DNA Secondary Structure Effects on DNA Synthesis Catalyzed by HIV-1 Reverse Transcriptase*

(Received for publication, May 20, 1998, and in revised form, August 5, 1998)

Zucai Suo‡ and Kenneth A. Johnson§
From the Department of Biochemistry and Molecular Biology, the Pennsylvania State University, University Park, Pennsylvania 16802

The effect of DNA secondary structure on polymerization catalyzed by human immunodeficiency virus (HIV-1) reverse transcriptase (RT) was studied using a synthetic 66-nucleotide DNA template containing a stable hairpin structure. Four RT pause sites were identified within the first half of the hairpin stem. Additionally, five weak pause sites within the second half of the stem and the loop of the hairpin were identified at low temperatures. These weak pause sites were relocated to the site of the first few stem base pairs of two new hairpins formed due to a change in DNA secondary structure. Each pause site was correlated with a high free energy barrier of melting the stem base pair. Pre-steady state kinetic analysis of single nucleotide incorporation showed that polymerization at each pause site occurred by both a fast phase (10–20 s⁻¹) and a slow phase (0.02–0.07 s⁻¹) during a single binding event. The reaction amplitudes of the fast phase were small (4–10% of enzyme sites), whereas the amplitudes of the slow phase were large (14–40%) at the pause sites. In contrast, only a single phase with a large reaction amplitude (32–50%) and a fast nucleotide incorporation rate (33–87 s⁻¹) was observed at the non-pause sites. DNA substrates at all sites had similar dissociation rates (0.14–0.29 s⁻¹) and overall binding affinity (16–86 nM). These results suggest that the DNA substrates at pause sites were bound in both productive and non-productive states at the polymerase site of RT. The non-productively bound DNA was slowly converted into a productive state upon melting of the next stem base pair without dissociation of the DNA from RT.

Reverse transcriptase (RT)¹ encoded by human immunodeficiency virus type 1 (HIV-1) is responsible for converting the viral single-stranded RNA genome into a double-stranded proviral DNA, which is then integrated into the host genome (1). HIV-1 RT has three activities as follows: RNA-dependent and DNA-dependent polymerases and RNase H (2). RT is a heterodimer of a 66-kDa subunit (p66) and a 51-kDa subunit (p51). The p51 subunit is a proteolytic fragment of the p66 subunit formed by the viral protease. The p66 subunit possesses polymerase and RNase H activities, whereas the p51 subunit lacks RNase H activity, and its polymerase active site is buried by the p66 subunit in the RT heterodimer (3).

During HIV-1 replication (1), RT first copies the viral RNA genome into a minus strand cDNA using a tRNA³⁰⁰⁻ as a primer. During polymerization of the cDNA strand, the RNase H activity hydrolyzes the entire RNA template with the exception of two short polypurine tracts. Then, RT utilizes the two polypurine tracts as primers to synthesize plus strand DNA and yields two discrete segments. Finally, both minus strand and plus strand strands are fully replicated through a strand displacement DNA synthesis mechanism.

During DNA synthesis, putative DNA secondary structures and template sequences have been found to arrest DNA polymerase α (4–6), Escherichia coli DNA polymerase III (7), and phage T4 DNA polymerase (8, 9). DNA secondary structures formed from palindromic and quasipalindromic sequences were assumed to facilitate interstrand misalignment of nucleotide repeats and may be responsible for replication-based mutations (10, 11). It has also been observed in vitro that HIV-1 RT pausing occurs not only in most runs (≥4 bases) of template adenosines and thymidines but also within the first few nucleotides of the predicted DNA secondary structures during primer elongation with large DNA templates derived from the gag, pol, and env genes of HIV-1 genome (12–14). HIV-1 RT pausing during plus strand synthesis was found to significantly enhance mutational frequency (14). Thus, potential DNA secondary structures may have significant effects on processivity and fidelity of DNA polymerases, but the effects of DNA secondary structures on DNA synthesis have not been directly studied to date.

Previously, we have utilized a 66-nucleotide RNA template containing a stable RNA hairpin to study effects of RNA secondary structure on polymerization catalyzed by HIV-1 RT in detail (15–17). The DNA analog of the 66-nucleotide RNA template should contain a similar, stable DNA hairpin structure. Therefore, we used the 66-mer DNA as template to determine the effect of DNA secondary structure on DNA synthesis catalyzed by HIV-1 RT. Our experiments not only confirmed the existence of a template hairpin but also revealed that the DNA hairpin strongly stalled RT at several sites. We also showed that the DNA hairpin structure changed during polymerization. The structural, thermodynamic, and kinetic bases for RT pausing and DNA secondary structure switching are established in this report.

EXPERIMENTAL PROCEDURES

Proteins—Wild-type HIV-1 RT was prepared as described previously (15). All concentrations of RT reported in this paper were determined spectrophotometrically at 280 nm, with an extinction coefficient of 260,450 M⁻¹ cm⁻¹. T4 polynucleotide kinase was purchased from New England Biolabs (Beverly, MA).

Chemicals—[γ³²P]ATP was purchased from ICN (Costa Mesa, CA). Acrylamide, dNTPs, EDTA, Tris acetate, magnesium acetate, sodium acetate, and urea were purchased from Sigma. Biospin columns were
purchased from Bio-Rad.

