Supporting Online Material for

Slide into Action: Dynamic Shuttling of HIV Reverse Transcriptase on Nucleic Acid Substrates

Shixin Liu, Elio A. Abbondanzieri, Jason W. Rausch, Stuart F. J. Le Grice, Xiaowei Zhuang*

*To whom correspondence should be addressed. E-mail: zhuang@chemistry.harvard.edu

DOI: 10.1126/science.1163108

This PDF file includes:

Materials and Methods
Figs. S1 to S12
References
Materials and Methods

**RT preparation.** Dye-labeled RT was derived from either the wildtype enzyme or an RNase H inactive mutant in which an E478Q mutation was introduced into the RNase H domain (1). To produce the H-labeled RT, the native cysteines on RT were changed to serine and a unique cysteine residue was introduced at the C-terminus of the p66 subunit to allow specific labeling of dye molecules. To create the F-labeled RT, one of the native cysteines at position 38 of the p66 subunit was retained for dye labeling. The p51 subunit of RT, which primarily serves a structural scaffolding role, was not labeled. Purified RT molecules were incubated with Cy3 maleimide (GE Healthcare) for 60 minutes in 100 mM pH 7.0 NaH₂PO₄/Na₂HPO₄ buffer. Cy3-labeled RT was then dialyzed for more than 48 hours to remove excess dye molecules.

**Nucleic acid preparation.** Synthetic DNA (Operon or IDT) and RNA (Dharmacon) oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis. When necessary, the strands were specifically derivatized with a biotin or an amino modifier during synthesis. Cy3 or Cy5 mono-reactive NHS ester (GE Healthcare) was post-synthetically conjugated to the primary amine group on DNA or RNA. Labeled oligonucleotides were HPLC purified by reverse phase chromatography on a C8 column (GE Healthcare).

To assemble a primer/template duplex of less than 60 bp in length, the two strands were annealed at a mixing ratio of 1:1, heated to 95 °C for 5 minutes, and slowly cooled to room temperature over 1 hour in annealing buffer (50 mM pH 8.0 Tris-HCl, 100 mM KCl, and 1 mM EDTA). To generate the ~550 bp duplex DNA, a short duplex (31 bp) modified with biotin and Cy5 was first annealed as above and then ligated by T4 DNA ligase (New England Biolabs) to a long duplex (~520 bp) generated by PCR. Both pieces contained an EcoRI restriction site and were digested by EcoRI (New England Biolabs) before ligation. The ligation product was gel purified.

To assemble a primer/template/non-template construct, the 3 strands were annealed with a 1:2:3 ratio in the annealing buffer using a Bio-Rad thermocycler with the following program: 1 minute at 95 °C, 10 minutes ramp to 70 °C, 30 minutes at 70 °C, 10 minutes ramp to 33 °C, 30 minutes at 33 °C, and finally cooling to and holding at 4 °C (2). The 1:2:3 mixing ratio was chosen to minimize the number of free primer molecules and primer/template complexes without the non-template strand. Annealed products were analyzed on 5% native polyacrylamide gels at
4 °C to confirm successful assembly. Removal of unannealed or partially assembled constructs was not necessary as they lacked biotin groups for surface immobilization or dye molecules for detection.

