High Processivity of the Reverse Transcriptase from a Non-long Terminal Repeat Retrotransposon*

Arkadiusz Bibillo and Thomas H. Eickbush‡
From the Department of Biology, University of Rochester, Rochester, New York 14627-0211

Received for publication, May 3, 2002, and in revised form, July 3, 2002
Published, JBC Papers in Press, July 5, 2002, DOI 10.1074/jbc.M204345200

R2 is a retrotransposable element that specifically inserts into the 28 S rRNA genes of arthropods. The element encodes a single protein with endonuclease activity that cleaves the 28 S gene target site and reverse transcriptase (RT) activity that uses the cleaved DNA to prime reverse transcription. Here we compare various properties of the R2 RT activity with those of the well characterized retroviral RT, avian myeloblastosis virus (AMV). In processivity assays using heterogeneous RNA templates, R2 RT can synthesize cDNA over twice the length of that synthesized by AMV RT and can synthesize cDNA over 4 times longer than AMV RT in assays with poly(rA) templates. The higher processivity of R2 RT compared with retroviral RTs is a result of the slower rate of dissociation of the enzyme from RNA templates. The elongation rates of the two enzymes are similar. Finally, a highly distinct property of the R2 RT, compared with retroviral enzymes, is its ability to displace RNA strands annealed to RNA templates during cDNA synthesis. We suggest that both the higher processivity and displacement properties of R2 RT compared with retroviral RT result from the greater affinity of the R2 protein for the RNA template upstream of its active site.

The processivity of an enzyme can be defined as the number of rounds of catalysis that can be performed before the enzyme dissociates from the substrate (1). The most processive nucleic acid polymerases are the large multisubunit polymerases involved in genomic DNA replication that can polymerize for tens of thousands of bp before dissociation. The high processivity of these complexes can be explained because they form a toroid around the DNA substrate (sliding clamp model) that prevents the complex from dissociation (2). Most other nucleic acid polymerases form a structure resembling a right hand, in which the palm contains the catalytic residues of the active site while the fingers and thumb grasp the template loosely to allow lateral movement along the length of the template (3).

Retroviral reverse transcriptases (RTs) have a hand structure (4) and exhibit low levels of processivity for a replicative DNA polymerase (5, 6). In typical in vitro assays with templates of random base composition, retroviral RTs can extend a cDNA strand for only a few hundred nucleotides before disso-

* This work was supported by National Institutes of Health Grant GM42790 (to T. H. E.). 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. Tel.: 585-275-7247; Fax: 585-275-2070; E-mail: eick@mail.rochester.edu.

1 The abbreviations used are: RT, reverse transcriptase; LTR, long terminal repeat; TPRT, target-primed reverse transcription; AMV, avian myeloblastosis virus; DTT, dithiothreitol; nt, nucleotide(s).

© 2002 by The American Society for Biochemistry and Molecular Biology, Inc.
known concentrations of bovine serum albumin using the fluoroimaging function of a Storm 860 PhosphorImager and ImageQuant. Avian myeloblastosis virus (AMV) reverse transcriptase was obtained from Promega, poly(rA)/oligo(dT)12–18 was from Amersham Biosciences, and poly(rA) was from Sigma.

Methods

Preparations of RNA Templates—All heterogeneous RNA templates were generated in vitro by run-off transcription using either T7 or T3 RNA polymerase (Fermentas Inc., Invitrogen). Templates for RNA synthesis were either restriction-digested pBSII(SK−) plasmid or PCR amplified products containing the T7 promoter. The 185-nt RNA was transcribed by T7 RNA polymerase from a template generated by PCR amplification of pBSII(SK−) (Stratagene) using primers AB.8 (GGGAGACAGCTATGACCATG) and AB.T7 (TAATACGACTCACTATAG). The 334-nt vector RNA was transcribed by T7 RNA polymerase from pBSII−SK− preincubated with PolII. The 590-nt RNA was transcribed like the 334-nt RNA except that a KpnI and BamHI fragment of the R1Dm element (positions 5020–5342) was cloned into the polylinker region of pBSII(SK−) (plasmid provided by D. Eichkub). The 1094-nt vector RNA was transcribed by T7 RNA polymerase using pBSII(SK−) preincubated with XmnI. The 117-nt RNA was transcribed from pBSII(SK−) preincubated with KpnI using T3 RNA polymerase. In vitro transcription resulted in 1–5 µg of 590-nt RNA template and 0.1 mg of poly(dI−dC), 6 mM MgCl2, 10 mM spermidine, 10 mM EDTA, 50 ng of poly(rA)/oligo(dT)12–18 template and 590-nt RNA. The 345-nt vector RNA was transcribed from T7 RNA polymerase using 125-nt plasmid DNA or gel-purified PCR fragments. 40 µm Tris-HCl (pH 7.9), 6 µm MgCl2, 10 mM DTT, 10 mM NaCl, 10 mM spermidine, 1 mM each rNTP, and 150 units of T7 or T3 RNA polymerase. Reactions were incubated at 37 °C until a pyrophosphate formation formed (~1.5 h). After synthesis, the DNA templates were removed by digestion with 10 units of Dnase I (Ambion Inc.) for 25 min at 37 °C. The products of transcription were ethanol-precipitated and separated on 5% urea-PAGE. Full-length RNA templates were excised from the gel, eluted at room temperature in 0.3 M sodium acetate, pH 5.2, 0.03% SDS for 1.5 h, extracted with phenol/chloroform, and ethanol-purified. Purified RNA transcriptions were dissolved in 50 mM NaCl to a final RNA concentration of 0.1 µg/µl.