**Synthetic Oligonucleotides**—The DNA 66-mer template (Fig. 1) and DNA primers that are complementary to the 3′ termini of the DNA 66-mer as well as control DNA 25/45-mer listed in Table I were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). The DNA oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis, and their concentrations were determined by UV absorbance at 260 nm. Before annealing, DNA primers were 5′-32P-labeled with T4 polynucleotide kinase. Unincorporated nucleotides were removed using a Bio-Spin-6 or BioSpin-30 column.

DNA duplexes were formed by annealing primers and templates at the appropriate ionic strength and ionic concentration. The mixtures were denatured at 90 °C for 5 min and then cooled slowly to room temperature in 2 h. To obtain DNA-product complexes, the DNA duplexes were admixed with T4 polynucleotide kinase. Unincorporated nucleotides were removed using a BioSpin-6 or BioSpin-30 column.

DNA duplexes were formed by annealing primers and templates at the appropriate ionic strength and ionic concentration. The mixtures were denatured at 90 °C for 5 min and then cooled slowly to room temperature in 2 h. To ensure that annealing had taken place the duplex mixtures were analyzed by nondenaturating polyacrylamide gel electrophoresis.

**Buffers**—Unless noted otherwise, all experiments using RT were carried out in RT-Mg2+ buffer containing 50 mM Tris acetate (pH 7.5 at 37 °C), 10 mM magnesium acetate, 100 mM potassium acetate, 0.1 mM EDTA. The pH of RT-Mg2+ buffer at 21 and 8 °C was 7.8 and 8.2, respectively. RT binding buffer is the same as RT-Mg2+ buffer except the pH was adjusted to 7.5 at 23 °C.

**Rapid Quench Experiments**—Rapid quench experiments were carried out in a apparatus designed by Johnson (18) and built by KinTek Corp. (State College, PA). Typically, the experiments were carried out by adding enzyme and DNA to preincubate in RT-Mg2+ buffer. An aliquot of this solution (15 μl) was rapidly mixed with an equal volume of solution containing nucleotide and Mg2+. The reactions were quenched with 90 μl of 0.3 M EDTA (final concentration) after time intervals ranging from 5 ms to seconds. All concentrations reported in this paper refer to concentrations during reactions after mixing.

**Pre-steady State Kinetic Analysis for Incorporation of Next Correct Nucleotide at 37 °C**—Pre-steady state kinetic analysis was conducted under the conditions where the DNA concentration was in slight excess relative to the enzyme concentration. The reactions were carried out by mixing a solution containing the preincubated complex of 50 nM HIV-1 RT and 100 nM 5′-labeled DNA duplex in RT-Mg2+ buffer with a solution of 100 μM dNTP in RT-Mg2+ buffer. Polymerization was quenched with 0.3 M EDTA at time intervals ranging from 6 ms to 7 s. DNA products were quantitated by sequencing gel analysis (see below). The data were fitted to a burst equation (see "Data Analysis").

**Nitrocellulose-DEAE Membrane Double-filter Binding Assay at 23 °C**—We used a previously published procedure to measure equilibrium dissociation constants of DNA substrates (19). Filter binding was carried out using a 96-well dot blot apparatus, nitrocellulose, and DEAE membranes from Schleicher & Schuell. This assay is based on the ability of the nitrocellulose membrane retaining RT-DNA and the DEAE membrane trapping all remaining DNA duplexes.

Wild-type RT (40 nM) was preincubated with increasing concentrations of DNA in RT binding buffer (50 mM Tris acetate, 10 mM magnesium acetate, 100 mM potassium acetate, 0.1 mM EDTA, pH 7.5 at 37 °C). Immediately prior to filtering each sample, a well of the dot-blot apparatus was washed with 100 μl of RT binding buffer. With the vacuum applied, 20 μl of each sample was added to the well followed by 100 μl of RT binding buffer. This was repeated three times for each DNA/RNA duplex concentration. Following titration, substrate binding was quantitated using a 445 SI PhosphorImager (Molecular Dynamics, Sunnyvale, CA).

**Processive Polymerization on 25/66-Mer**—We examined the kinetics of processive polymerization to test whether RT pausing occurs on the DNA hairpin of 25/66-mer. A preincubated solution of wild-type HIV-1 RT and 5′-32P-labeled 25/66-mer in RT-Mg2+ buffer was mixed with all four deoxynucleotides in RT-Mg2+ buffer at either 37, 21, or 8 °C. After various intervals, the reactions were stopped by 0.3 M EDTA, and the products were analyzed using denaturing sequencing gels as shown in Fig. 2. At 37 °C, the intermediate products 32-, 34-, 35-, and 36-mer accumulated slightly at 21 °C and 8 °C. At 37 °C, the intermediate products 34- and 35-mer accumulated significantly, and the full-length products formed rapidly. At 21 and 8 °C, these same intermediates accumulated more significantly (Fig. 2, B and C). Moreover, the 40-, 41-, 42-, and 48-mer accumulated slightly at 21 °C and particularly at 8 °C. The early intermediate 28-mer also accumulated slightly, which is likely a consequence of the run of template thymidines (13).