**Single-molecule FRET measurement.** Nucleic acid substrates were immobilized on PEG-coated quartz microscope slides through biotin-streptavidin linkage as previously described (3). The biotin group was placed at the 5' end of the DNA primer in all constructs unless otherwise mentioned. The surface-immobilized Cy5-labeled substrates were immersed in a solution containing Cy3-labeled RT (Fig. S2B). The fluorescence signals for the FRET donor Cy3 and acceptor Cy5 were detected using a prism-type TIRF microscope. Alternating 532 nm and 635 nm laser light were used to excite the sample. In 9 out of every 10 frames, the FRET donor Cy3 was excited by a 532 nm Nd:YAG laser (Crystal Laser), and every tenth frame, allowing us to probe FRET between the Cy3-labeled RT and Cy5-labeled substrates, a 635 nm laser (Coherent) was used to excite Cy5 directly to confirm the presence of active-Cy5 labeled substrate independent of RT binding. Emission from the donor and acceptor were separated using a dichroic mirror (Chroma Technology) and imaged onto the two halves of a back-illuminated electron multiplying CCD camera (Andor Ixon 887). Each RT binding event caused an increase in the total Cy3 and Cy5 fluorescence signals under the 532 nm excitation (Fig. S2C). The FRET value during the binding events was defined as $I_A/(I_A + I_D)$, where $I_A$ and $I_D$ were the fluorescence signals detected from the acceptor and donor channel, respectively. Single-step RT dissociation or Cy3 photobleaching indicated the binding of a single enzyme. Measurements were performed at room temperature (23 °C) unless otherwise specified. The imaging buffer contained 50 mM pH 8.0 Tris-HCl, 50 mM KCl, 6 mM MgCl₂, and 0.1 mg/ml BSA (New England Biolabs) except for experiments shown in Figure 2, where the imaging buffer contained 25 mM pH 8.0 Tris-HCl, 25 mM KCl, 3 mM MgCl₂, 30% v/v glycerol, and 0.1 mg/ml BSA. An oxygen scavenger system (10% w/v glucose, 300 μg/ml glucose oxidase, and 40 μg/ml catalase) and a reducing reagent (2 mM Trolox) was also included in the imaging buffer to minimize photobleaching and blinking (4).

During data acquisition, Cy3-labeled RT was added to surface-immobilized Cy5-labeled nucleic acid substrates. The concentration of Cy3-labeled p66 subunits used in the single-molecule experiments was 10 – 20 nM. A large excess of unlabeled wildtype p51 subunits were added such that the majority of Cy3-labeled p66 form heterodimers with p51. This strategy was
used to ensure that the concentration of the dye-labeled p66 in the solution did not overwhelm single molecule detection of RT-substrate complexes bound to the surface. As the p66 and p51 subunits alone exhibited much lower affinities to the nucleic acid substrates, the majority of the binding events observed involved a p66/p51 heterodimer. The observation of single-step RT dissociation or Cy3 photobleaching ensured the binding of a single enzyme. To monitor structural dynamics of the nucleic acid substrates, unlabeled RT was added to the Cy3 and Cy5 doubly labeled substrates.

**Kinetic analysis of RT sliding.** During the process of sliding, RT dwelt substantially longer at either end of the duplex than in the middle positions. We thus describe sliding kinetics by a simple model featuring four first-order rate constants as described in Fig. S7C: $k_{\text{front} \rightarrow \text{off}}$, dissociation rate constant from the front end of the duplex; $k_{\text{front} \rightarrow \text{back}}$, the transition rate constant from the front end to the back end; $k_{\text{back} \rightarrow \text{front}}$, the transition rate constant from the back end to the front end; $k_{\text{back} \rightarrow \text{off}}$, dissociation rate constant from the back end of the duplex. Assuming that the system was at equilibrium, the rate constants were calculated using the following formulae:

$$k_{\text{front} \rightarrow \text{off}} = \frac{1}{\tau_{\text{front}}}P_{\text{front} \rightarrow \text{off}}, \quad k_{\text{front} \rightarrow \text{back}} = \frac{1}{\tau_{\text{front}}}P_{\text{front} \rightarrow \text{back}},$$

$$k_{\text{back} \rightarrow \text{front}} = \frac{1}{\tau_{\text{back}}}P_{\text{back} \rightarrow \text{front}}, \quad k_{\text{back} \rightarrow \text{off}} = \frac{1}{\tau_{\text{back}}}P_{\text{back} \rightarrow \text{off}},$$