Preparation of the Template with Annealed Primer—DNA primers AB.34 (CGCTTCCTG-GGGCGACACCTC) and oligo(dT)12–18, were labeled at their 5′ termini with T4 polynucleotide kinase (Fermentas) using the forward reaction procedure. Labeling reactions were performed in a 20-µl final volume containing 100 mM Tris-HCl (pH 7.6), 10 mM MgCl2, 5 mM dithiothreitol, 0.1 mM spermidine, 0.1 mM EDTA, 50 µM (6–10 pmol) of DNA primer, 25 µCi of [γ-32P]ATP (3,000 Ci/mmol; PerkinElmer Life Sciences), and 10 units of T4 polynucleotide kinase. The annealed primer was added to the 345-nt RNA template and 590-nt RNA template and then heated at 80 °C and cooled slowly (3.5 °C/min). Fifty fmol of annealed RNA/DNA-primer or 10 µg of poly(rA)/oligo(dT) trap was used for each primer extension reaction.

Primer Extension Reactions—Unless otherwise specified, reverse transcription reactions were performed in 30-µl volumes. In the case of R2 RT, the preincubation and elongation mixture contained 50 µM Tris-HCl (pH 7.5), 0.2 mM NaCl, 10 mM MgCl2, 2.5 mM DTT, 0.01% Triton X-100. The amount of R2 protein and RNA template for each assay are specified in the figure legends. Primer extension reactions with AMV RT were conducted in a manner similar to that for R2 RT with the exception that the solution contained 50 µM Tris-HCl (pH 8.3), 50 mM KCl, 5 mM MgCl2, 5 mM DTT, 0.5 mM spermidine, and 0.5 µg of poly(rA)/oligo(dT) trap. RT reaction mixtures were preincubated for 5 min at 37 °C, and reverse transcription was started by the addition of dNTP to a final concentration of 250 µM. Only 250 µM dTTP was added in those reactions with poly(rA) templates. Polymerization reactions were for 4–7 min, depending on the length of the template. Reverse transcription reactions conducted under processive conditions were started by the simultaneous addition of 20 µg of heparin and 2 µg of poly(rA)/poly(dT)12–18 along with the dNTP at the start of the polymerization reaction (“trap” conditions). The trap with poly(rA) templates included 20 µg of heparin and 2 µg of poly(di-dC), 6 µg of E. coli RNA, and 0.1 µg of oligo(dT)12–18. In control reactions, the trap was added at the beginning of the preincubation to block initiation. Polymerization reactions were stopped by the addition of 20 µg/ml of 8% SDS. After precipitation, the products were separated on 6% denaturing polyacrylamide gels. The Gene Ruler 100-bp DNA ladder (Fermentas) was used as DNA length markers. The markers were end-labeled with T4 polynucleotide kinase using the protocol for labeling 5′-protruding termini of DNA by the exchange reaction (Fermentas).

Determination of Dissociation Rate Constants—The RT and template-primer complex were preincubated for 15 min at 37 °C to form an initiation complex before the addition of the heparin, poly(rA)/oligo(dT) trap. After various periods of time at 37 °C, elongation was started by the addition of dNTP. The moment of the trap addition was considered as time 0. As a control for the efficiency of the trap in blocking reinitiation, the trap was also added at the beginning of the preincubation period. The yield of cDNA was assumed to be proportional to the fraction of enzyme that had remained bound to the RNA template during the incubation with the trap before the dNTP addition. Products of the extensions were separated on 7% denaturing polyacrylamide gels, and the yield of cDNA products was determined with a PhosphorImager. The elongation at the 3′ end of the product, the cDNA accumulation at each time point was subtracted from that which accumulated in the control lane and then normalized to the cDNA yield at time 0. The data were then fitted to the single exponential decay: exp(−koff × t), where koff is the dissociation rate constant and t is the time of incubation in seconds with the trap (15).

Determination of Elongation Rates—Elongation rates were based on the increase of cDNA length as a function of time. For the heteropolymeric templates, 15 pmol of end-labeled DNA primers AB.34 and AB.23 were described for the processivity assays. For assays with the poly(rA) template, 1 µg of end-labeled oligo(dT)12–18 was annealed to 1.5 µg of poly(rA) (weight ratio 1:1.5). About 150 fmol of the 1094-nt RNA/AB.34 template and 0.1 µg of poly(rA)/oligo(dT)12–18 template were preincubated with 10 ng (80 fmol) of R2 RT or 25 ng of AMV RT. After preincubiation for 5 min at 37 °C, polymerization was initiated by the addition of dNTP to a final concentration of 250 µM and stopped by the addition of ethanol/SDS. The precipitated products were separated on 6% denaturing PAGE, and the length of the cDNA was determined relative to the RNA standards. In the case of the elongation on the homopolymeric templates, the longest polymerization products were used in the determination of the elongation rate. The elongation rate constant was then determined by fitting these data points to a linear function. The average elongation rate for R2 RT was also determined as described under “Results.” Elongation by R2 RT on the homopolymeric template revealed two populations of the cDNA length at each time point. Therefore, both a fast and slow elongation rate was determined for the R2 enzyme. Because of the uniformity in length of the cDNA generated on the poly(rA) template, the main peak of cDNA lengths as well as the leading edge of the peak was determined at each time point.

RNA Structure Predictions—Secondary structures of the 590- and 1094-nt RNA templates were predicted using the program RNAstructure, available on the World Wide Web at www.chem.rochester.edu/Faculty (16). Simulations were conducted on overlapping 50-nt regions of the RNA with a window size of 3, and all structures within 10% of the maximum energy difference were observed. In this approach, only the most stable short distance interactions (hairpins) are predicted.

