Published in final edited form as:
J Mol Biol. 2010 February 5; 395(5): 995. doi:10.1016/j.jmb.2009.11.072.

SINGLE-MOLECULE STUDY OF DNA POLYMERIZATION ACTIVITY OF HIV-1 REVERSE TRANSCRIPTASE ON DNA TEMPLATES

Sangjin Kim[1], **Charles M. Schroeder**[1,2], and **X. Sunney Xie**[1,*]

[1] Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA

[2] Department of Chemical & Biomolecular Engineering, University of Illinois at Urbana-Champaign, 600 S. Mathews Ave, Urbana, IL 61801, USA

Abstract

Human Immunodeficiency Virus-1 reverse transcriptase (HIV-1 RT) is a multifunctional polymerase responsible for reverse transcription of the HIV genome, including DNA replication on both RNA and DNA templates. During reverse transcription *in vivo*, HIV-1 RT replicates through various secondary structures on RNA and single-stranded DNA templates without the need for a nucleic acid unwinding protein, such as a helicase. In order to understand the mechanism of polymerization through secondary structures, we investigated the DNA polymerization activity of HIV-1 RT on long single-stranded DNA templates using a multiplexed single-molecule DNA flow-stretching assay. We observed that HIV-1 RT performs fast primer extension DNA synthesis on single-stranded regions of DNA (18.7 nt/s) and switches its activity to slow strand displacement synthesis at DNA hairpin locations (2.3 nt/s). Furthermore, we found that the rate of strand displacement synthesis is dependent on the GC content in hairpin stems and template stretching force. This indicates that the strand displacement synthesis occurs through a mechanism that is neither completely active nor passive, i.e. the opening of the DNA hairpin is driven by a combination of free energy released during dNTP hydrolysis and thermal fraying of base pairs. Our experimental observations provide new insight into the interchanging modes of DNA replication by HIV-1 RT on long single-stranded DNA templates.

Keywords

HIV-1 reverse transcriptase; single molecule; DNA flow-stretching assay; DNA hairpin; strand displacement synthesis

Introduction

Human Immunodeficiency Virus-1 reverse transcriptase (HIV-1 RT) is a 66/51-kDa heterodimeric DNA polymerase that synthesizes a double-stranded proviral DNA from a viral

*Corresponding author: Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA. Tel: (617) 496-9925 Fax: (617) 496-8709; xie@chemistry.harvard.edu.

RNA genome. The 66-kDa subunit is derived from the viral *pol* gene and contains enzymatic active sites for DNA polymerization and RNaseH activity, whereas the 51-kDa subunit is derived from proteolytic cleavage of the RNaseH domain in the 66-kDa subunit and lacks any catalytic role [1–3]. *In vivo*, HIV-1 RT possesses three distinct enzymatic activities: (1) polymerization of cDNA on RNA templates yielding RNA/DNA duplexes, (2) enzymatic degradation of RNA templates, and (3) synthesis of DNA using cDNA templates. HIV-1 RT also executes strand displacement synthesis on ~634-base-long duplex DNA region (long terminal repeat sequences) in order to complete replication of its genome [4,5]. Although multifunctional, HIV-1 RT is not an exceedingly efficient DNA polymerase. DNA polymerization by HIV-1 RT on both RNA and DNA templates exhibits a slow enzymatic rate (maximum rate of single nucleotide incorporation of ~33 s^{-1} at 37 °C [6]), low processivity (1–300 nucleotides depending on the template sequence [4,7]), and poor fidelity (error frequency of 1 in 5900 nucleotides polymerized on the DNA template [8]). Because of its pivotal role in the HIV-1 life cycle and as a drug target for the clinical treatment of HIV infection and acquired immune deficiency syndrome (AIDS), there have been extensive biochemical studies on HIV-1 RT, including several kinetic studies for single nucleotide addition [6,9].

An interesting open question for HIV-1 RT in terms of viral lifecycle is the effect of the large number of stable RNA or DNA secondary structures in the HIV genome during reverse transcription. During DNA replication, HIV-1 RT uses long stretches of single-stranded RNA or DNA molecules as templates. However, unlike most DNA polymerases that form a replisome, HIV-1 RT does not have customary accessory proteins, such as a helicase, sliding clamp or single-stranded DNA binding proteins (SSB), which accelerate destabilization of nucleic acid duplex on the template and increase processivity of DNA polymerase. Hence, secondary structures may serve as physical barriers for processive enzymatic synthesis, and HIV-1 RT must have a different mechanism to cope with secondary structures on the template. Based on the currently accepted HIV infection mechanism, stable secondary structures induce pausing of DNA synthesis, and enzymatic pauses during DNA replication may result in DNA recombination through template switching or base misincorporation [10,11]. In particular, recombination hot spots were found where stable stem-loop structures are present in the HIV genome, such as in the trans-activating response region [12,13]. At present, however, the molecular mechanism responsible for the ability of HIV-1 RT to replicate through successive secondary structures on a whole genome is not well understood because bulk studies mainly used short nucleic acids with a single hairpin for replication templates. The dynamic nature of HIV-1 RT can be easily obscured in bulk experiments when a population of RT molecules is studied because the asynchronous behavior of each enzyme molecule is not well represented by an ensemble average. Furthermore, it is experimentally challenging to study the sequence dependence of transient enzymatic pauses on templates comparable in length to the 9.7-kilobase HIV genome using bulk enzymology techniques.

In this work, we investigated the DNA polymerization activity of HIV-1 RT using a single molecule technique based on the hydrodynamic manipulation of DNA molecules. Single-molecule techniques offer the ability to monitor multiple turnovers of the polymerase in real-time without ensemble averaging and therefore present advantageous methods to study the unsynchronized, sequence-dependent dynamics during polymerization on a long DNA template [14–17]. Unlike short oligonucleotide DNA templates heretofore used for HIV-1 RT studies [7,11,18–21], bacteriophage λ (48.5 kilobases) contains variable DNA sequences and secondary structures in a row as commonly encountered in HIV genome *in vivo* [22]. We find that during polymerization on λ DNA, HIV-1 RT can switch between two distinct modes of DNA synthesis activity, fast primer extension and slow strand displacement, depending on whether downstream DNA is single-stranded or hairpin-structured. The rate of strand displacement synthesis is further analyzed as a function of the base sequence of hairpin stems and the template stretching force. The dependence of the strand displacement rate on GC-rich

hairpins and high template force led us to conclude that the free energy contribution of HIV-1 RT in basepair destabilization is significant, such that HIV-1 RT falls in between two limiting mechanisms—either the free energy release from dNTP hydrolysis drives opening of DNA duplex or the polymerase waits for thermal fraying of the DNA duplex. These new findings allow for a more quantitative understanding of strand displacement synthesis catalyzed by HIV-1 RT and offer new insights on hairpin-induced switching of HIV-1 RT activity along viral RNA or cDNA templates *in vivo*.

Results

Single-molecule trajectory of DNA synthesis catalyzed by HIV-1 RT on ssDNA template

In our single molecule experiments, individual single-stranded DNA templates are stretched by hydrodynamic flow in a microchannel (Fig. 1a). The termini of λ-phage DNA molecules are specifically end-functionalized with biotin and digoxigenin separately, such that the biotin-functionalized terminus of DNA is linked to a streptavidin-coated surface of a flow cell and the digoxigenin end is tethered to 2.8-μm magnetic beads functionalized with antidigoxigenin (Materials and Methods) [23–25]. Using the differential extension of single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA) at a constant stretching force (Fig. 1b), we are able to monitor the enzymatic conversion of ssDNA to dsDNA on a single molecule level based on the change in DNA extension over time, which is directly measured by observation of the tethered bead position as a function of time (Fig. 1c).

We studied the primer extension DNA replication activity of HIV-1 RT on flow-stretched ssDNA templates at room temperature (21 °C). An enzyme solution containing 11 nM HIV-1 RT and a saturating amount of dNTP (200 μM) was infused at a flow rate resulting in a template force of 3.7 pN (Materials and Methods). In the primer extension assay, we observed processive stretches of DNA synthesis occasionally interrupted by a slow mode of DNA replication, which appeared as a *plateau* ranging in time from a few to 200 seconds (an example trace is shown in Fig. 1d).

When instantaneous rates of the DNA polymerization trajectories are calculated and drawn into a histogram (Fig. 1e; Materials and Methods), there are a dominant peak (Gaussian shape) near 1 nucleotides per second (nt/s) due to contribution of the apparent plateaus and another peak near ~20 nt/s with a long tail. In contrast, when instantaneous rates are calculated from bead trajectories that do not show any DNA synthesis events, the histogram is a simple Gaussian with a center at ~0.05 nt/s (see Supporting Fig. 1). Therefore, slopes of the apparent plateaus are larger than those of the mechanical drift intrinsic to the assay, and we conclude that the plateaus arise from slow synthesis mode of HIV-1 RT.