**DNA Secondary Structure Switching Predicted by Mfold**— RT pausing after synthesis of the 32-, 34-, 35-, and 36-mer was expected because these strong pause sites correspond to GC base pairs in the stem of the DNA hairpin (Fig. 1). Slight RT pausing after synthesis of 40-, 41-, and 42-mer at

---

1 KinTek Corp. on-line address is as follows: www.kintek-corp.com.

2 KinTek Corp. on-line address is as follows: http://ibc.wustl.edu/~zuker/cgi.

3 Mfold is accessible on-line at the following address: http://ibc.wustl.edu/~zuker/cgi.
21 °C and 8 °C was unexpected, but plausible, since these weak pause sites were also located within the hairpin stem (Fig. 1). A slight accumulation of 47- and 48-mer was surprising because these two weak pause sites were located within the hairpin loop. RT pausing slightly at these two sites may be due to sequence dependence of the polymerase activity of RT during polymerization or DNA secondary structure change leading to relocation of the pause sites within a new hairpin. To test the
second possibility, we used Mfold to locate optimal structures of the remaining single-stranded template, whereas the 25-mer primer was gradually elongated. With primers shorter than 37 nucleotides, structure A (Fig. 3) was predicted to be the optimal structure, whereas structure B was the most stable structure with primers equal to or longer than 37 nucleotides. Structure C is the second most stable structure when primers are longer than 37 nucleotides and is only 1.5 kcal/mol less stable than structure B at 8 °C. This suggests that structure C may exist significantly to compete against structure B to stall RT once the primer was longer than 37 nucleotides. Interestingly, the five weak pause sites were within the bases of the hairpin stems of structures B and C rather than at the loop and the second half of the hairpin stem of structure A. Thus, DNA secondary structure switching is a likely reason for the slight accumulation of 40- to 42-mer and 47- to 48-mer.

Processive Polymerization on 25/61-Mer—To test further whether the slight accumulation of 47- and 48-mer is due to DNA secondary structure switching or sequence dependence of polymerization, we performed processive experiments with a 61-nucleotide DNA template (Fig. 4), which lacks the first five nucleotides of the 5' terminus of the DNA 66-mer. Mfold predicted that the DNA secondary structure change from structure A to B (Fig. 3) should not occur with the 61-mer template. Thus, the intermediates 47- and 48-mer would not accumulate with the 61-mer template if RT pausing at these two sites was due to DNA secondary structure change. In contrast, if sequence dependence of polymerization was partially or totally responsible for the slight accumulation of the 47- and 48-mer, RT pausing at the two sites should still occur with the 61-mer template because the 61- and 66-mer templates have identical sequences except for length. As shown in Fig. 4, the product accumulation pattern with the 61-mer template observed at 8 °C is similar to the one observed with the 66-mer template (Fig. 2C), except that the accumulation of the 47- and 48-mer did not occur. Thus, these results strongly support that the slight accumulation of the 47- and 48-mer with the 66-mer template was due to DNA secondary structure switching from structure A to B in Fig. 3 rather than sequence dependence of polymerization.

Free Energy Barrier of Base Pair Unwinding Predicted by Mfold—Once RT encounters a DNA hairpin, each stem base pair has to be melted prior to next nucleotide incorporation. Each unwinding step surpasses a certain energy barrier, the strength of which depends on the base pair and the nearest neighbors in the stem base pair (21). RT is expected to pause at each site where the next stem base pair is difficult to melt. Thus, the free energy barrier of melting the next stem base pair prior to the

![Image](https://example.com/figure3.png) **Fig. 3.** Three competing DNA secondary structures of 37/66-mer. The values of folding free energy of hairpins A, B, and C were predicted to be −6.3, −6.6, and −4.8 kcal/mol, respectively, at 8 °C using Mfold. The arrows represent the weak RT pause sites at 21 and 8 °C numbered from the 3' terminus of the DNA template.

![Image](https://example.com/figure4.png) **Fig. 4.** Processive polymerization on 25/61-mer at 8 °C. A preincubated solution of 100 nM HIV-1 RT and 100 nM 5'-32P-labeled 25/61-mer in RT-Mg2+ buffer was mixed with dATP, dGTP, dCTP, and dTTP (150 μM each) in RT-Mg2+ buffer (pH 8.2). The reactions were quenched with 0.3 M EDTA at the indicated times and analyzed by sequencing gel electrophoresis. The large and small arrows represent the strong (32-, 34-, 35-, and 36-mer) and weak RT pause sites (40-, 41-, and 42-mer), respectively, numbered from the 3' terminus of the DNA template.
incorporation of the 37th nucleotide. The 37th nucleotide incorporation associates with two different values of $\Delta G$. The positive $\Delta G$ is the free energy barrier of melting the next template base pair before the incorporation of the 37th nucleotide. The negative $\Delta G$ is the free energy change of the DNA secondary structure switching after the incorporation of the 37th nucleotide.

FIG. 5. Free energy barrier ($\Delta G$) of single template (the DNA 66-mer) base pair melting at 37 ($\bullet$) and 8 °C ($\square$) associated with the incorporation of the ith nucleotide. The 37th nucleotide incorporation associates with two different values of $\Delta G$. The positive $\Delta G$ is the free energy barrier of melting the next template base pair before the incorporation of the ith nucleotide. The negative $\Delta G$ is the free energy change of the DNA secondary structure switching after the incorporation of the 37th nucleotide.

incorporation of the ith nucleotide primed with an I–1 nucleotide primer. As shown in Fig. 5, the calculated energy barriers at 37 and 8 °C were plotted versus positions of nucleotides incorporated through the DNA hairpin. The energy barriers at 21 °C were also calculated (data not shown), and they are between the values of 37 and 8 °C.

