# RNA Secondary Structure Switching during DNA Synthesis Catalyzed by HIV-1 Reverse Transcriptase[†]

Zucai Suo and Kenneth A. Johnson*

*Department of Biochemistry and Molecular Biology, 106 Althouse Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802*

ABSTRACT: Changes in RNA secondary structure have been found to play important roles in translational regulation, protein synthesis, and mRNA splicing. In studies utilizing a 66 nucleotide RNA template with a stable hairpin structure, we have examined the effects of RNA secondary structure on HIV-1 reverse transcriptase activity. We identify several pause sites in the stem of the hairpin and show that these pause sites are correlated with the free energy of melting the next base pair in the stem. We also identify a pause site appearing in the loop of the hairpin and show that this is due to the rapid formation of a new hairpin structure occurring during the progress of DNA polymerization through the hairpin. The rapid change in RNA secondary structure to form the new hairpin selectively destabilizes the major hairpin and thereby accelerates the rate at which reverse transcriptase reads through RNA secondary structure.

During the replication of human immunodeficiency virus type 1 (HIV-1),[1] virally encoded reverse transcriptase (RT) copies single-stranded viral RNA genome into a minus-strand of DNA. The enzyme then uses the resultant cDNA to synthesize a plus-strand DNA, thereby allowing the double-stranded viral genome to be integrated into the DNA of the host (*1, 2*). The HIV-1 RNA genome forms complex secondary and tertiary structures which cause RT to pause during DNA synthesis. RT pausing has been observed in both stems and loops of predicted hairpin structures and in runs of template guanosines ($\geq 4$ bases) and cytidines ($\geq 3$ bases); therefore, the role of RNA secondary structure in pausing has been questioned (*3, 4*).

We present, in a series of three papers, the effect of RNA secondary structure on DNA synthesis and RNA cleavage catalyzed by HIV-1 RT. In this paper, a 66 nucleotide RNA template, derived from the HIV-1 genome (nucleotides 1636−1701) (*5*) and predicted to form a stable hairpin structure (Figure 1), was used to examine the effect of RNA secondary structure on DNA synthesis catalyzed by HIV-1 RT *in vitro* using rapid kinetic methods. This stable hairpin is seven nucleotides downstream from the *gag-pol* frameshift site of the HIV-1 genome and is essential for wild-type levels of translational frameshifting to occur *in vivo* (*6*). Our results reveal that HIV-1 RT pauses significantly at several positions in this RNA hairpin including one site in the loop of the predicted hairpin. We show that this unexpected pause site is due to a new RNA hairpin secondary structure which is formed rapidly during DNA synthesis. The second paper describes the effect of RNA secondary structure on the kinetics of polymerization catalyzed by HIV-1 RT (*7*). In the third paper of the series, we describe the effect of RNA
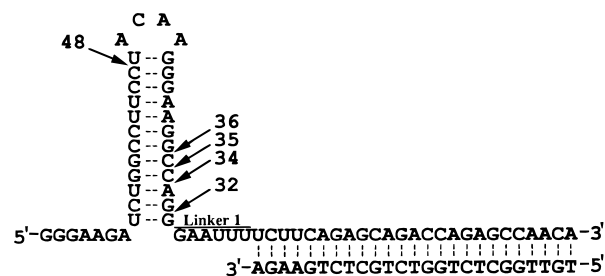


FIGURE 1: DNA/RNA 25/66-mer substrate. The folding free energy of this hairpin at 37 °C was predicted to be −20.3 kcal/mol using *Mfold*. The arrows represent the strong RT pause sites numbered from the 3′ terminus of the RNA template. Linker 1 denotes the single-stranded template region between the primer and the hairpin stem.

secondary structure on RNA cleavage catalyzed by the RNase H activity of HIV-1 RT (*8*). These three papers provide an understanding of the effect of RNA secondary structure on the polymerase and the RNase H activities of HIV-1 RT structurally, thermodynamically, and kinetically.

## MATERIALS AND METHODS

*Processivity Experiment.* Wild-type HIV-1 RT, DNA, and RNA oligonucleotides were prepared as described (*7*). The processive polymerization experiment was performed using a KinTek (State College, PA) rapid chemical quench flow as described previously (*9*). Typically, the experiment was carried out by rapidly mixing 15 $\mu$L of preincubated RT and duplex DNA/RNA in RT buffer (50 mM Tris−acetate, 100 mM potassium acetate, 0.1 mM EDTA, pH 7.5 at 37 °C) with an equal volume of a solution containing dNTPs and $Mg^{2+}$. This mixture was allowed to react for the specified times. The reactions were terminated by quenching with 90 $\mu$L of 0.5 M EDTA (pH 8.0).

*RNA Secondary Structure Prediction.* The *Mfold* computer program was used to predict RNA structure and calculate folding free energy. *Mfold* is based on the energy

---

* Author to whom correspondence should be addressed. Telephone: (814) 865-1200. Fax: (814) 865-3030. Email: kaj1@psu.edu.
[1] Abbreviations: HIV-1, human immunodeficiency virus type 1; RT, reverse transcriptase; RNase, ribonuclease.
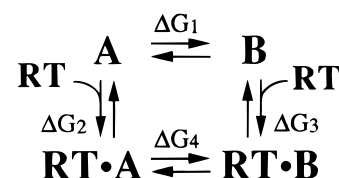
minimization method (*10*) and thermodynamic parameters measured for short oligonucleotides in the presence of 1 M $Na^+$ at pH 7.0 (*11*). Our reaction buffer is different and contains 10 mM $Mg^{2+}$ and 0.1 M $K^+$. However, it has been shown that thermodynamic parameters measured in 1 M $Na^+$ correspond closely to those measured in 0.15 M $Na^+$ and 10 mM $Mg^{2+}$ (*12*), and $K^+$ behaves similarly to $Na^+$ (*13*). The size of the single-stranded portion of our RNA template is in the same range as the oligonucleotides used for thermodynamic studies to obtain energy values (*11*). Moreover, the folding free energy of short oligonucleotides can be predicted to within 5%. Therefore, the results of structure prediction and thermodynamic calculation using *Mfold* provide a reasonable approximation for our experimental conditions. Furthermore, since we are interested in trends in the free energy of melting, the errors in the absolute values of the free energy changes are less important.