Hairpin-induced switching of enzymatic activity to slow synthesis mode of HIV-1 RT

We found that slow DNA synthesis events occur at similar template locations among experiments in which we varied the primer sequence, temperature (21 °C and 37 °C), and enzyme concentrations (0.5–54 nM). At 37 °C, we observed a significant decrease in the durations of the slow DNA synthesis events, but their template locations remained unchanged from those observed at 21 °C. We suspected two reasons for this sequence dependence. First, homopolymeric runs of specific bases may cause slippage and pausing of HIV-1 RT [4,7,18,26,27]. However, we analyzed the template sequence and did not observe any significant correlation between homopolymeric nucleotide stretches on λ DNA and the locations of slow DNA synthesis.

Sequence-dependent slow synthesis may also result from hairpin structures on the single-stranded DNA template [7,18,21,27–29]. We determined that hairpin locations on λ DNA showed

a strong correlation with the locations of slow DNA synthesis. For the correlation, we identified stable DNA hairpins at the experimental salt concentration and temperature by a rigorous search of the λ-DNA sequence using the Mfold algorithm [30]. We counted hairpins with a formation energy $\Delta G_{formation}$ greater than k_BT after $\Delta G_{formation}$ is corrected by the work done by hydrodynamic stretching force acting against hairpin formation. The selected hairpins range from 3 to 25-base stem hairpins with a small bulge. As a straightforward initial approach, we summed the number of bases in hairpin stems within 250-base moving window along λ DNA to express hairpin strength as a function of template position.

The duration of DNA polymerization events within a 250-base moving window was measured from 64 experimental trajectories (Fig. 2; also see Supporting Materials). The amplitude of the dwell time function accounts for durations of slow synthesis as well as the number of slow synthesis events within the moving window; therefore, the function statistically reflects the relative "strength" of slow synthesis along the template sequence. The Pearson correlation coefficient between hairpin strength and the duration of DNA polymerization was 0.75 up to 10,000 nt. Hence, the observed slow synthesis by HIV-1 RT is interpreted as hairpin-induced switching of HIV-1 RT activity to slow DNA synthesis, namely strand displacement synthesis.

Enzyme concentration-independent kinetics of slow DNA synthesis

Given that hairpins induce HIV-1 RT switching to a much slower mode of synthesis, one easily questions whether HIV-1 RT dissociates at the stem of hairpins and, if so, whether the rebinding time dominates the duration of the slow synthesis. Enzyme dissociation at hairpin stems is likely to occur considering the low processivity of the HIV-1 RT and the process of template switching *in vivo*, during which RT jumps between two copies of the viral genome. We investigate the contribution of enzyme dissociation and rebinding during slow synthesis steps by measuring the time duration of slow synthesis regions at various RT concentrations (Materials and Methods). If enzyme rebinding is the major rate-limiting process during slow DNA synthesis events, the time durations should be dependent on the enzyme concentration due to the waiting time between enzyme dissociation and rebinding.

For experiments with RT concentrations in the range of 1–54 nM, the distribution of slow synthesis durations is well-fitted with a single exponential decay with similar decay constants (see Fig. 3a-b; mean of the similar decay constants was $0.027 \pm 0.0008\ s^{-1}$). When the concentration of HIV-1 RT was less than 1 nM, the distribution is better fit with a double exponential function, suggesting that there is more than one major rate-limiting process and enzyme dissociation and rebinding processes likely play a role in the slow synthesis events observed in between fast stretches of DNA synthesis. Therefore, enzyme dissociation and rebinding are not rate-limiting processes during slow synthesis events if the enzyme concentration is in excess of 1 nM (our experiments are mainly conducted at enzyme concentrations of 11 nM). These observations are consistent with the reported dissociation constant for interaction of HIV-1 RT and primer/template DNA between 0.65 nM – 5 nM [6, 31] and with a bulk level experiment demonstrating that secondary structures on DNA templates affect neither RT enzyme dissociation rate nor the enzyme dissociation constant compared to those measured on a hairpin-free template [21]. Because crystal structures of HIV-1 RT have shown that HIV-1 RT has a large binding cleft [1–3], it is possible that the dissociation constant of HIV-1 RT on a DNA template may be tolerant to the hairpin structure.

We further examined the role of hairpin thermal opening in the slow synthesis events. We found that the complete opening of a hairpin by thermal fluctuations has a much longer timescale than the duration of the slow synthesis (see Supporting Materials). Considering these observations, we conclude that the rate limiting process underlying the observed slow synthesis is based solely on enzyme kinetics of strand displacement synthesis.

Chi-square fitting of single-molecule trajectories to determine the two synthesis rates

Assuming that primer extension and strand displacement synthesis are the rate-limiting processes during fast synthesis at open, non-hybridized bases and slow synthesis at hairpin stem bases, respectively, we performed a weighted least square fitting of our experimental trajectories with two parameters: the primer extension rate and the strand displacement rate (Materials and Methods). We determined a primer extension rate of 18.7 ± 6.0 nt/s and a strand displacement rate of 2.3 ± 0.56 nt/s. Within the error range, the primer extension rate is similar to 22.1 ± 2.4 nt/s acquired from histogram of instantaneous slopes in raw trajectories. The strand displacement rate is higher than 0.28 nt/s measured at 37 °C *in vitro* [5] or 0.05–0.10 nt/s measured using a single molecule FRET technique [32] and lower than 2.98–4.5 nt/s measured with Moloney murine leukemia virus RT [33,34]. Clearly, the sequence of the template and stretching force can affect the measured strand displacement rate, as described in the next section.

Active or passive mechanism for strand displacement synthesis of HIV-1 RT

Active and passive mechanisms have been used to describe the duplex unwinding activity of helicases [35,36]. Briefly, the active mechanism describes a helicase as a strong molecular motor such that it converts chemical energy from NTP hydrolysis to unzip duplexed template with high efficiency. On the contrary, the passive mechanism states that a helicase translocates to the next base after the immediately adjacent downstream base pair opens as a result of thermal fluctuations. Indeed, these mechanisms are not inherently contradictory, as they represent two extremes of an energetic continuum and also describe helicases that exhibit a partial passive mechanism of unwinding, such as T7 helicase [37]. While the speed of active unzipping can be as high as the velocity of translocation on single-stranded template, the rate of passive unzipping is limited by thermal breathing of a base pair at the duplex junction. Recent theoretical work predicted that the duplex unwinding rate of a strictly passive helicase with a kinetic step size of 1 base pair would be 7 times slower than the rate of translocation along single-stranded nucleic acid template [38].

Our correlation analysis shown in Figure 2 considered the number of bases in hairpin stems as a putative “strength” of each hairpin and is in reasonably good agreement with the active mechanism of strand displacement of HIV-1 RT. The active mechanism suggests that the enzyme may efficiently couple the energy released from dNTP hydrolysis (~20 k_BT) to unzip the duplex junction, such that the small base pair energy difference between AT and GC (~1 k_BT range) is not large enough to contribute to enzyme dynamics (Fig. 4a). Hence, the time required for RT to proceed through a hairpin is proportional to the number of bases in the hairpin stem and is negligibly affected by variation in base content in the hairpin stem.

However, the rate of HIV-1 RT strand displacement is nearly 10-fold slower than the rate of primer extension, which suggests that an additional duplex melting step in the strand displacement process may be rate limiting (Fig. 4b) [38,39]. If we consider thermal breathing of base pairs during strand displacement synthesis, we should weight GC over AT base pairs in hairpin stems to estimate hairpin strength in the correlation analysis. We estimated base pair opening energies using literature values of 2.9 k_BT for a GC pair (E_{GC}) and 1.3 k_BT for an AT pair (E_{AT}) [40]. These values include contributions from unpairing of hydrogen bonds, unstacking from the next base, and base rearrangements. Based on these free energies for base opening, we estimated transition state activation energies for base opening ($\Delta G^{\ddagger}$) using a proportionality factor β (i.e. Brønsted value) [41]. The ratio of time required for thermal melting of a GC base pair (τ_{GC}) versus an AT base pair (τ_{AT}) was calculated by invoking the Hammond postulate:

$$\frac{\tau_{GC}}{\tau_{AT}} \sim \frac{\exp(-\Delta G^{\ddagger}_{AT}/k_B T)}{\exp(-\Delta G^{\ddagger}_{GC}/k_B T)} = \exp(\beta(\frac{E_{GC}}{k_B T} - \frac{E_{AT}}{k_B T})) = \exp((2.9 - 1.3)\beta)$$

where the proportionality factor β may range between 0 and 1. If changes in the free energies for base opening are directly related to changes in transition state energy, β = 1 and the ratio τ_{GC}/τ_{AT} = 5, which suggests that the time required for RT to traverse a GC base pair at the duplex junction is approximately 5 times longer than for an AT base pair during passive strand displacement synthesis. Therefore, we assigned a weighting factor of 5 to GC base pairs in counting the number of bases in hairpin stems. NMR studies on base-pair opening of DNA duplexes showed that lifetimes of GC base pairs are approximately 3 times longer than those of AT base pairs [42], which is close to our estimated value for the weighting factor.

