Nucleic Acids Research, 1995, Vol. 23, No. 19 3901–3908

The orientation of binding of human immunodeficiency virus reverse transcriptase on nucleic acid hybrids

Jeffrey J. DeStefano

Department of Microbiology, Building 231, University of Maryland, College Park, MD 20742, USA

Received June 27, 1995; Revised and Accepted August 22, 1995

ABSTRACT

The binding of HIV reverse transcriptase (RT) to heteroduplexes was examined using a substrate consisting of a 42 nt chimeric nucleic acid composed (5′→3′) of 23 nt of RNA and 19 of DNA. This chimera was hybridized to an internal region of a relatively long complementary DNA or RNA. When the chimera was bound to DNA and conditions limiting cleavage to a single binding event between the enzyme and substrate were employed initial RNase H-directed cleavages occurred 19–21 nt from the chimera 5′-terminus. A 42 nt strand identical in sequence to the chimera and composed of only RNA was cleaved at the same locations. Reducing the length of the DNA portion of the chimera from 19 to 7 nt did not alter the cleavage positions, suggesting that cleavage was not coordinated by the DNA 3′-terminus. Under the same conditions cleavage was not detected when the chimera was bound to RNA. In contrast, addition of dNTPs to the DNA 3′-terminus of the chimera occurred only when the chimera was bound to RNA. The results support preferable binding of RT to RNA–DNA versus DNA–DNA hybrid regions and a model in which the orientation of binding to heteroduplexes is 5′→3′ (relative to the RNA strand), polymerase to RNase H active site, with sites associated with the DNA and RNA strand respectively.

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 reverse transcriptase (RT) (for a review see 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. An RT-associated, double-stranded, RNA-dependent RNase activity, termed RNase H* or RNase D, 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 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).

A number of reports addressing the spacial arrangement of the DNA polymerase and RNase H active sites of HIV-RT have been published (15–19). These indicate that the sites are arranged such that they contact polymer substrates ~17 nt apart. The actual estimates varied from 15 to 20 nt and the crystal structure of HIV-RT (20) predicts a separation of ~20 nt of duplex RNA–DNA hybrid. This work suggests that positioning of the RNase H active site on the RNA–DNA hybrid is coordinated by binding of the polymerase active site to the 3′-terminal nucleotide of a recessed DNA primer. In agreement with this explanation others (16,17) 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 primer 3′-terminus. 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 were also observed (16). Both polymerase-dependent and -independent cleavages were influenced to some degree by sequence preferences (16).

In the experiments cited above the substrates were relatively short segments of DNA hybridized to longer segments of RNA 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. Structures in which short RNA segments were hybridized to longer DNA have also been examined (21). With these substrates HIV-RT RNase H-mediated cleavage was 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, i.e. cleavage occurred at a fixed distance from the 5′-end of the RNA. These results support a model in which RT binds to the substrate with the polymerase active site proximal to the 5′-end of the RNA, but associated with the DNA strand (see Fig. 6 and 21). The RNase H active site would be nearer the RNA 3′-end and associated with the RNA strand. However, these experiments did not rule out a second model in which the most stable orientation of RT on the above heteroduplexes has the RT polymerase domain associated with the recessed RNA 3′-terminus. Since the orientation of RT in the above reports was assessed by RNase H or polymerase activity, if dNTP addition to RNA were inefficient extension may not have been detected even if the RT were properly oriented. Also, the RNase H domain of RT may not be properly oriented to mediate cleavage of the RNA strand when the polymerase domain is associated with the RNA 3′-terminus. The cleavage events observed could possibly have been catalyzed by a small proportion of the enzymes which bound

in an orientation allowing cleavage or by enzymes which originally bound at the RNA 3′-terminus and 'flipped' orientations.

In the current work substrates consisting of chimeric nucleic acid strands hybridized to relatively long complementary DNA or RNA were used to further evaluate the models. Results supported the previously suggested model (21) and confirmed the preference of RT binding to RNA–DNA versus DNA–DNA duplexes.

MATERIALS AND METHODS

Materials

Recombinant HIV-RT, having native primary structure, was graciously provided by the Genetics Institute (Cambridge, MA). This enzyme had a specific activity of ~40 000 U/mg. One unit of RT is defined as the amount required to incorporate 1 nmol 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. Superscript (RNase H minus reverse transcriptase) was from Betheseda Research Laboratories (BRL). T4 polynucleotide kinase was obtained from United States Biochemical Corp. T3 and T7 RNA polymerases, RNase T_1, placental RNase inhibitor (RNasin), rNTPs, dNTPs and all restriction enzymes were obtained from Boehringer Mannheim Biochemicals. Oligonucleotides were synthesized by Genosys Inc. Radiolabeled compounds were from New England Nuclear.

Methods

Standard cleavage and extension assays. HIV-RT (10 U) was preincubated with 4 nM substrate (see below) for 5 min in 10 µl 50 mM Tris–HCl, pH 8.0, 1 mM dithiothreitol, 0.1 mM EDTA and 5 mM KCl (buffer A). RNase H cleavage assays were initiated by addition of $MgCl_2$ (6 mM final concentration) in 2.5 µl buffer A, while polymerase extension assays were initiated with $MgCl_2$ and [α-^{32}P]dCTP (1 µM final concentration, 800 Ci/mmol) in buffer A. Reactions were run for 10 min at 37°C and were terminated by the addition of 1 vol 2× concentrated gel electrophoresis loading buffer [90% formamide (v/v), 20 mM EDTA, pH 8, 0.1% (w/v) xylene cyanol and bromophenol blue].

