Published online 20 January 2017

Nucleic Acids Research, 2017, Vol. 45, No. 6 ***3341–3352***
doi: 10.1093/nar/gkx004

Coordination between the polymerase and RNase H activity of HIV-1 reverse transcriptase

Małgorzata Figiel[1], Miroslav Krepl[2,3], Jarosław Poznański[4], Agnieszka Gołąb[1], Jiří Šponer[2,3] and Marcin Nowotny[1,*]

[1]Laboratory of Protein Structure, International Institute of Molecular and Cell Biology, 02-109 Warsaw, Poland, [2]Institute of Biophysics, Academy of Sciences of the Czech Republic, 612 65 Brno, Czech Republic, [3]Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacky University Olomouc, 77146 Olomouc, Czech Republic and [4]Institute of Biochemistry and Biophysics Polish Academy of Sciences, 02-106 Warsaw, Poland

Received December 02, 2016; Editorial Decision December 21, 2016; Accepted January 03, 2017

ABSTRACT

Replication of human immunodeficiency virus 1 (HIV-1) involves conversion of its single-stranded RNA genome to double-stranded DNA, which is integrated into the genome of the host. This conversion is catalyzed by reverse transcriptase (RT), which possesses DNA polymerase and RNase H domains. The available crystal structures suggest that at any given time the RNA/DNA substrate interacts with only one active site of the two domains of HIV-1 RT. Unknown is whether a simultaneous interaction of the substrate with polymerase and RNase H active sites is possible. Therefore, the mechanism of the coordination of the two activities is not fully understood. We performed molecular dynamics simulations to obtain a conformation of the complex in which the unwound RNA/DNA substrate simultaneously interacts with the polymerase and RNase H active sites. When the RNA/DNA hybrid was immobilized at the polymerase active site, RNase H cleavage occurred, experimentally verifying that the substrate can simultaneously interact with both active sites. These findings demonstrate the existence of a transient conformation of the HIV-1 RT substrate complex, which is important for modulating and coordinating the enzymatic activities of HIV-1 RT.

INTRODUCTION

The proliferation of human immunodeficiency virus 1 (HIV-1) strictly relies on the conversion of its single-stranded RNA to double-stranded DNA that can be integrated into the host cell genome (1–4). This process is called reverse transcription and involves RNA-dependent synthesis of the (-)DNA strand with concomitant degradation of the template RNA, followed by the DNA-dependent synthesis of (+)DNA. Both enzymatic activities that are required to complete this complex reaction derive from a single protein: HIV-1 reverse transcriptase (HIV-1 RT). DNA strands are synthesized by the polymerase domain of HIV-1 RT. The RNase H domain degrades the RNA strand of the RNA/DNA intermediate but also enables appropriate generation of the RNA primers, termed polypurine tracts (PPTs), which are necessary for the initiation of DNA synthesis.

Structural and biochemical studies of HIV-1 RT indicate spatial and temporal coordination of the two activities (5). The position of RNase H cleavage is determined by interactions between the nucleic acid and polymerase domain. For example, in DNA 3′-end-directed cleavage, the RNA strand of the RNA/DNA substrate is cleaved ∼18 nucleotides from the recessed 3′-end of the DNA that is bound by the polymerase domain (6,7). This cleavage is known to occur concomitantly with DNA polymerization, albeit less frequently than nucleotide incorporation, with approximately one RNase H cut for each 7 nucleotides that are added to the growing DNA chain (8,9). During this polymerization-dependent mode, RNase H cleavage occurs more efficiently when polymerization pauses (10–12). This was considered evidence that the substrate needs to be released from the polymerase active site for the RNase H cuts to occur. After polymerization is completed, longer RNA fragments that remain associated with DNA are removed by polymerization-independent RNA 5′-end-directed and internal cleavage (13–15). Moreover, RNase H activity strictly depends on the structural context of the whole enzyme, and the isolated RNase H domain exhibits minimal activity because it cannot bind the substrate on its own (16,17).

HIV-1 RT is a heterodimer that consists of two subunits. The larger subunit, termed p66, possesses polymerase and RNase H activities. The smaller subunit, termed p51, is a shorter version of p66 that lacks most of the RNase H do-

*To whom correspondence should be addressed. Tel: +48 22 597 07 17; Fax: +48 22 597 07 15; Email: mnowotny@iimcb.gov.pl

main sequence as a result of processing by HIV-1 protease. It has an altered conformation and plays a structural role in the heterodimer (4). In the available crystal structures of HIV-1 RT two conformations of the nucleic acid substrate are observed (18–20) (Figure 1). In polymerase mode, the 3′-end of the primer is bound at the polymerase active site and aligned for extension with the incoming nucleotide. However, the scissile phosphate of the RNA strand of the hybrid faces away from the active site of RNase H and is located too far from it for hydrolysis to occur (e.g. PDB ID: 4PQU) (20). RNase H mode is observed in structures that are solved in the presence of non-nucleoside RT inhibitors and with geometry-altering modifications of the substrate (19,20). In these structures, the unwound substrate does not interact with the polymerase active site, and the configuration of the RNA strand at the RNase H active site is nearly catalytic (e.g. PDB ID: 4B3O) (19). These structural data indicate that conformational changes in the substrate likely serve to restrict access of the substrate to the RNase H active site, thereby restricting the amount of RNase H activity which is crucial for the proper execution of reverse transcription (19). Although interactions with the polymerase and RNase H active sites may be mutually exclusive events in the ground state, the complex may be able to reach a high-energy conformation that is conducive to simultaneous catalytic interactions with both sites (i.e. simultaneous mode) as a minor populated state during genuine thermal fluctuations (Figure 1).

To investigate the mechanism of coordination between polymerase and RNase H activity, we used molecular dynamics (MD) simulations to determine whether RNA/DNA can simultaneously interact with both active sites of HIV-1 RT. We found that this can occur and that the corresponding substrate conformation involves mostly untwisting of the hybrid in the region of the substrate that interacts with the RNase H domain. To demonstrate the simultaneous binding of RNA/DNA at the two active sites experimentally, we used chemical cross-linking methodology to immobilize the substrate at the polymerase active site. We found that the substrate within cross-linked complexes effectively interacted with the RNase H active site. These findings reveal a mechanism of modulation and coordination of the enzymatic activity of HIV-1 RT.