*Calculation of Free Energy Barriers.* The free energy change $\Delta G$ (37 °C) of each RNA optimal structure relative to its unfolded state was calculated using *Mfold*. The parameter $\Delta G$ (37 °C) was then used in the calculations: $\Delta\Delta G_i = \Delta G_i - \Delta G_{i-1}$, which represents the difference in stability of template annealed to an $i$ and $i-1$ nucleotide primer. That is, the free energy difference $\Delta\Delta G_i$ (37 °C) is equal to the free energy barrier of melting the corresponding base pair of stem before the incorporation of the $i$th nucleotide primed with an $i-1$ nucleotide oligomer.

*RNA Secondary Structure Determination.* The RNase H-deficient HIV-1 RT mutant D443N was prepared as described (*8*). The duplex DNA/RNA was formed by annealing a 5-fold primer with 5′-$^{32}$P-labeled RNA 66-mer at 85 °C for 5 min followed by cooling the mixture slowly to room temperature over 1 h. In the absence of RT, 13 nM DNA/RNA duplex was treated with either 2.0 units of RNase T1 (Pharmacia), or 1.1 units of RNase I (Promega), or 1.05 units of RNase V1 (Pharmacia), or 2.0 units of RNase U2 (Pharmacia) for 15 s at 37 °C in buffer containing 10 mM magnesium acetate and 0.54 mg/mL tRNA (Sigma). In the presence of RT, 13 nM duplex DNA/RNA was preincubated with 1.3 $\mu$M D443N and 1 mM next correct dNTP in buffer containing 0.54 mg/mL tRNA and then treated with one of the RNases for 10 s at 37 °C: 1.5 units of RNases T1, 2.0 units of RNase I, 3.0 units of RNase U2. The cleavage reaction by 1.05 units of RNase V1 was performed similarly except that dNTP was excluded and 10 mM magnesium acetate was added. The large molar excess of D443N, shorter reaction times, and the excess of next correct nucleotides were used to minimize the dissociation of substrates from RT (*14*). The next correct nucleotide was not included in each RNase V1 cleavage reaction because RNase V1 requires $Mg^{2+}$ for activity. Cleavage by RNases T1, I, and V1 occurred in RT buffer. Cleavage reactions by RNase U2 were performed in 33 mM sodium citrate buffer (pH 3.5). We also mapped the RNA secondary structures of 32/66-mer and 35/66-mer in the presence and absence of D443N. These species are early pause intermediates in the reaction while the hairpin is still intact as confirmed by our RNase analysis (data not shown). At these sites, hairpins similar to that seen with the 36/66-mer were identified except that the stems were larger as expected.

*Calculation of $\Delta G_4°$ of Scheme 1.* In Scheme 1, $\Delta G_1$ and $\Delta G_4$ are the free energy changes of the structure switching from A to B in the absence and presence of RT, respectively.

Scheme 1



$\Delta G_2$ and $\Delta G_3$ are the free energy changes of RT binding to structure A and B, respectively. The binding affinity ($K_d$) of 37/66-mer and 36/66-mer to RT at 23 °C is 8.8 and 38.4 nM, respectively, measured by the nitrocellulose−DEAE double filter assay (*7*). The binding affinity of structure A to RT could not be measured directly. We assume that it is similar to the affinity of 36/66-mer (38.4 nM) because 36/66-mer exists predominantly as structure A of 37/66-mer based on the results of RNase mapping. Using the equation $\Delta G° = -RT \ln K = RT \ln K_d$, we calculated values of $\Delta G_2°$ and $\Delta G_3°$ of Scheme 1 equal to −10.53 and −11.44 kcal/mol at 23 °C, respectively. $\Delta G_1$ was calculated to be +1.8 kcal/mol at 37 °C using *Mfold*. The net free energy change of the thermodynamic cycle from structure A→B→RT•B→RT•A→A, $\Delta G_1° + \Delta G_3° - \Delta G_4° - \Delta G_2°$, should be equal to zero. Therefore, $\Delta G_4°$ is equal to $\Delta G_1° + \Delta G_3° - \Delta G_2°$. Accordingly, $\Delta G_4°$ was estimated to be −2.71 kcal/mol at 37 °C assuming that the temperature effects on $\Delta G_2°$ and $\Delta G_3°$ are equal and therefore offset each other. Using the affinity of 37/66-mer (3.95 nM) and 36/66-mer (44.5 nM) at the polymerase site of RT measured by active site titration at 37 °C (*7*), $\Delta G_4°$ was estimated to be −3.3 kcal/mol.

## RESULTS AND DISCUSSION

*Processive Polymerization on 25/66-mer.* We first examined processive DNA polymerization through the hairpin structure catalyzed by wild-type HIV-1 RT. Elongation of the 25/66-mer duplex DNA (25-base DNA primer annealed to RNA 66-base template) was tested by rapidly mixing the enzyme−DNA complex with the four deoxynucleoside triphosphates at either 37, 21, or 8 °C and then stopping the reaction at various times. The products were analyzed using a sequencing gel as shown in Figure 2A. Products of intermediate length (32-, 34-, 35-, 36-, and 48-mer) accumulated to a significant extent.