Interestingly, the variation in unwinding energy barrier (Fig. 5) correlates well with the product accumulation patterns (Fig. 2, A and C). Non-pause sites have either low or negligible free energy barriers. The RT pause sites after synthesis of the 32-, 34-, 35-, and 36-mer have high unwinding free energy barriers for next nucleotide incorporations. After synthesis of the 37-mer, the unwinding energy barrier changes from unfavorable (positive) to slightly favorable (negative) indicate that there is a potential DNA secondary structure change (Fig. 3). The slight accumulation of the 47- and 48-mer correlates to high unwinding free energy barriers for incorporations of the 48th and 49th nucleotides. Similarly, the relatively high free energy barriers of next stem base pair unwinding in structure C (data not shown) contributed to slight accumulation of 40- to 42-mer. Additionally, the base pair unwinding energy barriers are higher at 8 than at 37 °C.

Measurement of Kinetics of Nucleotide Incorporation at Pause and Non-pause Sites—Accumulation of intermediate products indicates that the kinetics of next nucleotide incorporation were significantly affected by the DNA secondary structure. To determine a kinetic basis for intermediate accumulation, we performed pre-steady state kinetic analysis of single nucleotide incorporation at three strong pause sites and two non-pause sites. Primers listed in Table I were synthesized, purified, and annealed to the DNA 66-mer template. The pre-steady state kinetics of next correct nucleotide incorporation into the annealed DNA duplexes were measured as described under “Experimental Procedures.” As a control, we also measured the kinetics of dATP incorporation into 25/45-mer (Table I), which was previously used as a model substrate to establish the kinetic pathway of HIV-1 RT (22).

The time courses (not shown) of dCTP incorporation into 37/66-mer, dGTP incorporation into 20/66-mer, and dATP incorporation into 25/45-mer are similar to each other and to previously published data (22). This suggests that kinetics of nucleotide incorporation at non-pause sites were not affected by DNA hairpin structure. Each time course displays a fast phase followed by a slower linear phase. The data were fitted to a burst equation, and the burst rates and reaction amplitudes were summarized in Table II. The rates of the linear phase (data not shown) were between 0.1 and 0.2 s⁻¹. The lower reaction amplitude observed with the 25/45-mer compared with the previously measured value using freshly prepared RT (16) indicates that enzyme lost about 20% activity with storage. However, this does not affect the overall kinetic analyses since we used the same pool of RT in all experiments presented in this paper, and we compared relative amplitudes.

The time courses of single nucleotide incorporation into 32/66-mer, 35/66-mer, and 36/66-mer fit poorly to the burst equation and yielded small burst amplitudes (4–10%) and high rates in the linear phase (0.6–0.9 s⁻¹). The fast linear phase could indicate that DNA dissociates from the enzyme more rapidly at the pause sites than at non-pause sites. To test this hypothesis, we directly measured the dissociation rate of RT/36/66-mer (Fig. 6). A molar excess of RT (350 nM) was preincubated with unlabeled 36/66-mer (50 nM) in RT-Mg²⁺ buffer in order for most of the 36/66-mer to be bound by RT. The RT/36/66-mer solution was then mixed with 5'-32P-labeled 25/45-mer (450 nM) in RT-Mg²⁺ buffer and incubated for 85 ms to 40 s. During incubation, free enzyme (both initially free enzyme and enzyme dissociated from 36/66-mer) was bound by the excess labeled 25/45-mer rather than the free 36/66-mer. Following incubation, a solution of Mg²⁺-dATP was added to initiate incorporation of dATP into 25/45-mer. Then the reaction was stopped with 0.3 M EDTA after a constant interval of 250 ms, sufficient to allow all RT/25/45-mer to be converted to RT/26/45-mer in a single turnover. By following nucleotide incorporation into 25/45-mer rather than 36/66-mer, we avoided the complication of a slow turnover of non-productively bound 36/66-mer (see below). The time course of 26/45-mer formation was fitted into a single exponential (see Experimental Procedures) to yield a dissociation rate of 0.17 s⁻¹ for RT/36/66-mer.

Similar experiments were performed to measure dissociation rates of other DNA substrates, and the results were summarized in Table II. The dissociation rates of 20/66-mer, 37/66-mer, and 25/45-mer agree with the rates of the linear phase obtained from the burst experiments and with the previously measured rate of 0.18 s⁻¹ (22). This indicates that the method we used can properly measure DNA dissociation rate with accuracy. However, the dissociation rates of 32/66-mer, 35/66-mer, and 36/66-mer are not only similar to DNA dissociation rates at these non-pause sites but also 2–3-fold slower than the rates of linear phase measured in the burst experiments. These suggest that the high rates of the slow phase observed at pause sites (0.6–0.9 s⁻¹) are not due to fast DNA dissociation. Rather, since only a small amount of DNA (4–10%) at each pause site was turned over to products in the fast phase, we speculated that other DNA molecules could be bound non-productively and somehow were slowly converted to productively bound substrates and then converted to products, contributing to the slow phase.

To test this hypothesis, we performed a trap experiment for incorporation of dCTP into RT/36/66-mer. A preincubated solution of RT and 5'-32P-labeled 36/66-mer in RT-Mg²⁺ buffer was mixed with a solution of Mg²⁺-dATP and a large molar excess of unlabeled 25/45-mer to start polymerization. The unlabeled 25/45-mer served as a trap for free RT to prevent rebinding of dissociated 36/66-mer. At increasing times, the polymerization was stopped by 0.3 M EDTA, and products were analyzed on a sequencing gel. The data (Fig. 7) were fitted to a double exponential equation and yielded a rate of 19 s⁻¹ and 7.8% reaction...
amplitude for the fast phase and a rate of 0.067 s⁻¹ and 14% reaction amplitude for the slow phase. Similar experiments were performed with 32/66-mer and 35/66-mer, and the results were summarized in Table II. The observed biphasic kinetics at all pause sites confirmed our hypothesis that DNA substrates were bound by enzyme both productively and non-productively at the polymerase site of RT. Productively bound substrates were more rapidly converted to products than non-productively bound substrates. Since we only measured a single enzyme binding event, DNA substrates bound in both states did not dissociate from enzyme before being converted to products.