MATERIALS AND METHODS

Molecular dynamics simulations

All of the simulations were performed with a ff99bsc0χ_{OL3} force field for RNA (21–23), ff99bsc0$\epsilon\zeta_{OL1}$ force field for DNA (21–22,24), and ff12SB force field for protein (21,25–26). The starting structure for the simulation was obtained by a series of short simulations of the 4PQU structure, in which we used molecular modeling to extend the substrate by 4 base pairs (bp), add a missing protein helix of the RNase H domain (amino acids [aa] 546–554; taken from the 1RTD structure), and revert the crystallographic mutation D498N. Utilizing this structure, we then used molecular modeling to alter the sequence of the nucleic acid substrate to match the CL3 hybrid. We also introduced two Mg^{2+} ions at the active site of the HIV-1 RNase H domain, based on the structure with the inhibitor β-thujaplicinol and Mn^{2+} ions (PDB ID 4QAG) (27). The standard simulation protocol for protein/RNA complexes was followed (28). After each modeling intervention in the structure, a standard minimization and equilibration protocol was used to stabilize the new structure before resuming the production simulation.

In the simulation, we imposed a series of flat-well restraints that are necessary to support a long-term catalytic interaction of the substrate with the RNase H domain. The restraints were based on the assumption that interactions between the substrate and the RNase H domain should be very similar to those observed in the structures of substrate complexes of cellular RNases H1 (29,30). The first restraint was the maximum distance of 7.2 Å between the C-α of Asp443 at the heart of the active site and the phosphorus atom of the hydrolyzed phosphate of the RNA (between nt +1 and −1). A pair of restraints was then imposed to recapitulate a critical interaction between the backbone of the DNA strand and an RNase H element, termed the phosphate-binding pocket (29,30). A non-bridging oxygen of the phosphate group that was 2 bp from the scissile phosphate (nucleotide −3) was restrained at maximum distances of 5.8 Å from C-α of Thr473 and 5.3 Å from C-α of Lys476. For the restraints, we used a function that imposes no energy penalty when the interatomic distance is shorter than the specified maximum distances. The energy penalty linearly increases for distances greater than the specified maximum distances. We emphasize that these restraints were fully justified and did not affect any of our conclusions because the goal of the simulations was to assess the compatibility of the substrate with this target geometry. The restraints thus helped obtain structural conditions that were equivalent to our cross-linking experiments within an affordable computational time and in the absence of a suitable experimental starting geometry. Currently, there are no experimental structures that capture simultaneous interactions of the substrate with the polymerase and RNase H domains of HIV-1 RT that could serve as suitable starting structures for MD simulations. The spontaneous, unaided formation of a catalytic site and binding of Mg^{2+} were beyond the timescale of conventional simulations. We also assumed that the appropriate catalytically active arrangement was not the dominantly populated 'ground-state' conformation of the complex.

Protein expression and purification

Genes that encode the p66 and p51 subunits of HIV-1 (isolate HXB2) RT were cloned into pET15 and pET28, respectively. The p51 subunit contained an N-terminal His-tag. To avoid non-specific cross-linking, Cys282 in both subunits was substituted with serine. Mutagenesis of the constructs was performed according to Stratagene's QuikChange protocol.

The two subunits were expressed in *Escherichia coli* BL21 Rosetta cells that were induced with 0.4 mM 1-thio-β-D-galactopyranoside at 16°C. The cells were then harvested and suspended in 40 mM NaH_2PO_4 (pH 7.0), 100 mM NaCl, 5% glycerol, 10 mM imidazole and 1.4 mM β-mercaptoethanol and incubated on ice in the presence of 1 mg/ml lysozyme and protease inhibitor cocktail. Follow-

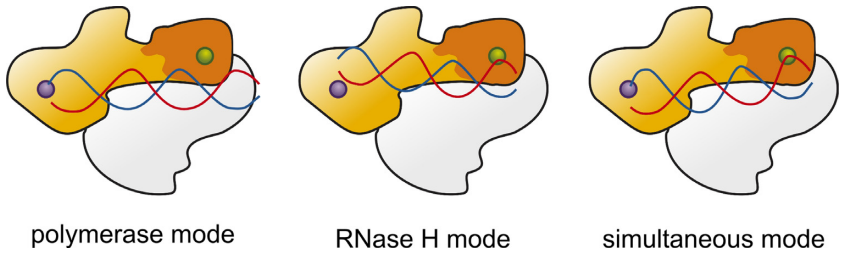

Figure 1. Three modes of HIV-1 interaction with an RNA/DNA substrate. The p66 subunit is shown in orange, with the RNase H domain indicated with a darker shade. The polymerase and RNase H active sites are represented by purple and green circles, respectively. DNA and RNA strands of the hybrid substrate are shown in blue and red, respectively.

ing sonication, the cleared lysate was loaded onto a HisTrap column (GE Healthcare) that was equilibrated with 40 mM NaH_2PO_4 (pH 7.0), 500 mM NaCl, 5% glycerol, 10 mM imidazole and 1.4 mM β-mercaptoethanol. After a wash step with 60 mM imidazole, the protein was eluted with 300 mM imidazole. The eluted fraction was dialyzed overnight against 40 mM NaH_2PO_4 (pH 7.0), 50 mM NaCl, 5% glycerol, 0.5 mM ethylenediaminetetraacetic acid (EDTA) and 0.5 mM dithiothreitol (DTT) and loaded on a Heparin column (GE Healthcare) that was equilibrated with the same buffer. Protein was eluted with a linear gradient of 50–500 mM NaCl. Selected fractions were pooled and applied to a Superdex 200 column (GE Healthcare) that was equilibrated with 20 mM Tris (pH 8.0), 100 mM NaCl, 5% glycerol, 0.5 mM EDTA and 0.5 mM DTT. Selected fractions were pooled and concentrated.

Chemically modified DNA oligonucleotides

DNA oligonucleotides that carried the 2-F-dI residue at the desired position of the through-base modification were ordered from Metabion (Martinsried, Germany). The oligonucleotides were reacted with cystamine ($NH_2CH_2CH_2S$-$SCH_2CH_2NH_2$) whose amine group substituted the fluorine atom in the base, producing thiol-modified DNA. Synthesis columns that contained the oligonucleotides were first deprotected with 3% dichloroacetic acid and washed with dichloromethane. The columns were incubated with cystamine dihydrochloride solution in trimethylamine/water (3:5) for 18 h with occasional agitation. The solution was collected from the column and dried using a speed vacuum, and the pellet was suspended in 1 ml of ammonia/40% methylamine (1:1; AMA). The columns were incubated with 1 ml of AMA solution for 2 h with occasional agitation. The washing solution was collected. Both fractions were incubated at 60°C for 20 h and dried using a speed vacuum. The pellets were resuspended in water, pooled and loaded on a DNAPac PA100 column (Dionex) that was equilibrated with 4 M urea, 20 mM MES (pH 6.5), 1 mM $NaClO_4$ and 0.2% acetonitrile at 65°C. Elution was performed with a linear gradient of 1–400 mM $NaClO_4$. DNA from the selected fractions was retrieved by ethanol precipitation and further purified on 12% denaturing TBE-urea polyacrylamide gel. Purified oligonucleotides were stored in ultrapure water at −20°C.