where $\tau_{\text{front}}$ and $\tau_{\text{back}}$ are the mean dwell times of RT at the front end and back end, respectively, which were derived from the FRET time traces. $P_{\text{front} \rightarrow \text{off}}$ is the probability of dissociation once RT is bound to the front end, $P_{\text{front} \rightarrow \text{back}}$ is the probability of sliding to the back end once RT is bound to the front end, $P_{\text{back} \rightarrow \text{front}}$ is the probability of sliding to the front end once RT is bound to the back end, $P_{\text{back} \rightarrow \text{off}}$ is the probability of dissociation once RT is bound to the back end. These probabilities were derived from FRET time traces using the following equations:

$$P_{\text{front} \rightarrow \text{off}} = \frac{\text{the number of front-end binding events that are followed by a dissociated event}}{\text{the total number of front-end binding events}},$$

$$P_{\text{front} \rightarrow \text{back}} = \frac{\text{the number of front-end binding events that are followed by a back-end binding event}}{\text{the total number of front-end binding events}},$$

$$P_{\text{back} \rightarrow \text{front}} = \frac{\text{the number of back-end binding events that are followed by a front-end binding event}}{\text{the total number of back-end binding events}},$$

$$P_{\text{back} \rightarrow \text{off}} = \frac{\text{the number of back-end binding events that are followed by a dissociated event}}{\text{the total number of back end binding events}}.$$
**Bulk primer extension assay.** The primer/template and primer/template/non-template substrates were annealed as described above. Cy5 (GE Healthcare) was conjugated to the primers to monitor DNA synthesis. Substrates (5 nM) were incubated with RT (20 nM) in RT reaction buffer (50 mM pH 8.0 Tris-HCl, 50 mM KCl, and 6 mM MgCl₂) at room temperature for 5 minutes in a final reaction volume of 100 μl. DNA synthesis was initiated by addition of the four dNTPs (each at 200 μM final concentration) and terminated by mixing 5 μl aliquot of the reaction mixture with 15 μl of stop buffer (96% formamide, 20 mM EDTA) at the time points indicated in the text. Reaction products were heated to 95 °C for 5 minutes, fractionated over a 10% denaturing polyacrylamide gel, and quantified using a Typhoon Trio Imager (GE Healthcare).
Fig. S1. Schematic of the HIV-1 reverse transcription pathway (5). RT initiates minus-strand DNA synthesis from a cellular tRNA hybridized to the primer binding site (PBS) located near the 5' terminus of viral RNA. During and following minus-strand synthesis, viral RNA is hydrolyzed via the RNase H activity of RT, leaving two specific RNA fragments, known as the polypurine tracts (PPTs), to serve as unique primers for initiation of plus-strand DNA synthesis. As plus-strand DNA synthesis proceeds, the linear replication intermediate circularizes via PBS homology, and RT performs DNA displacement synthesis to complete the genome conversion process. The final product of reverse transcription is a double-stranded proviral DNA flanked by long terminal repeat (LTR) sequences, which are subsequently recognized by the viral integrase protein for insertion into chromosomai DNA of the infected cell. TAR: trans-activation response element; cPPT: central polypurine tract; 3′PPT: 3′ polypurine tract; CTS: central termination sequence.
Fig. S2. Probing RT-substrate interaction in real time by single-molecule FRET. (A) The structure of HIV-1 RT bound to a primer/template duplex (6). The fingers, palm, thumb, connection, and RNase H domains of the p66 subunit are color-coded and the p51 subunit is in grey. The H- and F-labeling sites of the Cy3 on the enzyme are highlighted with green stars, while the front- and back-end labeling sites of Cy5 on the nucleic substrate are highlighted with red stars. (B) Single-molecule FRET assays for probing the enzyme-substrate interaction. The nucleic acid substrates were attached to a microscope slide surface, and binding/dissociation and motion of RT on the substrate were detected by FRET between Cy3 and Cy5. F and H on the enzyme mark the fingers and RNase H domains, respectively. (C) FRET analysis of individual RT binding events. (upper panel) Fluorescence time traces of a single enzyme-substrate complex. Cy5 fluorescence under 635 nm direct excitation (pink) indicates that an active Cy5-labeled substrate is present. Binding of Cy3-labeled RT leads to an increase in the total fluorescence in the Cy3 (green) and Cy5 (red) channels under 532 nm excitation. The binding event is highlighted in yellow. The FRET values during the binding event are shown in the lower panel.
