4330–4338 Nucleic Acids Research, 1993, Vol. 21, No. 18

Determinants of the RNase H cleavage specificity of human immunodeficiency virus reverse transcriptase

Jeffrey J.DeStefano[1], Lisa M.Mallaber[1], Philip J.Fay[1,2] and Robert A.Bambara[1,*]
Departments of [1]Biochemistry and [2]Medicine and the [3]Cancer Center, University of Rochester, Rochester, NY 14642, USA

Received May 10, 1993; Revised and Accepted August 3, 1993

ABSTRACT

We examined the ribonuclease H (RNase H) specificity of human immunodeficiency virus reverse transcriptase (HIV-RT) using heteropolymeric RNAs hybridized to complementary DNAs. Experiments were performed in the presence of excess challenger polymer (poly(rA)-oligo(dT)) to reveal cleavages resulting from single enzyme binding events. Previous results suggested that initial RNase H directed cleavages were a fixed distance from a DNA primer terminus recessed on an RNA template, *i.e.* determined by the binding position of the polymerase active site. The influences of recessed RNA termini were not evaluated. In current experiments, RNAs that were 30, 42, or 50 nucleotides long were hybridized to the same 88 nucleotide long complementary DNA, such that the 5′ terminal nucleotide of each RNA was hybridized to the 29th nucleotide from the 3′ end of the DNA. In all three cases the RNA was initially cleaved between the 19th and 21st nucleotides from its 5′ end. Thus, cleavage was not coordinated by the recessed 3′ terminus of the RNA. Subsequent cleavages in either direction on the RNA were also observed. An insertion within the RNA that moved the preferred initial cut sequence 10 nucleotides further from the 5′ end of the RNA decreased but did not abolish cleavage at the sequence. However, changing the nucleotide sequence in the region of the preferred cleavage either by the insertion experiment or mutagenesis did not significantly alter its capacity for cleavage. These results demonstrated a dominant position preference, plus a sequence priority. In another experiment, a 25 nucleotide long DNA was hybridized such that its 3′ terminal nucleotide was 9 nucleotides from the 5′ end of a 60 nucleotide complementary RNA. The preferred RNA cleavage sequence discussed above, was 10–14 nucleotides upstream of the 3′ end of the DNA. However, initial cleavages occurred 17–20 nucleotides from the DNA 3′ end, consistent with cleavage being coordinated by the recessed 3′ terminus of the DNA primer.

INTRODUCTION

The conversion of the single-stranded RNA genome of retroviruses to a double-stranded DNA requires several steps which are carried out by the multifunctional viral RT (for a review see ref. 1). This enzyme possesses both RNA- and DNA-dependent DNA polymerase and RNase H activities (2–5). The latter activity can cleave the RNA portion of an RNA–DNA hybrid. Recently, an RT associated double-stranded RNA-dependent RNase activity has also been reported (6).

The RNase H activity of the RT is proposed to be required at several stages of viral genomic replication. These include degradation of the RNA template after the synthesis of the first strand of DNA (7), generation of a specific oligopurine ribonucleotide primer from which second strand DNA synthesis will initiate, and subsequent removal of the oligopurine primer (8–14).

Several reports that address the coordination of the DNA polymerase and RNase H activities of HIV-RT have been published (15–19). These indicate that the polymerase and RNase H active sites of the RT are spatially arranged such that they contact polymer substrates approximately 17 nucleotides apart. The actual estimates varied from 15–20 nucleotides, and the crystal structure of HIV-RT (20) predicts a separation of about 20 nucleotides of duplex RNA–DNA hybrid. A conclusion of this work is that the positioning of the RNase H active site of the RT on the RNA–DNA hybrid is determined by binding of the polymerase active site to the 3′ terminal nucleotide a DNA primer active for synthesis on the hybrid. In agreement with this explanation, Furfine and Reardon (16) and Gopalakrishnan *et al.* (17) have shown that some RT mediated cleavages, which they termed 'polymerase dependent cleavages', could be advanced upon primer extension, and remained a fixed distance from the extended 3′ primer terminus. Furfine and Reardon (16) also observed cleavages that were not advanced upon primer extension (termed 'polymerase independent'), and therefore did not seem to be directed by the polymerase domain of the RT. Both polymerase dependent and independent cleavages were influenced to some degree by sequence preferences.

In the experiments cited above, the substrates were relatively short segments of DNA hybridized to longer segments of RNA

* To whom correspondence should be addressed at: Department of Biochemistry, PO Box 607, University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, NY 14642, USA

such that the 3′ terminus of the DNA was recessed on the RNA strand. This is a configuration that is present during first strand viral DNA synthesis. Another possible configuration likely to appear during viral replication consists of relatively small segments of RNA hybridized to larger segments of DNA. We previously examined both the synthetic and RNA degradative activities of a single enzyme on a given DNA−RNA hybrid substrate (21). Excess challenge polymer was used to limit the enzyme to a single binding event on a given primer-template. We found that the polymerization and RNase H activities of HIV-, MuLV-, and AMV-RT were not strictly coupled. Polymerization generally proceeded at a much faster rate than RNase H degradation. This resulted in many segments of the original RNA template remaining undegraded after a complete round of DNA synthesis. It is likely that this RNA is degraded after first strand synthesis, by a return of the RT, or by other RTs that follow behind the RT that is performing synthesis. During this process, the RNase H substrate would be short RNAs with 3′ and 5′ termini recessed on the longer newly synthesized DNA.

We have prepared RNA−DNA hybrid substrates with this latter general configuration. Our results indicate that HIV-RT RNase H-mediated cleavage of these substrates is not coordinated by the 3′ terminus of the RNA. Initial cleavage events were influenced to some extent by nucleotide sequence, and were generally in the proximity of the 5′ terminus of the RNA.