Assays with the heparin trap. In the trap assay reaction conditions were the same as for the standard assay except that the amount of HIV-RT used was decreased to 2 U (unless otherwise indicated) and heparin (1 µg/µl) was included along with the divalent cation at the start of the reaction. This modification limits the activity to preformed complexes of RT and hybrid substrate (22). The trapped reactions were terminated after 5 s, unless otherwise indicated.

RNA–DNA hybridization. Hybrids were prepared by mixing chimeras (see below) and RNA or DNA at an ~1:2 ratio of 3′-termini in buffer A. The mixture was heated to 65°C for 5 min and then cooled slowly to room temperature.

Quantitation of nucleic acids. Since the concentration of chimeric nucleic acid was too low for accurate spectrophotometric quantitation, the concentration of chimeras was determined by a native gel hybridization 'shift-up' assay (21). In this assay the chimera was hybridized to a complementary DNA nucleic acid of known concentration. Hybridizations were performed as described above. The hybridized samples were mixed with 6× 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 (23). A fixed level of 5′-end-labeled chimera was hybridized to various amounts of unlabeled DNA or a fixed level of 5′-end-labeled DNA was hybridized to various amounts of unlabeled chimera. The concentration of chimera was evaluated based on the amount of unlabeled nucleic acid required to 'shift-up' ~50% of the labeled nucleic acid (assuming 100% hybridization efficiency). The amounts of all other nucleic acids were determined from absorbance using a spectrophotometer.

Run-off transcript. Run-off transcription was carried out as described in the Promega Protocols and Applications Guide (24). Plasmid pBSM13+ was cleaved with various restriction enzymes (*Sma*I for the RNA portions of chimeras 23R–7D and 23R–19D, *Sal*I for the RNA portions of chimera 31R–15D, *Hae*III for 97R and *Mva*I for 111R; see Fig. 1). Either T7 (*Sma*I- and *Mva*I-cleaved plasmids) or T3 (*Sal*I- and *Hae*III-cleaved plasmids) RNA polymerase was used to prepare run-off RNA transcripts of various lengths, as indicated in Figure 1. The run-off transcripts of 97R and 111R (97 and 189 nt respectively; see below) were purified by electrophoresis on 8% polyacrylamide gels 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 and 0.1% sodium dodecylsulfate. The eluate was separated from the polyacrylamide by centrifugation in a microfuge and subsequent filtration through a 0.45 × 25 mm disposable syringe filter (Nalgene). The filtrate was then ethanol-precipitated with 3 vol ethanol.

RNA 111R was generated from the isolated 189 nt run-off transcript by hybridizing a DNA oligonucleotide complementary to nt 102–121 from the 5′-end of the transcript. After hybridization (described above) the mixture was supplemented with 10 mM $MgCl_2$ and 7.5 U *Escherichia coli* RNase H was added (50 µl final volume). The mixture was incubated at 37°C for 1 h and then extracted with phenol:chloroform:isoamyl alcohol (25:24:1) and precipitated with ethanol. Products were gel-purified as described above. RNA cleavage products derived from the 5′-end of the 189 nt transcript were ~111 nt in length and were isolated from the gel and processed as described above.

RNA transcripts for chimera preparation (23R and 31R) were 5′-end-labeled before gel purification. The transcript reactions was 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′ phosphates and subsequently the alkaline phosphatase was heat inactivated and removed by phenol extraction. The dephosphorylated RNA was then 5′-end-labeled with [γ-^{32}P]ATP (~3000 Ci/mmol) and T4 polynucleotide kinase as described in the USB Molecular Biology Reagents/Protocols Manual (25). The transcript reaction was run over a second Nensorb column and then purified by gel electrophoresis as described above.

Preparation of chimeric nucleic acids. RNA–DNA chimeras were prepared by template-directed extension of RNA transcripts with dNTPs. RNA 23R was hybridized in buffer A, supplemented with 6 mM $MgCl_2$, to a 55 nt DNA oligonucleotide of sequence 5′-TCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAA-

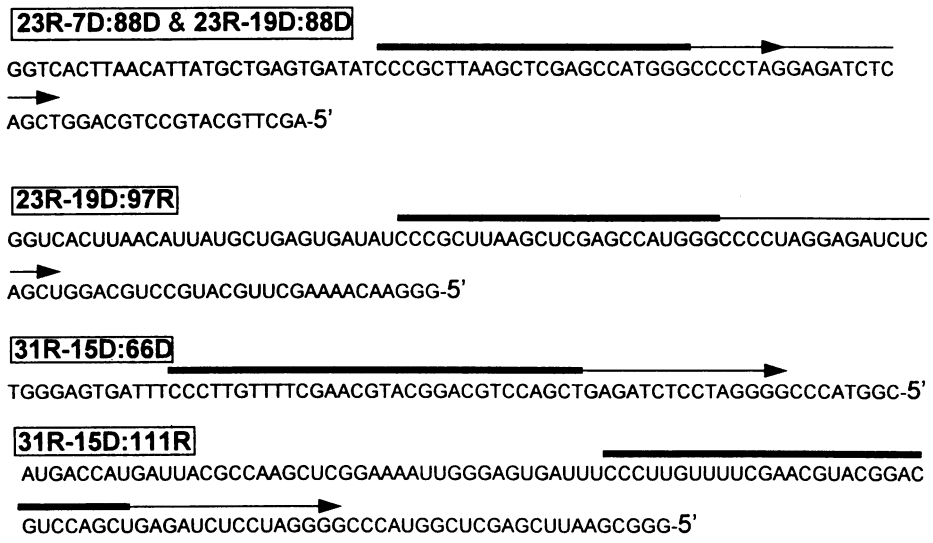

Figure 1. Configuration of substrates. The configuration of the experimental substrates is shown. The name of each substrate is given at the upper left. Each substrate consisted of two nucleic acid strands. The shorter strand is listed first, followed by a colon, then the longer strand. Chimeric strands consisting of RNA and DNA are indicated by a rule. The length of each homogeneous strand or the RNA and DNA portions of the chimeras is included in the name of the substrate. This is followed by the designation R or D for RNA or DNA respectively. The shorter strands are represented as thick (RNA) or thin (DNA) lines drawn in the 5′→3′ direction (left to right) with arrowheads at the 3′-terminus. The sequence of the longer strand is shown with the short strand positioned above the region of complementarity between the strands. Two different substrates with different lengths of DNA on the chimeric strand (as indicated by arrowheads) are shown in the uppermost illustration.