The accumulation of intermediates did not result from polymerizing to the ends of broken templates because full-length product was formed after long times and the RNA template was stable in the polymerization reactions. The stability of the RNA template was examined by incubation in the presence and the absence of the RNase H-deficient HIV-1 RT mutant D443N (*8*) in reaction buffer at 37 °C for 30 min. With the RNase H-deficient mutant, there was negligible formation of RNA cleavage products.

We considered the possibility that the accumulation of the intermediate 36-mer resulted from "jumping" over the hairpin and reading the overlap GA sequence once by RT during the primer elongation (Figure 1). To test this possibility, we performed a similar 15 min reaction as in Figure 2A but leaving out dGTP. Normal elongation would result in the product 34-mer, while jumping would have produced a 36-mer. DNA sequencing gel analysis showed an intense 34-mer band but not a 36-mer band, suggesting that 36-mer
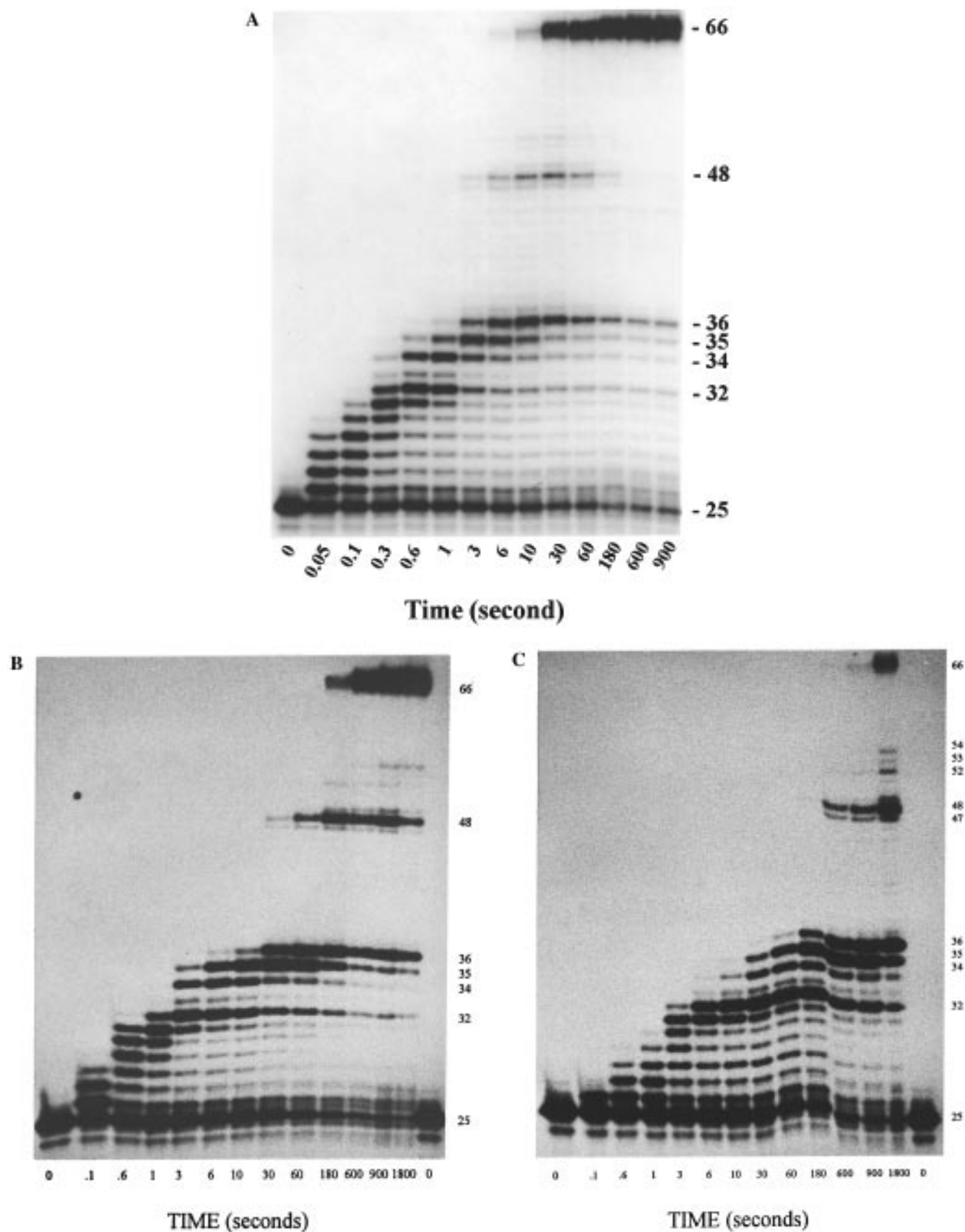
FIGURE 2: Processive polymerization on 25/66-mer. A preincubated solution of 100 nM HIV-1 RT and 100 nM 5′-$^{32}$P-labeled 25/66-mer in RT buffer was mixed with dATP, dGTP, dCTP, dTTP (150 $\mu$M each), and 10 mM magnesium acetate (final concentrations) in RT buffer. The reactions were quenched with 0.3 M EDTA at the indicated times and analyzed by sequencing gel electrophoresis. (A) pH 7.5, 37 °C; (B) pH 7.8, 21 °C; (C) pH 8.2, 8 °C.

pause product was synthesized by primer elongation into the hairpin. These results strongly suggest that the accumulation of intermediates is due to RT pausing caused by RNA secondary structure. This conclusion is further supported by the observation that the accumulation of intermediates is more significant at 8 °C and 21 °C than at 37 °C (Figure 2)
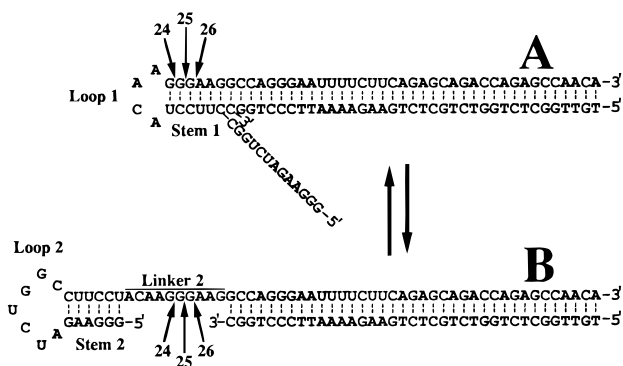
FIGURE 3: Two competing RNA secondary structures of 37/66-mer. The arrows indicate the cleavage sites numbered from the 5′ terminus of the RNA template.