Figure 5a shows the correlation between the weighted number of stem bases and the enzyme dwell time. The two curves demonstrate an improved correlation in the initial ~2500 base region compared to the active mechanism (Fig. 2; also see Supporting Material for discussions), and the overall Pearson correlation coefficient is improved to 0.87, which supports the idea that the synthesis rate of HIV-1 RT is sensitive to the base sequence in the hairpin stems.

We also found that the distribution of weighted number of bases in hairpin stems follows a single exponential distribution, similar to the histogram of time duration of slow synthesis (Fig. 3a); however, the distribution of the number of bases in hairpin stems does not show a single exponential (Fig. 5b–c). This result contradicts a purely active strand displacement by HIV-1 RT, whereby the duration of time spent in strand displacement mode should be proportional to the run length or the number of bases in hairpin duplex. Therefore, we conclude that HIV-1 RT does not play a purely active role in hairpin duplex unwinding.

Modeling strand displacement synthesis by HIV-1 RT

If HIV-1 RT were purely active, the strand displacement rate would be expected to be independent of the GC content in hairpin stems as well as template stretching forces; if the enzyme is passive, the strand displacement rate decreases with higher GC content in hairpin stems and increases with higher stretching forces. Previous single molecule studies on helicases developed a physical model to determine the active/passive character of helicases based on their DNA unzipping rate as a function of hairpin sequence and stretching forces [37,43]. We found the method can be applied to analyze the strand displacement rate of HIV-1 RT to quantify passiveness of HIV-1 RT. In order to apply the model developed for helicases to HIV-1 RT strand displacement data, we made two assumptions: (1) primer extension rate is sequence independent [6,44] and (2) base incorporation during strand displacement synthesis is preceded by duplex junction opening by *n* bases (Fig. 4), such that the observed strand displacement rate (k_{sd}) is lower than the primer extension rate (k_{pe}) by kinetic factors governing breathing of *n* base pairs including the base pair energy of DNA sequence, template stretching force, and the active/passive role of HIV-1 RT for base pair destabilization. In other words, the effective free energy to open 1 base pair (ΔG_{bp}) consists of the base pairing energy of the DNA duplex (ΔG_0) and the destabilization energy from template stretching force (ΔG_F) and enzyme (ΔG_{RT}).

$$k_{sd} = k_{pe}\exp(-n\frac{\Delta G_{bp}(F)}{k_B T}), \text{ where } \Delta G_{bp}(F) = \Delta G_0 - \Delta G_F - \Delta G_{RT}$$

The strand displacement rate, measured at a template stretching force of 3.7 pN, decreases with higher GC content in hairpin stems (Fig. 6a; Materials and Methods). The GC ratio in the hairpin stem bases yields ΔG_0 by using values 2.9 k_BT and 1.3 k_BT for GC and AT base pairs, respectively [40], and ΔG_F and k_{pe} can be obtained experimentally (see Supporting Materials). Fitting of data in Figure 6a yields $n = 2.6 \pm 0.30$ bases and $\Delta G_{RT} = 1.2 \pm 0.09$ k_BT (estimated standard deviation of the fit). For the closest integer value of $n = 3$ bases, $\Delta G_{RT} = 1.3$ k_BT.

In template force dependent measurements, we observed k_{pe} changes with template forces (Fig. 6b; Materials and Methods). In order to avoid sequence dependent variation of k_{sd}, we calculated k_{sd} around template position of 3700 nt where the strongest hairpin occurs in the first 10,000 nt of λ DNA (average GC ratio = 0.44). We observed that k_{sd} moderately increases with template forces (Fig. 6c). The best fit for $k_{sd}(F)/k_{pe}(F)$ was obtained with $n = 3.8 \pm 2.29$ bases and $\Delta G_{RT} = 1.4 \pm 0.30$ k_BT. For the closest integer value of $n = 4$ bases, $\Delta G_{RT} = 1.4$ k_BT. If HIV-1 RT were a purely passive enzyme, ΔG_{RT} should equal 0. Our results indicate that the contribution of the HIV-1 RT enzyme in destabilization of DNA duplex is comparable to 72 % of base-pairing energy. Hence, strand displacement of HIV-1 RT cannot be explained by either a completely passive or a purely active mechanism; however, HIV-1 RT stands somewhere in between these two extreme mechanisms.

Discussion

How other DNA polymerases respond to hairpins on the template?

The hairpin-induced switch in DNA polymerization activity has been observed in other DNA polymerases *in vitro* (e.g. DNA polymerase α [45,46], *E. coli* DNA polymerase II [47] and III [48], and T7 or T4 bacteriophage DNA polymerase [21,49]). However, addition of helicases or SSB oftentimes resolved the inhibitory effect of secondary structures on DNA replication by these DNA polymerases. With our single molecule assay, we also observed frequent pausing of T7 DNA polymerase spanning as long as a few tens of seconds during primer extension replication on flow-stretched λ DNA templates [25]. On the contrary, Lee, *et al* found no significant pauses of T7 DNA polymerase during leading-strand synthesis in the presence of helicase (gp4) [17]. With phi29 DNA polymerase, we observed a small number of less pronounced enzymatic pausing events during primer extension synthesis on ssDNA templates and little or no evidence of enzymatic pausing events during strand displacement replication on flow-stretched dsDNA templates (in all single molecule experiments, the enzyme concentration significantly exceeded the dissociation constant for the enzyme/primer-template pair) [23]. Enzymatic pauses of DNA polymerases at template secondary structures may be a general phenomenon in many DNA polymerases, but each DNA polymerase shows a different degree of sensitivity to the secondary structures. Moreover, most DNA polymerases work with a helicase or SSB in the context of a large macromolecular protein superstructure. Efficient strand displacement activity in some DNA polymerases, such as phi29 DNA polymerase, could also facilitate DNA polymerization through secondary structures. HIV-1 RT does not function in a replisome with customary accessory proteins that destabilize secondary structures and exhibits an intrinsically inefficient strand displacement activity; therefore, HIV-1 RT can be significantly affected by template secondary structures [7,11,18,21,50].

Comparison with previous kinetic studies

Our data suggests that HIV-1 RT follows a relatively passive mechanism during strand displacement synthesis on secondary structures. Kinetic measurements of single nucleotide incorporation by HIV-1 RT at pause or non-pause template sites provided a basis for understanding our data [20,21]. Suo and Johnson determined that polymerization at the template pause sites (generally occurring near the first base of a DNA hairpin stem) displays biphasic kinetics—small reaction amplitudes of fast phase (10–20 s^{-1}) and large amplitudes of slow

phase (0.02–0.07 s^{-1}); only single phase kinetics was observed at non-pause template DNA sites, comprised solely of a large amplitude fast phase (33–37 s^{-1}) [21]. To explain their data, the authors proposed a model whereby HIV-1 RT binds to DNA either in a productive or non-productive state at pause sites. A fast kinetic phase suggests that HIV-1 RT productively turns over to add a single nucleotide to the extending DNA chain; however, a slow kinetic phase results from the non-productive enzyme state which must be converted to the productive state upon melting of the next stem base pair. We can relate these two distinct kinetic schemes for pause and non-pause sites to two distinct kinetic patterns in our single molecule assay: replication trajectory regions exhibiting positive slopes represent fast single kinetics whereas slow synthesis regions represent the dominant slow phase in biphasic kinetics at pause sites. The timescale of the slow phase kinetics was approximately 30 seconds [21], which agrees with the range of our observed slow synthesis durations. While a conformational change of the polymerase is likely the rate-limiting step at non-pause sites [6,9,19], thermal melting of the next stem base pair, or conversion of the non-productive state to the productive state, may serve as the rate-limiting step at pause sites, which supports a relatively passive mechanism of strand displacement synthesis for HIV-1 RT.