DNA oligonucleotide that was modified with a thiol group on a two-carbon linker that was attached to the backbone phosphate was purchased from FutureSynthesis (Poznan, Poland). Depending on the target HIV-1 RT variant for cross-linking, the following variants of substrates with modifications on a particular base (indicated by underscore) or on a backbone phosphate (marked with an asterisk) were used:

Cl3Db (GTCAACATCGTGATGGTCGAGTG*T),
Cl3D.1 (GTCAACATCGTGATGGTCGAGTTG),
Cl3D.2 (GTCAACATCGTGATGGTCGAGTGT),
Cl3D (GTCAACATCGTGATGGTCGAGTGT).

Cross-linking of HIV-1 RT with RNA/DNA hybrids

Fluorescently labeled RNA oligonucleotides were ordered from FutureSynthesis (Poznan, Poland) and dissolved in ultrapure water with 0.5 mM EDTA. Complementary RNA and DNA oligonucleotides were hybridized in the presence of 100 mM NaCl. Cross-linking reactions that contained 4 μM HIV-1 RT, 5 μM substrate, 50 mM Tris (pH 7.4), 45 mM NaCl, 25 mM KCl and 30% glycerol were incubated at 37°C for 2 h, followed by 14 h at 24°C. For the cross-linking specificity test, the concentration of NaCl was increased to 125 mM, and 5 mM DTT was included in the reaction. Cross-linked samples were analyzed on NuPage pre-cast Bis-Tris gradient gels (ThermoFischer Scientific).

Purification of cross-linked complexes

The cross-linking reaction mixture was diluted 5-fold with 500 mM NaCl, 50 mM $NaPO_4$ (pH 7.8) and 10 mM imidazole and loaded on a His-Trap column. The column was washed with 50 mM NaCl, 50 mM $NaPO_4$ (pH 7.8) and 10 mM imidazole. The complex was then eluted with 250 mM imidazole. The collected fraction was directly applied to a heparin column that was equilibrated with 20 mM Tris (pH 8.0), 50 mM NaCl and 0.1 mM EDTA. The purified complex was collected in the flow-through. Excess free protein was eluted with 1 M NaCl. The complex was concentrated and stored in 10 mM MES (pH 6.5), 200 mM NaCl, 0.05 mM EDTA and 5 mM iodoacetamide.

RNase H activity assay for cross-linked complexes

Samples that contained 60 nM complexes were preincubated at 37°C for 30 min in standard solution (20 mM Tris [pH 8.0] and 100 mM NaCl), with the NaCl concentration increased to 500 mM or with the addition of 3 mg/ml heparin. The cleavage reaction was performed at 37°C for 0.5,

5 or 20 min in the presence of 8 mM $MgCl_2$ and stopped by the addition of 40 mM EDTA. Similar reactions were performed for 20 min on samples that were preincubated for 1 h at 37°C in the presence of 20 mM DTT. The DTT concentration was kept constant throughout the experiment. Hydrolysis products were analyzed on 20% denaturing TBE-urea polyacrylamide gels and visualized by fluorescence readout on a Typhoon Trio+ (GE Healthcare). To analyze the effect of incoming nucleotides on RNase H cleavage, the reactions were performed in the presence of heparin and additionally contained a 100-fold molar excess of 2′-deoxycytidine-5′-([α,β]-imido)triphosphate (dCpNHpp; Jena Biosciences, Jena, Germany).

Quenched flow experiments

Subsecond time-scale reactions were prepared using an SFM-400 apparatus (BioLogic, Seyssinet-Pariset, France) in quenched-flow mode. Solutions of 60 nM complex and 16 mM $MgCl_2$, both in 20 mM Tris (pH 8.0), 100 mM NaCl and 3 mg/ml heparin, were preincubated at 37°C for 30 min. The reactions were initiated by mixing the complex and magnesium solution at 37°C and stopped by the addition of 120 mM EDTA in a 1:1:1 volume ratio. Triplicate reactions at time-points of 15, 20, 30, 40, 60, 80, 120, 160, 240 and 320 ms were collected. Hydrolysis products were analyzed on 20% denaturing TBE-urea polyacrylamide gels and visualized by fluorescence readout. Cleavage efficiency was estimated based on the densitometric quantification of reaction products. To analyze the effect of incoming nucleotide on RNase H cleavage, the reactions additionally contained a 100-fold molar excess of dCpNHpp (Jena Biosciences).

Measurements of inhibition of complex formation by salt and heparin

The assays were conducted in 96-well, black, flat-bottom polystyrene NBS plates (Corning 3650) in a total reaction volume of 100 μl. The reaction buffer contained 100 mM NaCl, 50 mM Tris (pH 8.0) and 5 mM $CaCl_2$. For the measurement of binding inhibition by salt, 10 μl of HIV-1 RT in reaction buffer was added to the plate wells to obtain final concentrations of 1, 4, 7, 10, 25, 50, 120 and 250 nM, and the concentration of NaCl in the buffer was increased to 500 mM. For the measurement of binding inhibition by heparin, the final concentration of HIV-1 RT was fixed at 120 nM, but the reaction contained heparin at a final concentration of 0–5 mg/ml. Afterward, 10 μl of fluorescein-labeled RNA/DNA hybrid substrates in H_2O was added to the reactions to a final concentration of 5 nM. The reactions were prepared in triplicate. The reaction mixtures were mixed by shaking for 5 s and incubated for 30 min at 25°C. Immediately after incubation, fluorescence anisotropy was measured in a Tecan Infinite M1000 fluorescence microplate reader at an excitation wavelength of 470 nm and emission wavelength of 520 nm with a bandwidth of 5 nm. Binding curves were fit in Prism 5 software (GraphPad, San Diego, CA, USA) using the equation for one-site-specific binding.