TTCGCCCTATAGTGAGTCGT-3′. The 5′-terminal nucleotide of 23R was complementary to nt 14 from the 3′-end of the oligonucleotide. The RNA was extended to produce 23R–19D by adding 400 U Superscript RT and dNTPs to a final concentration of 500 μM in a total volume of 20 μl. The mixture was incubated for 30 min at 37°C and the chimera gel-purified as described above. Since the oligonucleotide was several nucleotides longer than the chimera these were easily resolved on the gel. Chimera 31R–15D was made using the same strategy but an oligonucleotide of sequence 5′-GGGGATCCTCTAGAGTCGACCTGCAGGC-ATGCAAGCTTTTGTTCCCTTTAGTGAGGGT-3′ was used. In this case the 5′-terminal nucleotide of 31R was complementary to nt 13 from the 3′-end of the oligonucleotide.

Gel electrophoresis. Denaturing and non-denaturing polyacrylamide gels (19:1 and 29:1 acrylamide:bis-acrylamide respectively) were prepared and subjected to electrophoresis as described (23).

Preparation of the RNA G-ladder and the base hydrolysis ladder. The ladders were prepared using 5′-end-labeled RNA which were homologous to the complete chimeric strand (see Fig. 1). The RNA G-ladder was prepared by limited digestion of 5′-end-labeled RNA with RNase T_1. Reactions (10 μl) contained: 7 M Urea, 50 mM sodium citrate, pH 5, 1 mM EDTA, 2 μg *E.coli.* tRNA, ~0.05 pmol 5′-labeled RNA and 2 U RNase T_1. The reaction was performed at 56°C for 15 min 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 (~0.05 pmol) in 0.1 N NaOH at 70°C for 1.5 min. The reaction was then neutralized with HCl. An equal volume of gel loading buffer was added to the sample before loading.

RESULTS

Construction of substrates for testing HIV-RT binding orientation

The substrates used in these experiments are shown in Figure 1. Each substrate consisted of a chimeric RNA–DNA strand hybridized to a complementary region of a longer RNA or DNA heteropolymer. The RNA portions of the chimeric strands were derived from the T7 (23R–7D and 23R–19D) or T3 (31R–15D) RNA polymerase promoters of pBSM13+. The DNA portion of the chimera was added using Superscript RT, as described under Materials and Methods. The longer RNA or DNA strands were produced as described under Materials and Methods. Most of the experiments were performed using substrates 23R–19D:88D or 23R–19D:97R. Note that these substrates possess a diversity of potential binding sites for RT. Each substrate has a region of RNA–DNA heteroduplex, while 23R–19D:97R also has an RNA–RNA duplex region. The 3′ hydroxyl of the DNA strand is recessed on either DNA (23R–19D:88D) or RNA (23R–19D:97R) and the 5′ phosphate of the RNA is recessed on DNA (23R–19D-88D) or RNA (23R–19D:97R).

Determination of the positioning of RT on the substrate was based on cleavage of the RNA portion or extension of the DNA portion of the chimera. These were assessed in separate reactions performed in the presence of heparin (1 μg/μl). Heparin effectively sequesters HIV-RT molecules which have dissociated from the nucleic acid substrate, but does not influence the binding or enzymatic activity of the enzyme on the substrate (22). In these experiments RT was prebound to the substrate in the absence of divalent cation and dNTPs for 5 min. During this time RT

molecules will bind based on their relative affinity for the different regions of the substrate. Reactions were initiated by adding Mg^{2+} in assays designed to assess the cleavage of 5′-^{32}P-labeled RNA–DNA chimeras or Mg^{2+} and the appropriate (see Materials and Methods) radiolabeled dNTP when extension of unlabeled chimeras was assessed. Heparin was included with these additions to 'trap' prebound RT molecules which dissociated from the substrate. This approach permitted analysis of enzymatic events occurring during a single binding event between the RT and substrate. Such events should depend on the position and orientation of RT on the substrate, provided that the rate constant for dissociation of RT from the substrate is less than the rate constant for catalysis, i.e. the enzyme must remain associated with the substrate long enough for polymerization or cleavage to occur. This would likely be the case, since RT dissociation rate constants, although varying depending on the substrate structure and sequence, are generally one to three orders of magnitude smaller than rate constants for the RNase H (26) and polymerase (26,27) catalysis. Note that the above substrates have a recessed DNA 3′-terminus. Deoxynucleotide addition by RT to such a terminus would be expected to occur efficiently if the enzyme were properly oriented on the substrate. This is as opposed to a recessed RNA 3′-terminus, where the efficiency of extension may be lower (28).

The presence of 7 or 19 nt of DNA on 23R–7D and 23R–19D respectively did not affect initial cleavage of the RNA portion when the chimera was bound to DNA

In a previous report (21) it was shown that the recessed 3′-terminus of an RNA hybridized to a longer DNA did not coordinate RT-mediated cleavage of the RNA strand. In contrast, the recessed 3′-terminus of a DNA bound to RNA coordinates cleavage of the RNA (see Introduction).