MATERIALS AND METHODS

Materials

Recombinant HIV-RT, having native primary structure was graciously provided to us by Genetics Institute (Cambridge, MA). This enzyme had a specific activity of approximately 21,000 U/mg. One unit of RT is defined as the amount required to incorporate 1 nmole of dTTP into nucleic acid product in 10 min at 37°C using poly(rA)-oligo(dT) as template-primer. Aliquots of HIV-RT were stored frozen at −70°C and a fresh aliquot was used for each experiment. T4 DNA ligase, T4 polynucleotide kinase, T7 RNA polymerase, *Escherichia coli* DNA polymerase I (Klenow fragment), and *Escherichia coli* ribonuclease H were obtained from United States Biochemical Corp. T3 RNA polymerase, RNase T_1, bovine pancreatic DNase (RNase free), placental RNase inhibitor, rNTPs and all restriction enzymes were obtained from Boehringer Mannheim Biochemicals; dNTPs were obtained from Pharmacia. The oligonucleotides for site-directed mutagenesis and substrate F (see below) were synthesized by Genosys Inc. (formerly Genetic Designs) (Houston, TX). The *Eco Rl−Sma 1* adapter, used in substrate E (see figure 1) was purchased from United States Biochemical Corp. The Double-Take double-stranded mutagenesis kit, used to create substrate G (see figure 1) and Supercompetent XL1-Blue cells were obtained from Stratagene Cloning Systems. All other chemicals were from Sigma Chemical Co. Radiolabelled compounds were from New England Nuclear or Amersham.

Methods

Unchallenged RNase H assays. HIV-RT (100 ng) was pre-incubated with 2 nM substrate for 3 min in 10 μl of 50 mM Tris−HCl (pH 8.0), 1 mM dithiothreitol, 2% glycerol (w/w), 0.1 mM EDTA, and 5 mM KCl (buffer A),. Assays were initiated by addition of $MgCl_2$ in 2.5 μl of buffer A to give a final concentration of 6 mM $MgCl_2$. Reactions were run for 10 minutes at 37°C and were terminated by the addition of 1 volume of 2 times concentrated gel electrophoresis loading buffer (90% formamide (v/v), 10 mM EDTA (pH 8), 0.1% (w/v) xylene cyanol and bromophenol blue).

Challenged RNase H assays. In the challenged assay, reaction conditions were the same as for the standard RNase H assay, except that the amount of HIV-RT used was decreased to 15 ng and 1.0 μg of poly(rA)-oligo(dT_{16}), the trapping polymer, was included along with the divalent cation at the start of the reaction. This modification limits the RNase H activity to preformed complexes of RT and hybrid substrate (21) The challenged reactions were terminated at various times ranging from two seconds to seven minutes as indicated on the Figures. Poly(rA)-oligo(dT_{16}) used in the reactions was prepared by mixing oligo(dT_{16}) with poly(rA) at a 1:8 ratio (w/w) in 10 mM Tris−HCl (pH 8.0) and 1 mM EDTA. The mixture was incubated for 30 min at 37°C, and then cooled slowly to room temperature.

RNA−DNA hybridization. Hybrids were prepared by mixing RNA and DNA at an approximately 1:2 ratio of 3′ termini in 10 mM Tris−HCl (pH 8.0), 1 mM EDTA and 80 mM KCl. The mixture was heated to 65°C for 10 min, and then cooled slowly to room temperature.

Quantitation of nucleic acids. RNAs that were labeled internally (see below) during run-off transcription were quantitated based on the specific activity of the labeled nucleotide in the transcription reaction, and the number of times that the nucleotide was present in the transcript. RNA that was 5′ end-labeled, and single stranded DNAs (see Figure 1), were quantitated by a native gel hybridization 'shift-up' assay. In this assay, the RNA or DNA being quantitated was hybridized to a DNA nucleic acid of known concentration which was complementary to a portion of that RNA or DNA. Hybridizations were performed as described above. The hybridized samples were mixed with 6 times concentrated native gel electrophoresis buffer (40% (w/v) sucrose, 0.25% (w/v) xylene cyanol and bromophenol blue) and loaded onto 12% native polyacrylamide gels. Gels were prepared and subjected to electrophoresis as described (22). When DNAs were quantitated, a fixed level of 5′ end-labeled complementary DNA was hybridized to various amounts of the DNA being quantitated. For quantitation of 5′ labeled RNAs, the RNA was kept at a constant level and the amount of unlabeled complementary DNA was varied. The concentration of the DNA being quantitated was evaluated from the amount required to 'shift-up' approximately 50% of the labeled complementary DNA (assuming 100% hybridization efficiency). The concentration of RNA was evaluated based on the amount of unlabeled complementary DNA required to 'shift-up' approximately 50% of the RNA.

Run-off transcript. Run-off transcription was done as described in the Promega Protocols and Applications Guide (1989). Plasmids pBSM13+ (substrates A−D, and G), pBSM13+ (*Eco* insertion at 885, substrate E), or pBSM13+ (GC, 895, 896) (substrate F) were isolated from transformed *Escherichia coli*. XL1 Blue cells by using a Qiagen plasmid kit as described by the manufacturer. The plasmid DNA was cleaved with various restriction enzymes (*Bam* HI for substrates A and E; *Sal* I for substrates B, D, and F; *Sph* I for substrate C; *Hind* III for substrate G (see Figure 1)). Either T7 (substrates A−C, E, and F) or T3 (substrate D) RNA polymerase was used to prepare

run-off RNA transcripts of various lengths, as indicated in Figure 1. For preparation of internally labeled transcripts, 500 μM each of ATP, GTP, UTP, and 18 μM α-[^{32}P]-CTP (approximately 44 Ci/mmol) were used in the transcription reactions. The DNA template was digested with bovine pancreatic RNase-free DNase I (1 unit/mg of DNA). The RNA transcript was then extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:24:1), and then ethanol precipitated twice using 3 volumes of ethanol and a final concentration of 1.25 M ammonium acetate. The run-off transcript was resuspended and purified by electrophoresis on 8, 10, or 12% polyacrylamide gels (depending on the size of the transcript) containing 7 M urea. The full-length transcript was located by autoradiography, excised from the gel, and eluted overnight in a buffer containing 0.5 M ammonium acetate, 1 mM EDTA, 0.1% sodium dodecyl sulfate. The eluate was separated from the polyacrylamide by centrifugation in a microfuge, and subsequent filtration through a 0.45 μm, 25 mm disposable syringe filter (Nalgene). The filtrate was then ethanol precipitated with 3 volumes of ethanol and 1/10th volume of 2 M NaCl.