Force dependence of primer extension rate

Template force dependence of HIV-1 RT activity of DNA polymerization allowed for a comparison of HIV-1 RT with other DNA polymerases that were studied at single molecule level. Although MD simulations have generated an interesting discussion regarding polymerization rate as a function of force [51], we focus our attention on the relatively low stall force of HIV-1 RT (Fig. 6b). The stall force for HIV-1 RT was ~ 7 pN whereas T7 DNA polymerase stalls at ~34 pN [14], *E. coli* Klenow fragment at ~20 pN [15], and phi29 DNA polymerase at ~ 37 pN [52]; in addition, this stall force is lower than that measured by atomic force microscopy (AFM) where HIV-1 RT molecules were adsorbed on the cantilever tip. The AFM experiment was conducted by applying a stretching force to both the enzyme molecule and ssDNA template, as opposed to application of a small stretching force to DNA templates in our assay [53]. We found that phi29 DNA polymerase shows robust DNA polymerization activity at template forces far in excess of 8 pN in the same experimental conditions, so the strong dependence of DNA synthesis on the template stretching force is an intrinsic property of HIV-1 RT. The strong sensitivity of HIV-1 RT replication to template stretching force can be understood when considering the low processivity of HIV-1 RT and its weak interaction of DNA template. In the DNA polymerase active site, amino acid residues of the DNA polymerase form a fine network with the template DNA base, incoming dNTP, two Mg^{2+} ions, and water molecules, and the correct coordination or interaction among these elements is required for the phosphodiester bond formation step [54,55]. As higher stretching force is applied to the ssDNA template, the probability of incoming dNTP to position itself correctly in the fine coordination network of active site decreases. Compared to fast and highly-processive T7 DNA polymerase [56], HIV-1 RT does not bind as strongly to the DNA template, and the protein surface area around the DNA polymerization active site in HIV-1 RT is more open compared to that in T7 DNA polymerase. Previous mutational studies also showed that the flexible active site may explain low fidelity of HIV-1 RT [57]. We conjecture that these loose structural characteristics may allow HIV-1 RT to tolerate secondary structures on single-stranded templates but render HIV-1 RT unable to replicate on templates stretched with even moderate forces.

HIV-1 RT strand displacement mechanism

The sequence dependence in correlation analysis (Fig. 5a) and sequence dependent strand displacement rate (Fig. 6a) rule out a purely active mechanism for HIV-1 RT. However, the energetic contribution of the enzyme (G_{RT}) near 1.4 k_BT opposes a purely passive mechanism for HIV-1 RT as well. Although the exact mechanism is not known, HIV-1 RT may undergo a conformational change after dNTP hydrolysis, which contributes nearly 72% of the average

base pairing energy and promotes template duplex unwinding. This suggests that HIV-1 RT cannot be explained by either a purely active or an entirely passive mechanism. HIV-1 RT is another example where the extreme active or passive classification cannot be applied for duplex unzipping processes [37].

A previous *in vitro* study using $KMnO_4$ oxidation revealed that HIV-1 RT melts two bases ahead of the primer terminus in a strand displacement construct [58,59]. The small difference between our results of $n = 3$ or 4 and 2 may be attributed to differences in experimental design and simple assumptions in the fitting model.

Note that fitting of the model to the sequence-dependent strand displacement rate (Fig. 6a) yields similar values of n and ΔG_{RT} as from force dependent rates (Fig. 6c). Analysis of the unzipping rate as a function of GC content in hairpin stems was not tried in previous studies where DNA templates with a single hairpin were used. By using λ DNA templates, we had various naturally occurring hairpins in one template molecule, and we were able to study strand displacement rates as a function of hairpin stem sequence as well.

Significance of hairpin-induced activity of HIV-1 RT for viral survival

Why does the HIV-1 genome have several hairpins [22], and why does HIV-1 RT remain an inefficient polymerase with high mutation rates? These are interesting and relevant questions. Frequent mutation of the HIV-1 genome is critical for viral proliferation and results in continued resistance to anti-retroviral drugs. In addition to lack of an intrinsic proofreading mechanism and low fidelity of HIV-1 RT [6,8], enzymatic pausing events at homopolymeric sequences or at hairpin structures have been proposed to enhance viral mutagenesis [10,11,26]. In particular, stable secondary structures on RNA and newly synthesized cDNA facilitate strand transfer and template switching, leading to genetic recombination in the retrovirus [7, 10,11,13]. Also, a recent study on secondary structures in the HIV-1 RNA genome reported the role of stem-loop structures in ribosomal pausing during translation, which may result in frame shifting [22]. Therefore, a large number of stable secondary structures in the HIV-1 genome may serve an important role for viral survival: HIV-1 *per se* may favor secondary structures in its genomic RNA (and intermediate cDNA) in order to increase overall mutations in future generations of the virus [60]. In this view, a moderately passive mechanism for strand displacement may allow HIV-1 RT to replicate through various hairpins while maintaining the ability to introduce mutations during slow strand displacement synthesis.

We should note that nucleocapsid proteins (NC) interact with HIV-1 RT in several steps during reverse transcription *in vivo*. NC is known to facilitate DNA duplex melting and reannealing *in vitro* [60–63], and previous *in vitro* studies have shown that complex formation between NC and template helps in the synthesis of a long DNA product by HIV-1 RT [29,64,65]. It is interesting to consider a molecular mechanism whereby NC affects secondary structure for the passage of HIV-1 RT during DNA replication. Future single molecule studies of HIV-1 RT polymerization activity on the HIV-1 viral genomic sequence in the presence of nucleocapsid proteins have the potential to reveal useful information regarding sequence-dependent DNA polymerization dynamics of HIV-1 RT.

Materials and Methods

Hydrodynamic flow-stretching assay

In order to observe slow HIV-1 RT activity, we used a previously described experimental single molecule assay with high mechanical stability and spatial resolution [23]. High degrees of mechanical stability and spatial resolution are crucial to observe the relatively slow DNA polymerization activity of HIV-1 RT over a long period of time; furthermore, a stable assay

enables direct identification of fast and slow modes of DNA synthesis from raw experimental trajectories without post-processing of the experimental data.

The hydrodynamic flow force exerted on the DNA-tethered beads was calibrated using the equipartition theorem, $F = k_B Tl/\langle \delta dy^2 \rangle$, where k_B is Boltzmann's constant, T is absolute temperature, l is the length of DNA molecule, and $\langle \delta dy^2 \rangle$ is the transverse mean square displacement of a dsDNA-tethered bead [66]. We confirmed that ssDNA and dsDNA templates experience the same force under a constant flow rate, regardless of degree of ssDNA/dsDNA conversion [23]. In addition to the hydrodynamic force exerted on DNA-tethered beads, a small rare-earth magnet positioned over the flow cell provides a weak vertical force (~0.5 pN) to gently levitate beads above the surface, thereby minimizing nonspecific bead adsorption to the coverslip surface. DNA-tethered beads are imaged using through-objective dark-field microscopy which results in a mechanically stable apparatus. Multiplexed images of 20–40 ssDNA-tethered beads are simultaneously recorded by a high-resolution charge-coupled device camera at a frame rate of 1 Hz. Trajectories of bead-tethered ssDNA molecules that exhibit DNA synthesis events are obtained by Gaussian centroid determination of bead positions in each frame by using Diatrack software (Semasopht, North Epping, Australia).

The number of bases synthesized by HIV-1 RT was calculated from displacement of bead centroid position by the following equation:

$$\#\text{ of nucleotides added} = \frac{\text{change in bead centroid position } (\mu\text{m})}{|L_{dsDNA}(F) - L_{ssDNA}(F)|} \times 48.5 \text{ kilobases}$$

where $L_{dsDNA}(F)$ and $L_{ssDNA}(F)$ are extension length of λ dsDNA and ssDNA at stretching force F.

DNA templates and enzymes

Bacteriophage λ DNA (48.5 kilobases) was purchased from New England Biolabs (Ipswich, MA) and was used as template DNA in primer extension experiments. Biotin or digoxigenin-labeled oligonucleotides and the DNA primer for the primer extension assay were purchased from Integrated DNA Technologies (Coralville, IA). Further information on DNA template construction is described in detail in Kim, *et al* [23]. Recombinant HIV-1 RT was purchased from Worthington Biochemical Corporation (Lakewood, NJ), and phi29 DNA polymerase was purchased from New England Biolabs. We confirmed purity of the HIV-1 RT stock by running SDS-polyacrylamide gel electrophoresis.

Primer extension assay

Single molecule experiments for primer extension DNA synthesis were conducted with ssDNA tethers, which were generated either by denaturing dsDNA with high pH treatment or by digestion of the non-tethered strand by λ exonuclease (experimental details may be found elsewhere [23,25]). DNA synthesis was monitored while HIV-1 RT reaction buffer (50 mM Tris-HCl pH 8.3, 40 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 100 μg/ml BSA) containing saturating amounts of substrate dNTP (200 μM) and HIV-1 RT (11 nM) was infused into the microfluidic flow cell. DNA replication initiates at the 3′ end of a 21-base primer (5′-AGG TCG CCG CCC CGT AAC CTG-3′) annealed near the surface end of the ssDNA tether. Unless otherwise stated, experiments were conducted at room temperature (21°C) with a solution flow rate of 2.9 mL/hr, which results in an equivalent template force of 3.7 pN.