Substrates 23R–7D:88D and 23R–19D:88D (Fig. 1) were used to determine if cleavage of the RNA portion of the chimeric strand was influenced by the position of the recessed DNA 3′-terminus. These substrates were identical, except that 23R–19D:88D had an additional 12 nt of DNA–DNA duplex. Therefore, the recessed DNA 3′-terminus of this substrate was 12 nt further downstream than that of 23R–7D:88D. Reverse transcriptase-mediated cleavage of 23R–7D:88D is shown in Figure 2. In these experiments the 5′-terminal nucleotide of the chimeric strand was labeled. Thus degradation products are RNAs derived from the 5′-end of the chimera. In the presence of the heparin trap (lanes 3–8) a major cleavage product of ~19 nt was present after 5 s (lane 3). The level of this initial product decreased as it was further cleaved, producing a smaller product of ~8 nt later in the reaction. There was also a small amount of a product ~20 nt in length at the earliest time point. Note that in the absence of heparin nearly all of the substrate was cleaved (lane 2). In a previous report (21) an analogous substrate in which the chimeric strand of 23R–7D:88D was substituted by a 30 nt homologous RNA strand was used as substrate. The cleavage products generated with this substrate were identical to those observed with 23R–7D:88D. The previous results, taken together with the current experiment, demonstrate that the presence of 7 nt of DNA on the chimeric strand did not affect the orientation of binding of RT to the substrate.

Cleavage of substrate 23R–19D:88D is shown in Figure 3 (lanes 1–3 and 6–9). In this experiment samples were incubated

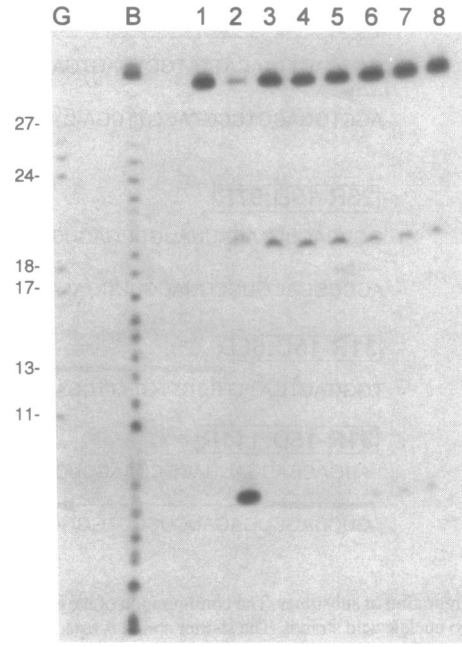

Figure 2. HIV-RT-mediated cleavage of substrate 23R–7D:88D. An autoradiogram of an experiment in which the chimeric portion of the substrate was 5′-end-labeled with ^{32}P is shown. Samples were digested with RT (see Materials and Methods) in the presence (lanes 3–8) or absence (lane 2) of the heparin trap for 5 (lane 3) or 10 s (lane 4) or 1 (lane 5), 3 (lane 6), 5 (lane 7), 7 (lane 8) or 10 min (lane 2). Lane 1 shows undigested substrate. Lanes labeled G and B were prepared by limited digestion with T1 RNase (G-ladder) or by base hydrolysis (B-ladder) as described under Materials and Methods. The numbers to the left designate the length of the ladder RNA (in nucleotides) as determined from the positions of some of the guanosine nucleotides in the RNA.

with RT for 10 min in the absence of heparin (lane 2) or for 5 s with heparin (lanes 3 and 6–9). Cleavage on this substrate occurred between nt 19 and 20, resulting in a major initial cleavage product of 19 nt. The same product was generated upon cleavage of 23R–7D:88D (Fig. 2). With both substrates cleavage was at a fixed distance from the 5′-end of the RNA. The position of the DNA terminus did not influence the position of RNase H-mediated cleavages.

Also shown in Figure 3 is cleavage of 23R–19D:88D in the presence of heparin and various amounts of RT (lanes 3 and 6–9). The position of the initial cleavage event was the same with all amounts of enzyme tested and even when 4 U RT was used (lane 9) only a fraction of the potentially degradable (see lane 2) substrate was acted upon. Two units of enzyme were used in standard assays to ensure that the vast majority of substrate was bound by a single RT molecule (assuming a Poisson distribution of enzyme on substrate and that most bound enzyme cleaved the substrate before dissociating). This permitted the observation of activities carried out by a single RT molecule bound at equilibrium to the substrate.

In order to assess the equilibrium binding position of RT on the substrate it was important that the assay was able to detect initial cleavage events. To determine this RT-mediated cleavage on a substrate (23R–19D:88D) in which the chimeric strand was uniformly labeled with radioactivity was performed. Cleavage of this substrate in the presence of heparin for 5 s or 7 min is shown