RNA transcripts that were 5′ end-labeled were prepared using 500 μM ATP, GTP, UTP, and CTP. The transcript reaction was then run directly over a Nensorb (Dupont) column according to the directions of the manufacturer. The isolated transcript was treated with alkaline phosphatase to remove the 5′ phosphate, and subsequently the alkaline phosphatase was heat inactivated and removed by phenol extraction as described (22). The dephosphorylated RNA was then 5′ end-labeled with γ-[^{32}P]-ATP (approximately 3000 Ci/mmol) and T4 polynucleotide kinase as described in the USB Molecular Biology Reagents/Protocols Manual (1992). The transcript reaction was run over a second Nensorb column, and then purified by gel electrophoresis as described above.

Preparation of pBSM13+ (Eco R1−Sma 1 adapter insertion 885). Substrate E (see Figure 1) was prepared by inserting a *Eco* R1−*Sma* 1 adapter with the following sequence: 5′-d(AATTCC-CGGG)-3′ at position 885 of pBSM13+. Ligation of the termini was performed according to Sambrook *et al.* (22). Following ligation, the plasmid with insert was used to transform competent XL1 Blue cells. Colonies were selected for loss of β-galactosidase activity by standard procedures (22). Plasmid DNA was isolated as described above (under *Run-off transcript*). The presence of the inserted DNA was verified by DNA sequencing of the appropriate region of the plasmid.

Preparation of pBSM13+ (GC, 895, 896). Substrate F (Figure 1) was prepared by changing the nucleotides of pBSM13+ at positions 895 and 896, from an A and T residue, respectively, to a G and C residue using Stratagene's Double-Take double-stranded mutagenesis kit according to manufacturers' directions. The mutant primer that was used in the protocol had the following sequence: 3′-AGCTCGAGCCGCGGGCCCTAG-5′, with the altered nucleotides underlined. After extension and recircularization of the mutagenized DNA, it was used to transform competent XL1 Blue cells. Several colonies were picked, and plasmid 'mini preps' were prepared (22). Plasmid DNA was isolated as described above. DNA from individual colonies was sequenced over the appropriate region to confirm the presence of the mutation. Plasmid DNA was then prepared from a single colony containing the mutated plasmid.

Gel electrophoresis. Denaturing 10% polyacrylamide sequencing gels (19:1 acrylamide:bis-acrylamide), containing 7 M urea, were prepared and subjected to electrophoresis as described (22).

Preparation of RNA G-ladder and base hydrolysis ladder. The RNA G-ladder was prepared by limited digestion of 5′ end-labeled RNA with RNase T1. Reactions contained in 10 μl: 7M Urea, 50 mM sodium citrate (pH 5), 1 mM EDTA, 2 μg *Escherichia coli.* tRNA, approximately 0.05 pmoles 5′-labeled RNA, and 2 units of RNase T_1. The reaction was performed at 56°C for 15 minutes and loaded directly onto the gel or frozen at −70°C. The base hydrolysis ladder was prepared by incubating in a volume of 5 μl, 5′ end-labeled RNA (approximately 0.05 pmoles) in 0.1 N NaOH at 70°C for 1.5 minutes. The reaction was then neutralized with HCl. An equal volume of gel loading buffer was added to the sample before loading.

RESULTS

The recessed 3′ terminus of the RNA on an RNA−DNA hybrid does not coordinate HIV-RT-directed RNase H mediated cleavage

Substrates A−D in Figure 1 were designed to test whether the positions of HIV-RT mediated cleavages were determined by the proximity of the potential cleavage site to the recessed 3′ terminus of an RNA primer on a longer DNA template. If this were the case, the RNase H active site of the RT is probably initially positioned by binding of the RT polymerase active site to the RNA 3′ terminus.

Since we wanted to determine the position of the initial RT-mediated cleavage events on each substrate, we designed

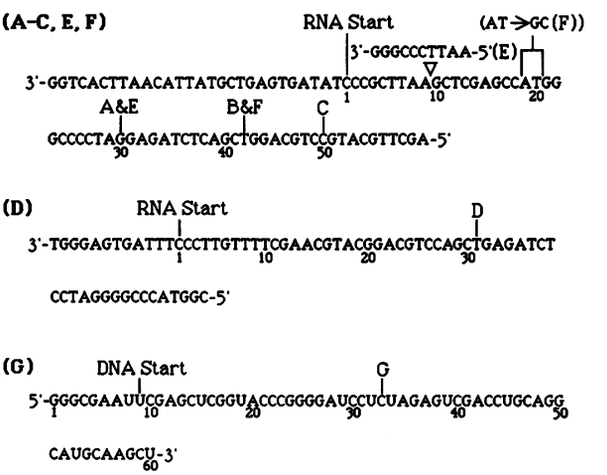

Figure 1. Configuration of substrates. The configuration of the experimental substrates is shown. The name of each substrate (A−G) is given at the upper left of the appropriate substrate(s). RNA is drawn in the 5′ to 3′ direction. DNA is drawn 3′ to 5′. In each case, the nucleotide sequence of the longer of the two strands of the heteroduplex is shown. The positions where the 5′ terminal nucleotide of the recessed RNA (substrates A−F) or 3′ terminal nucleotide of the recessed DNA (substrate G) are hybridized to the longer strand are indicated by RNA or DNA 'Start'. The position of the last nucleotide of the duplex is indicated by the letter denoting the substrate name. The sequence and position of the 10 base insertion in substrate E is indicated (▽) as are the positions of the AT to GC base change on the DNA of substrate F. The nucleotides of the RNA portion of each substrate are numbered (below each substrate) starting from the 5′ terminal residue.

