **271**, 16350–16356
- Gotte, M., Maier, G., Onori, A. M., Cellai, L., Wainberg, M. A., and Heumann, H. (1999) *J. Biol. Chem.* **274**, 11159–11169
- Rattray, A. J., and Champoux, J. J. (1987) *J. Virol.* **61**, 2843–2851
- Huber, H. E., and Richardson, C. C. (1990) *J. Biol. Chem.* **265**, 10565–10573
- Dina, D., and Benz, E. W., Jr. (1980) *J. Virol.* **33**, 377–389
- Furfine, E. S., and Reardon, J. E. (1991) *J. Biol. Chem.* **266**, 406–412
- Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) *Science* **282**, 1669–1675
- Rausch, J. W., and Le Grice, S. F. (1997) *J. Biol. Chem.* **272**, 8602–8610
- Ben Artzi, H., Zeelon, E., Amit, B., Wortzel, A., Gorecki, M., and Panet, A. (1993) *J. Biol. Chem.* **268**, 16465–16471
- DeStefano, J. J., Buiser, R. G., Mallaber, L. M., Bambara, R. A., and Fay, P. J. (1991) *J. Biol. Chem.* **266**, 24295–24301
- Ghosh, M., Howard, K. J., Cameron, C. E., Benkovic, S. J., Hughes, S. H., and Le Grice, S. F. (1995) *J. Biol. Chem.* **270**, 7068–7076
- Peliska, J. A., Balasubramanian, S., Giedroc, D. P., and Benkovic, S. J. (1994) *Biochemistry* **33**, 13817–13823
- Palaniappan, C., Kim, J. K., Wisniewski, M., Fay, P. J., and Bambara, R. A. (1998) *J. Biol. Chem.* **273**, 3808–3816
- Randolph, C. A., and Champoux, J. J. (1994) *J. Biol. Chem.* **269**, 19207–19215
- Fuentes, G. M., Rodriguez Rodriguez, L., Fay, P. J., and Bambara, R. A. (1995) *J. Biol. Chem.* **270**, 28169–28176
- Kung, H. J., Fung, Y. K., Majors, J. E., Bishop, J. M., and Varmus, H. E. (1981) *J. Virol.* **37**, 127–138
- Klarmann, G. J., Yu, H., Chen, X., Dougherty, J. P., and Preston, B. D. (1997) *J. Virol.* **71**, 9259–9269
- DesGroseillers, L., Rassart, E., Zollinger, M., and Jolicoeur, P. (1982) *J. Virol.* **42**, 326–330
- Finston, W. I., and Champoux, J. J. (1984) *J. Virol.* **51**, 26–33
- Fuentes, G. M., Fay, P. J., and Bambara, R. A. (1996) *Nucleic Acids Res.* **24**, 1719–1726
- Bowman, E. H., Pathak, V. K., and Hu, W. S. (1996) *J. Virol.* **70**, 1687–1694
- Lobel, L. I., Murphy, J. E., and Goff, S. P. (1989) *J. Virol.* **63**, 2629–2637
- Brown, P. O., Bowerman, B., Varmus, H. E., and Bishop, J. M. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 2525–2529
- Ellis, J., and Bernstein, A. (1989) *J. Virol.* **63**, 2844–2846
- Kulkosky, J., Katz, R. A., and Skalka, A. M. (1990) *J. Acquir. Immune Defic. Syndr.* **3**, 852–858
- Randolph, C. A., and Champoux, J. J. (1993) *Virology* **194**, 851–854
- Hong, T., Drlica, K., Pinter, A., and Murphy, E. (1991) *J. Virol.* **65**, 551–555
- Boone, L. R., and Skalka, A. M. (1981) *J. Virol.* **37**, 117–126