In phi29 DNA polymerase experiments, we used 15 nM enzyme solution in 50 mM Tris-HCl pH 7.5, 10 mM $(NH_4)_2SO_4$, 10 mM $MgCl_2$, 4 mM DTT, 200 μg/ml BSA.

Data analysis

Instantaneous polymerization rates from raw trajectories—We defined instantaneous polymerization rates as the current DNA synthesis rates at any moment in time along an experimental replication trajectory (e.g. Fig. 1d) and calculated them from raw trajectories by a least-square fitting approach with a fixed moving window. The window size was chosen large enough to avoid experimental noise taken as enzymatic rate and small enough to avoid slow synthesis phase underestimating the rate of primer extension polymerization events. We used window size of 20 for data acquired at 1 Hz (also see Supporting Materials).

Determining duration of slow synthesis events—When the instantaneous rate is under a certain threshold value, it is taken as a slow synthesis. The threshold for determining slow synthesis is critical, as low threshold values result in experimental noise underestimating the durations, whereas large threshold values result in failure to detect events in a trajectory. The first and last slow synthesis events recorded in each trajectory were discarded because they can include enzyme binding and complete dissociation, respectively. The calculations were performed with Matlab, and exponential curve fitting on the duration histogram was performed using Igor Pro (WaveMetrics, Lake Oswego, OR).

Chi square fitting—Our model predicts enzyme dwell time is composed of primer extension (rate of a) 250 $stem(x)$ $stem(x)$ and strand displacement synthesis (rate of b):

$T(x_i)=\frac{250-stem(x_i)}{a}+\frac{stem(x_i)}{b}$, where stem($x_i$) is the number of bases in hairpin stems within moving window x_i along the template. We found parameters (a, b), which minimizes

$\chi^2=\sum_{i=1}^{n}\left(\frac{t(x_i)-T(x_i)}{\sigma_i}\right)^2$, where $t(x_i)$ is enzyme dwell time in 250-bp moving window, and σ_i is standard error in $t(x_i)$ [67]. We used standard deviation in $t(x_i)$ in 64 traces for the standard error, σ_i. Further details and a fitting plot are provided in Supporting Materials.

Estimation of strand displacement rate vs. GC ratio—Based on the assumption that enzyme dwell time is composed of primer extension and strand displacement and that primer extension rate is not sequence dependent (18.7 ± 6.0 nt/s) [6,44], we estimated strand displacement rate from average dwell time with window of 250 bp. The error bars in Figure 6a are estimated from error in average dwell time and primer extension rate. See Supporting Materials for further details.

Primer extension rate as a function of template force—We calculated force dependent primer extension rate from a histogram of instantaneous polymerization rates from multiple traces of DNA synthesis (>20; Fig. 1e). We extracted the primer extension activity population by fitting a Gaussian function to the peak around 1 nt/s and subtracting the Gaussian fit from the original histogram. The remaining histogram for primer extension activity was fitted with a lognormal function in order to fit a relatively long tail of the primer extension rates. Primer extension rate was calculated from mean of the lognormal function (Fig. 6b). See Supporting Materials for further details.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Paul C. Blainey for discussions. This work is supported by Jane Coffin Childs Memorial Fund for Medical Research Fellowship and a NIH Pathway to Independence Award for C.M.S. and a NIH Director's Pioneer Award to X.S.X.

Abbreviations

HIV-1 RT	human immunodeficiency virus type-1 reverse transcriptase
SSB	single-stranded DNA binding protein
ssDNA	single-stranded DNA
dsDNA	double-stranded DNA
nt	nucleotides
bp	basepairs
k_{pe}	primer extension rate
k_{sd}	strand displacement rate

References

1. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science 1992;256:1783–1790. [PubMed: 1377403]
2. Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science 1998;282:1669–1675. [PubMed: 9831551]
3. Ding J, Das K, Hsio Y, Sarafianos SG, Clark ADJ, Jacobo-Molina A, Tantillo C, Hughes SH, Arnold E. Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody Fab fragment at 2.8A resolution. J Mol Biol 1998;284:1095–1111. [PubMed: 9837729]
4. Huber HE, McCoy JM, Seehra JS, Richardson CC. Human Immunodeficiency Virus 1 reverse transcriptase. J Biol Chem 1989;264:4669–4678. [PubMed: 2466838]
5. Fuentes GM, Rodriguez-Rodriguez L, Palaniappan C, Fay PJ, Bambara RA. Strand displacement synthesis of the long terminal repeats by HIV reverse transcriptase. J Biol Chem 1996;271:1966–1971. [PubMed: 8567645]
6. Kati WM, Johnson KA, Jerva LF, Anderson KS. Mechanism and fidelity of HIV reverse transcriptase. J Biol Chem 1992;267:25988–25997. [PubMed: 1281479]
7. Klarmann GJ, Schauber CA, Preston BD. Template-directed pausing of DNA synthesis by HIV-1 reverse transcriptase during polymerization of HIV-1 sequence in vitro. J Biol Chem 1993;268:9793–9802. [PubMed: 7683663]
8. Ji J, Loeb LA. Fidelity of HIV-1 reverse transcriptase copying RNA in vitro. Biochemistry 1992;31:954–958. [PubMed: 1370910]
9. Hsieh J, Zinnen S, Modrich P. Kinetic mechanism of the DNA-dependent DNA polymerase activity of HIV reverse transcriptase. J Biol Chem 1993;268:24607–24613. [PubMed: 7693703]
10. Negroni M, Buc H. Mechanisms of retroviral recombination. Annu Rev Genet 2001;35:275–302. [PubMed: 11700285]
11. Roda RH, Balakrishnan M, Hanson MN, Wohrl BM, Le Grice SFJ, Roques BP, Gorelick RJ, Bambara RA. Role of the reverse transcriptase, nucleocapsid protein, and template structure in the two-step transfer mechanism in retroviral recombination. J Biol Chem 2003;278:31536–31546. [PubMed: 12801926]
12. Balakrishnan M, Fay PJ, Bambara RA. The kissing hairpin sequence promotes recombination within the HIV-I 5′ leader region. J Biol Chem 2001;276:36482–36492. [PubMed: 11432862]

13. Moumen A, Polomack L, Roques B, Buc H, Negroni M. The HIV-1 repeated sequence R as a robust hot-spot for copy-choice recombination. Nucleic Acids Res 2001;29:3814–3821. [PubMed: 11557813]
14. Wuite GJL, Smith SB, Young M, Keller D, Bustamante C. Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. Nature 2000;404:103–106. [PubMed: 10716452]
15. Maier B, Bensimon D, Croquette V. Replication by a single DNA polymerase of a stretched single-stranded DNA. Proc Natl Acad Sci USA 2000;97:12002–12007. [PubMed: 11050232]
16. Shundrovsky A, Santangelo TJ, Roberts JW, Wang MD. A single-molecule technique to study sequence-dependent transcription pausing. Biophys J 2004;87:3945–3953. [PubMed: 15465875]
17. Lee JB, Hite RK, Hamdan SM, Xie XS, Richardson CC, van Oijen AM. DNA primase acts as a molecular brake in DNA replication. Nature 2006;439:621–624. [PubMed: 16452983]
18. Harrison GP, Mayo MS, Hunter E, Lever AML. Pausing of reverse transcriptase on retroviral RNA templates is influenced by secondary structures both 5′ and 3′ of the catalytic site. Nucleic Acids Res 1998;26:3433–3442. [PubMed: 9649630]
19. Pop MP, Biebricher CK. Kinetic analysis of pausing and fidelity of human immunodeficiency virus type 1 reverse transcription. Biochemistry 1996;35:5054–5062. [PubMed: 8664298]
20. Suo Z, Johnson KA. Effect of RNA secondary structure on the kinetics of DNA synthesis catalyzed by HIV-1 reverse transcriptase. Biochemistry 1997;36:12459–12467. [PubMed: 9376350]
21. Suo Z, Johnson KA. DNA secondary structure effects on DNA synthesis catalyzed by HIV-1 reverse transcriptase. J Biol Chem 1998;273:27259–27267. [PubMed: 9765249]
22. Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess JW Jr, Swanstrom R, Burch CL, Weeks KM. Architecture and secondary structure of an entire HIV-1 RNA genome. Nature 2009;460:711–716. [PubMed: 19661910]
23. Kim S, Blainey PC, Schroeder CM, Xie XS. Multiplexed single-molecule assay for enzymatic activity on flow-stretched DNA. Nat Methods 2007;4:397–399. [PubMed: 17435763]
24. van Oijen AM, Blainey PC, Crampton DJ, Richardson CC, Ellenberger T, Xie XS. Single-molecule kinetics of lambda exonuclease reveal base dependence and dynamic disorder. Science 2003;301:1235–1238. [PubMed: 12947199]
25. Schroeder, CM.; Blainey, PC.; Kim, S.; Xie, XS. Hydrodynamic flow-stretching assay for single-molecule studies of nucleic acid-protein interactions. In: Selvin, PR.; Ha, T., editors. Single molecule techniques a laboratory manual. Cold Spring Harbor Laboratory Press; Cold Spring Harbor: 2008. p. 461-492.
26. Ji J, Hoffmann JS, Loeb L. Mutagenicity and pausing of HIV reverse transcriptase during HIV plus-strand DNA synthesis. Nucleic Acids Res 1994;22:47–52. [PubMed: 7510388]
27. DeStefano JJ, Buiser RG, Mallaber LM, Fay PJ, Bambara RA. Parameters tha influence processive synthesis and site-specific termination by human immunodeficiency virus reverse transcriptrase on RNA and DNA templates. Biochim Biophys Acta 1992;1131:270–280. [PubMed: 1378301]
28. Olsen DB, Carroll SS, Culberson JC, Shafer JA, Kuo LC. Effect of template secondary structure on the inhibition of HIV-1 reverse transcriptase by a pyridinone non-nucleoside inhibitor. Nucl Acids Res 1994;22:1437–1443. [PubMed: 7514786]
29. Klasens BIF, Huthoff HT, Das AT, Jeeninga RE, Berkhout B. The effect of template RNA structure on elongation by HIV-1 reverse transcriptase. Biochim Biophys Acta 1999;1444:355–370. [PubMed: 10095059]
30. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 2003;31:3406–3415. [PubMed: 12824337]
31. Muller B, Restle T, Reinstein J, Goody RS. Interaction of fluorescently labeled dideoxynucleotides with HIV-1 reverse transcriptase. Biochemistry 1991;30:3709–3715. [PubMed: 1707667]
32. Liu S, Abbondanzieri EA, Rausch JW, Le Grice SFJ, Zhuang X. Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. Science 2008;322:1092–1097. [PubMed: 19008444]
33. Kelleher CD, Champoux JJ. Characterization of RNA strand displacement synthesis by Moloney murine leukemia virus reverse transcriptase. J Biol Chem 1998;273:9976–9986. [PubMed: 9545343]