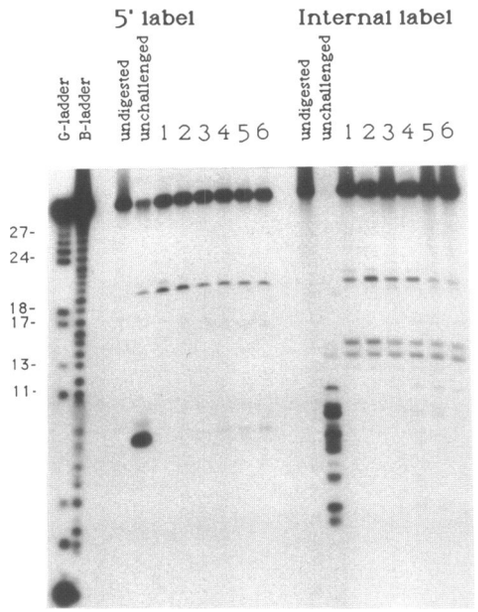

Figure 2. HIV-RT mediated cleavage of substrate A. An autoradiogram of an experiment in which the RNA portion of substrate A was 5′ end or internally labeled (as indicated) with ^{32}P is shown. The lanes labeled 'undigested' show the substrate incubated without enzyme for 7 minutes. The lanes labeled 'unchallenged' show the substrate incubated in the absence of poly(rA)-oligo(dT) challenger polymer for 10 minutes as described under **Methods**. Lanes 1–6 show challenged assays incubated for 2 and 10 seconds, and 1, 3, 5, and 7 minutes, respectively. Lanes labeled 'G-ladder' and 'B-ladder' were prepared using 5′ ^{32}P labeled substrate RNA by limited digestion with T1 RNase (G-ladder) or base hydrolysis (B-ladder) as described under *Methods*. The numbers to the left designate the length of the RNA (in nucleotides) as determined from the positions of some of the guanosine nucleotides in the substrate RNA.

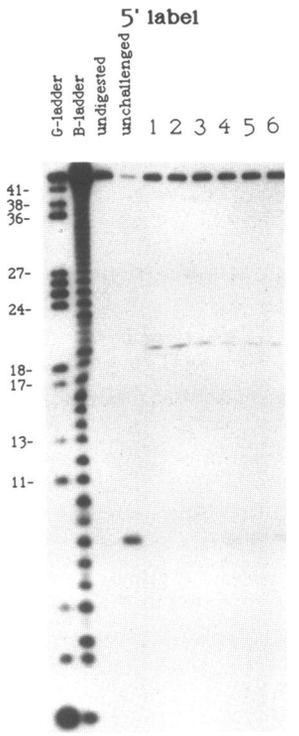

Figure 3. HIV-RT mediated cleavage of substrate B.- An autoradiogram of an experiment in which the RNA portion of substrate B was 5′ end labeled with ^{32}P is shown. Figure labels are as described in Figure 2.

experiments to allow substrate molecules to experience only a single enzyme binding event in the course of the reaction. This was accomplished in part by using excess substrate over enzyme. In general, less than 50% of the total substrate was cleaved by the enzyme used in these experiments. In addition, the enzymes were mixed with the substrate, then a large excess of a challenge polymer, over substrate, was added with the divalent cation used to start the reaction. The challenge polymer traps enzymes that dissociate from the substrate, preventing the return of any enzyme after it first dissociates (21). By using this approach, and performing the reactions over a time course, we were able to observe the initial cleavage events and any subsequent cleavages catalyzed by the RT prior to its dissociation.

The positions of the initial cleavages were determined by using 5′-^{32}P labeled RNA substrates. That the observed cleavages were actually the initial cleavages was confirmed by performing, in parallel, experiments with internally labeled RNA. Comparison of the gel profiles from the two types of experiments at early time points indicated that on most substrates a single major cleavage event occurred at a preferred position. RNA oligomers derived from the 3′ and 5′ portions of the RNA were initially observed. The position of the 5′-derived product in reactions with internally labeled substrate was determined based on comigration with the product in the reactions with the 5′ labeled substrate. The 3′-derived product, in the reactions with internally labeled substrates, was the other major product at early time points.

Shown in Figure 2 are the results from an experiment performed using substrate A. An RNA oligomer, 19 nucleotides in length (based on the migration of this product relative to the G- and base ladders), derived from the 5′ end of the substrate RNA (left side panel) was present at the two second time point (lane 1). There was also a faint band corresponding to a product 20 bases in length. At longer incubation times, a portion of the 5′-derived products were further degraded, and RNA oligomers of 15 and 8 nucleotides in length began to appear. These smaller products must have resulted from further degradation of the original 5′-derived product by a 3′ to 5′ (with respect to the RNA) directional nuclease activity. Such an activity has been reported to be associated with HIV-RT (15, 23). The right side panel of Figure 2 shows the results from parallel experiments using internally labeled RNA. At the two second time point (lane 1), a band corresponding to the 5′-derived RNA product was evident, and a pair of smaller bands were also seen. These were understood to be the 3′-derived products.

Using the G- and base ladders as standards, the major 5′-derived product in the above experiment migrated at a position corresponding to 19 nucleotides in length. The RNA portion of the substrate was originally 30 nucleotides long. Therefore, a major 3′-derived product of 11 nucleotides would be expected. Based on the standards, the presumed 3′-derived products were slightly longer than this. Since 3′-derived products could not be longer than 11 nucleotides if the 5′-derived product were 19 nucleotides, the discrepancy probably occurred because of a small difference in the migration of the ladders *vs.* the RT-derived products. This migration difference may have been caused by either the nucleotide sequence in the case of the 3′ products, or the presence of phosphates at the 3′ termini of the base ladder and G-ladder products. Because of this slight difference, the