FIGURE 4: Processive polymerization on 25/61-mer. A preincubated solution of 100 nM HIV-1 RT and 100 nM 5′-$^{32}$P-labeled 25/61-mer in RT buffer was mixed with dATP, dGTP, dCTP, dTTP (150 $\mu$M each), and 10 mM magnesium acetate (final concentrations) in RT buffer (pH 7.5) at 37 °C. The reactions were quenched with 0.3 M EDTA at the indicated times and analyzed by sequencing gel electrophoresis. The arrows represent the strong RT pause sites (32-, 34-, 35-, and 36-mer) and a weak RT pause site (41-mer) numbered from the 3′ terminus of the RNA template.

although RT was expected to be less active at lower temperatures and at slightly higher pH reaction conditions. Slight accumulation of 47-, 49-, 52-, 53-, and 54-mer was also observed more significantly at 8 °C (Figure 2C). However, the accumulation patterns at three temperatures are similar (Figure 2)

*RNA Secondary Structure Switching.* The sites of RT pausing after synthesis of products 32-, 34-, 35-, and 36-mer are expected because these sites correspond to GC base pairs in the stem of the predicted hairpin (Figure 1). However, the lack of 37-mer accumulation at 37 °C is surprising because the next base pair of the stem is also a GC base pair. RT pausing after synthesis of the 48-mer is another notable exception because the hairpin shown in Figure 1 should not exist when RT reaches this position. In order to understand the product accumulation pattern (Figure 2A), an RNA folding program, *Mfold*, was used to predict the optimal structures of the single-stranded portion of the template annealed to different primers. Using primers shorter than 37 nucleotides, the optimal RNA structures were predicted to correspond to structure A of Figure 3. However, utilizing primers from 37 to 49 nucleotides in length, the optimal RNA structures were as shown in structure B of Figure 3. The different optimal structures predicted by *Mfold* utilizing 36-mer and 37-mer primers suggest that the RNA template might rapidly change its secondary structure to form a new downstream hairpin after synthesis of the 37-mer (Figure 3). This postulate is supported by the kinetic studies of the next nucleotide incorporation and the studies of RNase H cleavage with the substrates 36/66-mer and 37/66-mer (*7, 8*).

According to this interpretation, the absence of pause sites between the 37-mer position and at sites leading up to the 48-mer is due to the open single-stranded region formed in the switch from hairpin A to hairpin B shown in Figure 3. Accordingly, all pause sites including the 48-mer pause site are located only in the predicted stems of stable hairpins. No pause sites are located in the single-stranded segments or loops. These conclusions are supported by the experimental results using a 61 nucleotide RNA template (Figure 4), which lacks the first 5 nucleotides of the 5′ terminus of the RNA 66-mer. *Mfold* predicted that there would be no RNA secondary structure change as shown in Figure 3 associated with this RNA 61-mer template. To test the hypothesis that the formation of hairpin B facilitates the unwinding of hairpin A, we examined processive DNA synthesis using the 61 nucleotide RNA template. As
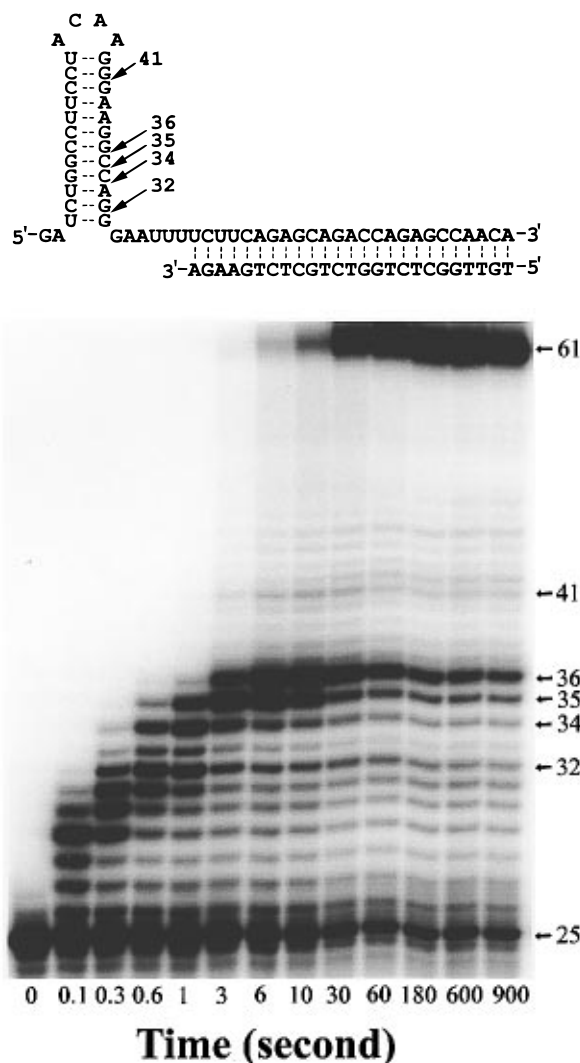
expected, the accumulation of 32-, 34-, 35-, and 36-mer (Figure 4) is similar to that shown in Figure 2A. However, intermediate products of 37, 38, and 41 nucleotides in length also accumulated slightly, and there was no accumulation of the 48-mer (Figure 4). This strongly suggests that the absence of 37-, 38-, and 41-mer products and the accumulation of DNA 48-mer with the 66-mer template are due to the formation of a new RNA hairpin.