RESULTS

Properties of the R2 Reverse Transcriptase—R2 non-LTR retrotransposons encode a single 120-kDa protein with both DNA endonuclease and reverse transcriptase activities (11, 17). Based on sequence similarity to characterized reverse transcriptases, the R2 RT domain spans a central 450-amino acid region of the protein. This RT domain is similar in size and sequence to the RT domain of all other non-LTR retrotransposons but larger than the RT domain of retroviruses (18). The R2 protein does not contain sequences similar to RNase H domains (19), and no RNase H activity has been detected in in vitro assays (11). Located at the amino-terminal end of the R2 protein are two DNA binding motifs: a CCHH type zinc finger motif, and a c-Myb-like motif (20). This amino-terminal domain has been shown to specifically bind the DNA target site of the R2 element (20) (Fig. 1). The carboxyl-terminal domain of the R2 protein contains an endonuclease activity that is responsible for the DNA cleavage that initiates the TPRT reaction (20). The 120-kDa R2 protein exists as a monomer at low concentrations in the absence of RNA but becomes multimeric, possibly forming a dimer, in the presence of its RNA template during the DNA cleavage/TPRT reaction (20, 21).

2 S. Christensen and T. Eickbush, manuscript in preparation.
All previous studies of the DNA cleavage and TPRT reaction by the R2 protein have been conducted in 200 mM NaCl, because lower salt concentrations result in lower rates and less specific cleavage of the DNA target site (21). In a similar manner, the reverse transcriptase activity of the R2 protein, as assayed in a TPRT reaction or in a primer extension reaction, is most active near 200 mM NaCl (data not shown). Therefore, all assays in this report have been conducted in 200 mM NaCl. Retroviral enzymes, on the other hand, are typically assayed in the range of 50–75 mM KCl. In this report, 75 mM KCl has been used for all retroviral enzyme assays.

**Processivity of R2 RT**—AMV RT is one of the better characterized and most processive of the retroviral RTs (15, 22–24). Fig. 1 compares the length of cDNA products synthesized by AMV RT with that synthesized by R2 RT. The template for this reaction was a 590-nt RNA molecule with a $^{32}$P-labeled DNA oligonucleotide preannealed to its 3′-end. AMV RT generated a wide range of cDNA products, the majority of which were less than 590 nt in length (AMV RT; lane 1). R2 RT also generated a range of cDNA lengths; however, full-length (590 nt) products were clearly the most prominent (R2 RT; lane 1).

To directly confirm that the reaction products seen in Fig. 1 represented the length of cDNA synthesized before the RT dissociated from the RNA template, and not multiple rounds of elongation, the reactions were also conducted under conditions that did not allow the reinitiation of RT. These single round reactions, frequently referred to as RT “trap assays,” involved the addition of heparin and an excess of poly(rA)/oligo(dT) (25). Heparin inhibits reinitiation by reverse transcriptases, whereas any reinitiation that may still occur will be associated with the more abundant, unlabeled poly(rA)/oligo(dT) template. To demonstrate the efficiency of this trap, lane 2 for each enzyme represented the addition of heparin and the poly(rA)/oligo(dT) at the same time that the 590-nt RNA-primer complex was added to the RT. When the trap was added at the beginning of the assay, neither enzyme synthesized detectable levels of cDNA primed by the oligonucleotide annealed to the RNA template. In lane 3 for each enzyme, the RTs were first bound to the 590-nt RNA template-primer complex followed by the addition of the trap along with the dNTPs to start reverse transcription.

To quantify the difference in the accumulation of cDNA products generated by the R2 and AMV RTs, the radioactivity throughout lane 3 for each enzyme was scanned on a PhosphorImager, and the results are plotted in Fig. 2A. Shown in Fig. 2B are similar processivity assays conducted in the presence of the trap with the R2 and AMV RTs using a second RNA template, which is 1094 nt in length. Again, the cDNA products generated by R2 RT on this RNA were longer than those generated by AMV RT. However, comparison of the cDNA products generated on the 590- and 1094-nt templates indicated that both enzymes had higher processivity on the 590-nt RNA template. For example, the AMV RT generated significant cDNA products longer than 300 nt on the 590-nt RNA template but few products longer than 300 nt on the 1094-nt RNA template. In the case of R2 RT, cDNA fragments of 100, 185, 210, 230, and 265 nt accumulated from the 1094-nt RNA template, whereas few short cDNA fragments were generated on the 590-nt RNA template.

To describe the difference in processivity between the R2 and AMV RTs, we have calculated the probability of (premature) termination per nucleotide for each enzyme. The fraction of the total cDNA products generated on the 1094-nt RNA template that are greater than various specific lengths was determined from the scans in Fig. 2A and plotted in Fig. 3A. These same values were also plotted on a log scale in Fig. 3B. The curve fitted to these points can be described by an exponential function, $\exp(-P_{off} \times n)$, where $n$ represents the number of incorporated nucleotides, and $P_{off}$ is the probability of premature termination per nucleotide. $P_{off}$ for R2 RT on the 1094-nt RNA was $4.0 \times 10^{-3}/nt$, whereas $P_{off}$ for AMV RT was $9.3 \times 10^{-3}/nt$. The probability of termination for these enzymes on the 590-nt RNA template was $2.6 \times 10^{-3}/nt$ for R2 RT and $5.4 \times 10^{-3}/nt$ for AMV RT. Similar processivity assays were also conducted with the homopolymeric template, poly(rA), using end-labeled oligo(dT)$_{18}$ as primer (Fig. 3C). Both AMV and R2 RTs exhibited significantly lower termination on this homopolymeric template.
$P_{\text{off}}$ was determined to be $0.4 \times 10^{-3}/\text{nt}$ and $1.5 \times 10^{-3}/\text{nt}$ for R2 and AMV RT, respectively. The greater processivity of AMV and other retroviral RTs on poly(rA) templates has been previously reported (5, 22).