Fig. S3. Surface immobilization and dye labeling did not significantly perturb the DNA polymerase activity of RT. (A) To monitor the polymerase activity of the RT with single-molecule FRET, an RNA template (orange) labeled with Cy3 (green star) was annealed to a DNA primer (black) labeled with Cy5 (red star). Addition of RT (yellow) and dNTPs led to primer extension and the conversion of the single-stranded RNA region to DNA/RNA duplex,
which should increase the distance between the Cy3 and Cy5 dyes with a corresponding decrease in FRET. (B) FRET histograms of the surface immobilized substrates before (left) and 40 min after (right) the addition of RNase H-inactive (RNase H-) RT (20 nM) and dNTPs (200 μM each). As expected, the FRET peak decreased from 0.76 to 0.24 upon primer extension. Primer extension kinetics were measured by recording the fraction of molecules converted from 0.76 FRET to 0.24 FRET as a function of time, as shown in black filled squares in (D). (C) Ensemble gel electrophoresis assay for DNA polymerase and RNase H activities of four different RT constructs: unlabeled wildtype RT, unlabeled RNase H- RT, H-labeled RNase H- RT and F-labeled RNase H- RT. To initiate synthesis, RT (20 nM) and dNTPs (200 μM each) added to the Cy3 and Cy5 doubly labeled substrates as shown in (A) and the reaction products were analyzed by gel electrophoresis. Primer extension was monitored by fluorescence of the Cy5-labeled primer (top panels) and the RNase H cleavage reaction was monitored by fluorescence of the Cy3-labeled RNA template (bottom panels). All four RT constructs showed substantial DNA polymerase activities, but only wildtype RT showed RNase H activity. Primer extension kinetics were measured by recording the fraction of fully-extended primer as a function of time and displayed in (D). (D) Primer-extension kinetics measured by single-molecule FRET and ensemble gel electrophoresis. The final extent of DNA synthesis in the gel assay is slightly lower than that obtained in the single-molecule assay, likely because the unannealed primers present in the starting material would show up in the gel, but not in single-molecule assay due to the lack of FRET donor. Fitting these curves with single exponential gave extension rate constants of 0.018 sec\(^{-1}\) for unlabeled wildtype RT (ensemble), 0.018 sec\(^{-1}\) for unlabeled RNase H- RT (ensemble), 0.019 sec\(^{-1}\) for unlabeled RNase H- RT (single-molecule), 0.015 sec\(^{-1}\) for H-labeled RNase H- RT (ensemble), and 0.026 sec\(^{-1}\) for F-labeled RNase H- RT (ensemble). The primer extension kinetics are similar in all these cases, indicating that surface immobilization and dye labeling did not perturb DNA polymerase activity substantially.
Fig. S4. Photophysical properties of the FRET dyes are not significantly altered upon RT binding. Unlabeled RT was added to the 38 bp DNA/RNA hybrid doubly labeled with Cy5 at the template 3' end spatially close to the RNase H domain of RT when the enzyme binds to the back end of the duplex, and Cy3 toward the middle of the primer spatially close to the RNase H domain of RT when it binds to the front end of the duplex. The FRET distributions were similar before (blue bars) and after (red line) the addition of 50 nM unlabeled RT, a concentration at which substrates would be in complex with RT most of the time according to a native gel shift assay. This result suggests that the photophysical properties of the dyes were not changed appreciably by RT binding.