As HIV-1 RT traverses RNA secondary structure, each stem base pair is melted in a stepwise fashion prior to each nucleotide incorporation. Our kinetic data suggest that the binding of RT to DNA exists in an equilibrium between a nonproductive binding state ($E \cdot D_n^N$) and a productive binding state ($E \cdot D_n^P$) at the strong RT pause sites (*7*). The conversion of the nonproductive binding state to productive binding state has to surpass a certain energy barrier. Our analysis suggests that the major contributor of this energy barrier is the free energy barrier of base pair unwinding when RT reads through the stem of the RNA hairpin. However,
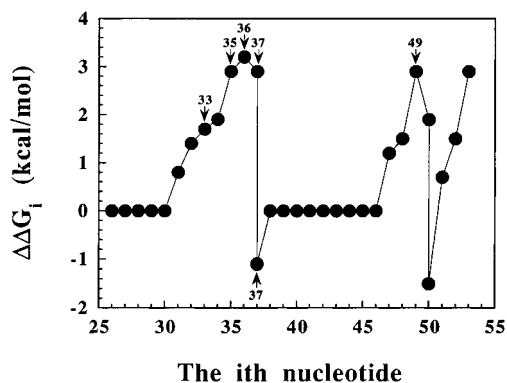
**The ith nucleotide**

FIGURE 5: Free energy barrier ($\Delta\Delta G_i$) of single-template (the RNA 66-mer) base-pair melting at 37 °C associated with the incorporation of the *i*th nucleotide. The 37th and the 50th nucleotide incorporations both have two different values of $\Delta\Delta G_i$ associated with each of them. The positive $\Delta\Delta G_i$ is the free energy barrier of melting the next template base pair before the incorporation of the *i*th nucleotide. The negative $\Delta\Delta G_i$ is the free energy change of the RNA secondary structure switching after the incorporation of the 37th or the 50th nucleotide.

at nonpause sites, RT reads single-stranded DNA and binds DNA only productively and, therefore, there is no energy barrier of base pair unwinding. Once DNA is bound productively by RT, there are no differences in the processes of nucleotide incorporations at pause and nonpause site because the fast rates of nucleotide incorporations are similar at all sites (*7*). This suggests that the major difference of free energy change associated with DNA polymerization through the single-stranded and double-stranded regions of the RNA template is from the base pair unwinding processes. The different affinity of RT toward the single-stranded and the double-stranded DNA should have a minor effect on the overall free energy change of DNA polymerization because the nitrocellulose−DEAE double filter binding assay has revealed that all DNA substrates have similar affinity toward RT (*7*). Therefore, the free energy barrier of base pair unwinding should have a correlation to RT pausing.

The predicted free energy barrier of each melting step, $\Delta\Delta G_i$ (37 °C), varies with the position of the incorporated nucleotide within the RNA 66-mer template (Figure 5). It is notable that the pattern of the free energy of unwinding as a function of position through the hairpin (Figure 5) correlates well with the product accumulation pattern (Figure 2A) in that the pause sites occur at positions with predicted high energy barriers for unwinding. For example, the free energy barriers for the unwinding of 35th, 36th, 37th, and 49th nucleotides are notably high (∼3 kcal/mol), leading to a large accumulation of 34-, 35-, 36- and 48-mer products. The melting steps predicted to occur after the incorporation of the 50th nucleotide are not important at 37 °C because the RNA secondary structure predicted by the program may not exist in the presence of RT, although they may have contributed to the slight accumulation of 52-, 53-, and 54-mer at 8 °C (Figure 2C).

It is interesting to note that $\Delta\Delta G$ (at 37 °C) is calculated to be −1.1 kcal/mol following synthesis of the 37-mer product. This favorable free energy change results from the formation of hairpin B since it is predicted to be more stable than hairpin A after the synthesis of the 37 nucleotide product. Thus, the formation of the second hairpin downstream facilitates the unwinding of the first hairpin (see discussions below).

The overall free energy change for single-nucleotide incorporation catalyzed by T7 DNA polymerase at 20 °C was estimated to be approximately −4 kcal/mol (*15*), and we expect a similar overall free energy change for single-nucleotide incorporation catalyzed by HIV-1 RT. Although not directly coupled, this is the free energy available as the net driving force to unwind the next stem base pair. Although the mechanism of stepwise base pair melting while reading through the hairpin has not been established, the most plausible mechanism supported by our data is the thermal breathing of individual stem base pairs followed by the rapid incorporation of the next nucleotide. This simple model accounts for the direct correlation between the free energy of unwinding and the magnitude of the pause, without the need to invoke any special assumptions relating equilibria and kinetics. At the pause sites, it is the amplitude of the fast reaction, not the rate, which is affected. Therefore, the free energy of unwinding determines the fraction of enzyme molecules that can bind in the productive mode. Although this simple model qualitatively accounts for the pausing, this work does not rule out the possible involvement of the protein in facilitating the melting of the RNA secondary structure by either stabilizing the productively bound template or destabilizing the secondary structure.

von Hippel and Yager (*16*) have developed a theoretical model to account for the changes in the termination frequency from less than 1% to greater than 99% during RNA synthesis. They have shown how small changes in the energy barriers (3−4 kcal/mol) cause a switch in kinetics from elongation to termination. In addition, they define the net free energy for protein binding to a transcription bubble in terms of the sum of the elementary contributions from changes in DNA and RNA structure and protein binding. We have applied similar analysis to the present system by defining the free energy changes associated with RNA secondary structure unwinding. However, in this case, we show that the kinetic effect is indirect by altering the fraction of RNA in the productively bound state during reverse transcription.