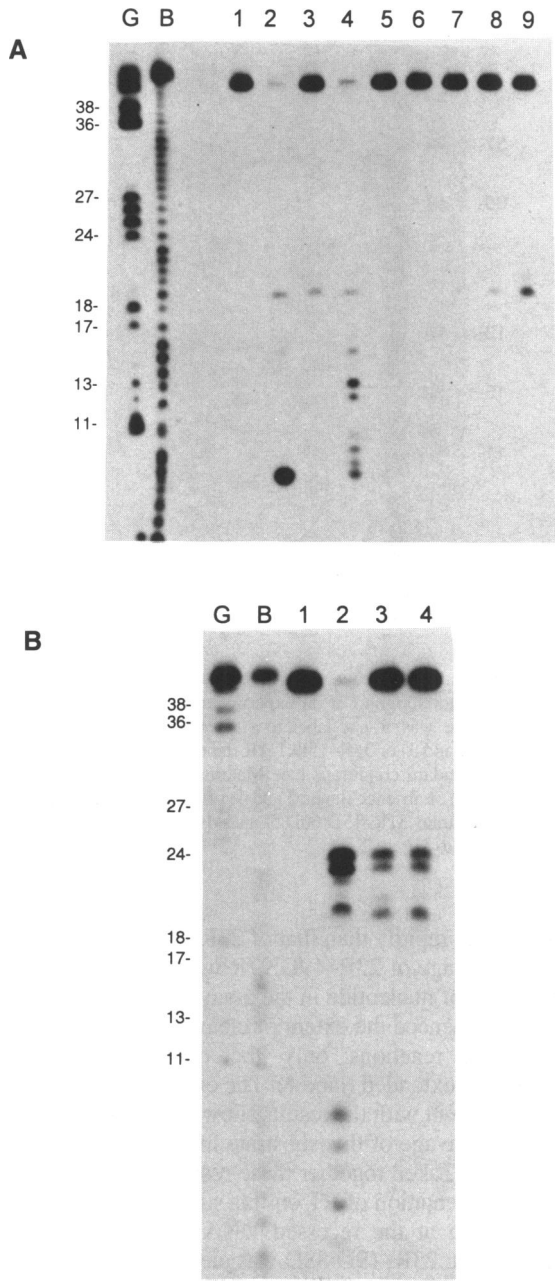

Figure 3. HIV-RT-mediated cleavage of substrates 23R–19D:88D and 23R–19D:97R. (**A**) An autoradiogram of an experiment in which the chimeric portion of the substrate was 5′-end-labeled with ^{32}P is shown. Substrate 23R–19D:88D (lanes 2, 3 and 6–9) or 23R–19D:97R (lanes 4 and 5) was digested with HIV-RT in the presence (lanes 3 and 5–9) or absence (lanes 2 and 4) of the heparin trap. The amounts of HIV-RT used in the assays shown in lanes 6–9 were 0.5, 1, 2 or 4 U respectively, while standard levels were used for other assays (see Materials and Methods). Lane 1 shows undigested substrate 23R–19D:88D. Lanes labeled G and B are as described in the legend to Figure 2. (**B**) An autoradiogram of an experiment in which the chimeric portion of substrate 23R–19D:88D was internally labeled with ^{32}P (~1 μCi/pmol chimera) is shown. Cleavage assays were in the presence (lanes 3 and 4) or absence (lane 2) of the heparin trap under standard conditions (see Materials and Methods), with the exception of the assay shown in lane 4, in which the reaction was terminated 7 min after initiation. Undigested substrate is shown in lane 1.

in lanes 3 and 4 respectively of Figure 3B. Cleavage resulted in three prominent products of lengths calculated as 19, 22 and 23 nt. The two larger products differed in length from the smaller by 3 and 4 nt, based on small amounts of two intermediate products observed in lane 2. This would make the larger products 22 and 23 nt in length. However, the largest appeared to co-migrate with the 24 nt G-ladder band. The discrepancy may be due to the slight difference in the structure of the G-ladder and base hydrolysis ladder products versus the cleavage products (see below) and the chimeric composition of the cleavage products. The smallest product migrated at the same position as the 5′-derived products that resulted from cleavage of the 5′-end-labeled substrate (Fig. 3). This product migrated slightly slower than the 19 nt base hydrolysis ladder product. This would be expected, since the products of base hydrolysis and RNase T1 cleavage have 3′ phosphate groups, compared with 3′ hydroxyls generated by RNase H. The additional negative charges likely lead to a slightly increased mobility in comparison with nucleic acids of equivalent length with 3′ hydroxyls. Given that the 5′-derived product was 19 nt in length, if the products resulted from a single cleavage event a 3′-derived product of 23 nt would be expected. There was also a low level of a 5′-derived product of 20 nt (see Figs 2 and 3). The corresponding 3′-derived product would be 22 nt long. All of these products are observed in Figure 3B at the 5 s time point (lane 3), although the proportions of each product were not clearly consistent with each resulting from an initial cleavage event. The level of 5′-derived 19 nt product is much greater than 5′-derived 20 nt product, while the level of 23 nt 3′-derived product is only about twice that of the 22 nt product. One possibility is that a portion of the 23 nt product is rapidly processed to produce 22 nt product and this event is too rapid to be detected within the time course of the assay. Clearly, however, the assay does detect initial cleavage events and this result indicated that most of the cleavages on this substrate occurred between nt 19 and 20 from the 5′-end of the chimeric strand.

Reverse transcriptase was unable to cleave the 23R–19D chimeric strand in the presence of heparin when it was bound to RNA (23R–19D:97R)

Also shown in Figure 3 are the results of cleavage assays using 23R–19D:97R as substrate. Reactions were in the presence (lane 5) or absence (lane 4) of heparin. In the presence of the trap cleavage of the chimeric strand was not detected. Even upon longer incubations (up to 7 min) no product was observed when heparin was included at the start of the reaction (data not shown). In contrast, in this particular experiment essentially all the substrate was cleaved when the trap was omitted and a higher enzyme concentration was used (see Materials and Methods). The presence of small (<23 nt) products in lane 4 indicated that some cleavages occurred in the RNA–RNA hybrid region of the duplex. These were likely catalyzed by the reported RNase D activity of HIV-RT (6). Note that these cleavages were not detected during single enzyme binding events (lane 5). This suggests that the dissociation rate constant for the binding orientation of RT required to produce these cleavages is greater than the rate constant for RNase D-mediated cleavage. Therefore, RT does not usually remain in this orientation long enough for cleavage to occur. However, since cleavage does occur in the absence of heparin, then at least a portion of the enzymes that bind in this orientation remain associated long enough to cleave the

substrate. A second possible explanation is that at equilibrium a small fraction of the enzymes are bound to the substrate in the proper orientation for RNase D-mediated cleavage of the chimera. If the fraction were very small the number of chimeras cleaved during a single binding event with the enzyme may be below the detection limits of the assay.