Interestingly, only the fast phase of nucleotide incorporation was observed when the same experiments were performed with 20/66-mer, 37/66-mer, and 25/45-mer (data not shown). This suggests that DNA substrates were bound only productively at non-pause sites and that the kinetics of single nucleotide incorporation at pause and non-pause sites are different. The data were fitted to a single exponential equation (see "Experimental Procedures"), which gave a dissociation rate of 0.067 ± 0.01 s⁻¹ with an amplitude of (7.8 ± 0.5%), and a slow rate of 0.067 ± 0.009 s⁻¹ with an amplitude of (14 ± 1%), relative to the enzyme concentration.

FIG. 6. Dissociation of HIV-1 RT-36/66-mer at 37 °C. A preincubated solution of 350 nM RT and 50 nM unlabeled 36/66-mer in RT-Mg²⁺ buffer was mixed with 450 nM 5²⁻²²P-labeled 26/44-mer in RT-Mg²⁺ buffer for periods of 0.085–40 s. At the end of the period, 60 μM Mg²⁺·dCTP in RT-Mg²⁺ buffer was added, and the reaction was quenched with 0.3 M EDTA after a constant interval of 250 ms. The amount of 5²⁻²²P-labeled 26/44-mer formed was analyzed by sequencing gel analysis. The data (●) represent the plot of concentrations of 26/44-mer versus mixing time. The solid line is the fit of the data to a single exponential equation (see "Experimental Procedures"), which gave a dissociation rate of 0.17 ± 0.01 s⁻¹ for RT-36/66-mer, a free RT concentration of 93.7 ± 0.4 nM, and a concentration of 20.7 ± 0.5 nM for the active RT-36/66-mer complex in the preincubated solution of RT and 36/66-mer.

FIG. 7. Biphasic kinetics of dCTP incorporation into HIV-1 RT 36/66-mer at 37 °C. A solution of RT (50 nM) was preincubated with 36/66-mer (50 nM) in RT-Mg²⁺ buffer and then was mixed with Mg²⁺·dCTP (100 μM) and unlabeled 25/45-mer (5 μM) in RT-Mg²⁺ buffer. The reactions were quenched at time intervals ranging from 6 ms to 60 s by the addition of 0.3 M EDTA. DNA products were quantitated by sequencing gel analysis. The data (●) were fitted to a double exponential equation (see "Experimental Procedures"), which gave a fast rate of 19 ± 5 s⁻¹ with an amplitude of (7.8 ± 0.5%), and a slow rate of 0.067 ± 0.009 s⁻¹ with an amplitude of (14 ± 1%), relative to the enzyme concentration.
is a fit of the data (line Kd Procedures”), which gave a of nitrocellulose and DEAE double-filter dot-blot apparatus. The mM Mg2 RT binding buffer (50 mM Tris acetate, 100 mM potassium acetate, 10 mM Mg2+, 0.1 mM EDTA, pH 7.5 at 23 °C) and then applied to the wells of nitrocellulose and DEAE double-filter dot-blot apparatus. The solid line is a fit of the data (○) to a quadratic equation (see “Experimental Procedures”), which gave a Ka value for the RT-36/66-mer complex of 46 ± 5 nM and an amplitude of (31 ± 1.0)%.

RT binding buffer, the nitrocellulose membrane retained the RT-36/66-mer complex, whereas DEAE membrane trapped all free 36/66-mer. After correction for nonspecific retention of DNA by the nitrocellulose membrane, the average concentrations of the RT-36/66-mer were plotted versus the total concentration of 36/66-mer (Fig. 8). The data were fitted to a hyperbola (see “Experimental Procedures”) and yielded an equilibrium dissociation constant of 46 nM. Experiments performed with other DNA substrates (Table II) indicate that DNA hairpin did not dramatically affect DNA binding to RT.

DISCUSSION

Several polymerases including DNA polymerase α (4–6), E. coli DNA polymerase III (7), phage T4 DNA polymerase (8, 9), and HIV-1 RT (12–14) have been hypothesized to stall at potential DNA secondary structures of large single-stranded DNA templates. As a template for the synthesis of the plus strand DNA during HIV-1 replication (1), the single-stranded minus strand DNA is several thousand nucleotides in length and was predicted to form complex DNA secondary structures at several sites using the computer program Mfold. HIV-1 RT could pause strongly at potential DNA secondary structures of the minus strand DNA template during plus strand synthesis in vivo since no accessory proteins possessing helicase-like activity have been found to increase the processivity of HIV-1 RT (15). In this paper, a DNA 66-mer template containing a stable hairpin structure, confirmed by enzymatic footprinting analysis, was used as a model to study directly the effects of DNA secondary structure on polymerization catalyzed by HIV-1 RT.