*RNA Secondary Structure Determination.* To examine the role of RT in the observed change in template secondary structure during polymerization, we used RNases T1, U2, I, and V1 to map the secondary structure of the RNA 66-mer annealed with different primers in the absence (Figure 6A) and presence (Figure 6B) of the RNase H-deficient RT mutant D443N. As shown, the RNase H activity is not important for RT to traverse RNA secondary structure since we observed an identical product accumulation pattern obtained with D443N by performing the same processivity experiment described for Figure 2A (*8*). RNases T1 and U2 cleave single-stranded RNA specifically on the 3′ side of G and A residues, respectively. RNases I and V1 cleave predominantly single-stranded and double-stranded RNA, respectively. The major difference between hairpins A and B is that nucleotides 24−29 including G24, G25, and G26 are located in stem 1 of structure A, but in a single-stranded segment of structure B. Thus, the sensitivity toward digestion by RNase T1 at sites G24, G25, and G26 serves as a signal for the change in RNA secondary structure from hairpin A to hairpin B.

In the absence of RT (Figure 6A), RNases T1, U2, and I recognized "linker 1" of the 25/66-mer, loop 1 of 25/66-mer and 36/66-mer of our predicted structure A, and loop 2 of 48/66-mer of our predicted structure B (Figure 3).
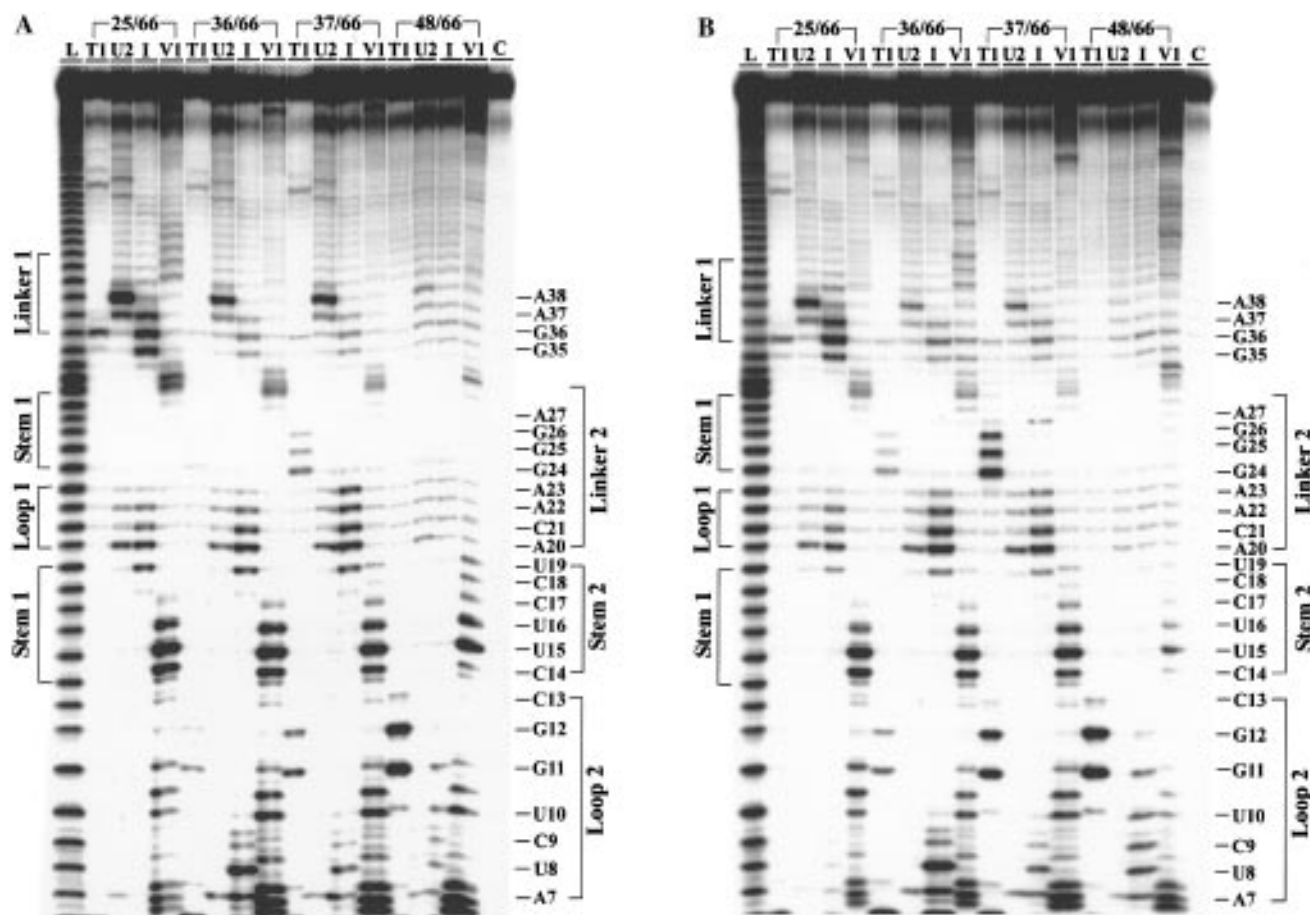
FIGURE 6: RNase-probing analysis at 37 °C of the 5′-³²P-labeled RNA 66-mer annealed to different DNA primers. Lanes L and C denote the alkaline hydrolysis ladder and the control RNase reaction without ribonuclease, respectively. (A) In the absence of HIV-1 RT. RNases T1, U2, I, and V1 were 2.0, 2.0, 1.1, and 1.05 units, respectively. Cleavage time was 15 s. (B) In the presence of HIV-1 RT. RNases T1, U2, I, and V1 were 1.5, 3.0, 2.0, and 1.05 units, respectively. Cleavage time was 10 s.