An alternative measure for the processivity of an enzyme is a determination of the probability of chain elongation (26). The probability of chain elongation can be readily calculated for R2 and AMV RT by subtracting the $P_{\text{off}}$ value we have determined from 1.0 ($1 - P_{\text{off}}$). The probabilities of chain elongation for R2 are therefore 0.9960, 0.9974, and 0.9994 on the 1094-nt, 590-nt, and poly(rA) templates. The corresponding values for AMV RT are 0.9907, 0.9946, and 0.9985.

**R2 RT Has a Low Dissociation Rate from RNA Templates**—To determine whether the higher processivity of R2 RT compared with AMV RT is a result of its lower rate of dissociation from the RNA templates or its faster rate of polymerization, we next compared the dissociation rates of R2 RT and AMV RT. A 183-nt RNA template with a $^{32}$P-labeled DNA primer annealed to its 3′-end was preincubated with each RT to allow formation of an initiation complex. To this complex was then added the heparin, poly(rA)/oligo(dT) trap for different periods of time before the addition of dNTPs to initiate polymerization. Because the trap prevented any released polymerase from reassociating with the RNA template, the proportion of cDNA products observed after dNTP addition represented the level of polymerase that remained bound to the template during incubation with the trap. The separated cDNA products are

**FIG. 2.** PhosphorImager scans of the processivity assays for AMV and R2 RTs on different RNA templates. A, processivity assay (i.e., in the presence of the trap) on a 590-nt RNA template. Scans are of lane 3 of the gel shown in Fig. 1, B, a similar processivity assay to that conducted in A but using a 1094-nt RNA template (see "Experimental Procedures"). The approximate size of the cDNA products was interpolated relative to 100-nt size standards shown below each scan. Peaks in the cDNA profiles correspond to "pause sites," locations where the enzyme dissociated from the template or was blocked from further elongation. Major peaks are labeled with capital letters (AMV RT) or lowercase letters (R2 RT). Presented below the gel scans are the nucleotide sequences of a portion of the RNA template. Predicted hairpin regions of the RNA templates are indicated by underlining (dotted lines connect the paired region of each hairpin). The locations of the major pause sites for each enzyme are indicated above each sequence.

**FIG. 3.** Graphic representation of R2 and AMV RT polymerization as a function of distance along a RNA template. $\bullet$, R2 RT data points; ◦, AMV RT data points. A, plotted are the fractions of the cDNA that were of greater length than an arbitrarily defined length of zero ($A_{\text{all}}/A_T$). $A_T$ is equal to the total cDNA of length greater than 80 nt, and $A_{\text{all}}$ is equal to the total cDNA greater than the length indicated. The data are derived from the experiment in Fig. 2B, B, the same data as in A but transformed to a logarithmic scale, $\ln(A_{\text{all}}/A_T)$, and fitted to a linear function with the slope represented by $P_{\text{off}}$ (see "Results"). C, a similar analysis to that in B but using a poly(rA) template.
shown in Fig. 4A, and the total amount of cDNA synthesized at each time point (irrespective of size) was plotted in Fig. 4B. A dramatic difference in the rate of dissociation by the two RTs was found. The level of cDNA products generated by the AMV-RT decreased ~10-fold after a 2-min preincubation with the trap, whereas even after a 45-min incubation the level of cDNA products generated by R2 RT decreased only 2-fold. Note also that the greater processivity of the R2 RT could even be observed on this short RNA template. Virtually all of the cDNA products generated by R2 RT on the 1094-nt template are truncated. The decreasing fraction of enzyme that remained associated with the template over time (Fig. 4B) was fit using a decreasing exponential function ($\exp(-k_{off} \times s)$), where $s$ represents seconds, and the dissociation rate for R2 RT was determined ($k_{off} = 3.8 \pm 2.4 \times 10^{-4}$ s$^{-1}$). Because the absence of substrate dNTP may change the properties of the protein-template complex and thus affect its stability, an additional dissociation experiment was conducted similar to that in Fig. 4 but with the addition of 5 $\mu$M dATP to the preincubation mixture (data not shown). dATP is the first nucleotide to be incorporated on the 183-nt RNA template; thus, its presence should permit complexes that more closely represent those that have initiated reverse transcription. The $k_{off}$ determined for R2 RT under these conditions ($k_{off} = 2.7 \times 10^{-4}$ s$^{-1}$) was not significantly different from that determined in the absence of dATP.

A similar determination of the dissociation rate for AMV RT can only be crudely estimated, because most of the enzyme had already dissociated at the shortest time point tested. However, this rate would be ~50-fold faster than R2 RT, or similar to values previously reported (15, 23). Somewhat lower rates of dissociation have been reported for other retroviral enzymes (5, 27), but we have not found a report of a dissociation rate as low as with R2 RT.

The Elongation Rate of R2 RT Is Similar to That of AMV RT—To determine the elongation rate of R2 RT, the enzyme was incubated with either the 590- or the 1094-nt RNA template preannealed with the end-labeled DNA primer to allow ternary complex formation. The four dNTPs were then added to start the reaction, and after short time periods the reaction was rapidly stopped by the addition of SDS and ethanol. The cDNA products generated by R2 RT on the 1094-nt template are shown in Fig. 5A (left panel). The maximum lengths of cDNA synthesized at each time point on the two templates are plotted in Fig. 5B and used to determine an elongation rate. An identical set of experiments has also been conducted with AMV RT
using the 590-nt template, and the maximum cDNA lengths for this RT are plotted in Fig. 5B. The elongation rates for R2 were found to be 14.2 nt/s and 11.0 nt/s on the 590- and 1094-nt templates, respectively, similar to the rate of 12.9 nt/s determined for AMV RT on the 590-nt template. These values represent maximum rates of elongation, because they were based on the longest cDNA products generated at each time point. The average elongation rate for each enzyme was more difficult to determine because of the accumulation of truncation products throughout the gel. However, we have estimated the average elongation rate for R2 RT on the 1094-nt RNA by determining the accumulation of cDNA greater than specific lengths as a function of time (data not shown). Based on the time required for one-half of the total cDNA to accumulate at each length, we obtained an average elongation rate for R2 RT of 6.8 nt/s. We therefore suggest that the average elongation rate for R2 RT is about 60% of the maximum elongation rate.