**Fig. S5.** Different FRET labeling scheme reveal the front-end and back-end binding states of RT on a DNA/RNA hybrid. **(A)** H-labeled RT was added to the 38 bp back-end labeled DNA/RNA hybrid as shown in Fig. 1B. The FRET histogram (blue bars) were fit with two Gaussian peaks (red line). **(B)** F-labeled Cy3-RT was added to the 38 bp DNA/RNA hybrid with Cy5 attached to the front end of the duplex. The two binding positions at opposite ends of the hybrid would predict two FRET peaks: one peak at a high FRET value corresponding to RT bound to the front end of the hybrid with the Cy3 dye on the fingers domain close to the Cy5 dye at the substrate front end, and the other peak at a medium FRET value corresponding to RT bound to the back end of the hybrid with the Cy3 dye near the middle of the duplex. This prediction was indeed confirmed experimentally, as the measured FRET histogram (blue bars) showed high and medium FRET peaks centered at 0.88 and 0.52, respectively. **(B)** H-labeled RT was added to the front-end labeled 38 bp DNA/RNA hybrid. Again, enzyme binding at opposite ends of the hybrid would predict two FRET peaks: one at a medium FRET value corresponding to RT bound to the front end of the hybrid with the Cy3 dye on the RNase H domain near the middle of the duplex, and the other near zero FRET corresponding to RT bound to the back end of the hybrid with the Cy3 and Cy5 dyes separated by roughly the entire 38 bp length of the duplex. This prediction was also confirmed experimentally: The FRET histogram (blue bars) was fit with two Gaussian peaks centered at 0.31 and 0 (red line). In all three labeling schemes, the equilibrium constants between the front- and back-end binding states were measured to be ~3:1. While these data were obtained using the RNase H inactive RT mutant, wildtype RT yielded similar results. In the latter case, measurements were taken shortly after addition of the enzyme, before significant cleavage of the DNA/RNA hybrid could occur.
**Fig. S6.** Sliding dynamics of RT on DNA/RNA hybrids with different end structures. H-labeled RT was added to the back-end labeled 38 bp DNA/RNA hybrid with several distinct end structures. (A) A 5′ overhang and 3′ recess of the RNA template. A 3-nt unpaired flap is also present at the 3′ terminus of the RNA template. (B) A 5′ overhang and 3′ recess of the template without a flap. (C) A template in which both 5′ and 3′ termini are recessed. In all cases, FRET was observed to transit between 0.39 and 0.95, corresponding to RT bound to the front and back ends of the hybrid region, respectively. The ratios between the areas under the 0.39 and 0.95 peaks are 3.3:1, 3.5:1, and 1.4:1, respectively. These results suggested that RT could slide on DNA/RNA hybrids with different end structures, and the relative stability of the two end binding states differed moderately depending on the end structures. The relatively higher stability of the front-end binding state in (A) and (B) suggest that the enzyme favors a recessed 3′ terminus of the primer.
**Fig. S7.** Effects of duplex structures and small molecule ligands on sliding kinetics. (A) Sliding of RT on a 38 bp DNA/RNA hybrid. (Left) H-labeled RT bound to back-end labeled DNA/RNA hybrid. (Middle) FRET histogram. (Right) A representative FRET time trace showing transitions between the 0.39 and 0.95 states. (B) Sliding of RT on a 38 bp DNA duplex of the same sequence with the same labeling strategy. Note that the low FRET peak obtained with duplex DNA (FRET = 0.29) was lower than that of the DNA/RNA hybrid (FRET = 0.39), due to the larger inter-base distance in a duplex DNA (7, 8). The representative FRET time trace showing transitions between the 0.29 and 0.95 states on duplex DNA. (C) Sliding kinetics of RT on the DNA/RNA hybrid, on the DNA/RNA hybrid in the presence of 1 mM dGTP or 100 µM nevirapine, or on the DNA duplex. (Left) A simple kinetic model of sliding and dissociation featuring four first-order rate constants. $k_{\text{back} \rightarrow \text{off}}$: dissociation rate constant from the back end of the duplex. $k_{\text{back} \rightarrow \text{front}}$: transition rate constant from the back to the front end. $k_{\text{front} \rightarrow \text{back}}$: transition rate constant from the front to the back end. $k_{\text{front} \rightarrow \text{off}}$: dissociation rate constant from the front end. (Right) Kinetic