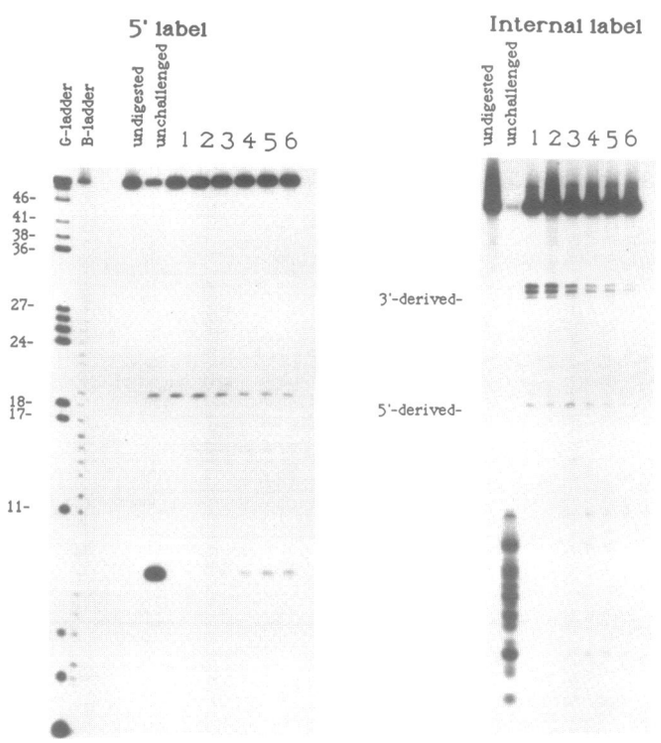

Figure 4. HIV-RT mediated cleavage of substrate C. An autoradiogram of an experiment in which the RNA portion of substrate C was 5′ end or internally labeled (as indicated) with ^{32}P is shown. Figure labels are as described in Figure 2. In addition, the positions of the 3′ and 5′ derived RNA cleavage products from the substrate with internally labeled RNA are indicated.

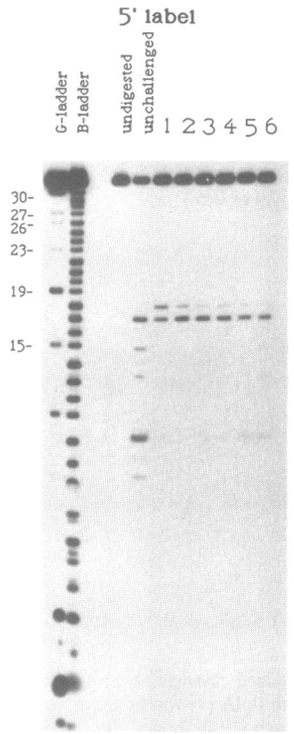

Figure 5. HIV-RT mediated cleavage of substrate D. An autoradiogram of an experiment in which the RNA portion of substrate D was 5′ end labeled with ^{32}P is shown. Figure labels are as described in Figure 2.

lengths of the 5′ and 3′-derived products may not be exactly the values indicated by the standards.

It appears from these results that the predominant initial cleavage event was between the 19th and 20th nucleotide from the 5′ end of the RNA molecule. However, the experiment does not rule out the possibility that a substantial portion of initial cleavages occurred between nucleotides 20 and 21, followed quickly by the cleavage of one additional nucleotide before the first time point was taken. This was possible since, at the two second time point, bands corresponding to each of the two cleavages were observed in both the 5′ labeled and internally labeled substrate lanes.

We note that 5′-derived cleavage products from both 5′-end and internally-labeled substrates migrate to similar positions on the gel, even though the 5′-ends of the internally-labeled products are triphosphorylated. Evidently, the additional charges, when part of the triphosphate structure, do not significantly increase mobility.

Figures 3 and 4 show the results of experiments performed using substrates B and C, respectively. These substrates are similar to substrate A (Figure 2), differing only in that the RNA portion of the substrates has been extended in the 3′ direction by an additional 12 and 20 nucleotides in substrates B and C, respectively. Thus, the 3′ termini of the RNA primers are 12 and 20 nucleotides further downstream than those of substrate A. Despite this, the initial cleavages on these substrates occurred at the same positions as with substrate A. At the two second time points, with 5′-labeled RNAs (lanes 1, Figures 3 and 4, substrates B and C, respectively) a major 5′-derived RNA and a minor RNA, 19 and 20 nucleotides in length, respectively, were observed. A portion of these products were further degraded as the reaction proceeded. When the internally labeled substrate was used (substrate C only, Figure 4), the 19 and 20 nucleotide long 5′-derived products were also observed, accompanied by major 3′-derived products approximately 30−31 nucleotides in length. It is interesting to note that a portion of the 3′ derived products were also further degraded as the reaction proceeded. This would have resulted from a 5′ to 3′ (with respect to the RNA) directional nuclease activity. Such an activity has also been noted by others (24). However, unlike the 3′ to 5′ activity, it has not been extensively characterized. The results clearly demonstrate that the position of the initial cleavages on these substrates was not determined by binding of the RT polymerase domain to the recessed 3′ terminus of the RNA.

We performed an identical experiment to that shown in Figure 2 using internally labeled RNA, except that 50 μM each of the four dNTPs was included in the reaction. Addition of nucleotides had no effect on the amount or distribution of the cleavage products at any given time point (data not shown). There was also no indication of any DNA synthesis initiated from the RNA strand of the substrate in the presence of trapping polymer. However, in experiments performed in the absence of trapping polymer, a low level of DNA synthesis was observed (data not shown). This demonstrates that the RT can use the RNA, or RNA cleavage products, to prime DNA synthesis, albeit inefficiently. These results suggest that the RNAs used in our assays were not recognized well as primers for the initiation of DNA synthesis, but instead were recognized as substrates for RNase H mediated degradation.

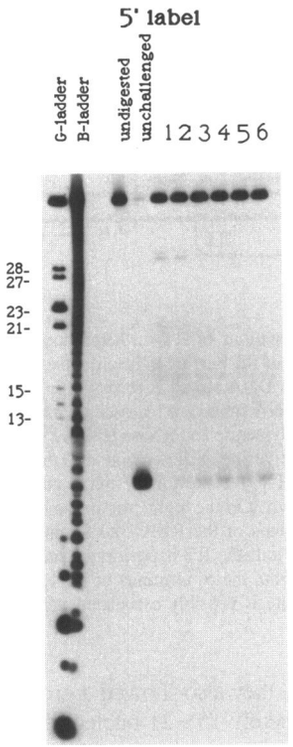

Figure 6. HIV-RT mediated cleavage of substrate E. An autoradiogram of an experiment in which the RNA portion of substrate E was 5′ end labeled with ^{32}P is shown. Figure labels are as described in Figure 2.