Table 1: Total Molar Percentage of Three Cleavage RNA Products after Nucleotides G24, G25, and G26 by RNase T1 in the Presence of RT (+RT) or Absence of RT (−RT) and the Ratio between Them (+RT)/(−RT)

| substrate | −RT (%) | +RT (%) | (+RT)/(−RT) |
|---|---|---|---|
| 37/66-mer | 2.49 | 9.37 | 3.8 |
| 36/66-mer | 0.62 | 1.84 | 3.0 |

Digestion by RNase V1 shows the products characteristic of the predicted stems of the 25/66-mer, 36/66-mer, and 48/66-mer. The RNA cleavage pattern of 37/66-mer is similar to that of 36/66-mer except for the appearance of the weak bands of G24, G25, and G26. This suggests that the majority of 37/66-mer exists as structure A, implying that structure A is more stable than structure B in the absence of RT. However, even at lower concentrations of RNase T1 (1.5 units instead of 2.0 units) and a shorter reaction time (10 s instead of 15 s), the total percentage of the three bands of G24, G25, and G26 in the presence of RT (Figure 6B) is almost 4 times greater than in the absence of RT (Table 1). Additionally, the intensity of bands G11 and G12 increased. These two positions are relatively more sensitive to RNase T1 when they are in loop 2 of structure B than when they are in the 5′ dangling end of structure A (Figure 3). Interestingly, the total percentage of these five bands (G11, G12, G24, G24, and G25) of 37/66-mer is approximately equal to the total percentage of cleavage at G11 and G12 of 48/66-mer, which are the only sites sensitive to RNase T1. These results strongly suggest that 37/66-mer existed in

solution predominantly as structure B in the presence of RT. In the presence of the RNase H-deficient mutant of RT, the RNA cleavage patterns of 25/66-mer, 36/66-mer, and 48/66-mer (Figure 6B) are similar to those in the absence of RT (Figure 6A).

The intensity of bands G24, G25, and G26 as products of RNase T1 digestion provides markers to quantify the change in RNA secondary structure in forming hairpin B. These bands are not evident after digestion of the RNA in the presence of DNA primers 32 or 35 nucleotides in length (32/66-mer and 35/66-mer), indicating stable formation of hairpin A with little formation of hairpin B (data not shown). However, bands G24, G25, and G26 are prominent in the digestion products of the 36/66-mer or 37/66-mer. The intensity of these three bands is approximately 5-fold greater with the 37/66-mer substrate, suggesting that perhaps only 20% of the 36/66-mer has a secondary structure similar to the 37/66-mer in the presence of RT. This agrees with the observation that the fast phase reaction amplitude with 37/66-mer is about 6-fold larger than with 36/66-mer (*7*). However, the addition of RT to the 36/66-mer increases the intensity of these bands by approximately 3-fold, directly showing that the binding of RT induces the change in RNA secondary structure from hairpin A to hairpin B with DNA primers of 36 or 37 nucleotides in length.

Hydroxyl radical and DNase I footprinting analyses have previously shown that HIV-1 RT protects the template nucleotides from +3 to −15 and from +7 to −23, respec-

tively (+3 represents three nucleotides downstream from the 3′ terminus of the primer) (*17*, *18*). Our RNase footprinting analysis demonstrates that the downstream extremity of the region protected by RT extends only to position +3 or nucleotide A27. The cleavage after nucleotides G24, G25, and G26 by RNase T1 is probably due to the downstream template overhang close to the RT binding site being partially protected by the $\beta3-\beta4$ hairpin of the p66 finger subdomain of HIV-1 RT (*17*, *18*) and thereby accessible to RNase T1 molecules (11 kDa) and hydroxyl radical, which are smaller than DNase I molecules (31 kDa).

*HIV-1 RT Facilitates RNA Secondary Structure Change.* Scheme 1 shows our model for the RNA secondary structure switching from structure A to B (Figure 3) in the presence and absence of RT. The isolated hairpin of structure B is calculated to be 1.1 kcal/mol more stable than the lone hairpin of structure A at 37 °C. However, the coaxial stacking of adjacent helices between the hairpin stem and the primed double-stranded region in structure A is expected to stabilize the hairpin of structure A by 2.9 kcal/mol (*19*, *20*). Structure B does not have such coaxial stacking because its two helices are separated by 10 nucleotides. Therefore, $\Delta G_1$ of Scheme 1 was calculated to be 1.8 kcal/mol at 37 °C. The positive value indicates that structure A is more stable than structure B in the absence of RT, which agrees with the result of RNase mapping. However, the coaxial stacking in structure A could be eliminated by the binding of RT, thereby triggering a switch of RNA secondary structure to form hairpin B. This is reasonable because the hairpin of structure A and its large 5′ dangling end may have physical contact with the RT binding cleft, disrupting the coaxial stacking of the two stems of structure A. Considering this factor, $\Delta G_4$ was predicted to be −1.1 kcal/mol. Moreover, our direct measurements indicate that RT has higher affinity to substrate B than to substrate A (*7*) which further shifts the equilibrium from RT•A to RT•B such that $\Delta G_4$ will be more negative (−2.7 to ∼−3.3 kcal/mol at 37 °C). Thus, the RNA structure change is thermodynamically favorable due to the binding of RT. This conclusion is further supported by the RNase mapping result showing that structure B is more stable than structure A in the presence of RT.

