Mechanism of DNA Strand Transfer Reactions Catalyzed by HIV-1 Reverse Transcriptase

James A. Peliska and Stephen J. Benkovic

Two DNA strand transfer reactions occur during retroviral reverse transcription. The mechanism of the first, minus strand strong-stop DNA, transfer has been studied in vitro with human immunodeficiency virus 1 reverse transcriptase (HIV-1 RT) and a model template-primer system derived from the HIV-1 genome. The results reveal that HIV-1 RT alone can catalyze DNA strand transfer reactions. Two kinetically distinct ribonuclease (RNase) H activities associated with HIV-1 RT are required for removal of RNA fragments annealed to the nascent DNA strand. Examination of the binding of DNA · RNA duplex and single-stranded RNA to HIV-1 RT during strand transfer supports a model where the enzyme accommodates both the acceptor RNA template and the nascent DNA strand before the transfer event is completed. The polymerase activity incorporated additional bases beyond the 5′ end of the RNA template, resulting in a base misincorporation upon DNA strand transfer. Such a process occurring in vivo during retroviral homologous recombination could contribute to the hypermutability of the HIV-1 genome.

Retroviral reverse transcription is catalyzed by the virally encoded enzyme reverse transcriptase (RT). This multifunctional enzyme can catalyze DNA-dependent and RNA-dependent DNA polymerization as well as the RNA hydrolysis of DNA·RNA hybrids (RNase H activity) (*1, 2*). In a series of complex events, the RT as part of a replication complex facilitates the synthesis of double-stranded DNA from a diploid RNA (Fig. 1) (*3, 4*). In a widely accepted model for reverse transcription, minus strand DNA synthesis is initiated at a specific transfer RNA (tRNA) primer complementary to the primer-binding site (PBS) located near the 5′ end of the viral RNA and continues to the end of the viral genome, generating what has been termed minus strand strong-stop DNA (Fig. 1A). This short nascent DNA strand is translocated to the 3′ end of the same or a second viral RNA molecule, guided by the direct repeat sequence (r) on both ends of the RNA (Fig. 1, B and C), and minus strand DNA synthesis is completed. Plus strand DNA synthesis is then initiated at an RNase H–resistant RNA oligonucleotide known as the polypurine tract located near the U3 region at the 5′ end of the minus strand DNA and proceeds through an intramolecular DNA strand transfer facilitated by RNase H removal of the tRNA primer (*5, 6*) (Fig. 1, D to F). DNA strand transfer also apparently occurs during the process of retroviral recombination where the RT transfers to a second strand of viral RNA on encountering a break in the RNA template during minus strand DNA synthesis (*7*) as has been shown in both in vivo and in vitro studies (*6, 8–11*).

Several important questions surround the steps involved in strand transfer during retroviral reverse transcription, including the overall mechanism of the process, whether there is a key intermediate, such as a three-stranded (RNA·DNA·RNA), formed in the transfer process as implied in Fig. 1B, which is accommodated by a binding site on the RT and the timing of the RNA cleavage relative to DNA synthesis. Two kinetically distinct RNase H activities have been attributed to HIV-1 RT: a polymerase-coupled activity that cleaves the RNA template as the DNA is actively synthesized, and a polymerase-independent RNase H activity that functions in the absence of DNA elongation (*12, 13*), but whether both are required in HIV strand transfer reactions has not been determined.

We now address these issues by describing the molecular and kinetic characteristics of a model system designed to examine strand transfer reactions catalyzed by HIV-1 RT in vitro. As would be expected from earlier reports, our data reveal that RT alone is capable of catalyzing strand transfer reactions (*9–11*). We find that RNase H activity is required for an efficient strand transfer process and that it acts in concert with the polymerase of the enzyme. The latter, furthermore, manifests an unanticipated ability to extend by a single base the DNA of a DNA · RNA duplex past the 5′ end of the RNA template prior to strand transfer. If this same reaction occurs in vivo, it could provide a means for introducing nucleotides that are not template-encoded, possibly contributing to the hypermutability of the HIV-1 genome. Kinetic studies and time-dependent crosslinking experiments imply the formation of RNA · DNA strand transfer intermediates facilitated by the binding of RNA template at a site spanning both the p51 and p66 subunits of HIV-1 RT that correlates with a possible site visualized in the crystal structures of HIV-1 RT (*14, 15*). Understanding the characteristics of these reactions catalyzed by HIV-1 RT in vitro is critical in determining its role in the transcription complex in vivo and for possible therapeutic use.

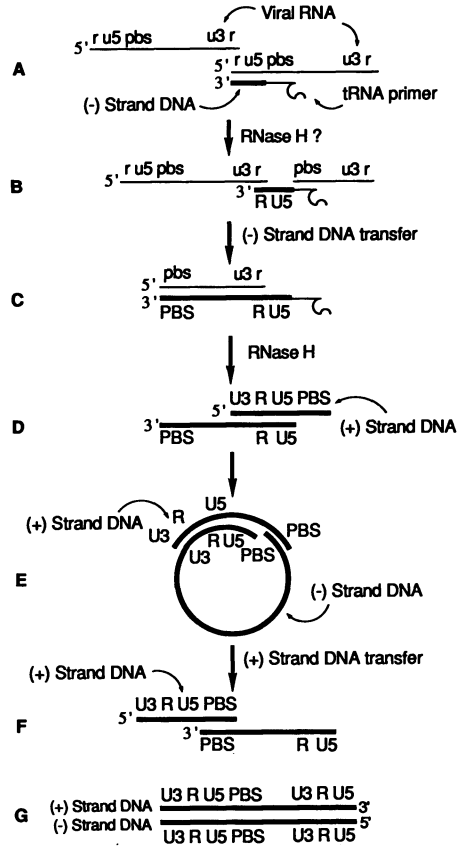

Fig. 1. Model for the processing of diploid RNA to double-stranded proviral DNA during retroviral reverse transcription. For clarity, processing of only one of the viral RNA strands is shown. Capital letters and bold lines represent DNA while lower case and thin lines represent RNA. The terminal repeat sequence (r) is thought to aid in the minus strand strong-stop DNA transfer (A → B). An intermolecular transfer is shown, but both inter- and intramolecular transfer may occur (*5, 6*). The functional role of RNase H activity in this reaction has not been conclusively determined. Plus strand DNA synthesis is initiated at a polypurine-rich RNA primer located near the u3 region after RNase H hydrolysis of the RNA template (D). After removal of the tRNA primer by RNase H, a plus strand DNA transfer is thought to occur in an intramolecular fashion (E → F) aided by the complementarity phosphate-buffered saline sequences located on the 3′ ends of both plus and minus strand DNA.

The authors are in the Department of Chemistry, the Pennsylvania State University, University Park, PA 16802.