Measurements of the affinity between cross-linked complexes of HIV-1 RT and incoming nucleotides

The assay was conducted in 96-well, black, flat-bottom polystyrene NBS plates (Corning 3650) in a total reaction volume of 100 μl. First, 10 μl of cross-linked complexes in reaction buffer (100 mM NaCl, 50 mM Tris [pH 8.0], 5 mM $CaCl_2$ and 3 mg/ml heparin) was added to the plate wells to obtain final concentrations of 1, 2.5, 5, 10, 25, 50 and 100 nM for saturation curves, followed by the addition of 10 μl of Texas Red-labeled dCTP or dATP in H_2O to a final concentration of 5 nM. The reactions were prepared in triplicate. The reaction mixtures were mixed by shaking for 5 s and incubated for 30 min at 25°C. Immediately after incubation, fluorescence anisotropy was measured in a Tecan Infinite M1000 fluorescence microplate reader at an excitation wavelength of 590 nm and emission wavelength of 612 nm with a bandwidth of 5 nm. Binding curves were fit in Origin software (OriginLab, Northampton, MA, USA) using the equation for one-site-specific binding, and the K_d was estimated.

The binding affinity between cross-linked complexes and dCpNHpp was measured by thermophoresis. The reaction buffer contained 100 mM NaCl, 50 mM Tris (pH 8.0), 5 mM $CaCl_2$ and 3 mg/ml heparin. A series of 1:1 dilutions of dCpNHpp were prepared starting from 5 μM. The dCpNHpp dilutions were then mixed in a 1:1 ratio with 50 nM solutions of the cross-linked complex. The samples were incubated for 20 min at room temperature and loaded into capillaries. The measurements were performed using Monolith-NT.115 reader (NanoTemper Technologies, München, Germany), based on the fluorescent signal from the Cy5 label on the 5′ end of the RNA strand of the cross-linked substrate. The data were analyzed using NanoTemper Analysis software, and the K_d was calculated.

RESULTS

Molecular dynamics simulations

We used MD simulations to test whether the hybrid substrate can adopt a conformation that allows for the simultaneous interaction with both active sites of HIV-1 RT and to determine the corresponding conformation of the nucleic acid. For our MD simulations, we used the 4PQU structure (20), in which the RNA/DNA substrate interacts only with the polymerase active site and is thus in the polymerase mode. We added a missing protein helix of the RNase H domain (aa 546–554; taken from the 1RTD structure) to the structure and reverted the D498N substitution. Finally, we extended the RNA/DNA substrate to 26 bp. Using this model, we performed a 100 ns MD simulation, in which we imposed a series of distance restraints to promote the catalytic interaction between the substrate and the HIV-1 RT RNase H domain. The restraints were based on the structures of substrate complexes of cellular RNases H1 (29,30) and included catalytic interactions between the scissile phosphate of the RNA and the active site and a crucial interaction between a phosphate-binding pocket element and the DNA phosphate group of nt −3. In all of our analyses, we used nucleotide numbering relative to the preferred RNase H cleavage site 18 nt from the 3′-end of the

DNA primer. The details of the restraints are given in the Materials and Methods. We emphasize that the use of restraints was not meant to polish the simulations but rather to obtain structural conditions that correspond to the catalytic interaction of the hybrid with the RNase H domain within an affordable computational time and in the absence of a suitable experimentally determined geometry.

The simulations showed that the substrate could be neatly brought to the catalytic interaction with the RNase H active site, equivalent to the one that was conserved for cellular RNases H1 (Figure 2A–C, Supplementary Movie 1). Importantly, although no restraints were used to preserve the contact between the substrate and the polymerase domain, the substrate was stably bound at the polymerase active site throughout the simulation, including preservation of the incoming nucleotide and metal ion binding. This implies that the interaction with the RNase H active site did not perturb the binding of the primer end at the polymerase active site. Similar behavior was observed for simulations with RNA/DNA hybrids of various sequences (M. Figiel *et al*., in preparation). To test our simulation results on a longer timescale, we extended the simulation length from 100 ns to 1 μs. We also conducted four additional simulations, each 300 ns long, to verify convergence. The extended simulation time fully confirmed the initial result.

To characterize the substrate conformation change that was required to bring the nucleic acid in the catalytic interaction with the RNase H domain, we compared the values of the parameters that described the helical geometry of the RNA/DNA hybrid in the polymerase mode structure (4PQU) with average values of the same parameters throughout the MD simulation. In the simulation, we did not observe any significant global substrate bending. Overall, the DNA strand underwent much larger conformational changes than RNA, in agreement with higher rigidity of the RNA. All of the major conformational changes in the hybrid occurred in the vicinity of the RNase H domain, and the substrate at the polymerase domain remained virtually unchanged (Figure 2). The main conformational change in the nucleic acid that was observed was untwisting of the double helix. In the crystal structure of the polymerase mode (4PQU), the phosphate of nt −4 was located in the vicinity of the phosphate-binding pocket. Bringing the phosphate of nt −3 into the pocket required untwisting the substrate, which occurred in the segment of the hybrid that comprised nucleotides −9 to −1 (Figure 2D). This section of the hybrid was unwound by a total of 24°. Another change that we observed during the MD simulation was narrowing of the minor groove width, particularly for nucleotides −3 and −2 (Figure 2E). The requirement of minor groove narrowing was in agreement with the results of previously published biochemical experiments that used substrates with doublet insertions of locked nucleic acid in the DNA strand (31). Locked nucleic acids lock the substrate in the A-form conformation with a wide minor groove, and their introduction in the −4 to −1 region blocked RNase H cleavage.

In the course of the simulation, we also observed conformational changes in the protein, including movements of the thumb and finger subdomains of the polymerase domain, which are known to be mobile (19), and movement of the RNase H domain toward the RNA backbone (Figure 2C). In fact, it was this movement that positioned the RNA scissile phosphate at the active site of the RNase H domain rather than changes in the position of the RNA strand.

We compared the conformational changes that were observed during the course of our simulation with the differences that were observed between crystal structures of HIV-1 RT substrate complexes in polymerase and RNase H modes. The global changes in the RNA/DNA configuration in our MD simulations and between the crystal structures were different, which was expected because of the different behavior of the hybrid at the polymerase active site. However, a common element was the untwisting of the hybrid that was observed both between the pairs of crystal structures and in our simulation. Therefore, we conclude that hybrid untwisting is a critical element of the RNA/DNA conformational change that allows productive interactions with the RNase H active site without perturbing hybrid binding at the polymerase domain. Most importantly, removal of the substrate from the polymerase active site is not required for catalytic interaction of the nucleic acid with the RNase H domain.