Similar and Different Effects on Polymerization by DNA and RNA Secondary Structures—With the 66-mer DNA template, RT paused after synthesis of 32-, 34-, 35-, and 36-mer at 37, 21, and 8 °C and paused slightly after synthesis of 40–42-mer and 47–48-mer at 21 and 8 °C (Fig. 2). RT paused strongly after synthesis of 32-, 34-, 35-, 36-, 47-, and 48-mer with the RNA analog of the DNA 66-mer as the template (17). Similar pause sites were observed with both RNA and DNA since both templates formed similar hairpins shown in structures A and B (Fig. 3). The two templates displayed similar secondary structure changes (from structure A to B) after synthesis of the 37-mer, although DNA secondary structure switching was only observable at low reaction temperatures, especially at 8 °C. RT did not pause after synthesis of 40-, 41-, and 42-mer with the RNA template even at 8 °C. The accumulation of 40-, 41-, and 42-mer with the DNA 66-mer template is due to formation of structure C which competes against the slightly more stable structure B (Fig. 3). Using Mfold, structure B at 8 °C was predicted to be only −1.8 kcal/mol more stable than structure C with the DNA 66-mer template, but −6.6 kcal/mol more stable than structure C with the RNA 66-mer template. Thus, structure C was much less stable than structure B with the RNA template at 8 °C and did not exist significantly to stall RT. Additionally, RT pausing was stronger with the RNA 66-mer template than with its DNA analog at the same reaction temperatures. For example, the 48-mer accumulation was very clear with the RNA template (17), whereas it was not observed with the DNA template at 37 °C. This is because DNA secondary structures are less stable than their RNA analogs at the same temperatures. For example, the folding free energy for the DNA hairpin shown in Fig. 1 and its RNA analog was predicted by Mfold to be −12.9 kcal/mol and −20.3 kcal/mol at 37 °C, respectively.

Structural Basis for RT Pausing and DNA Secondary Structure Switching—At all three temperatures 37, 21, and 8 °C, RT paused after synthesis of 32-, 34-, 35-, and 36-mer (Fig. 2). The pause sites were located within the first half of the hairpin stem (Fig. 1) and correspond to melting of GC base pairs; however, RT did not pause after synthesis of the 37-mer, where the next stem base pair was also a GC base pair. A potential DNA secondary structure change (Fig. 3) was predicted, which leads to a 10-nucleotide single-stranded template gap between the primer 37-mer and the new hairpin in structure B, where RT did not pause (Fig. 2). Once RT reached the new hairpin, slight pausing was observed after synthesis of the 47- and 48-mer at 21 and 8 °C (Fig. 2). The two weak pause sites appearing within the loop in structure A were relocated to the base of the hairpin in structure B. Similarly, RT pausing only slightly after synthesis of 40-, 41-, and 42-mer at 21 and 8 °C (Fig. 2) can be explained by a second DNA secondary structure change from structure A to structure C (Fig. 3). This change also relocated the three weak pause sites from unreasonable positions at the end of the hairpin stem in structure A to the base of the new hairpin in structure C. Thus, all RT pause sites were located at the first several base pairs of hairpin stems due to DNA secondary structure switching. Moreover, the relocation of pause sites was due to DNA secondary structure switching rather than sequence dependence. This conclusion was supported by the product accumulation pattern with the DNA 61-mer template, which lacked the first 5 nucleotides on the 5’-terminus of the 66-mer and thereby prevented the secondary structure change and accumulation of 47- and 48-mer (Fig. 4). Interestingly, the two weak pause sites after synthesis of 40- and 47-mer correspond to melting of two potential wobble base pairs GG and GT, respectively, in structures B and C (Fig. 3). If deoxynucleotides GG and GT did not form wobble base pairs as nucleotides rGs and rUs in the hairpin of RNA 66-mer (16), they could still stack on the base pairs of the hairpin stems to stall RT. Additionally, the accumulation of both 40- to 42-mer and 47- to 48-mer during polymerization indicates that the secondary structure change from structure A to B competed against the change from structure A to C. This is reasonable because structure C is slightly less stable than structure B and slightly more stable than structure A. The intermediates 40- to 42-mer and 47- to 48-mer did not accumulate at 37 °C because none of the DNA secondary structures shown in Fig. 3 were stable enough to stall RT and were rapidly read through by the strand-displacement activity of RT (23). However, we predict that DNA secondary structure switching will be observed at
37 °C with DNA templates containing more stable DNA hairpins than the ones described in this paper.

We also examined the processive polymerization at 37 and 8 °C with T7 DNA polymerase (exo−), and the patterns of product accumulation (data not shown) are similar to the patterns shown in Fig. 2, A and C. Thus, our results unambiguously prove that DNA secondary structures stall polymerases as proposed previously by numerous groups with different enzymes (4–9, 12, 13). Moreover, all pause sites were located within the first several base pairs of hairpin stems rather than at hairpin loops or the ends of hairpin stems. The pause sites at the hairpin loops are relocated due to DNA secondary structure switching. Interestingly, the pause sites that were previously located in the predicted loops of DNA templates (4, 13) can be relocated within the stems of new hairpins of shortened single-stranded template predicted by Mfold after DNA primer extension approaches the pause sites.

Thermodynamic Basis for RT Pausing and DNA Secondary Structure Switching—RT pausing within the hairpin stems of DNA secondary structures suggests that the overall energy barriers for nucleotide incorporation are higher at hairpin stems than at single-stranded regions. When RT traverses a DNA hairpin, an adjacent stem base pair has to be melted prior to incorporation of the next nucleotide. At non-pause sites, RT reads single-stranded DNA and the melting of the next stem base pair is not required. Thus, the difference in the overall free energy for nucleotide incorporations at the single-stranded and double-stranded regions of the DNA template is from the base pair unwinding processes. This was supported by our kinetic mechanism (see discussion below). Thus, the energy barrier of base pair melting correlates with RT pausing.