We designed a two-template RNA system to examine the mechanism of DNA strand transfer (Fig. 2). The primary RNA template was derived from the first 40 bases (reading 5′ to 3′) of the HIV-1 terminal repeat (r) (*16*) and was primed at its 3′ end with the complementary 20-base DNA oligonucleotide. The second strand of RNA shares the same sequence as the last 20 bases (5′ end) of the primary RNA template and includes an additional 21 bases derived from the u3 region of the viral genome. This generates an overlap in the two RNA templates that effectively simulates conditions occurring during minus strand strong-stop DNA transfer (Fig. 2) (*17*). To detect the DNA strand transfer we examined the DNA products after DNA polymerization for the appearance of full-length 61-base DNA.

The kinetics of DNA strand transfer catalyzed by HIV-1 RT. To demonstrate the mandatory involvement of RNase H activity in strand transfer, we performed this assay with both wild-type (*18*) and an RNase H–deficient mutant of HIV-1 RT that displays no detectable RNase H activity but retains wild-type polymerase activity (*19*). The time-dependent formation of DNA products resulting from strand transfer (Fig. 3) reveals that even at reaction times exceeding eight half-lives, less than 1.5 percent of the transfer product (the limit of detection) was found in the reaction containing the RNase H–deficient mutant, whereas the wild-type protein effectively processed more than 85 percent of the primary DNA product. Although some other property of HIV-1 RT necessary for catalyzing strand transfer may have been compromised by the protein mutation, we do not favor this interpretation for two reasons. First, the relatively conservative point mutation incorporated into the HIV-1 RT mutant, unlike a deletion mutation, is not expected to alter protein structure and thereby affect other protein functions. Second, the kinetics of RNA degradation in the wild-type reaction further emphasize the role of RNase H in the transfer reaction (as discussed below).

This result is consistent with earlier work demonstrating RNase H hydrolysis of the 5′-terminal portion of the Rous sarcoma virus genome during reverse transcription (*20*) and in strand transfer reactions catalyzed in vitro by Moloney murine leukemia virus (MLV) reverse transcriptase (*10*). In contrast, RNase H may not be a mandatory requirement for DNA strand transfer if homopolymer templates such as poly(A) · oligo(dT) are used (*11*) or if the enzyme, as in the case of MLV RT, is able to unwind the 3′ end of the primary duplex before hybridization of the nascent DNA with the RNA acceptor template (*9*). For HIV-1 RT, the latter must be an inefficient process. However, the hybridization of 40-base DNA (primary product) and 41-base RNA (acceptor template) (Fig. 2) is accelerated in our reaction by the presence of HIV-1 RT. The observed second-order rate constant of hybridization increased linearly with HIV-1 RT concentration and was more than doubled (2467 M^{-1} s^{-1}) above the background (995 M^{-1} s^{-1}) at 200 nM HIV-1 RT while 400 nM HIV-1 RT increased this constant further (3533 M^{-1} s^{-1}) (*21*).

The DNA template transfer reaction catalyzed by HIV-1 RT was assayed with the DNA and RNA of the primary template-primer alternatively labeled with ^{32}P. On initiation of the transfer reaction, the ^{32}P-labeled 20-base DNA primer of the 20-base DNA · 40-base RNA duplex was rapidly elongated to the corresponding 40-base DNA (Fig. 4A). Effectively, all of the DNA primer was extended within the first 5 minutes. This 40-base product was further elongated within 5 minutes by HIV-1 RT via a blunt-end nucleotide addition to a 41-base product at an apparent rate of approximately 0.1 min^{-1} and subsequent incorporation of additional nucleotides at a slower rate. Similar blunt-end addition has been observed for other DNA polymerases, but again at a slower rate (*22*). This blunt-end addition was unaffected by the presence of 41-base RNA acceptor template, demonstrating that these products did not arise from a transfer process. Further analysis of the data shows that both the 40-base and 41-base primary DNA products can transfer onto the 41-base RNA acceptor template, with subsequent rapid elongation to the 61-base full-length product, which appears more than 10 minutes later (*23*).

The identity of the base added to the blunt end was determined by incubating the blunt end 40-base DNA · 40 base RNA with each individual nucleotide triphosphate and HIV-1 RT. On the basis of the rates at which each nucleotide was incorporated, we determined that HIV-1 RT shows a preference for addition—dAMP > dGMP > dTMP > dCMP—onto this particular substrate in a ratio of 1:0.5:0.2:0.07 (*24*). Because this addition was not influenced by the presence of RNA acceptor template, it is an intrinsic property of the reverse transcriptase with a blunt-end duplex substrate. For this duplex, the addition of any base other than dATP would result in the incorporation of an incorrect base at

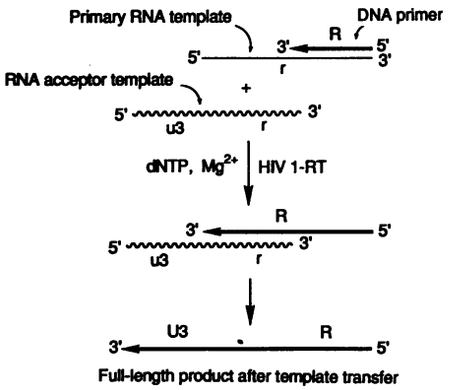

Fig. 2. DNA strand transfer assay system. The 40-base primary RNA template derived from the repeat sequence (r) located at the 5′ end of the HIV-1 genome (Fig. 1) was primed with a complementary DNA oligonucleotide. The acceptor RNA template was homologous to the last 20 bases (5′ end) of the primary RNA template (r) and included an additional 21 bases derived from the u3 genomic sequence (*17*). In the standard assay, 5′ end-labeled [^{32}P]DNA is used and, after reaction initiation and termination (EDTA quench), the products were analyzed by polyacrylamide gel electrophoresis (PAGE) under denaturing conditions. The occurrence of strand transfer is indicated by the appearance of 61-base DNA product.