rate constants derived from the FRET time traces. Nucleotide binding reduced $k_{\text{front} \rightarrow \text{back}}$ substantially without significantly affecting $k_{\text{back} \rightarrow \text{front}}$, while nevirapine substantially increased $k_{\text{front} \rightarrow \text{back}}$ without affecting $k_{\text{back} \rightarrow \text{front}}$. Compared to the DNA/RNA hybrid case, the rate constants for escaping the back end ($k_{\text{back} \rightarrow \text{front}}$ and $k_{\text{back} \rightarrow \text{off}}$) were much higher for duplex DNA. Detailed analysis methods were described in Materials and Methods.
**Fig. S8.** RT shuttles between the two ends of the DNA/RNA hybrid with preferred intermediate states. (A) Additional FRET time traces showing RT sliding on a 56 bp DNA/RNA hybrid at 12 °C. Another trace is shown in Fig. 1D. These traces reveal gradual transitions between the 0 FRET (front-end binding) and 0.95 FRET (back-end binding) states and preferred intermediates near FRET ~ 0.3 – 0.5. (B) FRET histogram constructed from multiple transition regions. The histogram also includes 25 frames before and after each transition region to show the two end-bound states (FRET = 0 or 0.95). A predominant intermediate peak at FRET ~ 0.3 – 0.5 was observed. (C) Histograms for the transition time between the 0 and 0.95 FRET states at 23 °C (red bars) and 12 °C (green bars). The mean transition time was 0.88 sec at 23 °C, and 2.6 sec at 12 °C.
**Fig. S9:** The histogram of the time delay $\tau$ between binding of RT and placing of RT at the polymerization site as observed in Figure 2.
**Fig. S10.** The FRET distributions of the RNA strand displacement substrates p\(X/T/NT\), each consisting of an RNA template (T) to which a complementary DNA primer (p) and RNA non-template strand (NT) were simultaneously hybridized, with the Cy3 and Cy5 dyes flanking the T/NT duplex region. Overlapping regions between the DNA primer and the RNA non-template strand are colored in light grey and orange, respectively. We use the notation p\(X/T/NT\) to represent a substrate whose primer has been extended by \(X\) nucleotides. Here \(X = 0, 4,\) and 10. The FRET measurements were obtained in the absence of RT.
**Fig. S11.** Structural dynamics of RNA displacement substrates in the presence of RT. RT and dNTPs were added to the p4/T/NT substrate doubly labeled with Cy3 and Cy5. The primer strand was chain-terminated to prevent extension. DNA and RNA strands are colored in black and orange, respectively. Overlapping regions between the DNA primer and the RNA non-template strand are colored in light grey and orange, respectively. The FRET histogram (blue bars) shows two peaks at 0.3 and 0.6 and FRET time traces (not shown) reveal transitions between the two states. The histogram agrees quantitatively with the one obtained for the p0/T/NT substrate once dCTP, dGTP, dATP, and ddTTP were added to allow 4-nucleotide extension of the p0 primer (red line). The latter histogram is also shown in Fig. 3B.
Fig. S12. Comparison of normal and strand displacement synthesis. (A) Primer extension assays for RNA-directed DNA synthesis (lanes 1-4), RNA strand displacement synthesis (lanes 5-9), DNA-directed DNA synthesis (lanes 10-13), and DNA strand displacement synthesis (lanes 14-18). The extended primer was monitored by Cy5 fluorescence. As shown in the gel, both RNA-directed (lanes 1-4) and DNA-directed (lanes 10-13) DNA synthesis are highly efficient with rapid accumulation of full length product. In the case of strand displacement synthesis, RT supported limited RNA strand displacement synthesis, terminating predominantly after extension of the primer by 5 nt (lanes 5-9). The exact termination site was sequence dependent (data not shown). In contrast, the primer was fully extended during DNA strand displacement synthesis (lanes 14-18). (B) Primer extension kinetics of DNA displacement synthesis measured by ensemble gel electrophoresis (as shown in A) and single-molecule FRET (Fig. 5C). The fraction of primers fully extended by RT was plotted as a function of time. Similar primer extension kinetics were observed with the two methods.
REFERENCES