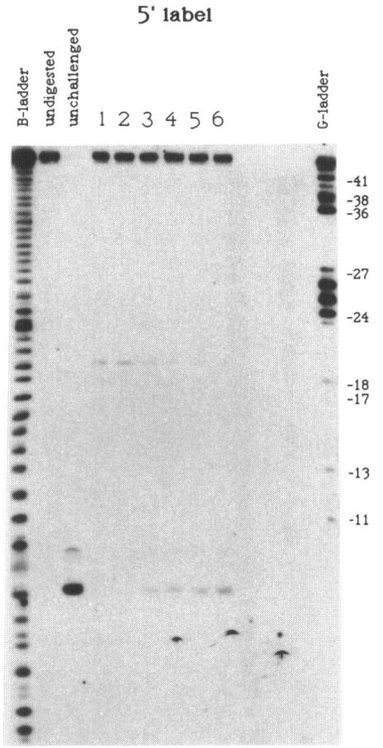

Figure 7. HIV-RT mediated cleavage of substrate F. An autoradiogram of an experiment in which the RNA portion of substrate F was 5′ end labeled with ^{32}P is shown. Figure labels are as described in Figure 2.

It was possible that, although RNA cleavage was not directed by the 3′ terminus of the RNA on substrates A, B, and C it may have been on another substrate. To test this, we prepared a similar substrate but with a different sequence (Figure 1, substrate D). Figure 5 shows the cleavage of this substrate, which was labeled with ^{32}P at the 5′ end. Cleavage of substrate D resulted in 5′-derived RNAs that were 17 and 18 nucleotides in length. As with substrates A, B, and C, a portion of the 5′-derived product was further degraded as the reaction proceeded, generating some smaller products. We also performed this experiment with a substrate that was identical to substrate D except that the RNA portion of the substrate was 12 nucleotides longer (in the 3′ direction). The initial cleavage position on this substrate was the same as that on substrate D (data not shown). Thus, even on a substrate with a thoroughly different sequence, the 3′ terminus of the RNA did not determine the positions of RNA cleavages.

The effect of sequence and position on RNA degradation

The initial cleavage positions on all of the substrates described above did not vary when the position of the RNA 3′ terminus was changed. That is, they remained fixed with respect to the 5′ terminus of the RNA. The initial cleavages may have been directed by a particular preferred sequence in this region. Possibly the proximity to the 5′ end of the RNA was important. It was clear from the two sets of substrates tested that there is not a set distance from the 5′ end of the RNA at which initial cleavages occur. With one set of substrates (A, B, and C) the major initial cleavage occurred 19 nucleotides from the 5′ end while with the substrate of a different sequence (D) cleavage was 17 nucleotides from the end.

Substrate E was employed to test whether the preferred cleavage site from substrates A, B, and C could direct cleavage from another position on the substrate (see Figure 1). This substrate was prepared by inserting a 10 base pair *EcoR* I linker into the *EcoR* I site on substrate A. This manipulation moved the preferred cleavage sites that were 19 and 20 nucleotides from the 5′ end of the RNA on substrate A, to positions 29 and 30 on substrate E. HIV-RT mediated cleavage of 5′ end labeled substrate E is shown in Figure 6. In this case multiple cleavages were observed at the two second time point (lane 1). There was a major group of cleavages 15−21 nucleotides from the 5′ end of the RNA. Many of these were within the AAUU region of the insert (positions 16−19). Also evident were cleavages at positions 29 and 30. The result indicated that even when the preferred cleavage site was 10 nucleotides further from the 5′ end of the RNA it was still one of the initial cleavage sites. However, the level of cleavage at this site was substantially diminished relative to all the cleavages observed. This conclusion presumes that the sites observed at the two second time point in Figure 6 were initial cleavages. It was possible that some of the smaller products resulted from secondary cleavages of the original 29 and 30 nucleotide long products. However, this seems unlikely since on the other substrates tested (Figures 2 and 4) it appeared that only single cleavages generally occurred by two seconds. Also, on all the substrates tested, the time frame of subsequent cleavages of the initially observed products indicated

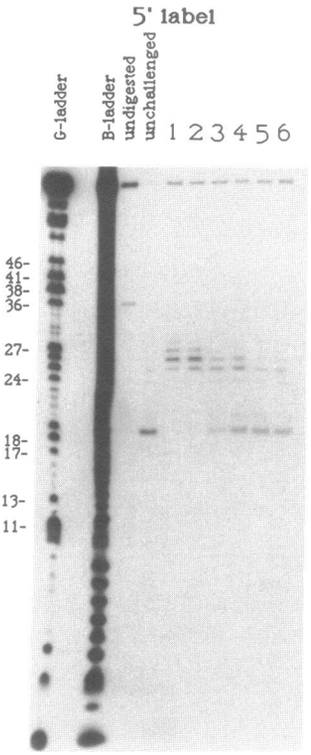

Figure 8. HIV-RT mediated cleavage of substrate G. An autoradiogram of an experiment in which the RNA portion of substrate G was 5′ end labeled with ^{32}P is shown. Figure labels are as described in Figure 2.

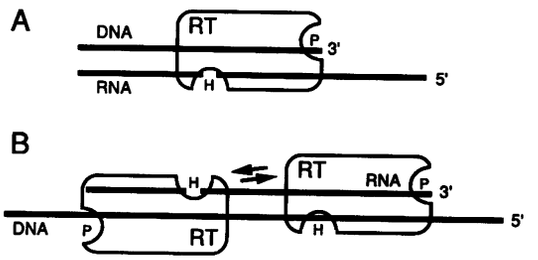

Figure 9. Possible orientations of RT on heteroduplex substrates. Shown is a schematic representation of RT binding to heteroduplex substrates configured such that the 3′ terminus of a DNA strand is recessed on an RNA strand (**A**) or the 3′ and 5′ termini of an RNA strand are recessed on a DNA strand (**B**). The letters 'P' and 'R' denote the polymerase and RNase H active sites of the RT, respectively. In A, the polymerase active site is shown bound to the 3′ terminus of the DNA, positioning the RNase H active site for the initial cleavage. In B, two possible binding modes are shown. On the right, the polymerase active site of the RT is bound to the 3′ terminus of the RNA, positioning the RNase H active site along the DNA. On the left, the RT is positioned such that the RNase H active site is a fixed distance from the 5′ terminus of the RNA. In this configuration, the polymerase active site is possibly associated with the DNA.

that this process is relatively slow. A substantial increase in products derived from multiple cleavages was generally not observed until at least 1 minute into the reactions.