Elongation rates of polymerases are easier to interpret and more accurately determined on homopolymeric templates, because the simple structures of these templates give rise to more uniform rates of nucleotide polymerization by the enzyme. The lengths of cDNA generated by R2 RT on a poly(rA) template are shown in Fig. 5A (right panel). Surprisingly, we found that two different populations of cDNA fragments were produced that represented significantly different elongation rates (the “fast” and “slow” fractions). The leading edge of the “fast” cDNA products generated by R2 RT was synthesized at a rate of 28.8 ± 8.7 nt/s, or about twice the rate determined on heterogeneous RNA templates. Using the average lengths of the cDNA in this fast cDNA fraction, the R2 elongation rate was 19.5 ± 4.0 nt/s. The “slow” fraction were synthesized at a maximum rate of 12.6 ± 2.2 nt/s (longest cDNA lengths) and an average rate of 9.5 ± 1.8 nt/s (average cDNA lengths). This experiment was conducted a number of times with different R2
protein preparations. Because older R2 protein preparations produced higher percentages of the slower fraction, this slow fraction appeared to be a result of structural changes in the R2 protein, which prevented it from assuming a more active conformation. Presumably, these two rates of elongation seen with the poly(rA) templates were also occurring on heterogeneous templates but were not readily observed because of the broader range of cDNA lengths generated on complex RNAs. The elongation rate of AMV RT on the poly(A) template was determined to be 22.2 nt/s (maximum rate) (13.6 nt/s average rate). This rate is similar to that reported for human immunodeficiency virus RT on a poly(rA) template (5).

These results indicate that the elongation rates of R2 RT are 2–3 times higher on poly(rA) templates compared with heterogeneous RNA templates. These rates are similar to those determined for AMV RT, suggesting that the increased processivity of R2 RT compared with AMV RT cannot be attributed to a faster rate of elongation.

Features of the RNA Template That Promote R2 RT Dissociation—Whereas the difference in $P_{off}$ for R2 RT relative to AMV RT is best explained by the reduced rate of dissociation of the R2 RT complex, the difference in $P_{off}$ between the two RTs is only 2–4-fold, not the over 50-fold difference seen in Fig. 4. The predicted $P_{off}$ for R2 RT based on our determined rate of elongation and rate of dissociation was over 10-fold lower than our observed value for $P_{off}$. One possible explanation for why the dissociation rate of a static RT-RNA complex is a significant underestimate of its dissociation rate as the enzyme transverses the RNA template is that enzyme dissociation is highly sensitive to the structure of the RNA.

As shown in Fig. 2, both R2 and AMV RT have higher probabilities of dissociation at specific sites along the RNA template. Such sites have been referred to in the literature as “pause sites.” The locations of the AMV RT and R2 RT pause sites in the region of maximum resolution on our gels are shown in Fig. 2 below the gel profile scans of the cDNA products. Sites marked with a capital letter correspond to AMV RT pause sites, whereas those indicated with a lowercase letter are the R2 RT pause sites. We have also used RNA secondary structure algorithms to predict potential secondary structures for the RNA templates in these regions (see “Experimental Procedures”). Regions of the template that were predicted to form hairpin loops are underlined and connected by dotted lines.

Consistent with previous studies of retroviral enzymes (24, 28, 29), pause sites for AMV RT were sometimes found either before or after RNA hairpin structures. For example, AMV pause sites C–F on the 590-nt RNA template surround the three hairpin structures predicted to form by this RNA. R2 RT on the other hand exhibited only low levels of pausing at these sites (sites e and d are potentially associated with the first two hairpins). In contrast, R2 RT exhibited significant pausing near two weak AT-rich hairpin structures on the 1094-nt RNA (sites b–d), whereas AMV did not pause at these locations. Clearly, R2 and AMV RT respond differently to the secondary structure of the RNA templates. In particular, the R2 pause sites do not correlate with the predicted strength of the RNA secondary structures.

A larger proportion of the pause sites by both RTs appear to be associated with runs of A and U nucleotides. Indeed, both enzymes reverse transcribed the RNA template with less stable secondary structures (the 1094-nt RNA) more poorly than the RNA template with the greater secondary structure. This finding is consistent with previous reports demonstrating that pause sites for different retroviral RTs are most frequently associated with homopolymeric runs of A and U (24, 28, 29).

Unfortunately, the features of the RNA sequence that give rise to the strongest pauses by AMV RT and R2 RT are not easy to identify in these experiments. For example, whereas all of the major AMV RT pause sites on 1094-nt templates were associated with runs of A and U from 4 to 6 nucleotides in length, there was no pausing at a run of five U nucleotides around position 270. Meanwhile, R2 RT pause sites were associated with short runs of A and U, not with the longer homopolymeric runs where the AMV RT had dramatic pause sites.

These findings suggest that R2 RT is not highly sensitive to the structure of the RNA template. Similar to retroviral RT, dissociation by R2 RT is most likely to occur at regions of the template containing short runs of A and U residues. The specific features of these RNA sequences that result in R2 RT dissociation are complex and thus likely to involve higher order structures of the RNA template.