Preparation of cross-linked complexes of HIV-1 RT and RNA/DNA hybrid substrates

We next experimentally tested whether the substrate can simultaneously interact with both active sites of HIV-1 RT. We immobilized the substrate at the polymerase active site of HIV-1 RT using chemical cross-linking and evaluated whether RNase H cleavage can still occur within such a covalently tethered complex. We chose a cross-linking method that relies on the formation of a disulfide linkage between a cysteine residue that is introduced to the protein by site-directed mutagenesis and a thiol group that is attached to either the backbone or base of the DNA (32,33) (Supplementary Figure S1). For cysteine substitution, we selected two residues at and in the vicinity of the polymerase active site that are sufficiently close to the RNA/DNA substrate in the polymerase mode (structure 4PQU) for chemical cross-linking to occur, but in the RNase H mode (structure 4B3O) are too far from the substrate for disulfide formation (Figure 3, Supplementary Figure S2). The first residue that was selected for through-backbone cross-linking was Asp186. In the 4PQU structure, the sidechain of Asp186 is located close to the DNA phosphate between the last and penultimate nucleotides of the primer. The second residue, Met184, which is located in the vicinity of the last and penultimate bases of the primer in polymerase mode, was selected for the through-base reaction (described in detail in the Supplementary Data). Both Asp186 and Met184 are part of the highly conserved 183-YMDD-186 polymerase active site motif. As a reference residue, we selected Gln258, which is located in the thumb subdomain of the polymerase domain and close to the base of the 6th nucleotide of the primer (from the 3′-end) in both the polymerase and RNase H modes (Figure 3, Supplementary Figure S2E and F). In fact, HIV-1 RT substrate complexes that contain this cross-link have been crystallized in both modes (18,20). Therefore, we used Q258C cross-linked complexes as a reference in which both modes are allowed. As a control for the speci-

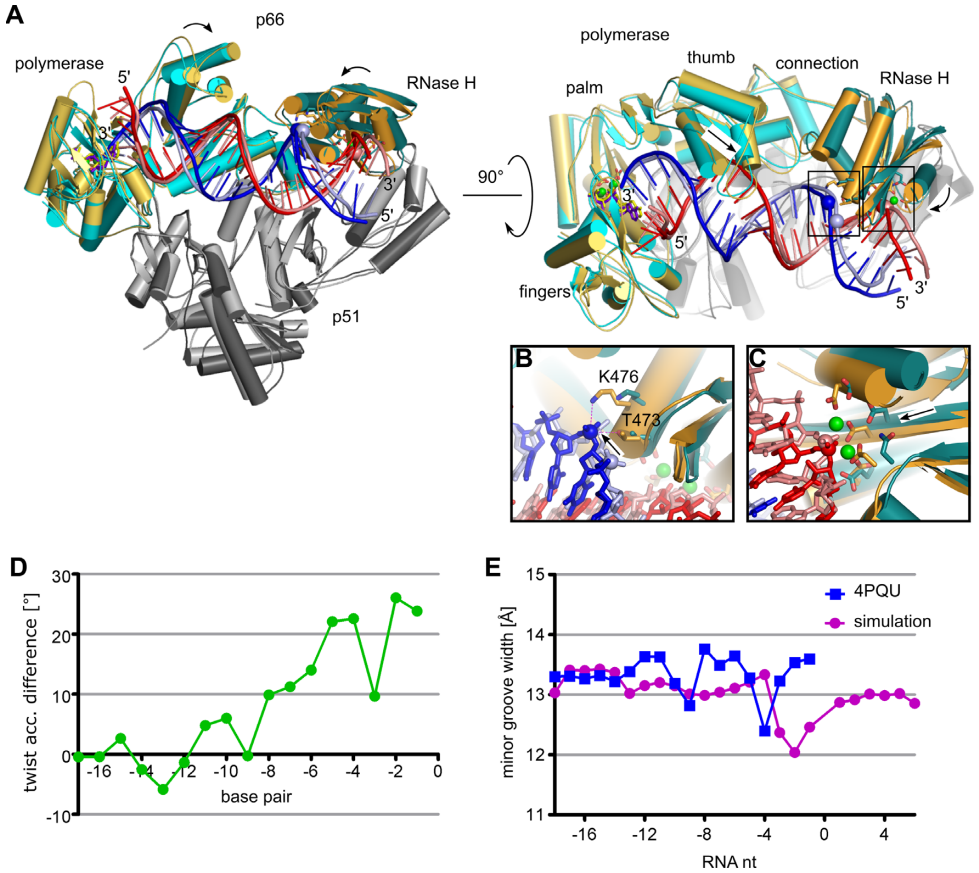

Figure 2. Molecular dynamics (MD) simulation of HIV-1 RT–hybrid substrate complex. (**A**) Superposition of starting model (p66 in cyan; p51 in light gray; RNA/DNA in light shades of red and blue, respectively) and final model in MD simulations (p66 in orange; p51 in darker gray; RNA/DNA in darker shades of red and blue, respectively). The starting model was based on the crystal structure of HIV-1 RT bound to RNA/DNA substrate in polymerase mode (PDB ID: 4PQU), modified by extending the RNA/DNA substrate by 4 bp. Residues that form the active sites are shown as sticks. Scissile phosphates and phosphates bound in the phosphate binding pocket are shown as spheres. Incoming nucleotides are shown in purple and yellow for the 4PQU structure and the MD model, respectively. Magnesium ions are shown as dark green spheres for the starting MD model and green spheres for the final MD model. Movements of the thumb and RNase H domains in the MD simulation are indicated with arrows. (**B**) Close-up of the phosphate-binding pockets of the starting and final MD models. Residues that form the phosphate-binding pocket are shown as sticks and labeled. The positions of the phosphate group of nt −3 in the beginning of the simulation and at its end are indicated by spheres. The direction of movement of the DNA strand is indicated with an arrow. (**C**) Close-up of the RNase H active sites of the starting and final MD models. Active site residues are shown as sticks. Scissile phosphates are shown as spheres. The direction of movement of the RNase H domain is indicated with an arrow. (**D** and **E**) Analysis of conformational changes in the RNA/DNA substrate in the course of the MD simulation. (**D**) Plot of helical twist difference. Helical twist values were summed along the helix for the 4PQU structure and for average values during MD simulation. The difference of these sums between the MD model and 4PQU structure is plotted for each residue (twist accumulation difference). Base pairs are numbered according to their position relative to the scissile phosphate. (**E**) Minor groove width of the RNA/DNA substrate in the 4PQU structure (blue squares) and average minor groove width during the MD simulation (purple circles).

ficity of our approach, we used Tyr183. Although located immediately next to Met184, according to the distance analysis, Y183C should not form a through-base cross-link with a hybrid in a non-strained conformation (Supplementary Table S1).