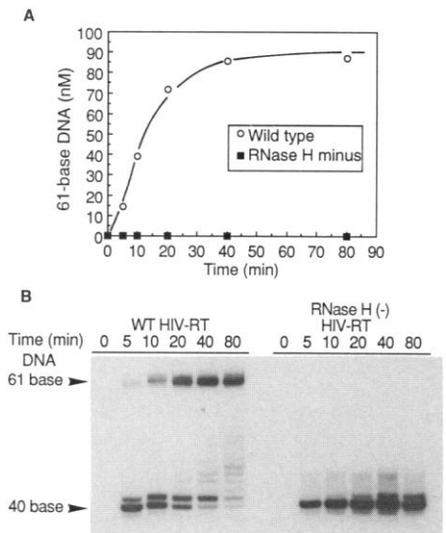

Fig. 3. RNase H requirement for strand transfer. Both wild-type (WT) and RNase H–deficient [RNase H (−)] HIV-1 RT were tested (Fig. 2, assay) for their ability to support DNA strand transfer. (**A**) Quantitative analysis of strand transfer reaction products (**B**) isolated by PAGE on 20 percent acrylamide, 8 M urea, TBE gels. Reaction mixtures contained assay buffer [50 mM tris-HCl, pH 8, 75 mM KCl, 0.1 mM EDTA, 1 mM DTT (dithiothreitol), 0.1 percent Triton X-100], 200 nM 5′ end-labeled 20-base [^{32}P]DNA · 40-base RNA, each dNTP at 100 μM, 7 mM $MgCl_2$, 800 nM 41-base RNA and wild-type HIV-1 RT (200 nM) or RNase H (−) (200 nM) were incubated at 37°C. At the times indicated, reaction samples were withdrawn, and the reaction was terminated by addition to EDTA (110 mM, final). The product bands were visualized by autoradiography, and the products were quantified by band excision and liquid scintillation counting. Only elongated products are shown.

position 41 after DNA synthesis (*17*). On the basis of the determined ratios above, this would represent approximately 43 percent of the strand-transferred DNA containing a blunt-end addition.

The synthesis of DNA on competent template primers by HIV-1 RT in vitro is processive and occurs at single-base incorporation rates of 14 s^{-1} (*25*). This is consistent with the rapid 20-base DNA primer elongation (Fig. 4A) where production of 40-base DNA was complete within the first 5 minutes. Likewise, the putative strand transfer intermediate 40-base DNA·41-base RNA template primer was independently shown to rapidly polymerize to 61-base DNA under the reaction conditions (Fig. 4). This indicates that some step other than DNA synthesis must be limiting the formation of the 40-base DNA·41-base RNA strand transfer intermediate.

To determine the nature of this limiting step, we designed experiments to monitor the fate of the 40-base RNA during template transfer. Concomitant with the elongation of the 20-base DNA primer to the 40-base DNA primary product, the template RNA was degraded by the RNase H activity associated with HIV-1 RT to form an initial 14-base RNA product (Fig. 4B). This process was detectable within the first 5 minutes, reflecting the ability of HIV-1 RT to rapidly cleave RNA templates. Generation of the 40-base DNA · 14-base RNA intermediate was thus too rapid to be the rate-limiting step in DNA strand transfer.

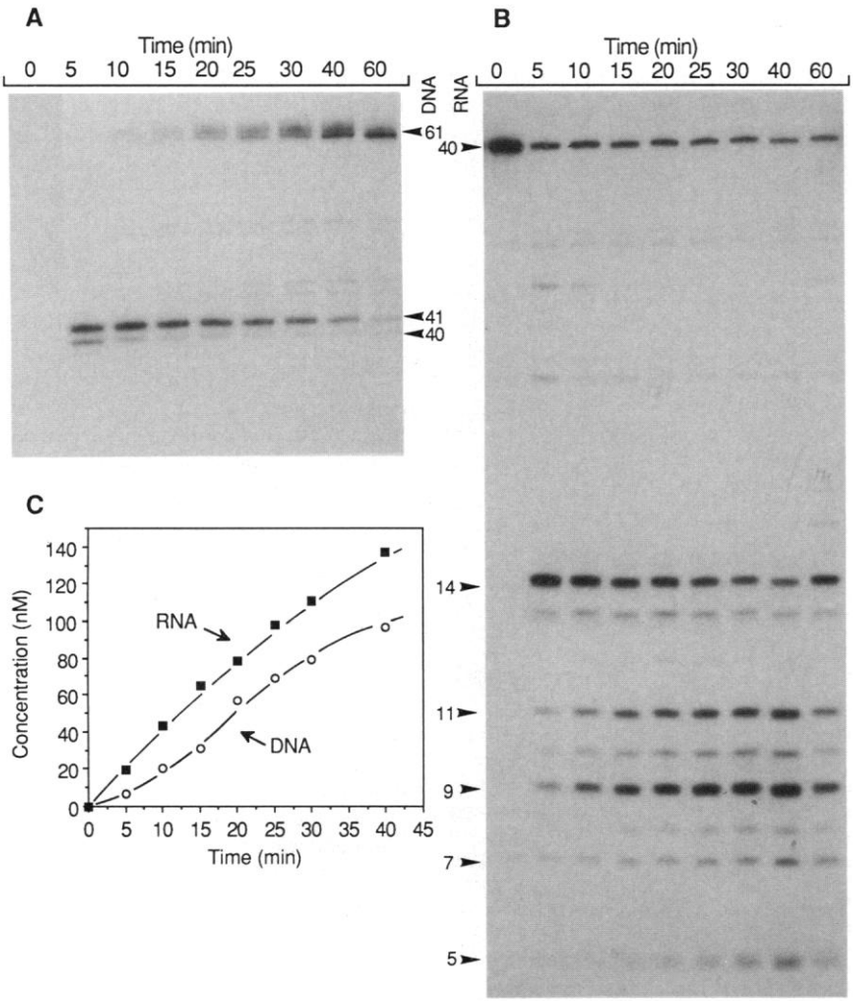

Fig. 4. Analysis of substrate processing during DNA strand transfer. Reactions contained 300 nM 5′ end-labeled 20-base [^{32}P]DNA · 40-base RNA (**A**) or 300 nM 20-base DNA · 5′ end-labeled 40 base [^{32}P]RNA (**B**), 500 nM 41-base RNA, each dNTP at 150 μM, 7 mM $MgCl_2$, and 100 nM HIV-1 RT in assay buffer at 37°C. Reactions were terminated with EDTA, and the products were analyzed as in Fig. 3. The lengths of the DNA products are shown at the right. Only elongated DNA products are shown. (**C**) The data in (A) and (B) were analyzed after product band excision and scintillation counting. The yield of DNA strand transfer product (○) is correlated with that of RNA fragments of less than ten bases (■). The experimental evidence presented in the text suggests that the yield of these RNA fragments should coincide with that of primary DNA product capable of transferring to the RNA acceptor template.

The lack of RNA fragments of intermediate length suggests that the polymerase and RNase H activities were coupled, that is, the RNA template was cut progressively and during the synthesis of DNA (*12*). This result is consistent with the recently reported crystal structures of HIV-1 RT, which showed that the nucleic acid duplex lies in a cleft spanning both the polymerase and RNase H active site of HIV-1 RT (*14, 15*). When the polymerase encounters the end of the template, RNase H activity is temporarily halted 14 base pairs from the 5′ end of the RNA (Fig. 4B), as with an earlier reported RNase H cleavage 15 base pairs from the 3′ terminus of the DNA catalyzed by HIV-1 RT under polymerization conditions (*13*). Our studies at shorter times demonstrated that the 14-base RNA fragment was produced from a 19-base precursor, suggesting that the RNase H action may proceed to rapidly remove four to five additional bases after the termination of DNA synthesis (*12*). This result would be in accord with a 18-base separation between the polymerase and RNase H sites, on the basis of modeling an A-form DNA · RNA hybrid into the substrate binding cleft of the crystal structure of the p66-p51 heterodimer of HIV-1 RT (*15*). Thus, DNA synthesis by the polymerase activity of HIV-1 RT appears to translocate the newly synthesized DNA·RNA hybrid into the RNase H active site where the RNA was enzymatically processed.