The Effects of Temperature on the Reverse Transcription Reaction—To further study the factors affecting the dissociation of R2 RT from the RNA template, the ability of the R2 and AMV RTs to reverse transcribe the 590-nt RNA template were compared at temperatures ranging from 25 to 45 °C (Fig. 6). The reactions were conducted under the conditions of a heparin, poly(rA)/oligo(dT) trap to permit the synthesis of cDNA derived from only a single round of enzyme association. The cDNA produced at each temperature was separated on a polyacrylamide gel (Fig. 6A), the amount of cDNA larger than specific lengths was determined on a PhosphorImager, and the probability of termination ($P_{off}$) for each enzyme was calculated at each temperature (Fig. 6B).

In the case of the AMV-RT, $P_{off}$ gradually decreased when the temperature of the reaction was raised from 25 to 45 °C, consistent with previous reports that the processivity of retroviral RTs increases with temperature (30). Meanwhile, temperature had minimal effect on the $P_{off}$ for R2 RT over the range from 25 to 35 °C, at which point further temperature increases caused an increase in termination. This loss of processivity at higher temperatures was presumably a result of partial denaturation of the R2 protein, which has evolved to function at the ambient temperatures of most insects (20–30 °C). This contrast with retroviral RTs, which have evolved to function at the higher temperatures associated with their vertebrate hosts.

It is interesting to note that the general pattern of pause sites for R2 and AMV RT did not show substantial differences with temperature. Because the strength of the RNA secondary structure will decrease with increasing temperature, this result is consistent with the findings that it is usually primary sequence rather than the secondary structure of the RNA that determines these pause sites.

Displacement Reactions of the R2 RT on RNA Templates—Because experiments such as those in Fig. 2 with naturally occurring RNA templates suggested that R2 RT may be more efficient at polymerizing through secondary structures of the RNA template, we also compared the ability of R2 RT and AMV RT to transcribe RNA templates in which longer hairpin loops had been engineered into the template. These experiments generally suggested that R2 RT was better able than AMV RT to polymerize through highly stable hairpin structures; however, the efficiency with which R2 RT could polymerize through a hairpin was dependent upon the sequence of the hairpin (data not shown). Because RNA templates with long hairpin sequences were difficult to generate, we also tested RNA templates in which we simply annealed two complementary RNA molecules. Whereas these duplex RNA templates would not mimic the possible reformation of a hairpin following passage of the polymerase, they would mimic the requirements of a polymerase to initially denature a RNA hairpin. These studies
revealed that R2 RT has the unusual ability to displace RNA molecules annealed to an RNA template.

Shown in Fig. 7 is one such experiment involving a 334-nt RNA template to which has been annealed a 117-nt RNA strand complementary to the 5'-end of the template (see the drawing at the bottom of Fig. 7). Reverse transcription was initiated from an end-labeled DNA primer annealed to a location near the middle of the 334-nt template. R2 RT and AMV RT were both tested in the presence (lane 2) and the absence (lane 1) of the annealed 117-nt RNA under conditions that allow only a single round reaction to occur (presence of the trap) and under conditions of multiple rounds of initiations (absence of the trap). In the case of the AMV RT, the presence of 117-nt RNA annealed to the template completely prevented cDNA synthesis more than a few nucleotides past the begin-

**Fig. 6. Effects of temperature on the processivity of R2 and AMV RT.** A, the RNA template (590-nt vector RNA) and reverse transcription assays are like those conducted in Fig. 1 for a processive reaction (i.e., with the trap), except that preincubations were for 5 min at 35°C followed by a 2-min equilibration at the new temperature before the addition of dNTPs to start elongation. Elongation reactions were conducted for 4 min at the temperature shown at the top of each lane. B, effects of temperature on the probability of premature termination, \( P_{off} \). The quantitation was based on scans of the gel in A. \( \bigcirc \), \( P_{off} \) values for R2 RT; \( \bigcirc \), \( P_{off} \) values for AMV RT.

**Fig. 7. R2 can displace an RNA strand annealed to a RNA template.** A, the RNA template for this reaction was a 334-nt RNA with an end-labeled oligonucleotide primer annealed near the middle of the RNA. In the other reaction, this same template contained a 117-nt-long RNA strand annealed to the 5'-end of the template (see "Experimental Procedures" for a description of these components). Lane 1, 300 fmol of the annealed RNA-primer complex was preincubated with either 2 ng (20 fmol) of R2 or 5 ng of AMV RT for 5 min at 37°C. After preincubation, reverse transcription was started by the addition of dNTP and stopped after 5 min at 37°C. Lane 2, reactions were identical to that in lane 1 except that 300 fmol of the RNA-primer complex with the annealed RNA block was used in the reaction. In the processive runs (left panel) 2.5 µl of trap was added along with the dNTP at the start of the reaction. B, diagrams describing the reaction assayed in A. Gray lines, RNA strands; black lines, end-labeled DNA primers; dotted lines, synthesized cDNA.
ning of the RNA duplex region. In contrast, R2 RT was able to synthesize full-length cDNA products irrespective of whether the RNA template contained the duplex region. Indeed, there was not even a significant "pause site" generated by the annealed RNA, and pause sites typically seen within the 334-nt RNA template in the absence of the annealed RNA were reduced in the presence of the annealed RNA. Further characterization of the ability of R2 RT to displace annealed RNA and DNA strands associated with RNA and DNA templates is currently in progress.