We also considered an inverse approach, in which the complex is cross-linked at the RNase H domain and polymerase activity is monitored. However, inspection of the structural data indicated that although the cross-linking strategy could be designed, it would not discriminate between the polymerase and RNase H modes. In such a case, the cross-link would form both when the substrate was engaged at the active site of RNase H and when the RNA was located too far from it for catalysis to occur, as observed in the 4PQU structure (polymerase mode). Therefore, we followed only the approach of immobilizing the substrate at the polymerase active site.

We prepared a set of RNA/DNA hybrids with a random sequence and a double-stranded region of 24 bp. Schemes of the substrates are shown in Figure 3. The RNA strand of the hybrids contained Cy5 at its 5′-end and fluorescein at the 3′-end to visualize the cleavage products. HIV-1 RT is known to preferentially position the recessed 3′-end of the DNA at the polymerase active site (3′-end-directed binding). Therefore, to better define the register of substrate binding, we introduced 6-nt 5′ overhangs to the RNA strand of the hybrids to create a recessed 3′ terminus of the DNA. Four variants of the DNA strand were prepared that differed in the type and position of the modification. In the first variant of the substrate, CL3.b, which was intended for experiments with D186C, the phosphate group between the last and

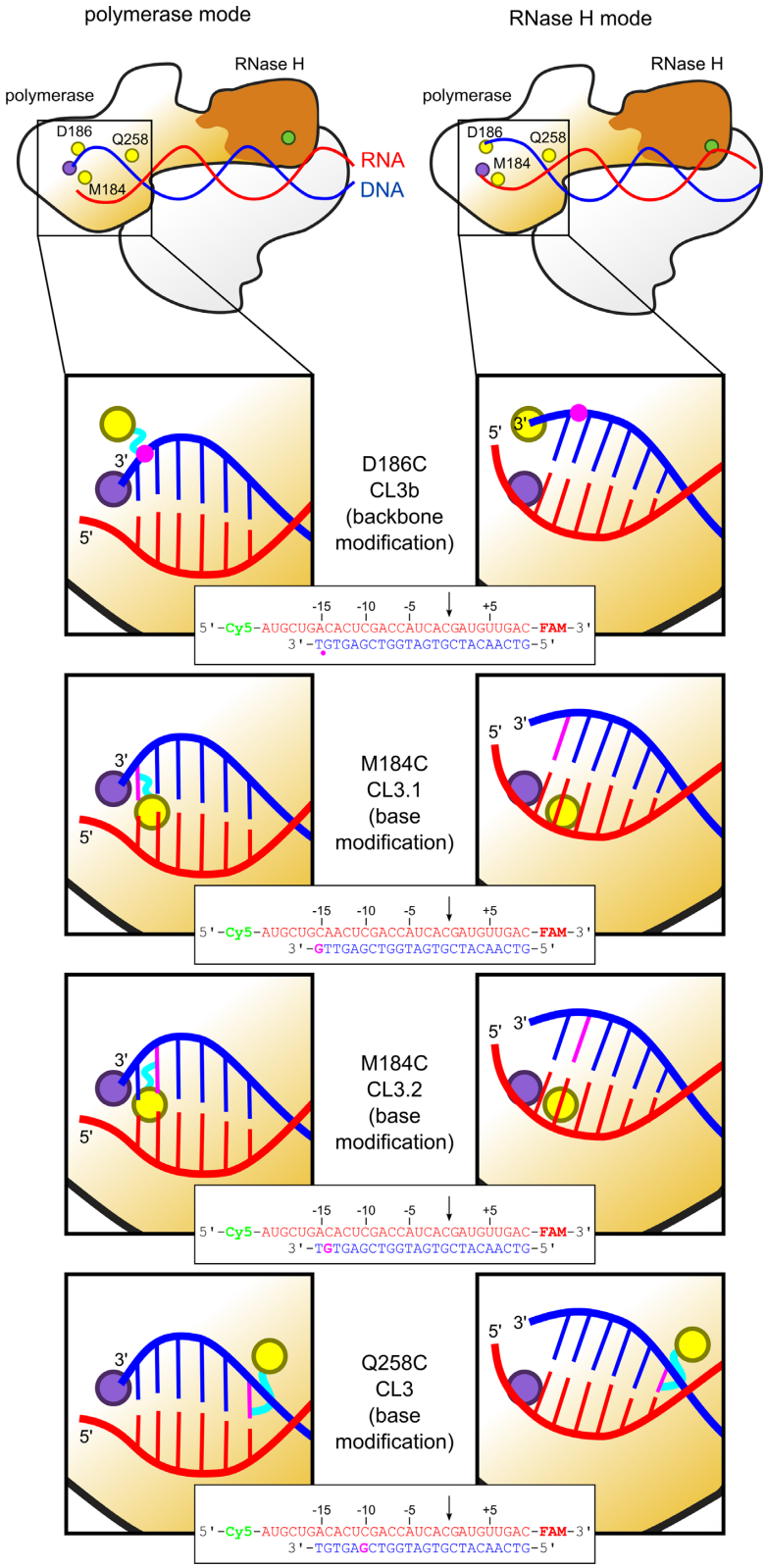

Figure 3. Schematic of the design of cross-linking chemistry. The p66 subunit is shown in orange, with the RNase H domain indicated with a darker shade. The polymerase and RNase H active sites are represented by purple and green circles, respectively. DNA and RNA strands of the hybrid substrate are shown in blue and red, respectively. Yellow circles indicate residues that were individually substituted with cysteines for cross-linking. Their position relative to the relevant DNA modification is shown in the close-ups. The chemical modification that contained a thiol group is indicated in magenta (sticks represent modified bases; circles represent modified backbone phosphates). The disulfide linkage between the thiol groups of the cysteine and modified phosphate/base is indicated as a cyan line. Sequences of the substrates are shown for each cross-linking reaction. Base-modified nucleotides are shown in magenta. The magenta circle indicates the position of the modified backbone phosphate.