Under the conditions described in Fig. 4, the production of the DNA transfer product then occurred after an initial lag phase owing to the RNase H processing of the primary template (Fig. 4C). Examination of the reaction products by nondenaturing gel electrophoresis indicated that the 14-base RNA fragment was still annealed to the nascent 40-base DNA and required removal before the 41-base RNA acceptor template could form a functional template-primer conformation. This was accomplished by a kinetically distinct polymerase-independent RNase H activity and gave rise to the shorter RNA cleavage products (Fig. 4B), a reaction that occurs even in the absence of the 41-base RNA acceptor template. Processing of the RNA strand apparently continued until a DNA · RNA hybrid of less than ten bases was formed whose melting temperature (T_m < 35°C) was below that of the reaction (37°C) (*26*).

Quantitation of RNA fragments of nine bases or fewer should then correlate with the production of DNA capable of transfer onto the acceptor RNA template, subsequently giving rise to full-length transfer product. The rate of formation of the transferable DNA, obtained from the quantitation of these RNA fragments, although somewhat faster than the formation of

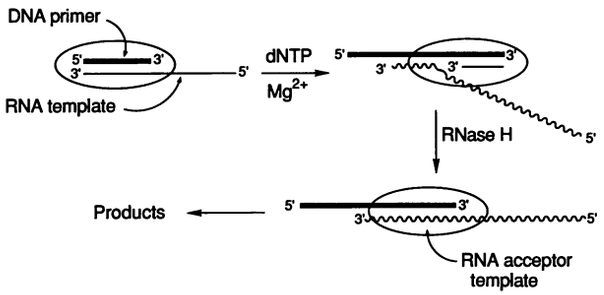

Fig. 5. A model for DNA strand transfer proceeding through an RNA · DNA · RNA intermediate accommodated by HIV-1 RT.

strand transfer product, does track the formation of full-length strand transfer product (Fig. 4C). The discrepancy is probably due to some nine-base RNA fragments remaining annealed to the primary DNA since its melting temperature was comparable to the reaction temperature.

The molecularity and polarity of the strand transfer reaction. Several lines of evidence suggest that HIV-1 RT can accommodate a strand transfer reaction through a putative RNA·DNA·RNA enzyme-bound species (Fig. 5). Variation of the concentration of 41-base RNA acceptor template in the presence of fixed amounts of enzyme and 20-base DNA · 40-base RNA showed a concentration dependence for 61-base DNA product formation. The individual progress curves are characterized by an invariant lag in product formation consistent with the mandatory removal of terminal RNA fragments before DNA transfer (Fig. 6A). A replot of the observed steady-state rates and RNA concentration is hyperbolic (Fig. 6B), indicative of but not demanding a mechanism where the acceptor template forms a Michaelis-Menten complex prior to template transfer.

Physical evidence derived from protein–DNA · RNA crosslinking experiments further defines the location of the binding of the transfer intermediates to HIV-1 RT. Previous studies have shown that duplex and single-stranded DNA molecules can be crosslinked to HIV-1 RT upon exposure to ultraviolet radiation. Crosslinking of $A_{(12-18)}dT_{10}$ to heterodimer p66-p51 gave