34. Whiting SH, Champoux JJ. Properties of strand displacement synthesis by Moloney Murine Leukemia Virus reverse transcriptase: mechanistic implications. J Mol Biol 1998;278:559–577. [PubMed: 9600839]
35. Lohman TM, Tomko EJ, Wu CG. Non-hexameric DNA helicases and translocases: mechanisms and regulation. Nat Rev Mol Cell Biol 2008;9:391–401. [PubMed: 18414490]
36. Pyle AM. Translocation and unwinding mechanisms of RNA and DNA helicases. Annu Rev Biophys 2008;37:317–336. [PubMed: 18573084]
37. Johnson DS, Bai L, Smith BY, Patel SS, Wang MD. Single-molecule studies reveal dynamics of DNA unwinding by the ring-shaped T7 helicase. Cell 2007;129:1299–1309. [PubMed: 17604719]
38. Betterton MD, Julicher F. Opening of nucleic-acid double strands by helicases: active versus passive opening. Phys Rev E 2005;71:011904.
39. Whiting SH, Champoux JJ. Strand displacement synthesis capability of moloney murine leukemia virus reverse transcriptase. J Virol 1994;68:4747–4758. [PubMed: 7518525]
40. Bockelmann U, Essevaz-Roulet B, Heslot F. Molecular stick-slip motion revealed by opening DNA with piconewton forces. Phys Rev Lett 1997;79:4489–4492.
41. Fersht, A. Structure and mechanism in protein science. W.H. Freeman and Company; New York: 1999.
42. Leroy JL, Kochoyan M, Huynh-Dinh T, Gueron M. Characterization of base-pair opening in deoxynucleotide duplexes using catalyzed exchange of the imino proton. J Mol Biol 1988;200:223–238. [PubMed: 2836594]
43. Lionnet T, Spiering MM, Benkovic SJ, Bensimon D, Croquette V. Real-time observation of bacteriophage T4 gp41 helicase reveals an unwinding mechanism. Proc Natl Acad Sci USA 2007;104:19790–19795. [PubMed: 18077411]
44. Thrall SH, Krebs R, Wohrl BM, Cellai L, Goody RS, Restle T. Pre-steady-state kinetic characterization of RNA-primed initiation of transcription by HIV-1 reverse transcriptase and analysis of the transition to a processive DNA-primed polymerization mode. Biochemistry 1998;37:13349–13358. [PubMed: 9786651]
45. Villani G, Fay PJ, Bambara RA, Lehman IR. Elongation of RNA-primed DNA templates by DNA polymerase alpha from Drosophila melanogaster embryos. J Biol Chem 1981;256:8202–8207. [PubMed: 6790535]
46. Weaver DT, DePamphilis ML. The role of palindromic and non-palindromic sequences in arresting DNA synthesis in vitro and in vivo. J Mol Biol 1984;180:961–986. [PubMed: 6098692]
47. Sherman LA, Gefter ML. Studies on the mechanism of enzymatic DNA elongation by E.coli DNA polymerase II. J Mol Biol 1976;103:61–76. [PubMed: 60493]
48. LaDuca RJ, Fay PJ, Chuang C, McHenry CS, Bambara RA. Site-specific pausing of deoxyribonucleic acid synthesis catalyzed by four forms of E.coli DNA polymerase III. Biochemistry 1983;22:5177–5188. [PubMed: 6360204]
49. Hacker KJ, Alberts BM. The rapid dissociation of the T4 DNA polymerase holoenzyme when stopped by a DNA hairpin helix. A model for polymerase release following the termination of each Okazaki fragment. J Biol Chem 1994;269:24221–24228. [PubMed: 7929078]
50. Suo Z, Johnson KA. RNA secondary structure switching during DNA synthesis catalyzed by HIV-1 reverse transcriptase. Biochemistry 1997;36:14778–14785. [PubMed: 9398198]
51. Andricioaei I, Goel A, Herschbach D, Karplus M. Dependence of DNA polymerase replication rate on external forces: a model based on molecular dynamics simulations. Biophys J 2004;87:1478–1497. [PubMed: 15345530]
52. Ibarra B, Chemla YR, Plyasunov S, Smith SB, Lazaro JM, Salas M, Bustamante C. Proofreading dynamics of a processive DNA polymerase. EMBO J 2009;28:2794–2802. [PubMed: 19661923]
53. Lu H, Macosko J, Habel-Rodriguez D, Keller RW, Brozik JA, Keller DJ. Closing of the fingers domain generates motor forces in the HIV reverse transcriptase. J Biol Chem 2004;279:54529–54532. [PubMed: 15385563]
54. Steitz TA. DNA- and RNA-dependent DNA polymerases. Curr Opin Struct Biol 1993;3:31–38.
55. Steitz TA. DNA polymerases: structural diversity and common mechanisms. J Biol Chem 1999;274:17395–17398. [PubMed: 10364165]

56. Doublie S, Tabor S, Long AM, Richardson CC, Ellenberger T. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution. Nature 1998;391:251–258. [PubMed: 9440688]
57. Harris D, Kaushik N, Pandey PK, Yadav PNS, Pandey VN. Functional analysis of amino acid residues constituting the dNTP binding pocket of HIV-1 reverse transcriptase. J Biol Chem 1998;273:33624–33634. [PubMed: 9837947]
58. Winshell J, Champouz JJ. Structural alterations in the DNA ahead of the primer terminus during displacement synthesis by reverse transcriptases. J Mol Biol 2001;306:931–943. [PubMed: 11237609]
59. Winshell J, Paulson BA, Buelow BD, Champoux JJ. Requirements for DNA unpairing during displacement synthesis by HIV-1 reverse transcriptase. J Biol Chem 2004;279:52924–52933. [PubMed: 15465813]
60. Bampi C, Jacquenet S, Lener D, Decimo D, Darlix J. The chaperoning and assistance roles of the HIV-1 nucleocapsid protein in proviral DNA synthesis and maintenance. Curr HIV Res 2004;2:79–92. [PubMed: 15053342]
61. Williams MC, Rouzina I, Wenner JR, Gorelick RJ, Musier-Forsyth K, Bloomfield VA. Mechanism for nucleic acid chaperone actiity of HIV-1 nucleocapsid protein revealed by single molecule stretching. Proc Natl Acad Sci USA 2001;98:6121–6126. [PubMed: 11344257]
62. Cosa G, Harbron EJ, Zeng Y, Liu HW, O'Connor DB, Eta-Hosokawa C, Musier-Forsyth K, Barbara PF. Secondary structure and secondary structure dynamics of DNA hairpins complexed with HIV-1 NC protein. Biophys J 2004;87:2759–2767. [PubMed: 15454467]
63. Levin, JG.; Guo, JH.; Rouzina, I.; Musier-Forsyth, K. Progress in Nucleic Acid Research and Molecular Biology, Vol 80. Vol. 80. Elsevier Academic Press Inc; San Diego: 2005. Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: Critical role in reverse transcription and molecular mechanism; p. 217-286.
64. Ji X, Klarmann GJ, Preston BD. Effect of human immunodeficiency virus type 1 nucleocapsid protein on HIV-1 reverse transcriptase activity in vitro. Biochemistry 1996;35:132–143. [PubMed: 8555166]
65. Anthony RM, DeStefano JJ. In vitro synthesis of long DNA products in reactions with HIV-RT and nucleocapsid protein. J Mol Biol 2007;365:310–324. [PubMed: 17070544]
66. Strick TR, Allemand JF, Bensimon D, Croquette V. The elasticity of a single supercoiled DNA molecule. Science 1996;271:1835–1837. [PubMed: 8596951]
67. Press, WH.; Teukolsky, SA.; Vetterling, WT.; Flannery, BP. Numerical recipes in C. Cambridge University Press; New York: 1992.