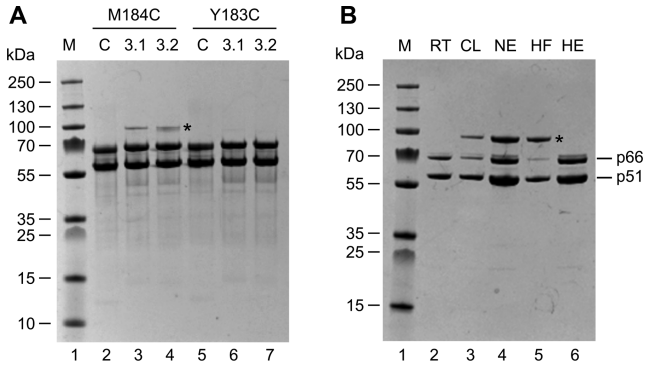

Figure 4. Specificity of the cross-linking approach and purification of cross-linked complexes. (**A**) Cross-linking of M184C and Y183C variants of HIV-1 RT to base-modified substrates in the presence of 125 mM NaCl and 5 mM DTT. Lane 1, molecular weight marker; lanes 2–4, reactions with M184C; lanes 5–7, reactions with Y183C. Substrates in the cross-linking reaction are indicated above the gel. C, no substrate added. The band that corresponds to the cross-linked product is indicated with an asterisk. (**B**) Example of cross-linked complex purification (HIV-1 RT D186C with CL3b). M, molecular weight marker; RT, HIV-1 RT; CL, sample after cross-linking reaction in 45 mM NaCl and without DTT; NE, nickel elution; HF, heparin column flow-through; HE, heparin column elution. The band that corresponds to the cross-linked product is indicated with an asterisk.

penultimate nucleotides of the primer was modified with a thiol group on a two-carbon linker (Figure 3, Supplementary Figure S2). The other two versions of this substrate, CL3.1 and CL3.2, which were intended for cross-linking with M184C and Y183C variants, contained the same modification on a guanine in the last nucleotide (with an appropriate change in sequence) or penultimate nucleotide (Figure 3, Supplementary Figure S2). The fourth variant, CL3, which was intended for experiments with Q258C, contained a base modification in the 6th nucleotide from the 3′-end.

To obtain cross-linked complexes, we incubated HIV-1 RT D186C, M184C, Y183C and Q258C variants with appropriately modified oligonucleotides, and formation of the complex was monitored by non-reducing sodium dodecyl sulphate-polyacrylamide gel electrophoresis. As explained above, our analysis of the structures predicted that cross-linked complexes should form for the M184C variant with CL3.1 and CL3.2 substrates but not for the same hybrids with Y183C protein. Indeed, in the presence of 5 mM DTT and 125 mM NaCl, we observed cross-linked species only for M184C and not for Y183C (Figure 4A), which confirmed the specificity of our approach.

RNase H activity within the cross-linked complexes

To study RNase H activity within the cross-linked complexes, we purified HIV-1 RT D186C, M184C and Q258C variants that were tethered to appropriately modified CL3 substrates (Figure 4B). Each protein variant was mixed with the corresponding RNA/DNA hybrid at a 1:1.25 molar ratio in a buffer without DTT to increase the reaction yield, and the mixture was incubated overnight. After the reaction, ∼50% of the p66 subunit was cross-linked with the hybrid (Figure 4B). The purification protocol for the cross-linked complex comprised two steps. First, the sample was loaded onto a nickel column in the presence of 0.5 M NaCl. All protein-containing species were captured on the column and the buffer with high salt concentration washed away any nucleic acid that was not covalently linked to the protein. Second, the fraction that was eluted from the nickel column with a buffer with high imidazole concentration and that contained both free and cross-linked protein was passed through a heparin column. The free protein bound to the resin, and the cross-linked protein-substrate complexes were collected as the flow-through fraction (Figure 4B).

Purified complexes were used in RNase H activity assays. We sought to exclude the possibility that some of the observed activity resulted from a reaction that was catalyzed by non-covalent complexes that either co-purified with the cross-linked species or appeared over time because of cross-link breakage. We used heparin as a substrate competitor (trap) or a high NaCl concentration, which prevents protein–RNA/DNA interactions. Using fluorescence anisotropy measurements we verified that 3 mg/ml heparin or 0.5 M NaCl completely inhibited substrate binding by HIV-1 RT (Supplementary Figure S3). For each cross-linked complex, RNase H cleavage was initiated by the addition of Mg^{2+}, the essential co-factor. Aliquots of reactions with and without high salt or heparin trap were withdrawn at time points between 30 s and 20 min. The reactions were terminated with EDTA, and the reaction products were analyzed on TBE-urea gels with fluorescence visualization. The results showed efficient RNase H cleavage for all four cross-linked complexes: M184C-CL3.1, M184C-CL3.2, D186C-CL3b and Q258C-CL3 (Figure 5A). At the shortest time of 30 s, up to 55% of the substrate was already converted to cleavage products. As a control, similar reactions were performed with complexes that were pretreated with 20 mM DTT, which caused breakage of the disulfide bond either in the whole population of complexes or a part of it, depending on the protein variant that was used. In such samples, new cleavage products were observed. These cleavages were performed by the free enzyme that dissociated from the substrate and bound it with a different register. The addition of salt or heparin did not affect RNase H cleavage in the cross-linked complexes. However, in the samples that were pretreated with DTT, the addition of salt or heparin strongly inhibited the reaction and completely eliminated off-register cleavage. This confirmed that the addition of either a high salt concentration or heparin trap restricted the observed activity exclusively to chemically cross-linked complexes. In conclusion, immobilization of the substrate at the active site of the polymerase did not preclude RNase H cleavage.

To exclude the possibility that the observed cleavage resulted from rare rearrangements of the enzyme–substrate complex that occurred only during the prolonged times of our experiment, we studied the reaction with the addition of heparin at shorter times, ranging from 15 to 320 ms, using a quenched-flow apparatus (Figure 5B). The product was already observed at the shortest time of 15 ms, indicating that a conformation that allowed simultaneous interactions with the two active sites was not a rare event. The reaction proceeded with similar kinetics for all four complexes: D186C-CL3b, M184C-CL3.1 and M184C-Cl3.2 in which