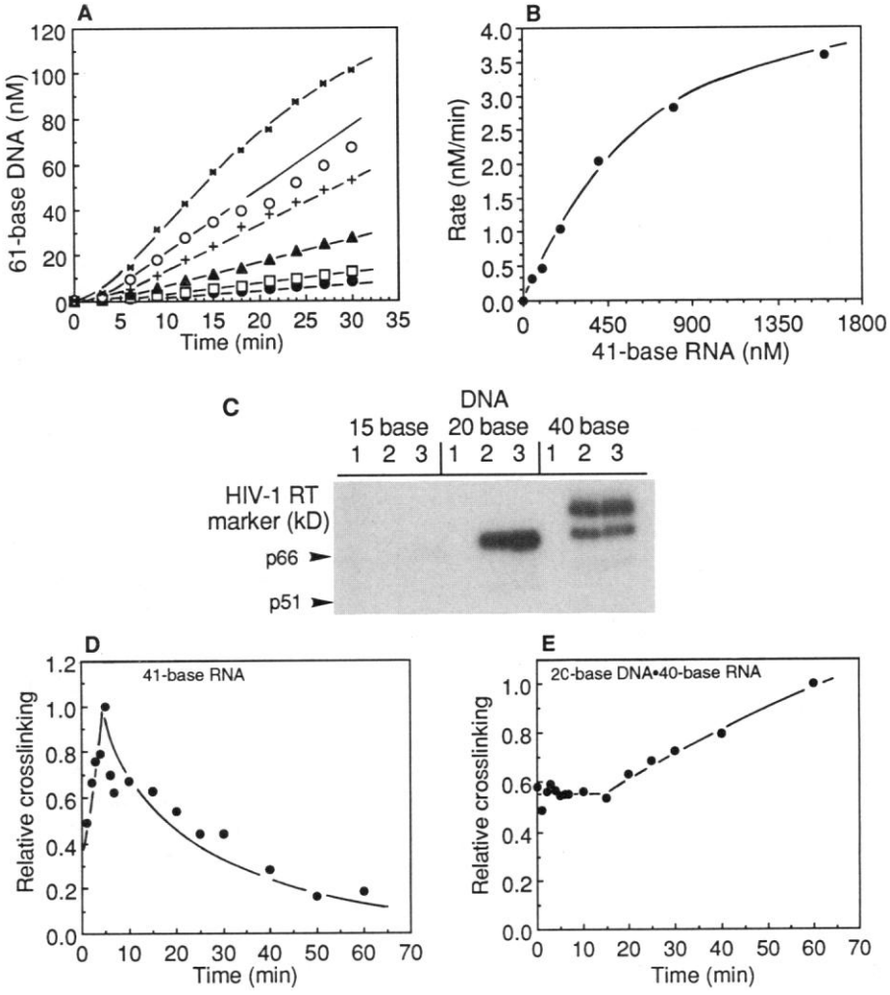

Fig. 6. Kinetic and physical evidence for formation of enzyme-bound DNA strand transfer intermediates. (**A**) The dependence of DNA strand transfer on 41-base RNA acceptor template concentration. Reactions contained 300 nM 5′ end-labeled, 20-base [^{32}P]DNA · 40-base RNA, each dNTP at 150 μM, 7 mM $MgCl_2$, 80 nM HIV-1 RT and 50 nM (●), 100 nM (□), 200 nM (▲), 400 nM (+), 800 nM (○), or 1600 nM (*) of 41-base RNA acceptor template. Reactions were analyzed as in Fig. 3. (**B**) A replot of the observed strand transfer rates and 41-base RNA concentration derived from (A). The rates plotted were obtained from the post-lag linear portion of each progress curve. The data were fitted to the equation rate = [41-base RNA] $V/(K_m$ + [41-base RNA]) to yield a K_m = 731 nM for 41-base RNA and V = 5.3 nM/min. (**C**) Characteristic ultraviolet crosslinking of various DNA oligonucleotides to the p51-p66 heterodimer of HIV-1 RT. 5′-Oligonucleotides labeled with ^{32}P at the 5′ end were incubated with 200 nM HIV-1 RT in reaction buffer and irradiated with ultraviolet as described below: (1) 450 nM DNA, no irradiation; (2) 450 nM DNA, with irradiation; (3) 900 nM DNA, with irradiation. The arrows on the left designate the position of non-crosslinked HIV-1 RT p51 and p66 determined by Coomassie blue staining. The decrease in the electrophoretic mobility of the crosslinked enzyme was due to the increased size of the protein-DNA complex. (**D and E**) Crosslinking of substrates during DNA strand transfer. (**D**) The time dependence of 41-base RNA acceptor template binding to HIV-1 RT under strand transfer conditions was examined by crosslinking reaction samples by means of ultraviolet light. The reaction mixture contained 300 nM 20-base DNA · 40-base RNA, 100 nM HIV-1 RT, each at 0.5 μM dNTP, 7 mM $MgCl_2$, and 500 nM 5′ end-labeled 41-base RNA in assay buffer, 37°C. Reaction samples were quenched with EDTA (20 mM) and then simultaneously irradiated with two 15-W bulbs (254 nm) at a distance of 3 cm for 15 seconds. In the absence of 20-base DNA · 40-base RNA, the 41-base RNA crosslinked with the same efficiency as the maximum values obtained in (D) (5 minutes). (E) The reaction was identical to (D) except 5′ end-labeled 41-base [^{32}P]RNA was replaced with 5 end-labeled 20-base [^{32}P]DNA · 40 base RNA. Crosslinked products were heated to 100°C for 5 minutes and then subjected to electrophoresis under denaturing conditions through 12 percent acrylamide, 0.1 percent SDS. The crosslinked HIV-1 RT was detected by autoradiography, and quantitated by liquid scintillation. All data were at least five times higher than that of background. The same reaction as (E) performed in the absence of 41-base RNA showed that crosslinking of the DNA is not significantly affected by the RNA acceptor template.

exclusive crosslinking to the p66 subunit (*27*), as would be expected from comparison with biochemical (*28, 29*) and structural (*15*) studies that suggest that the p66 subunit is the only functional subunit in the heterodimer. However, $tRNA^{Lys}$ is crosslinked to both p66 and p51 subunits of heterodimer HIV-1 RT (*30*), which is in accord with a model based on crystallographic evidence that the p51 subunit forms part of the tRNA binding site (*15*). We find that short DNA or RNA oligonucleotides crosslink only to the p66 subunit, but strands longer than 25 nucleotides appear to contact both p51 and p66 subunits (Fig. 6C) (*31*). The ratio of RNA or DNA crosslinking to the two subunits (approximately 1:1, Fig. 6C) is independent of oligonucleotide concentration and the time of ultraviolet irradiation (*32*). Examination of the crosslinking of the ^{32}P-labeled 20-base DNA · 40 base RNA duplex substrate revealed that the 20-base DNA primer initially crosslinks to the p66 subunit alone, but after DNA synthesis to the corresponding 40-base product, the DNA crosslinks with the p51 subunit. Collectively, the results indicate a common binding pocket spanning both subunits.

The implied simultaneous binding of both the primary DNA elongation product and the RNA acceptor template was further tested in a series of competition crosslinking experiments. In the presence of increasing amounts of 20-base DNA · 40 base RNA template-primer, binding of the 41-base RNA acceptor template to HIV-1 RT was competitively inhibited. The estimated inhibition constant (K_i = 27 nM) for 20-base DNA · 40-base RNA is in agreement with its independently determined dissociation constant (K_d) of 28 nM (*33*). Thus the acceptor template is initially excluded from binding to the enzyme during strand transfer reactions. However, on addition of nucleotide triphosphates and $MgCl_2$, the crosslinking of the 41-base RNA acceptor template to both p51 and p66 subunits rapidly increases while the DNA binding remains relatively unchanged (Fig. 6, D and E). The onset of 41-base RNA binding has a time dependence similar to that of the lag-phase associated with the production of full-length strand transfer product (Fig. 6A). This is consistent with an intermediate species in which the 41-base RNA acceptor template occupies a binding site left vacant by RNase H hydrolysis of the primary template during DNA synthesis. At longer times (>20 minutes) there is a decrease in 41-base RNA due to RNase H cleavage of the RNA after strand transfer and a concomitant increase in elongated DNA, which accounts for the trend in the data.

The putative strand transfer intermediate is a 40-base DNA · 41 base RNA hybrid with both strands having 3′-hydroxyl groups (Fig. 2). Since HIV-1 RT has been shown to support polymerization off either DNA or RNA primers, this intermediate could partition in two directions (Fig. 7). DNA synthesis with the DNA serving as the primer would result in the previously observed 61-base transfer product (Fig. 3), whereas synthesis off the RNA primer would result in a mixed DNA-RNA intermediate. The RNA in turn would be susceptible to RNase H degradation resulting in a 20-base DNA product. Partitioning of the presumed strand transfer intermediate was assayed by incorporating radioactively labeled nucleotide triphosphates (NTP's) and examining the products by polyacrylamide gel electrophoresis (Fig. 7B). Only the production of a 61-base DNA resulting from synthesis from the DNA primer was detectable under these conditions. In addition, the putative DNA · RNA intermediate was generated through hybridization of purified oligonucleotides and added to HIV-1 RT, and its partitioning was determined (Fig. 7B). Both reactions produced identical results, suggesting that HIV-1 RT has a strong preference (>98 percent) to elongate DNA rather than RNA primers, thus ensuring the correct elongation polarity of strand transfer intermediates.