A chimeric substrate of a different sequence yielded similar results

In order to ensure that the results determined for the above chimeric substrates were not influenced by sequence-specific parameters unique to these substrates a cleavage assay on a second substrate was performed. A chimeric strand of different sequence (see Materials and Methods and Fig. 1), termed 31R–15D, was hybridized to the complementary regions on a 66 nt DNA or 111 nt RNA. These substrates had the same basic configuration as those used in the above experiments (Fig. 3). Cleavage experiments using 31R–15D:66D (lanes 2 and 3) or 31R–15D:111R (lanes 4 and 5) as substrate are shown in Figure 4. Reactions were performed in the presence (lanes 3 and 5) or absence (lanes 2 and 4) of the heparin trap. The starting material (lane 1) contained a low level of degraded substrate, as was evident from the appearance of small amounts of 31 and 24 nt contaminants. In the absence of trap both substrates were cleaved, although a portion of 31R–15D:66D (lane 4) remained undegraded, consistent with a lower cleavage efficiency for RNA–RNA versus RNA–DNA (see below). In contrast, in the presence of the trap cleavage occurred only when 31R–15D was bound to DNA (lane 3). The cleavage products generated were of 17 and 18 nt and were identical to those generated when a strand consisting of only RNA was bound to 66D (21). These results are in agreement with those for 23R–19D:88D and 23R–19D:97R and suggest that binding of RT to these substrates is a consequence of their basic structure and not a unique aspect related to the sequence of the particular substrate.

In the presence of the trap reverse transcriptase extended 23R–19D when bound to RNA, but not DNA

Figure 5 shows primer extension assays performed as described under Materials and Methods using 23R–19D:88D (lanes 1 and 2) and 23R–19D:97R (lanes 3 and 4) in the presence (lanes 2 and 4) or absence (lanes 1 and 3) of heparin. In these assays only extended products are visible, since the starting material was not radiolabeled. Label was acquired via extension of the chimeric strand at the 3′-end with [α-^{32}P]dCTP. Both substrates were extended in the absence of the trap. Substrate 23R–19D:88D showed extension products ranging from ~22 to 26 nt. These are likely 3′-derived cleavage products (see Fig. 3B) extended by 1–2 nt. Since dCTP was used for extension and the next two nucleotides on the template strand (long strand) were guanosines, extension of the chimeric strand by 1 or 2 nt could occur. The chimeric strand of 23R–19D:97R was also extended. Full-length strands were extended and some smaller extended cleavage products were also observed. The presence of extended uncleaved products is inconsistent with the result shown in Figure 3A (lane 4), in which nearly all of this substrate was cleaved in the absence of the trap. However, in some assays only a portion of the 23R–19D:97R substrate was cleaved and in a time course experiment similar to that shown in Figure 2, but performed in the absence of the trap, the chimeric strand of 23R–19D:88D was cleaved much more rapidly than that of 23R–19D:97R (data not shown). Thus cleavage of 23R–19D:97R was relatively inefficient. The presence of nucleotide in the assays shown in Figure 5 may also have influenced the extent of cleavage.

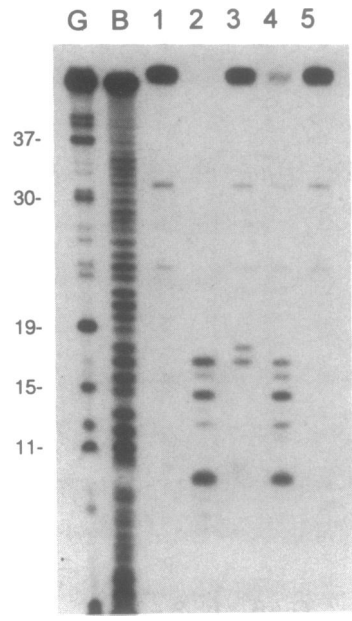

Figure 4. HIV-RT-mediated cleavage of substrates 31R–15D:66D and 31R–15D:111R. An autoradiogram of an experiment in which the chimeric portion of the substrate was 5′-end-labeled with ^{32}P is shown. Substrate 31R–15D:66D (lanes 2 and 3) or 31R–15D:111R (lanes 4 and 5) was digested with HIV-RT under standard conditions (see Materials and Methods) in the presence (lanes 3 and 5) or absence (lanes 2 and 4) of the heparin trap. Lane 1 shows undigested substrate 31R–15D:66D. Lanes labeled G and B are as described in the legend to Figure 2.

In the trapped reactions only the chimeric strand of 23R–19D:97R was extended (lane 4). The extended product was not cleaved, consistent with the results shown in Figure 3A (lane 5), in which no cleavage of this substrate in the presence of the trap was observed. Taken together these results suggest that the favored binding orientation of RT on this substrate positions the polymerase domain at the recessed DNA 3′-terminus of the chimera. In contrast, 23R–19D:88D was cleaved (Fig. 3A., lane 3), but not extended (Fig. 5, lane 2) in the presence of the trap. The cleavage products were identical to those generated with 23R–7D:88D (Fig. 2). These results suggest that RT binds to this substrate at a fixed distance from the 5′-end of the chimera and that the RNase H domain of RT is productively associated with the RNA portion of the chimera.