**DISCUSSION**

The experiments in this report demonstrate that the reverse transcriptase from the R2 non-LTR retrotransposon has greater processivity than one of the most processive of the retroviral reverse transcriptases, AMV RT. CDNA products generated by R2 RT processivity were over 2-fold longer on heterogeneous RNA templates and nearly 4-fold longer on a poly(rA) template. R2 RT was not significantly affected by weak secondary structures associated with natural RNA templates. Instead, like retroviral RTs, R2 enzyme pausing (dissociation) was most frequently induced by short homopolymeric runs of A or U residues. Consistent with this finding, R2 RT was more processive at lower temperatures (25°C) than at higher temperatures (45°C).

Despite their low processivity, synthesis of full-length reverse transcripts by retroviral RTs are still efficient because polymerization occurs within viral particles (8). This small compartment effectively increases the concentration of substrates, promoting rapid reassocation of the enzyme with the template. Unlike retroviral RT, high processivity of a non-LTR retrotransposon’s RT may be required for retrotransposition. In the TPRT model of non-LTR retrotransposition, reverse transcription occurs free in the nucleus primed by a cleaved chromosomal DNA strand (11). Final integration of the element does not depend upon sequences at the 5’ end of the element; thus, if the enzyme dissociates from the RNA template before reaching the 5’-end of the transcript, integration of a truncated copy will result. Because 5’-truncated copies cannot contribute to future generations of active elements, there should be selective pressure on the non-LTR retrotransposon to evolve reverse transcriptases with high processivity. The frequency with which 5’ truncation occurs for different non-LTR elements differs widely. For example, only about one-third of all copies of the highly abundant L1 elements in the human genome are full-length insertions (31), and only one-half of the R2 elements are full-length in many species of Drosophila (32, 33). On the other hand, non-LTR elements like the R2 element of B. mori appear highly efficient at generating full-length copies, since 5’ truncation has never been observed (17, 18). It will be interesting to determine whether the frequency of 5’ truncations for a non-LTR retrotransposon can be directly correlated with the processivity of its encoded RT.

The greater processivity of the R2 enzyme appears to be due to a lower dissociation rate from the RNA template compared with that of AMV RT, not a higher rate of polymerization. However, the 2-4-fold difference in cDNA lengths is not as great as the 50-fold difference determined for the dissociation rates of these two enzymes (Fig. 4). We suggest that this discrepancy results from R2 being a multifunctional enzyme with highly specific RNA and DNA binding capabilities in addition to its RT activity. Indeed, in the endogenous reverse transcription catalyzed by R2 RT, cDNA synthesis is primed by a DNA target site, and only RNA transcripts containing the 3’-untranslated region of the R2 element are used as templates (11, 14, 34, 35). We suggest that when the R2 protein is stationary on the RNA template, other domains of the protein in addition to the active site of the RT can participate in binding, giving rise to high stability. However, when the enzyme is actively transversing the RNA template during polymerization, binding by these additional domains is disrupted, significantly raising the rate of dissociation.

What is the structural basis for the greater processivity of R2 RT? Many of the most processive nucleic acid polymerases form multivesicular toroids that completely encircle the template (2, 36). However, such toroids are not required for a polymerase to be highly processive, since the phi29 DNA polymerase is a single subunit enzyme that can synthesize DNA strands over 70 kb in length (37). Based on the extensive sequence similarity between the reverse transcriptases of non-LTR retrotransposons and that of retroviruses, we have conducted homology modeling predictions of the R2 RT structure (18, 38). These predictions suggest that the R2 enzyme can assume right-handed structures similar to the structures formed by retroviral RTs (4, 39). However, the size of the RT domain encoded by non-LTR retrotransposons is larger than that of retroviruses. For example, in the fingers subdomain, there are two additional segments (referred to as segments A and B in Ref. 18), which suggests that the R2 enzyme contains either additional or longer fingers. Whereas sequence similarities within the thumb subdomain are too limited to allow direct modeling between retroviral and non-LTR element proteins, a comparison between the RT domains of many non-LTR retrotransposons reveals a conserved region downstream of the palm subdomain that is nearly twice the length of the thumb domain of retroviruses (19). We suggest that the greater processivity of R2 RT results from its more extensive fingers and thumb subdomains that allow the enzyme to more completely wrap around the RNA template. This greater area of contact between the R2 RT and the RNA substrate allows a more stable interaction of the enzyme with the template (slower dissociation) but does not limit the flexibility of the enzyme for linear movement along the template. The elongation rate of the R2 enzyme is similar to that of AMV RT (Fig. 5).

A second property of R2 RT identified in this report is its ability to displace RNA strands annealed to the RNA template. R2 RT can readily displace a 117-nt RNA strand annealed to the RNA template, whereas reverse transcription by AMV RT was completely blocked by this annealed RNA strand (Fig. 7). It is not known if this displacement by R2 RT is simply a passive process, in which thermal breathing of the duplex in front of the enzyme allows cycles of nucleotide addition by the RT or whether the R2 enzyme itself contributes to the process of strand displacement. In either event, this displacement property is likely to be associated with the manner in which R2 RT binds an RNA template. We have previously postulated that the larger finger subdomains of the R2 RT enable more extensive binding of the enzyme to the RNA template upstream of the active site (14). This upstream binding is greatest with the RNA corresponding to the 3’-untranslated region of the R2 element; however, the free 3’-end of any RNA can be bound by the enzyme. The unusual ability of R2 RT to bind the 3’-end of an RNA molecule upstream of the active site gives rise to its remarkable ability to jump onto the 3’-end of a second RNA template after completing synthesis of a first template (14). This postulated ability of R2 RT to bind single-stranded RNA upstream of the active site may also be able to drive the helix-coil equilibrium of a duplex region upstream of the active site toward the melted state and thus contribute to the RNA displacement reaction described in this report. Clearly, more experiments on the rate of displacement and the influences of temperature and ionic strength are needed to demonstrate whether R2 RT is actively involved in strand displacement.
Acknowledgments—We thank Bill Burke, Shawn Christensen, and Junqiang Ye for assistance and helpful discussions. We thank Danna Eickbush for comments on the paper.