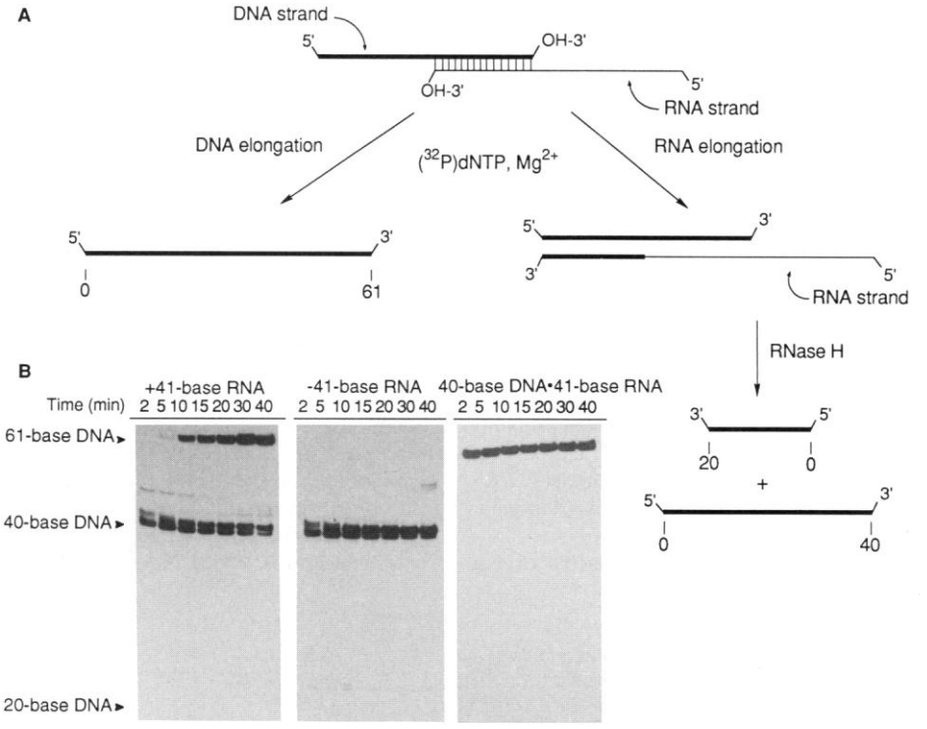

Fig. 7. Determination of the polarity of DNA strand transfer. (**A**) The 40-base DNA · 41-base RNA strand transfer intermediate could partition in two directions; either DNA elongation or RNA elongation. DNA elongation results in the production of 61-base DNA product while RNA elongation is expected to produce a 20-base DNA product after RNase H processing. (**B**) The assay shown in (A) was performed at 37°C in the presence of 300 nM of 20-base DNA · 40-base RNA, 100 nM HIV-1 RT, each dNTP (including [α-^{32}P]dGTP at 100 μM), 7 mM $MgCl_2$, and 500 nM 41-base RNA (left panel) or no 41-base RNA (center panel). The partitioning of thermally annealed 40-base DNA · 41-base RNA was performed as a control under identical conditions (right panel). The identity of the DNA products is shown at the left. Reactions were performed as in Fig. 3.

Detailed DNA strand transfer model. The kinetic and physical data were used to construct a detailed model for DNA strand transfer catalyzed by HIV-1 RT. The primary template-primer initially binds tightly to the HIV-1 RT with the 3′-hydroxyl of the DNA primer positioned for DNA synthesis in the polymerase active site. In this configuration, the second RNA acceptor template is precluded from binding to the enzyme. Upon addition of $MgCl_2$ and nucleotide triphosphates, the DNA strand is rapidly and processively elongated to the 40-base primary product; at the same time, the template RNA is degraded by the RNase H activity at a locus 18 or 19 bases away from the polymerase site, producing a 40-base DNA · 14-base RNA hybrid intermediate. With most of the 40-base RNA template now removed by RNase H action, the 41-base RNA acceptor template can bind to a site spanning both p66 and p51 subunits of HIV-1 RT. Before the 41-base RNA can serve as a competent template for further DNA synthesis, the 14-base RNA fragment must be removed by a kinetically

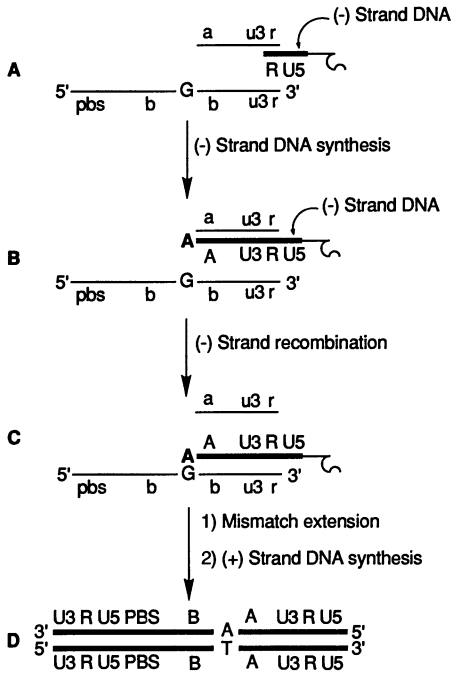

Fig. 8. Model for recombination-induced misincorporation by HIV-1 RT. For clarity, the RNA degradation by RNase H is not shown but is mandatory for strand transfer. During forced copy-choice recombination, the minus strand DNA is synthesized until a nick in the template RNA (template a) is reached (**A** → **B**). The reverse transcriptase adds an additional base (shown here as dAMP) onto the blunt end of the DNA · RNA primary product (**B**). After DNA strand transfer onto the second RNA template (template b), depending on the sequence of the RNA template, a mismatch could result (shown here as a G · A mispair) (**C**). In the absence of any proofreading capability of the RT, the terminal mismatch is incorporated upon DNA synthesis. Subsequent plus strand DNA synthesis resulted in base conversion (**D**).

distinct polymerase-independent RNase H activity. Since the spatial distance between the RNase H and polymerase active sites appears to be approximately 18 or 19 bases (*12*, *15*), cleavage of the residual primary template RNA probably requires an alteration of the duplex binding where the DNA 3′-hydroxyl is no longer positioned in the polymerase pocket. Removal of the RNA fragment allows the final hybridization of the acceptor template; after the transferred primer is repositioned into the polymerase site, rapid DNA synthesis to the 61-base strand transfer product ensues.

The genetic variability of the HIV genome may allow it to evade the host's immune system. The mutation rates of the *env* and *gag* genes of HIV-1 are so high that a nucleotide sequence change occurs essentially every replicative cycle (*34*). A major source of genetic variation in HIV-1 may be due to the inherent inaccuracy of the RT (*35*), with error rates in copying DNA and RNA estimated to be as high as one misincorporation per 5000 to 6000 nucleotides polymerized. Our in vitro results suggest yet another way in which HIV-1 might generate variation during reverse transcription. During DNA strand transfer reactions that simulate minus strand forced copy-choice recombination in vitro, more than 50 percent of the primary DNA is extended past the primary RNA template prior to DNA strand transfer (Figs. 3 and 8). Because the addition at the nucleotide blunt end occurs in the absence of excess acceptor template, it is not expected to be directed by the sequence of the next base of the acceptor template. Therefore, potentially any base-pair mismatch could be generated after DNA strand transfer resulting in a point mutation at the recombination site (Fig. 8). If such a reaction occurs in vivo, the degree to which this could contribute to the variation of the retroviral genome would in turn depend on the integrity of the packaged viral RNA. Because so much recombination occurs during reverse transcription (*6*–*8*), the potential for recombination-induced mutation could be quite high.