Mutation in the region of the preferred cleavage site did not alter the position of the initial cleavage

If the nucleotide sequence at the preferred cleavage site was the major factor in directing RNase H mediated cleavage, then altering the sequence should change the cleavage site. To test this we prepared substrate F by site-directed mutagenesis of the plasmid used to produce substrates A, B, and C. Substrate F is identical in sequence to substrate B at all positions except for 19 and 20, from the 5′ end of the RNA. The T and A residues in the plasmid at these positions were changed to C and G, respectively. Thus, the RNA transcript produced from the mutated plasmid had the sequence 5′-GCGC-3′ at positions 18−21 while the RNA from the unaltered plasmid read 5′-GUAC-3′ over the same span. As was indicated earlier, there was a slight uncertainty in defining the length of initial cleavage products from the standards. Therefore, the mutated positions in substrate F either span or are adjacent to the preferred cleavage site of substrate B. Figure 7 shows an experiment performed with 5′ end-labeled substrate F. The initial cleavage event occurred at position 19 (lane 1, two second time point) on this substrate just as it did on substrate B. The result indicates that HIV-RT displays a strong preference to cleave the RNA approximately 19 nucleotides from its 5′ end.

This observation is also consistent with the results in Figure 6. Those indicated that there was some sequence preference to cleavage. However, they also showed a group of prevalent initial cleavages approximately 15−21 nucleotides from the 5′ end of substrate E, even when the preferred cleavage sequence was moved and replaced by another sequence. Clearly there is a strong positional preference for HIV-RT-directed cleavages on substrates of this general configuration.

We also note that the G-ladder produced from the RNA portion of substrate F did not indicate that a G residue was present at position 20. The existence of a G at this position was verified by DNA sequencing of the site-directed mutant. The inability to detect this G residue with T-1 RNase may be the result of a change in the structure of the RNA caused by the mutation.

Can cleavage be directed away from the preferred cleavage site by the recessed 3′ terminus of a DNA primer?

On substrates having a recessed 3′ terminus of a DNA primer on a complementary RNA, many of the cleavages are directed by the binding of the RT to the DNA 3′ terminus. There are, however, some sites, referred to in the **Introduction**, at which cleavage does not appear to be influenced by the RT binding to the 3′-DNA terminus (16, 17). We wanted to determine whether the preferred cleavage site sequence evident from our experiments could override the positioning of cleavage directed by the 3′ terminus of a recessed DNA.

Substrate G (see Figure 1) was used to explore this possibility. This substrate consisted of a 25 nucleotide long DNA oligonucleotide hybridized to a complementary RNA that was 60 nucleotides long, and identical to the RNA of substrate C for the first 50 nucleotides from the 5′ end. The 3′ terminal nucleotide of the DNA was bound to the 9th nucleotide from the 5′ terminus of the RNA. This positioned the preferred cleavage sites recognized in substrates A, B and C 11−12 nucleotides downstream of the 3′ terminus of the DNA. Note that these sites were 10 and 11 nucleotides from the 3′ terminus of the RNA in substrate A. An experiment using substrate G, with the RNA labeled at the 5′ end, is shown in Figure 8. Cleavages producing RNAs from 25−28 nucleotides in length were observed at the two second time point (lane 1). As the reaction proceeded,

degradation of a portion of these initial products, by the 3′ to 5′ nuclease activity of the RT, was also observed. The most prominent initial cleavage product was about 26 nucleotides in length. There were also smaller amounts of products 1−3 nucleotides longer and 1 nucleotide shorter. The amount of the shorter product increased with time up to about 3 minutes, while the amounts of the larger products decreased over time. This suggests that smaller products in the group are generated, at least in part, by additional cleavages of the larger ones.

The lengths of the initial products implied that cleavage occurred 17−20 nucleotides downstream of the 3′ terminus of the DNA. These lengths are consistent with cleavage having been directed by binding of the RT polymerase domain to the DNA 3′ terminus. This result suggests that the positions of the cleavages were coordinated by the binding if the RT polymerase active site to the 3′ terminus of the recessed DNA. However, further experimentation would be required to prove coordination unequivocally. It was also interesting that much of the initially cleaved RNA was degraded over time to a product that was 19 nucleotides long, consistent with cleavage at the preferred cleavage site from substrates A−C.

DISCUSSION

We have examined HIV-RT directed RNA cleavage on substrates with relatively short RNAs hybridized to longer complementary DNAs or short DNA hybridized to a longer complementary RNA. In both cases the 3′ and 5′ termini of the shorter nucleic acid were recessed on the longer one. Our results indicated that the recessed 3′ termini of DNAs but not RNAs can direct the cleavage of the RNA on RNA-DNA hybrids. On substrates with RNA 3′ termini recessed on DNA, cleavage specificity was influenced by the proximity of the sequence to the 5′ end of the RNA, while some sequence specificity was also observed.

Since the RNAs used in these experiments possessed recessed 3′ termini, it was possible that the RT could use the RNAs as primers for DNA synthesis. During retrovirus replication RNAs are used to prime both first and second strand DNA synthesis (for a review see ref. 1). It was possible that the polymerase domains of some of the RTs in the above experiments were preferentially binding to the 3′ terminus of the RNA. However, this configuration may not position the RNase H domain in a manner that allows RNase H mediated cleavage to occur (see Figure 9). Since dNTPs were not included in the reactions, extension of the RNA would not have been possible. Therefore, only RTs bound in the proper configuration to mediate cleavage would have been detected. It is possible that the RT binds preferentially to the 3′ terminus of the RNA but can shift positions and mediate cleavage. Thus, the above experiments do not necessarily prove that the RT has greater affinity for the cleavage site region than the RNA 3′ terminus. However, when dNTPs were included in the cleavage assays, there was no indication of any DNA synthesis initiated from the RNA primer on the substrate in the presence of trapping polymer. These results suggest that the RNAs used in our assays were not recognized well as primers for the initiation of DNA synthesis, but instead were recognized as substrates for RNase H mediated degradation. The reason for the inefficient priming is not known. One possibility is that the RT does not bind to the substrate in an appropriate configuration to position the polymerase active site at the RNA 3′ terminus (see Figure 8 and discussion below). This would suggest that the RT has relatively higher affinity for regions of the substrate other than the RNA 3′ terminus. A second possibility is that extension of RNA primers by RT is inefficient in comparison to extension of DNAs. Thus, although the RT polymerase active site may bind to the RNA 3′ terminus the enzyme has difficulty extending the RNA.