As shown in Fig. 5, Mfold predicted that the energy barriers for melting the next stem base pair were high prior to incorporations of the 33rd, 35th, 36th, and 37th nucleotides which lead to accumulation of 32-, 34-, 35-, and 36-mer, although there is no energy cost at the non-pause sites. The high melting free energy barriers associated with the incorporations of the 48th and 49th nucleotides indicate that the 47- and 48-mer could also accumulate. We did observe RT pausing at the two sites at 21 and 8 °C (Fig. 2). RT pausing after synthesis of the 40- to 42-mer also correlates with relative high free energy barriers of melting stem base pairs of structure C (Fig. 3). Therefore, the patterns of the predicted free energy barriers of next stem base pair unwinding as a function of position through the hairpin structure correlates very well with product accumulation patterns.

As shown in Fig. 5, the free energy barrier becomes slightly favorable after synthesis of 37-mer. This is due to DNA secondary structure switching from structure A to the slightly more stable structure B (Fig. 3). However, if free in solution, the hairpin in structure A should be more stable than the lone hairpin in structure B due to contribution from coaxial stacking between adjacent hairpin stem and the primed double-stranded region in structure A (24–26). Once bound by RT, the observed DNA secondary structure switching suggests that the coaxial stacking in structure A was eliminated by a potential kink at the interface between the hairpin and the primer. The crystal structure of human DNA polymerase β complexed with a nicked DNA has shown that there is a 90° kink occurring precisely at the nick (27). Thus, HIV-1 RT could play a role in facilitating DNA secondary structure switching as in facilitating RNA secondary structure switching (17).

Kinetic Basis for RT Pausing—Pre-steady state kinetic analyses have revealed that incorporation of the next correct nucleotide (dATP) into the substrate 25/45-mer by HIV-1 RT occurs in a fast phase at a rate of 33 s⁻¹, which is rate-limited by a protein conformational change. Under single nucleotide incorporation conditions, the burst of polymerization is followed by a slow steady state phase, which is limited by slow DNA dissociation at a rate of 0.2 s⁻¹ (22). With both 20/66-mer and 37/66-mer, the kinetics of incorporation of next correct nucleotide are similar to the control substrate 25/45-mer. As shown in Table II, fast polymerization (33–87 s⁻¹), high reaction amplitude in the fast phase (32–45%), tight binding of RT-DNA (16–33 nM), and slow dissociation of RT-DNA (0.14–0.18 s⁻¹) contribute to fast nucleotide incorporations at the two sites. Thus, the DNA hairpin structures did not affect kinetics of polymerization at the non-pause sites.

The incorporation of next nucleotides into 32/66-mer, 35/66-mer, and 36/66-mer are significantly different from the control 25/45-mer. In a single binding event, polymerization displayed biphasic kinetics at the three strong pause sites. Table II shows that slightly slower polymerization (10–19 s⁻¹) in the fast phase at the pause sites compared with the non-pause sites indicates that DNA hairpin (Fig. 1) might inhibit the protein conformational change. More significantly, small reaction amplitudes (4–10%) in the fast phase indicate that only a small fraction of the RT-DNA was rapidly elongated to products at the RT pause sites. Large reaction amplitudes (14–37%) and slow polymerization rates (0.015–0.067 s⁻¹) in the slow phase suggest that a large fraction of the RT-DNA slowly turned over to products. Similar overall DNA binding affinity and DNA dissociation rates at all sites indicate that DNA did not dissociate from enzyme in the slow phase at the strong pause sites. Thus, the large reaction amplitudes and slow turnover rates in the slow phase at the pause sites rather than fast DNA dissociation or weak DNA binding accounts for the stalling of RT. Moreover, the kinetic data with 36/66-mer are similar to 35/66-mer but very different from 37/66-mer, suggesting either a DNA secondary structure switching or complete disruption of the DNA hairpin structure shown in Fig. 1 after synthesis of the 37-mer at 37 °C.

No Significant Effect of DNA Secondary Structure on Tight Binding of DNA to RT—As shown in Table II, five DNA substrates with the 66-mer template and the control substrate 25/45-mer have similar overall binding affinity to RT (16–86 nM) measured by nitrocellulose-DEAE double-filter binding assay. This suggests that the DNA secondary structure did not dramatically affect the tight binding of DNA at the RT binding cleft. Similar dissociation rates of all DNA substrates (Table II) also support this conclusion. It is not surprising that the RT binding cleft can tolerate a large DNA or RNA secondary structure because previous crystal structure data (28, 29) and chemical footprinting data (30) have revealed that RT has a flexible binding cleft larger than 19 base pairs. However, high overall binding affinity of DNA to RT does not mean that DNA binds productively to the polymerase-active site. The large amplitudes in the slow phase (see discussion above) indicate that DNA substrates primarily bind non-productively at RT pause sites.

Kinetic Mechanism of Pausing Due to DNA Secondary Structure—We previously determined a mechanism shown in Scheme I of RT pausing due to RNA secondary structure (16). Similar biphasic kinetics of nucleotide incorporation in the presence of excess DNA to trap free enzyme indicate that polymerization at strong pause sites should follow the same mechanism. At pause sites, DNA substrates bind to RT in two modes as follows: a small fraction of DNA binds productively (EDₜ) and is converted rapidly to products at a rate of (10–19 s⁻¹); a large fraction of DNA binds non-productively (EDₜ) and is somehow slowly converted to the productively binding mode at a rate of 0.015–0.067 s⁻¹ without dissociation from enzyme
and then turned over to products. The non-productively bound mode is a function of the next stem base pair, which blocks the incoming nucleotide to pair with the template nucleotide. Because the slow interconversion between the two binding modes is the rate-limiting step for non-productively bound DNA to be turned over to products, the process of melting the next stem base pair primarily contributes to the free energy barrier and turned over to products, the process of melting the next stem base pair primarily contributes to the free energy barrier and turned over to products. The slow interconversion between the two binding modes is the rate-limiting step for non-productively bound DNA to be cause the slow interconversion between the two binding modes (13) may be due to a function of fast DNA dissociation and slow polymerization. This requires further study.