REFERENCES AND NOTES

1. D. Baltimore, *Nature* **226**, 1209 (1970); H. M. Temin and S. Mizutani, *ibid.*, p. 1211.
2. S. P. Goff, *J. Acquired Immune Defic. Syndr.* **3**, 817 (1990).
3. E. Gilboa, S. W. Mitra, S. Goff, D. Baltimore, *Cell* **18**, 93 (1979).
4. H. Varmus and R. Swanstrom, in *RNA Tumor Viruses*, R. Weiss *et al.*, Ed. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, ed. 2, 1982), chap. 5, pp. 369–512.
5. A. T. Panganiban and D. Fiore, *Science* **241**, 1064 (1988).
6. W. Hu and H. M. Temin, *ibid.* **250**, 1227 (1990).
7. J. M. Coffin, *J. Gen. Virol.* **42**, 1 (1979).
8. F. Clavel *et al.*, *J. Virol.* **63**, 1455 (1989); W. Hu and H. M. Temin, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1556 (1990); D. W. Goodrich and P. H. Duesberg, *ibid.*, p. 2052; H. Stuhlmann and P. Berg, *J. Virol.* **66**, 2378 (1992).
9. G. Luo and J. Taylor, *J. Virol.* **64**, 4321 (1990).
10. J. Garces and R. Wittek, *Proc. R. Soc. London Ser. B* **243**, 235 (1991).
11. H. E. Huber, J. M. McCoy, J. S. Seehra, C. C. Richardson, *J. Biol. Chem.* **264**, 4669 (1989); R. G. Buiser, J. J. DeStefano, L. M. Mallaber, P. J. Fay, R. A. Bambara, *ibid.*, **266**, 13103 (1991).
12. V. Gopalakrishnan, J. A. Peliska, S. J. Benkovic, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
13. E. S. Furfine and J. E. Reardon, *J. Biol. Chem.* **266**, 406 (1991).
14. E. Arnold *et al.*, *Nature* **357**, 85 (1992).
15. L. A. Kohlstaedt, J. Wang, J. M. Friedman, P. A. Rice, T. A. Steitz, *Science* **256**, 1783 (1992).
16. L. Ratner *et al.*, *Nature* **313**, 277 (1985).
17. The sequences of the oligonucleotides are:
    20-base DNA · 40-base RNA:
    5′-AGAGCTCCCAGGCTCAGATC-3′
    3′-UCUCGAGGGUCCGAGUCUAGACCAGAUU GGUCUCUCUGGG-5′
    41-base RNA:
    5′-ACCAGAUUGGUCUCUCUGGGUCAUGUCC GUUUUUCGUCGAG-5′
    RNA oligonucleotides were synthesized by runoff transcription with T7 RNA polymerase [J. F. Milligan and O. C. Uhlenbeck, *Methods Enzymol.* **180**, 51 (1989)] and purified by electrophoresis through 20 percent acrylamide–8 M urea and TBE (tris, borate, EDTA). RNA quantitation was facilitated by the inclusion of trace amounts of [α-^{32}P]UTP in the transcription reaction. Oligonucleotides were 5′ end-labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase [as indicated in the protocol of USB]. Before RNA oligonucleotides were labeled, they were dephosphorylated with calf intestine alkaline phosphatase. The labeled RNA oligonucleotides were subjected to gel purification as described above. The DNA · RNA hybrids were formed from a mixture of equal amounts of each oligonucleotide in 50 mM tris-HCl, pH 7.5, that were heated to 70°C, and then slow cooled to room temperature. Hybrids were then purified by nondenaturing gel electrophoresis through 20 percent acrylamide and TBE.
18. The recombinant HIV-1 RT used in our study was expressed as described by V. Mizrahi *et al.* [*Arch. Biochem. Biophys.* **273**, 347 (1989)] and purified according to a large-scale adaptation of the method described therein by D. Woolf and Dr. J. Culp (SmithKline Beecham Pharmaceuticals).
19. V. Mizrahi, T. M. Usdin, A. Harington, L. R. Dudding, *Nucleic Acids Res.* **18**, 5359 (1990).
20. M. S. Collett, P. Dierks, J. T. Parsons, A. J. Faras, *Nature* **272**, 181 (1978); J. L. Darlix, P. A. Bromley, P. F. Spahr, *J. Virol.* **23**, 659 (1977).
21. The hybridization of 40-base DNA to 41-base RNA was examined by equilibrating 5′ end-labeled 40-base [^{32}P]DNA (300 nM) and 41-base RNA (500 nM) at 37°C in assay buffer (Fig. 3) containing 7 mM $MgCl_2$ and 150 μM each dNTP. At intervals up to 30 minutes, reaction samples were removed and added to HIV-1 RT (255 nM). After 40 seconds, the reaction was quenched with EDTA, and the elongated DNA product (representing formed hybrid) was determined by denaturing gel electrophoresis through 20 percent acrylamide, 8 M urea, and TBE. For hybridizations in the presence of HIV-1 RT the enzyme was included in the reaction mixture, extending the DNA primer as it is hybridized to the 41-base RNA. The reaction products were analyzed as a second-order reaction by standard procedures [W. P. Jencks, *Catalysis in Chemistry and Enzymology* (Dover, New York, 1987)].
22. J. M. Clark, C. M. Joyce, G. P. Beardsley, *J. Mol. Biol.* **198**, 123 (1987); J. M. Clark, *Nucleic Acids Res.* **16**, 9677 (1988).
23. The rapid elongation of the 40-base DNA · 41-base RNA putative strand transfer intermediate by HIV-1 RT was independently confirmed by thermally annealing 40-base DNA to 41-base RNA as in Fig. 7B.
24. The relative rates of blunt-end addition of the four possible nucleotide triphosphates was determined by incubating each of the dNTP's (150 μM) in the presence of 7 mM $MgCl_2$, 100 nM HIV-1 RT, and 300 nM (5′ end)–labeled 40-base [^{32}P]DNA · 40-base RNA hybrid in reaction buffer. At times from 0 to 40 minutes, reaction samples were withdrawn and quenched by the addition of EDTA (120 mM). The samples were diluted with buffer (90 percent formamide), heated to 90°C for 5 minutes, and then placed directly onto a 20 percent acrylamide, 8 M urea, TBE gel. The substrate and product bands were detected by autoradiography and the individual bands were excised and subjected to scintillation counting. The concentration of 41-base DNA was plotted as a linear function of time and the observed rate constants were obtained by linear regression analysis. The resulting rate constants were (nM/min): dATP, 4.4; dGTP, 2.3; dTTP, 0.84; and dCTP, 0.29.
25. J. E. Reardon, *Biochemistry* **31**, 4473 (1992).
26. The T_m for the DNA · RNA hybrid was estimated to be approximately 5° to 10°C higher than that calculated for the DNA · DNA hybrid (30°C); J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, ed. 2, 1989), p. 9.51.
27. N. Cheng, G. R. Painter, P. A. Furman, *Biochem. Biophysical Res. Commun.* **174**, 785 (1991).
28. S. F. J. LeGrice, T. Naas, B. Wohlgensinger, O. Schatz, *EMBO J.* **10**, 3905 (1991); Z. Hostomsky, Z. Hostomska, T. Fu, J. Taylor, *J. Virol.* **66**, 3179 (1992).