DISCUSSION

In this report HIV-RT-directed RNA cleavage and primer extension on substrates with relatively short chimeric (5′-RNA–DNA-3′) strands hybridized to a longer complementary DNA or RNA was examined. In all cases the 3′- and 5′-termini of the shorter nucleic acid were recessed on the longer one. When the chimeric strand was bound to DNA results indicated that RT interacted with these substrates in the same manner as when a complete RNA homolog of the chimeric strand was bound (Fig. 6B; 21), i.e. RT catalyzed cleavage of the chimera at a fixed

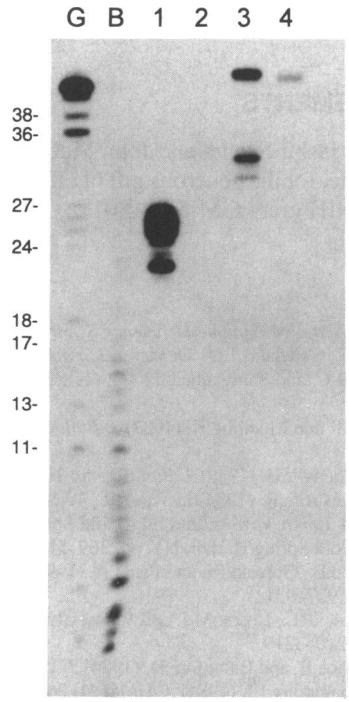

Figure 5. HIV-RT-mediated nucleotide extension on substrates 23R–19D:88D and 23R–19D:97R. An autoradiogram of an experiment in which [α-^{32}P]dCTP was used to attempt to extend the chimeric strand of substrate 23R–19D:88D (lanes 1 and 2) or 23R–19D:97R (lanes 3 and 4) is shown. Assays were performed in the presence (lanes 2 and 4) or absence (lanes 1 and 3) of the heparin trap as described under Materials and Methods. Lanes labeled G and B are as described in the legend to Figure 2.

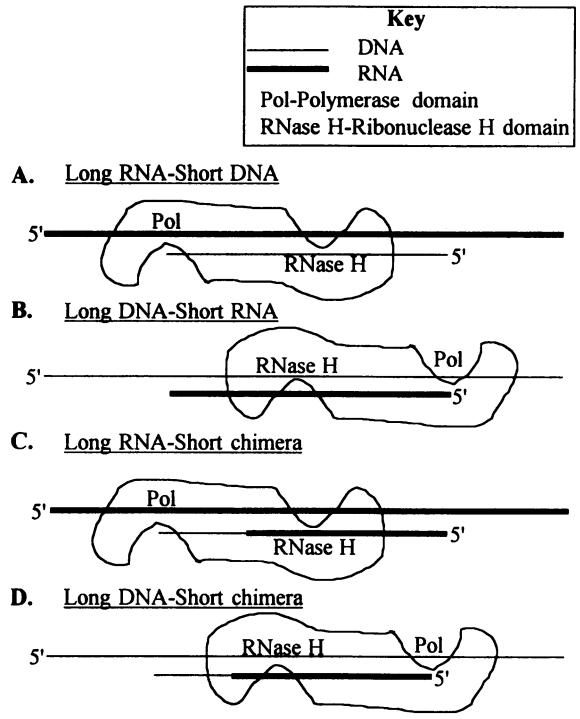

Figure 6. Possible orientations of RT on heteroduplex substrates. Shown is a schematic representation of RT binding to heteroduplex substrates of the indicated configurations. In each case the depicted orientation is that inferred from the activity of the RT during a single binding event with the substrate. Other, presumably less stable (see Discussion), orientations may also occur.

distance from its RNA 5′-end and independent of the position of the recessed DNA 3′-terminus (Figs 2 and 3A). These results suggest that RT is oriented on this substrate as shown in Figure 6D. The enzyme is bound proximal to the 5′-end of the chimeric strand such that the RNase H domain associates with the RNA on the chimeric strand.

It is unlikely that the orientation shown in Figure 6D represents a relatively low affinity binding position and that an alignment in which the polymerase domain is associated with the chimera 3′-terminus is more stable. Such an argument could be made for substrates with a short RNA bound to DNA (21). If the rate constant for dNTP extension of an RNA primer is smaller than the dissociation rate constant for RT on the substrate, extension would occur infrequently, even if the polymerase domain was properly configured for dNTP incorporation (see Results). However, 23R–19D:88D possesses a recessed DNA 3′-terminus, on which the rate constant for dNTP addition is likely much larger than the dissociation constant (26,27). Despite this, in the presence of the trap no addition was detected (Fig. 5, lane 2). It should be noted that rate constants for dissociation and dNTP extension or RNase H catalysis on chimeric substrates of the type used in these experiments have not been measured. Clearly, however, the dissociation constant for RT on 23R–7D:88D was much smaller than the rate constant for RNase H-mediated catalysis. This can be inferred from the observation that RT cleaved the substrate in the presence of the trap and catalyzed additional cleavages for several minutes after the initial cleavage (see Fig. 2). A similar result was obtained in a time course cleavage using 23R–19D:88D as substrate (data not shown). It could be argued that incorporation was not detected with 23R–19D:88D because RT incorporates inefficiently at this particular site. However, we detected no incorporation in the presence of the trap using 23R–7D:88D, a substrate with the 3′-terminus of the chimera at a different position (data not shown). Also, RT clearly extended the DNA 3′-terminus of the cleaved chimera of 23R–19D:88D in the absence of the trap (Fig. 5, lane 1). Taken together with previous results (21) the data strongly suggest that a relatively short RNA or homologous chimera, when bound to DNA, is recognized by RT as a substrate for degradation and not as a primer for DNA synthesis.