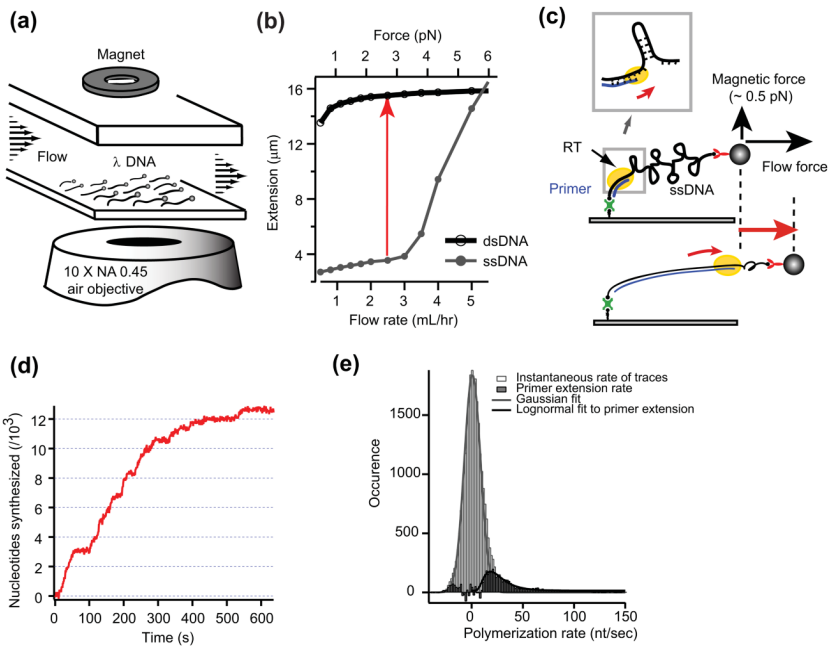

Figure 1. DNA replication on flow-stretched ssDNA by HIV-1 RT

(a) Schematic diagram of the experimental setup. (b) Force-extension curves for ssDNA and dsDNA. The polymerase activity of ssDNA conversion to dsDNA can be monitored as a change in the extension length of each DNA molecule at a given force (red arrow). (c) Schematic of DNA polymerization by HIV-1 RT resulting in a lengthening of a ssDNA tether. (Inset) As HIV-1 RT encounters hairpin duplex during DNA polymerization, what is the mechanism for the replication on this region? (d) Example trajectory from tracking a DNA-tethered bead over time. (e) Histogram of HIV-1 RT instantaneous polymerization rates from raw polymerization trajectories at a stretching force of 3.0 pN. Effective plateaus are reflected by a Gaussian peak at around 0 nt/s. Original distribution subtracted with the Gaussian fit yields a population for the primer extension activity, which is fitted with a lognormal function (black line).

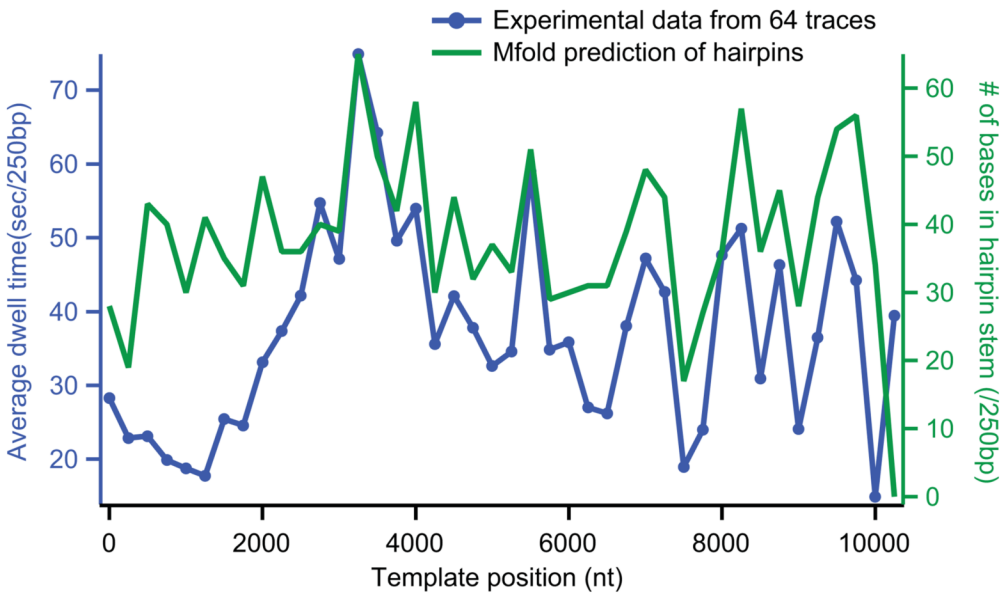

Figure 2. Sequence dependent strand displacement synthesis of HIV-1 RT

Average dwell time (blue, dot) measured along a λ-DNA template (average of 64 traces) and the number of bases involved in hairpin stems in 250-base moving window (green, line). Correlation analysis between the average dwell time and hairpin stem number density yields a Pearson coefficient of 0.75 up to 10,000 nucleotides.

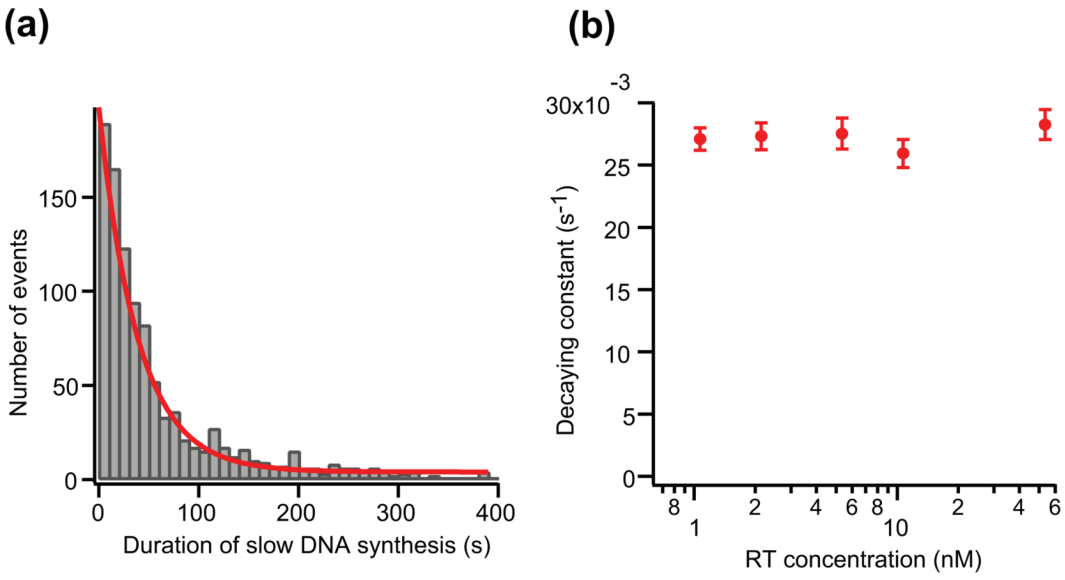

Figure 3. Enzyme concentration dependence of durations of slow DNA synthesis

(a) histogram of durations of slow DNA synthesis regions observed during DNA polymerization of 11 nM HIV-1 RT. The histogram is well-fitted to a single exponential function with a decay constant $0.026 \pm 0.001\ s^{-1}$. (b) decay constants for the single exponential fit to histograms of slow synthesis durations measured at enzyme concentrations between 1–54 nM.