29. R. W. Sobol *et al.*, *Biochemistry* **30**, 10623 (1991); A. Basu, K. K. Ahluwalia, S. Basu, M. J. Modak, *ibid.* **31**, 616 (1992).
30. C. Barat *et al.*, *EMBO J.* **8**, 3279 (1989); N. J. Richter-Cook *et al.*, *J. Biol. Chem.* **267**, 15952 (1992).
31. Size dependence was determined as follows. DNA oligonucleotides of 15, 20, 25, 36, 37, 40, and 57 bases were crosslinked to heterodimer HIV-1 RT, and the products were analyzed by PAGE through 8 percent acrylamide, 0.1 percent SDS. Autoradiography was used to locate crosslinked protein. Unlabeled HIV-1 RT heterodimer on the same gel was located by staining with Coomassie blue. Two crosslinked bands were observed with oligonucleotides exceeding 25 bases in length, whereas the smaller DNA strands showed only a single band (Fig. 6C). The R_f values of the single bands of the smaller DNA crosslinks and the upper bands of the doublets show a linear correlation with the log of the molecular mass, which extrapolates to the observed R_f for unlabeled HIV-1 RT p66. Likewise, the lower bands of the doublet crosslinks show a similar correlation with HIV-1 RT p51. The slopes of the two lines are the same as would be expected if DNA crosslinking affected p51 and p66 migration equally. Because of the low efficiency of the crosslinking, no doubly crosslinked protein was observed—that is, between DNA and p51-p66 (~120 kD)—since statistically this would at maximum represent 0.04 percent of the total protein (*32*).
32. The maximal crosslinking of DNA and RNA complexes with RT was observed to be 2 percent. This was achieved at irradiation times of 120 seconds (see Fig. 6 for details), while significant protein photodegradation was observed at longer irradiation times. However, crosslinking efficiency was sacrificed slightly if we used irradiation times of less than 20 seconds in order not to measurably disturb the binding equilibrium RT + DNA $\rightleftharpoons$ RT · DNA. This ensures that the observed crosslinking represents a thermodynamic process (*28*).
33. The K_d for the complex of enzyme to the 20-base DNA · 40-base RNA was determined kinetically with the use of a heparin trap to sequester free HIV-1 RT and limit polymerization to a single turnover. The HIV-1 RT (25 nM) with varying amounts of 5′ end-labeled 20-base [^{32}P]DNA · 40-base RNA (5 to 160 nM) was first equilibrated at 37°C for 5 minutes, and the reaction was then initiated by the addition of a mixture containing dTTP (150 μM) (the next correct base to incorporate (*17*)], 7 mM $MgCl_2$, and heparin at 400 μg/ml. At times up to 2 minutes, samples were withdrawn and quenched into EDTA (110 mM). The products were analyzed by denaturing gel electrophoresis and the 21-base DNA concentration was determined by band excision and scintillation counting. The concentration of elongated primer [representing the concentration of primer-template to the enzyme and in the preliminary equilibration mixture (*25*)] was plotted as a function of template-primer and the data were fitted to the quadratic [RT:20 base DNA · 40-base RNA] = $0.5(K_d + [RT] + [20\text{-base DNA} \cdot 40\text{-base RNA}]) - \{0.25(K_d + [RT] + [20\text{-base DNA} \cdot 40 \text{ base RNA}])\}^2 - ([RT][20\text{-base DNA} \cdot 40 \text{ base RNA}])^{1/2}$ to yield a $K_d = 28 \pm 7$ nM. The K_i for the inhibition of 41-base RNA crosslinking to HIV-1 RT by 20-base DNA · 40 base RNA was determined by incubating 5′ end-labeled ^{32}P 41-base [^{32}P]RNA (400 nM) with HIV-1 RT (200 nM) and varying amounts of 20-base DNA · 40-base RNA (0 to 860 nM) and subjected to ultraviolet as in the legend of Fig. 6. The irradiated samples were added to SDS loading buffer, heated to 100°C for 5 minutes, and then subjected to electrophoresis through 10 percent acrylamide, 0.1 percent SDS. The bands corresponding to crosslinked 41-base RNA · HIV-1 RT were located by autoradiography, excised, and quantitated by scintillation counting. The K_i was estimated as the concentration of 20-base DNA · 40-base RNA required to reduce the crosslinking by 50 percent.
34. M. Goodenow *et al.*, *J. Acquired Immune Defic. Syndr.* **2**, 344 (1989).
35. J. D. Roberts, K. Bebenek, T. A. Kunkel, *Science* **242**, 1171 (1988); B. D. Preston, B. J. Poiesz, L. Loeb, *ibid.*, p. 1168; J. Weber and F. Grosse, *Nucleic Acids Res.* **17**, 1379 (1989); K. Bebenek, J. Abbotts, J. D. Roberts, S. H. Wilson, T. A. Kunkel, *J. Biol. Chem.* **264**, 16948 (1989); M. Ricchetti and H. Buc, *EMBO J.* **9**, 1583 (1990); V. K. Pathak, H. M. Temin, *Proc. Natl. Acad. Sci., U.S.A.* **87**, 6019 (1990); J. Ji and L. Loeb, *Biochemistry* **31**, 954 (1992); K. Bebenek, J. D. Roberts, T. A. Kunkel, *J. Biol Chem.* **267**, 3589 (1992); H. Yu and M. F. Goodman, *ibid.*, p. 10888; J. C. Boyer, K. Bebenek, T. A. Kunkel, *Proc. Natl. Acad. Sci., U.S.A.* **89**, 6919 (1992).
36. We thank C. Debouck for the wild-type HIV-1 RT and V. Mizrahi for the RNase H (−) mutant. Supported in part by NIH grant GM13306 and by NIH fellowship AI08275 (J.A.P.).

22 June 1992; accepted 28 September 1992