Although RTs can use some RNAs as primers for DNA synthesis, it is not surprising that the recessed 3′ terminus of DNA but not RNA can direct RNA cleavage. Perhaps the spacial arrangement of the polymerase and RNase H active sites of the RT precludes cleavage of the RNA strand when the polymerase active site is bound to the 3′ terminus of the RNA (see Figure 9). It was interesting that the preferred cleavage sites on the recessed RNA substrates used in these studies were 17−20 nucleotides downstream of the 5′ terminus of the RNAs. This is approximately the distance between the polymerase and RNase H active sites of HIV-RT. One possibility is that the RT polymerase active site is positioned near the 5′ terminus of the RNA but is 'recognizing' the DNA strand. That is, the positioning of the RT is the same as would be expected if a DNA primer with a recessed 3′ terminus were bound to an RNA template (as in substrate G). This configuration would position the RNase H catalytic site over the RNA strand of the hybrid and allow cleavage (see Figure 9).

During the course of viral replication it is likely that several potential RNA primers are generated. The RT may either initiate synthesis from these RNAs, degrade them or both. Results from this study and others (10, 25) suggest that many of these RNAs are recognized as substrates for RNase H mediated degradation, and are not efficient primers for DNA synthesis. The RNA is degraded by the RNase H activity of the RT rather than used as a primer. A special case is the polypurine tract, a portion of the viral genome that is used to prime second strand DNA synthesis. It is possible that the RT has high affinity for, and can efficiently initiate synthesis from the 3′ terminus of the tract. Alternatively, the inability of the RT to degrade this tract allows it to be used as a primer, although inefficiently in comparison to DNA primers.

Our results suggest that recessed 3′ termini of DNA on RNA templates, but not RNA on a DNA templates, can coordinate the HIV-RT RNase H mediated cleavage of the RNA strand. Both configurations represent structures that are present during viral replication. It is clear that the specificity's involved with determining the cleavage of these two types of structures are quite different. Initial cleavage of the latter type of structure appears to be influenced by the binding of the RT near the recessed 5′ terminus of the RNA strand. Further experimentation will be required to determine what directs the binding of the RT to this region, and how the RT assumes the proper conformation to cleave the recessed RNA strand.

ACKNOWLEDGEMENTS

We wish to thank Drs Jasbir Seehra and John McCoy, representing Genetics Institute, for the generous gift of HIV-RT. This work was supported by NIH grant AI 27054 and in part by core grant CA 11198 to the University of Rochester Cancer Center. P.J.F. is an Established Investigator of the American Heart Association.

REFERENCES

1. Goff, S. P. (1990) *Journal of Acquired Immune Deficiency Syndromes* **3,** 817–831.
2. diMarzo Veronese, F., Copeland, T. D., DeVico, A. L., Rahman, R., Oroszlan, S., Gallo, R. C., and Sarngadharan, M. G. (1986) *Science* **231,**1289–1291.
3. Hansen, J., Schulze, T., and Moelling, K. (1987) *J. Biol. Chem.* **262,** 12393–12396.
4. Starnes, M. C. and Cheng Y. C. (1989) *J. Biol. Chem.* **264,** 7073–7077.
5. Varmus, H. and Swanstrom, R. 1984. in *RNA Tumor Viruses,* (Weiss, R., Teich, N., Varmus, H., and Coffin, J., eds.) 2nd Ed., pp. 369–512, Cold Spring Harbor Laboratory Publications, Cold Spring Harbor, N.Y.
6. Ben-Artzi, H., Zeelon, E., Gorecki, M., and Panet, A. (1992) *Proc. Natl. Acad. Sci. USA,* **89,** 927–931.
7. Watson K. F., Schendel, P. L., Rosok, M. J., and Ramsey, L. R. (1979) *Biochemistry* **18,** 3210–3219.
8. Champoux, J. J., Gilboa, E. and Baltimore D. (1984) *J. Virol.* **49,** 686–691.
9. Finston, W. I. and Champoux, J. J. (1984) *J. Virol.* **51,** 26–33.
10. Huber, H. E. and Richardson, C. C. (1990) *J. Biol. Chem.* **265,** 10565–10573
11. Mitra, S. W., Goff, S., Gilboa, E., and Baltimore, D. (1982) *J. Biol. Chem.* **257,** 5983–5986.
12. Resnick, R., Omer, C. A., and Faras, A. J. (1984) *J. Virol.* **51,** 813–821.
13. Smith, J. K., Cywinski, A. and Taylor, J. M. (1984a) *J. Virol.* **49,** 200–204.
14. Smith, J. K., Cywinski, A. and Taylor, J. M. (1984b) *J. Virol.* **52,** 314–319.
15. DeStefano, J. J., Buiser, R. G., Mallaber, L. M., Bambara, R. A., and P. J. Fay. (1991a) *J. Biol. Chem.* **266,** 24295–24301.
16. Furfine, E. S. and Reardon, J. E. (1991) *J. Biol. Chem.* **266,** 406–412.
17. Gopalakrishnan, V., Peliska, J. A., and Benkovic, S. J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 10763–10767.
18. Fu, T. and Taylor, J. (1992) *J. Virol.* **66,** 4271–4278.
19. Wöhrl, B. M., and Moelling, K. (1990) *Biochemistry* **29,** 10141–10147.
20. Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992) *Science* **256,** 1783–1790.
21. DeStefano, J. J., Buiser, R. G., Mallaber, L. M., Myers, T. W., Bambara, R. A., and Fay, P. J. (1991b) *J. Biol. Chem.* **266,** 7423–7431.
22. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. Schatz, O., Mous, J., and Le Grice, S. F. J. (1990) *EMBO J.* **9,** 1171–1176.
24. Wöhrl, B. M., Volkmann, S. and Moelling, K. (1991) *J. Mol. Biol.* **220,** 801–818.
25. Peliska, J. A. and Benkovic, S. J. (1992) *Science* **258,** 1112–111.