Note that this proposition does not preclude RNAs from being used as primers after first strand DNA synthesis. Evidence conclusively shows that second strand DNA synthesis is initiated from a unique RNA primer termed the 'polypurine tract' (8–14) and perhaps from other RNAs (29). The results shown here and those of others (10,22,28) suggest that the use of RNAs as primers is likely inefficient. The polypurine tract may represent a special case in which the inability of the RT to degrade this tract allows it to be used as a primer, although inefficiently in comparison with DNA primers. Indeed, a DNA homolog of this tract primes RT-directed DNA synthesis more efficiently than the polypurine tract (G.Fuentes and R.Bambara, personal communication).

Interestingly, when the chimeric strand was bound to a long RNA (23R–19D:97R) incorporation at the 3′-terminus of the

chimera in the presence of the trap did occur (Fig. 5, lane 4). This suggests that RT binds this substrate such that the polymerase domain associates with the DNA 3′-terminus of the chimera (see Fig. 6C). This binding orientation is consistent with RT binding this substrate in the same manner as if a DNA homolog of the chimera were bound to RNA (Fig. 6A; 21). In such a configuration the binding position of the enzyme is determined by the position of the recessed 3′-terminus of the DNA. Note that this orientation positions the RNase H domain such that cleavage of the longer strand can occur. In fact, in experiments in which the 97 nt RNA strand of 23R–19D:97R was radiolabeled cleavage of this strand was detected in the presence and absence of heparin (data not shown).

Although specific details concerning interaction of RT with the numerous potential substrates generated during retroviral replication remain to be determined, there are some common themes which can be drawn based on the orientations shown in Figure 6. In all cases the polymerase active site region associates with DNA such that the DNA strand runs 5′→3′ in the direction from the RNase H to the polymerase active site. The RNase H active site region associates with RNA such that the RNA strand or RNA portion of the chimera runs 5′→3′ in direction from the polymerase to the RNase H active site. In addition, in those substrates with a chimeric strand RT preferentially binds in the heteroduplex region, as opposed to DNA–DNA or RNA–RNA regions. This is consistent with RT binding with higher affinity to RNA–DNA versus DNA–DNA substrates (28). It is still not clear what role the length of the RNA regions may play in determining binding. Perhaps RT orients differently on shorter RNAs which cannot span the distance between the polymerase and RNase H active sites. Such RNAs would likely be generated frequently during first strand DNA synthesis. Also, the relative role of termini versus duplex type in determining binding affinity and positioning of RT on substrates remains to be explored.

Although these experiments suggest that the binding orientations shown in Figure 6 represent the most stable orientations on these substrates, other orientations are definitely possible. For example, on 23R–19D:97R cleavage of the chimeric strand in the absence of the trap occurred (Fig. 3A, lane 4). Such cleavages could not have occurred if the orientation shown in Figure 6C was the only possible configuration of RT on this substrate. The configuration which allows for chimera cleavage may represent a relatively unstable binding orientation, for which the dissociation rate constant for RT is relatively large. Due to the magnitude of this constant catalysis may occur only rarely, when RT assumes this orientation. Such may also be the case for RNAs which prime second strand DNA synthesis. The favored binding orientation on such RNAs may be similar to that shown in Figure 6B. In this case the RNA is recognized as a substrate for degradation. However, RNAs like the polypurine tract, which are resistant to degradation, may be extended by RTs that bind in a less favorable and likely less stable orientation that allows extension of the RNA with dNTPs.

ACKNOWLEDGEMENTS

I wish to thank Drs Jasbir Seehra and John McCoy, representing the Genetics Institute, for the generous gift of HIV-RT. This work was supported by NIH grant GM-51140-01

REFERENCES

1 Goff,S.P. (1990) *J. 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 Weiss,R., Teich,N., Varmus,H. and Coffin,J. (eds), *RNA Tumor Viruses*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 369–512.
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. (1984) *J. Virol.*, **49**, 200–204.
14 Smith,J.K., Cywinski,A. and Taylor,J.M. (1984) *J. Virol.*, **52**, 314–319.
15 DeStefano,J.J., Buiser,R.G., Mallaber,L.M., Bambara,R.A. and Fay.,P.J. (1991) *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 Kolstaedt,L.A., Wang,J., Friedman,J.M., Rice,P.A. and Steitz,T.A. (1992) *Science*, **256**, 1783–1790.
21 DeStefano,J.J., Mallaber,L.M., Fay,P.J. and Bambara,R.A. (1993) *Nucleic Acids Res.*, **21**, 4330–4338.
22 Peliska,J.A. and Benkovic,S.J. (1992) *Science*, **258**, 1112–111.
23 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
24 Promega (1989) *Protocols and Applications Guide*.
25 United States Biochemical Corp. (1992) *Molecular Biology Reagents/Protocols Manual*.
26 Kati,W.M., Johnson,K.A., Jerva,L.F. and Anderson,K.S. (1992) *J. Biol. Chem.*, **267**, 25988–25997.
27 Reardon,J.E. (1993) *J. Biol. Chem.*, **268**, 8743–8751.
28 Lokhova,I.A., Nevinsky,G.A., Gorn,V.V., Veniaminova,A.G., Repkova,M.V., Kavsan,V.M., Rudenko,N.K. and Lavrik,O.I. (1990) *FEBS Lett.*, **274**, 156–158.
29 Charneau,P., Alizon,M. and Clavel,F. (1992) *J. Virol.*, **66**, 2814–2820.
30 DeStefano,J.J., Bambara,R.A. and Fay,P.J. (1993) *Biochemistry*, **32**, 6908–6915.