REFERENCES
1. Hippel, P. H., Fairfield, F. R., and Dalejsi, M. K. (1994) Ann. N. Y. Acad. Sci. 726, 118–131
2. Kelman, Z., Hurwitz, J., and O'Donnell, M. (1998) Structure 6, 121–125
3. Joyce, G. M., and Steitz, T. A. (1995) J. Bacteriol. 177, 6321–6329
4. Kohli斯塔d, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992) Science 256, 1783–1790
5. Huber, H. E., McCoy, J. M., Seehra, J. S., and Richardson, C. C. (1989) J. Biol. Chem. 264, 4669–4675
6. Katz, R. A., and Skalka, A. M. (1994) Annu. Rev. Biochem. 63, 133–173
7. Katz, R. A., and Skalka, A. M. (1990) Annu. Rev. Genet. 24, 409–445
8. Whitcomb, J. M., and Hughes, S. H. (1992) Annu. Rev. Cell Biol. 8, 375–306
9. Boeke, J. D., and Stoye, J. P. (1997) in Retroviruses (Coffin, J. M., Hughes, S. H., and Varmus, H. E., eds) pp. 343–435, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
10. Malik, H. S., and Eickbush, T. H. (2001) Genome Res. 11, 1187–1197
11. Luan, D. D., Korman, M. H., Jakubezak, J. L., and Eickbush, T. H. (1993) Cell 72, 585–605
12. Eickbush, T. H., and Malik, H. S. (2002) in Mobile DNA, 2nd Ed. (Craig, N. L., Craigie, R., Gellert, M., and Lambowitz, A. M., eds) pp. 1111–1144, American Society for Microbiology Press, Washington, D. C.
13. Eickbush, T. H. (2002) in Mobile DNA, 2nd Ed. (Craig, N. L., Craigie, R., Gellert, M., and Lambowitz, A. M., eds) pp. 813–835, American Society for Microbiology Press, Washington, D. C.
14. Bihilo, A., and Eickbush, T. H. (2002) J. Mol. Biol. 316, 459–473
15. Yu, H., and Goodman, M. F. (1992) J. Biol. Chem. 267, 10888–10896
16. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) J. Mol. Biol. 21, 911–940
17. Burke, W. D., Calalang, C. C., and Eickbush, T. H. (1987) Mol. Cell. Biol. 7, 2221–2230
18. Burke, W. D., Malik, H. S., Jones, J. P., and Eickbush, T. H. (1999) Mol. Biol. Evol. 16, 502–511
19. Malik, H. S., Burke, W. D., and Eickbush, T. H. (1999) Mol. Biol. Evol. 16, 793–805
20. Yang, J., Malik, H. S., and Eickbush, T. H. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 7847–7852
21. Yang, J., and Eickbush, T. H. (1998) Mol. Biol. Evol. 15, 3455–3465
22. Gregerson, D. S., Albert, J., and Reid, T. W. (1980) Biochemistry 19, 301–306
23. DeStefano, J. J., Buiser, R. G., Mallaber, L. M., Myers, T. W., Bambara, R. A., and Fay, P. J. (1991) J. Biol. Chem. 266, 7425–7431
24. Harrison, G. P., Mayo, M. S., Hunter, E., and Lever, A. M. L. (1998) Nucleic Acids Res. 26, 3433–3442
25. Kim, B., Ayran, J. C., Sagar, S. G., Adman, E. T., Fuller, S. M., Tran, N. H., and Horrigan, J. (1999) J. Biol. Chem. 274, 27666–27673
26. McClure, W. R., and Chow, Y. (1980) Methods Enzymol. 64, 277–297
27. Beard, W. A., Minnick, D. T., Wade, C. L., Prasad, R., Won, R. L., Kumar, A., Kunkel, T. A., and Wilson, S. H. (1996) J. Biol. Chem. 271, 12213–12220
28. Klarmann, G. J., Schaubler, C. A., and Preston, B. D. (1993) J. Biol. Chem. 268, 9793–9802
29. Sue, Z., and Johnson, K. A. (1997) Biochemistry 36, 12459–12467
30. Malboeuf, C. M., Isaacs, S. J., Tran, N. H., and Kim, B. (2001) BioTechniques 30, 1074–1084
31. Kazarzian, H. H., Jr., and Moran, J. V. (1996) Nat. Genet. 19, 19–24
32. Jakubezak, J. L., Zenni, M. K., Woodruff, R. C., and Eickbush, T. H. (1992) Genetics 131, 129–142
33. George, J. A., Burke, W. D., and Eickbush, T. H. (1996) Genetics 142, 853–863
34. Luan, D. D., and Eickbush, T. H. (1995) Mol. Cell. Biol. 15, 3882–3891
35. Luan, D. D., and Eickbush, T. H. (1996) Mol. Biol. Cell. 19, 4726–4734
36. Breyer, W. A., and Matthews, B. W. (2001) Protein Sci. 10, 1699–1711
37. Blainvo, L., Bernard, A., Lazaro, J. M., Martin, G., Garmentia, C., and Salas, M. (1989) J. Biol. Chem. 264, 8935–8940
38. Xiong, Y., and Eickbush, T. H. (1990) EMBO J. 9, 3351–3362
39. Sarafianos, S. G., Das, K., Tartillo, C., Clark, A. D., Jr., Ding, J., Whitcomb, J. M., Boyer, P. L., Hughes, S. H., and Arnold, E. (2001) EMBO J. 20, 1449–1461