Interestingly, the pause sites previously located in the predicted loops of RNA template derived from the HIV-1 *gag*, *pol*, and *env* genes (*3*, *4*) can be relocated in the stems of new hairpin structures of shortened single-stranded template predicted by *Mfold* after DNA primer extension approaches the pause sites. Surprisingly, the predicted template secondary structure rearrangements only occur after extension of the primers into the hairpin, suggesting that template secondary structure is not static during DNA synthesis and that changes in structure play a major role in RT pausing. The switch in RNA secondary structure from hairpin A to hairpin B may provide a pathway to help HIV-1 RT traverse the complex secondary structures of the viral RNA genome during minus-strand DNA synthesis. Moreover, this may be a generally applicable phenomena involving the existence of tandem hairpin structures whereby formation of the smaller downstream hairpin facilitates the unwinding of the larger upstream hairpin.

Dynamically, RNA secondary structure switching must be faster than the rate-limiting step leading to the incorporation of the next nucleotide (38-mer at a rate of 70 s$^{-1}$) (*7*), to account for the lack of accumulation of DNA 37-mer product at 37 °C (Figure 2A). This indicates that RNA secondary structure switching occurs in milliseconds. This fast switching is not without precedent. RNA structure switching of *Leptomonas collosoma* spliced leader RNA was seen to occur at a similar rate (*21*, *22*). The rapid time scale of RNA structure switching suggests that it may occur by a branch migration pathway (*21*). For example, the possible interaction between the 5′ dangling end and stem 1 of structure A could provide a nucleus such as G1U19, which may then grow into structure B at the expense of structure A. Cleavage products shorter than 12 nucleotides in length produced by RNase V1 (Figure 6) support the existence of such an interaction. The alternative pathway for RNA structure change involving complete disruption of structure A followed by formation of structure B is unlikely because the disruption of structure A occurs with a large positive enthalpy change (59.6 kcal/mol) (*23*).

RNA secondary structure switching has been found to play a significant role in several biological processes, including translational regulation (*24*, *25*), protein synthesis (*26*), and mRNA splicing (*27*). We demonstrate here that RNA secondary structure switching also plays an important role in replication catalyzed by HIV-1 RT *in vitro*. Additionally, we show that RT pausing in the stems of RNA hairpins correlates with the predicted energy barrier of melting next-stem base pair. These results could apply to other reverse transcriptases and RNA polymerases.

## REFERENCES

1. Hoffman, A. D., Banapour, B., and Levy, J. A. (1985) *Virology* 147, 326−335.
2. Huber, H. E., Macoy, J. M., Seehra, J. S., and Richardson, C. C. (1989) *J. Biol. Chem. 264*, 4669−4678.
3. Dudding, L. R., Nkabinde, N. C., and Mizrahi V. (1991) *Biochemistry 30*, 10498−10506.
4. Klarmann, G., Schauber, C. A., and Preston, B. D. (1993) *J. Biol. Chem. 268*, 9733−9802.
5. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R., Jr., Pearson, M. L., Lautenberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C., and Wong-Staal, F. (1985) *Nature 313*, 277−284.
6. Parkin, N. T., Chamorro, M., and Varmus, H. E. (1992) *J. Virol. 66*, 5147−5151.
7. Suo, Z., and Johnson, K. A. (1997a) *Biochemistry 36*, 12459−12467.
8. Suo, Z., and Johnson, K. A. (1997b) *Biochemistry 36*, 12468−12476.
9. Johnson, K. A. (1986) *Methods Enzymol. 134*, 677−705.
10. Zuker, M. (1989) *Science 244*, 48−52.
11. Jaeger, J. A., Turner, D. H., and Zuker, M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7706−7710.
12. Williams, A. P., Longfellow, C. E., Freier, S. M., Kierzek, R., and Turner, D. H. (1989) *Biochemistry 28*, 4283−4291.
13. SantaLucia, J., Jr., Allawi, H. T., and Seneviratne, P. A. (1996) *Biochemistry 35*, 3555−3562.
14. Kati, W. M., Johnson, K. A., Jerva, L. F., and Anderson, K. S. (1992) *J. Biol. Chem. 267*, 25988−25997.
15. von Hippel, P. H., and Yager, T. D. (1992) *Science 255*, 809−812.

16. Patel, S. S., Wong, I., and Johnson, K. A. (1991) *Biochemistry 30*, 511−525.

17. Metzger, W., Hermann, T., Schatz, O., and Le Grice, S. F. J. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5909−5913.

18. Wöhrl, B. M., Tantillo, C., Arnold, E., and Le Grice, S. F. J. (1995) *Biochemistry 34*, 5343−5350.

19. Walter, A. E., and Turner, D. H. (1994) *Biochemistry 33*, 12715−12719.

20. Walter, A. E., Turner, D. H., Kim, J., Lyttle, M. H., Muller, P., Mathews, D. H., and Zuker, M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 9218−9222.

21. LeCuyer, K. A., and Crothers, D. M. (1993) *Biochemistry 32*, 5301−5311.

22. LeCuyer, K. A., and Crothers, D. M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 3373−3373.

23. Craig, M. E., Crothers, D. M., and Doty, P. (1971) *J. Mol. Biol. 62*, 383−401.

24. Putzer, H., Gendron, N., and Grunberg-Manago, M. (1992) *EMBO J. 11*, 3117−3127.

25. Babitzke, P., and Yanofsky, C. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 133−137.

26. Wool, I. G., Glück, A., and Endo, Y. (1992) *Trends Biochem. Sci. 17*, 266−269.

27. Madhani, H. D., and Guthrie, C. (1992) *Cell 71*, 803−817.

BI971963M