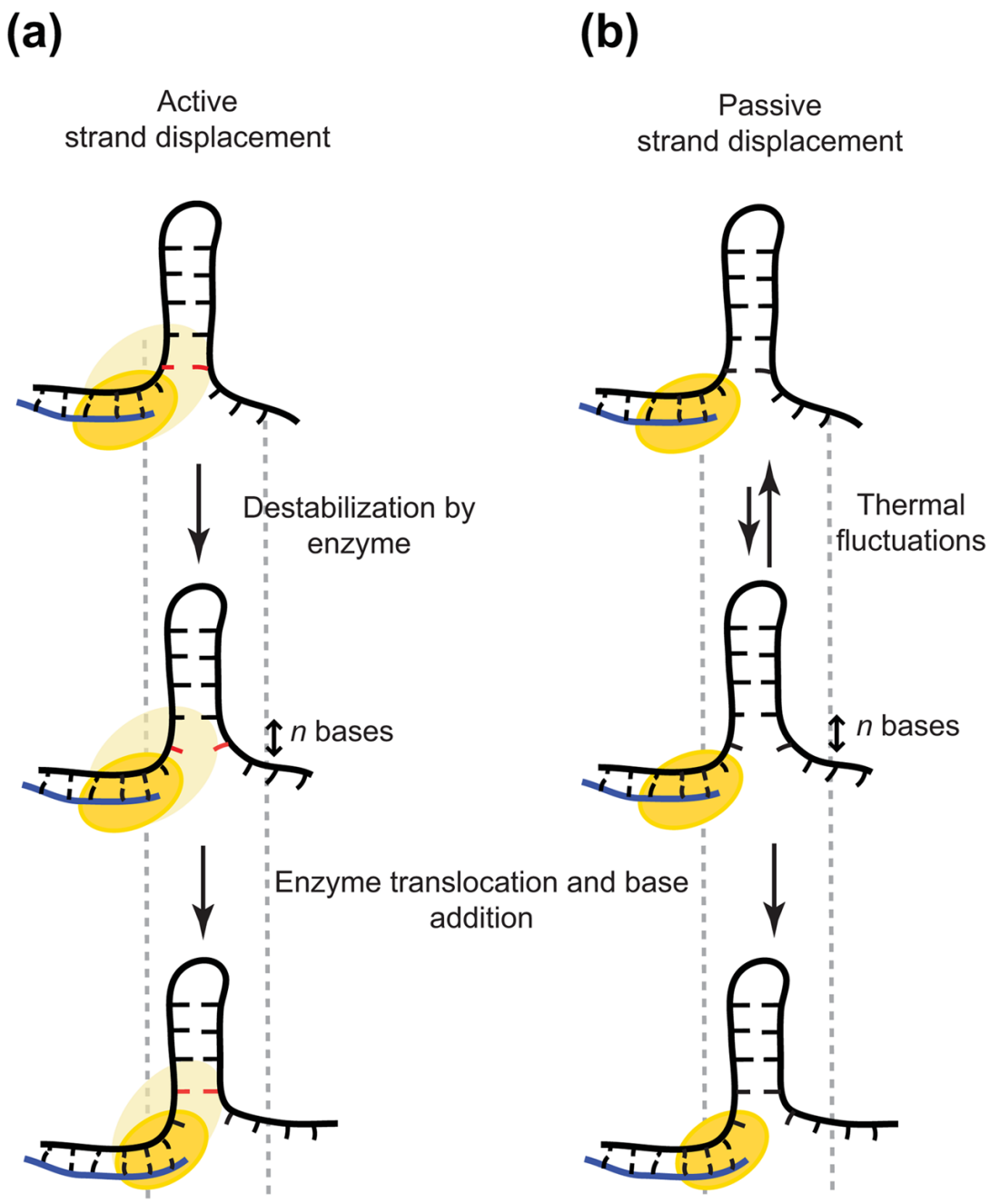

Figure 4. Active and passive mechanisms for strand displacement DNA synthesis by HIV-1 RT near hairpin locations

(a) Active strand displacement. Yellow imprint demonstrates the interaction between the enzyme and DNA, which destabilizes DNA junction (red base). (b) Passive strand displacement. The enzyme waits for junction opening by thermal fluctuations. In both cases, *n* bases are the number of bases open in front of the enzyme in each turnover of strand displacement synthesis.

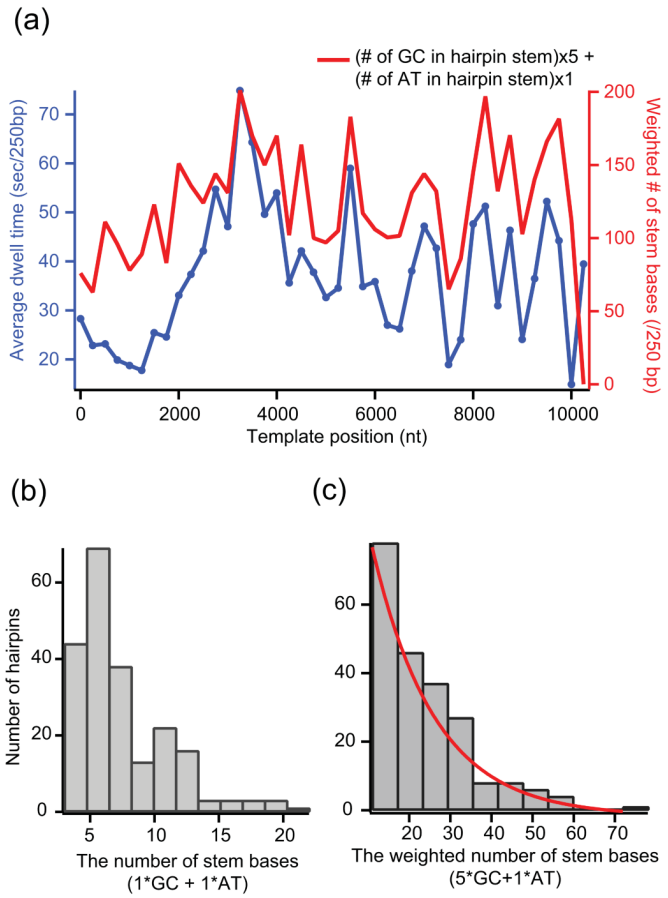

Figure 5. Relatively passive mechanism for strand displacement synthesis of HIV-1 RT

(a) The average dwell time (blue, dot) is plotted with the weighted number of bases in hairpin stems in 250-base moving window (red, line). The number of hairpin stem bases is weighted according to the base content in order to account for different timescale for thermal opening in GC and AT base pairs. Correlation analysis between the two curves yields a Pearson coefficient of 0.87. (b) Histogram of the number of bases in hairpin stems found in λ DNA template. (c) Distribution of the weighted number of stem bases in the DNA template. This is a histogram version of the red curve in (a).

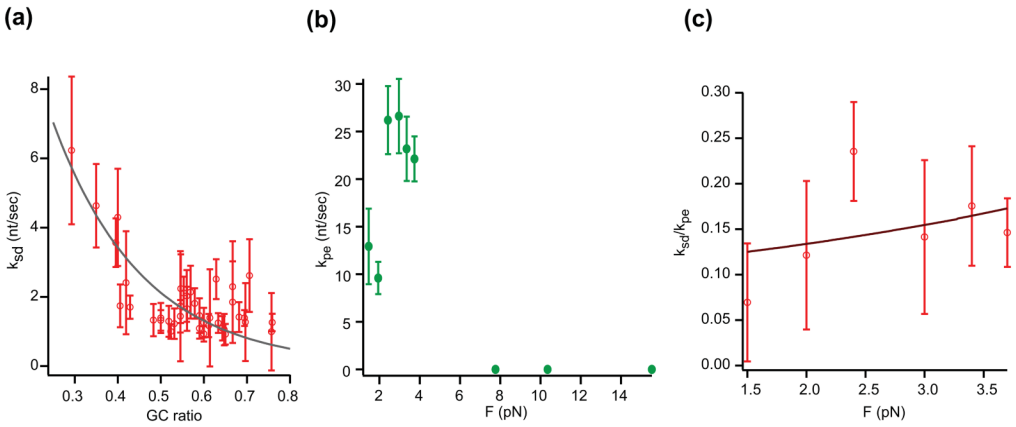

Figure 6. Sequence and stretching force dependence of k_{sd} and k_{pe}

(a) k_{sd} as a function of GC content in hairpin stems. The exponential fit based on the simple model for helicase unzipping yields n = 3 bases, ΔG_{RT} = 1.3 k_BT. (b) Force-dependent k_{pe} of HIV-1 RT. (c) Ratio between k_{sd} and k_{pe} is shown as a function of template stretching force. The best fit is done with n = 4 bases, ΔG_{RT} = 1.4 k_BT (see Supporting Materials for analysis).