**Potential Biological Significance of Polymerase Pausing and DNA Secondary Structure Switching**—The ability of HIV to evade host immune response during chronic infections is related to the high mutation frequency of the viral genome. Since HIV-1 RT is an error-prone enzyme (22, 31) and polymerase pausing may also facilitate other biological functions. RT pausing may enhance template strand switching during HIV replication (23). In eukaryotic and prokaryotic systems, RT pausing may also enhance template strand switching during HIV replication. In addition to pro-mutagenesis, DNA secondary structures from trinucleotide repeats. Switching could play a role in formation of different DNA secondary structures during DNA synthesis. Our results with HIV-1 RT could be applied to other poly-ranucleases.

**Acknowledgments**—We thank Dr. Roger Goody (Max-Planck Institute, Heidelberg, Germany) for providing the clone of wild-type HIV-1 RT (p66/p51). We thank Allison A. Johnson for carefully reading this manuscript.

**REFERENCES**

1. Hottiger, M., and Hubscher, U. (1996) Biol. Chem. Hoppe-Seyler 377, 97–120
2. Goff, S. (1990) J. Acquired Immune Defic. Syndr. 3, 817–831
3. Jacobo-Molina, A., and Arnold, E. (1991) Biochemistry 30, 6351–6361
4. Kagni, L. S., and Clayton, D. A. (1982) Proc. Natl. Acad. Sci. U. S. A. 79, 983–987
5. Weaver, D. T., and DePamphilis, M. L. (1982) J. Biol. Chem. 257, 2075–2086
6. Weaver, D. T., and DePamphilis, M. L. (1984) J. Biol. Chem. 260, 961–986
7. LaDuca, R. J., Fay, P. J., Chuang, C., McHenry, C. S., and Barbara, R. A. (1983) Biochemistry 22, 5177–5188
8. Bedinger, P., Munn, M., and Alberts, B. M. (1989) J. Biol. Chem. 264, 16880–16886
9. Hacker, K. J., and Alberts, B. M. (1994) J. Biol. Chem. 269, 24221–24228
10. Trinh, T. Q., and Sinden, R. R. (1991) Nature 352, 544–547
11. Freudenreich, C. H., Stavenhagen, J. B., and Zakian, V. A. (1997) Mol. Cell. Biol. 17, 2090–2096
12. Dudding, L. R., Nkabinde, N. C., and Mirzali, V. (1995) Biochemistry 30, 10498–10506
13. Klarmann, G. J., Schaubauer, C. A., and Presten, B. D. (1993) J. Biol. Chem. 268, 9793–9802
14. Ji, J., Hoffmann, J., and Loeb, L. (1994) Nucleic Acids Res. 22, 47–52
15. Sato, T., and Johnson, K. A. (1995) Biochemistry 34, 12468–12476
16. Sato, Z., and Johnson, K. A. (1997) Biochemistry 36, 12459–12467
17. Sato, Z., and Johnson, K. A. (1997) Biochemistry 36, 14778–14785
18. Johnson, K. A. (1988) Methods Enzymol. 134, 677–705
19. Wong, L., and Lehman, T. M. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 5428–5432
20. Zeker, M. (1989) Science 244, 48–52
21. Santalucia, J., Jr., Allawi, H. T., and Seneviratne, P. A. (1996) Biochemistry 35, 3555–3562
22. Kati, W. M., Johnson, K. A., Jerva, L. F., and Anderson, K. S. (1992) J. Biol. Chem. 267, 25968–25977
23. Huber, H. R., McCoy, J. M., Seehra, S. J., and Richardson, C. C. (1989) J. Biol. Chem. 264, 4669–4678
24. Leontis, N. B., Kwok, W., and Newman, J. J. (1991) Nucleic Acids Res. 19, 759–766
25. Lu, M., Guo, Q., Marky, C. A., Seeman, N. C., and Kallenbach, N. R. (1992) J. Mol. Biol. 233, 781–789
26. Kielczewski, J., Dunn, J. J., and Studier, F. W. (1992) Science 256, 1787–1791
27. Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J., and Pelletier, H. (1997) Biochemistry 36, 11205–11215
28. Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992) Science 256, 1783–1790
29. Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, P., Hizi, A., Hughes, S. H., and Arnold, E. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 6320–6324
30. Wohrle, B. M., Tantillo, C., Arnold, E., and Le Griece, S. F. J. (1995) Biochemistry 34, 5343–5350
31. Johnson, K. A. (1993) Annu. Rev. Biochem. 62, 685–713
32. Kunkel, T. A. (1990) Biochemistry 29, 8003–8011
33. Papanicolaou, C., and Ripley, L. S. (1991) J. Mol. Biol. 221, 805–821
34. Bebenek, K., Abbotts, J., Roberts, J. D., Wilson, S. H., and Kunkel, T. A. (1989) J. Biol. Chem. 264, 16943–16950
35. Fry, M., and Loeb, L. A. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 763–767
36. Trinh, T. Q., and Sinden, R. R. (1993) Genetics 134, 409–422