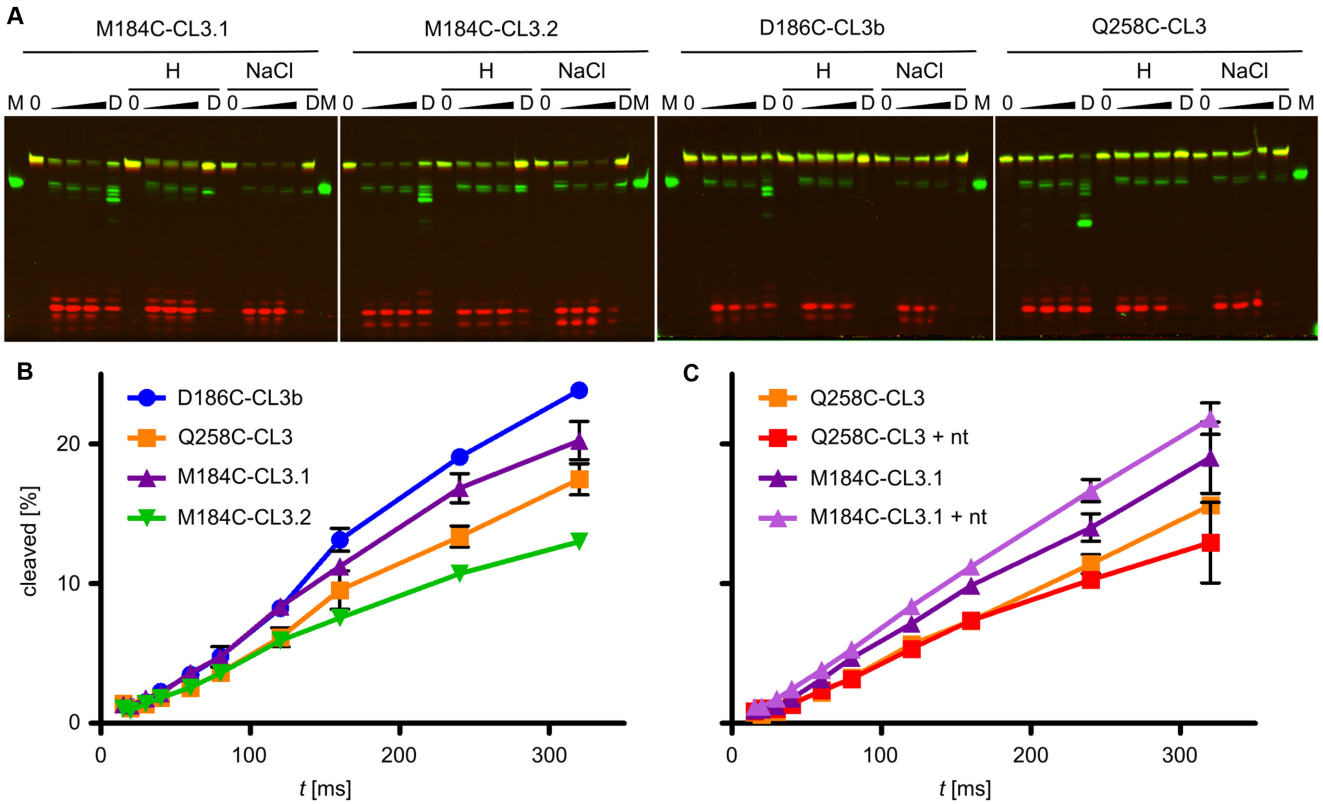

Figure 5. RNase H cleavage within cross-linked HIV-1 RT–substrate complexes. (**A**) RNase H cleavage within cross-linked complexes on a minute time-scale. The protein variant and cross-linked hybrid are indicated on top of the gels. For M184C, two substrates were used with a modification in the last nucleotide (CL3.1) or penultimate nucleotide (CL3.2) of the DNA strand. Reactions were performed for 0.5, 5 or 20 min in the presence of 3 mg/ml heparin (H) or 0.5 M NaCl (NaCl). Reactions in the presence of 20 mM DTT (D) were performed for 20 min. The reactions were stopped with the addition of EDTA and analyzed on 20% TBE-urea denaturing gels. M, marker (Cy5-labeled 24-mer RNA, corresponding to the product of the cleavage 18 nt from the 3′-end of the primer); 0, unreacted complex. The triangle indicates increasing time from 30 s to 20 min. Yellow bands correspond to the intact substrate with both fluorescent labels. Green and red bands correspond to cleavage products that comprised 5′ Cy5 or 3′ fluorescein, respectively. (**B**) RNase H cleavage within cross-linked complexes on a sub-second time-scale. The reactions were performed using a quenched-flow apparatus. The cleaved fraction of the substrate was quantified by densitometry. Data from three independent experiments were averaged and plotted for each time-point. (**C**) RNase H cleavage of CL3 within cross-linked complexes in the presence of a non-hydrolyzable incoming nucleotide. The analysis was performed as in (B).

the nucleic acid was locked at the polymerase active site and Q258C-CL3 in which the substrate could be displaced from the polymerase active center. These results indicate that displacement of the substrate from the polymerase active site was not required for RNase H cleavage and thus the substrate could interact with both active sites simultaneously.

Organization of the polymerase active site

We then sought to provide additional support for the simultaneous interaction of the substrate with both active sites. In particular, we wanted to exclude the possibility that RNase H activity within cross-linked complexes was attributable to transient substrate displacement from the polymerase active site. Within cross-linked complexes, the primer end is tethered to the protein, so such displacement would only be possible with structural distortion of the RNA/DNA helix and protein structure at the polymerase active site. We decided to determine whether the incoming nucleotide can bind to the cross-linked complexes and whether this binding affects RNase H activity. We reasoned that binding of the nucleotide would confirm that the polymerase active site is properly organized, including the protein residues and substrate end. Moreover, nucleotide binding would stabilize the substrate at the polymerase active site. Therefore, if substrate displacement from the polymerase active site is necessary for RNase H activity, then bound nucleotide would inhibit RNase H cleavage.

We first tested whether our Q258C-CL3 and M184C-CL3.1 cross-linked complexes bind dNTP. Fluorescence anisotropy was used to measure the affinity of incoming nucleotides that were labeled with Texas Red for these complexes (Supplementary Figure S4). Ca^{2+} ions that mimic Mg^{2+} at the active site but do not promote polymerization were used to block incorporation of the nucleotide in these assays. Fluorescent dCTP analog, which would base pair with the template G at the polymerase active site, was bound by both complexes ($K_d = 3.6 \pm 2.0$ nM and 3.3 ± 2.0 nM for Q258C-CL3 and M184C-CL3.1, respectively). This interaction was specific because no binding within the analyzed concentration range was observed for the mismatched dATP analog. As an alternative approach to test nucleotide binding, we measured the affinity between cross-linked substrate complexes of HIV-1 RT and a non-hydrolyzable analogue of dCTP (dCpNHpp) in the presence of Ca^{2+} ions

using a thermophoresis method. The K_d values were 12.0 ± 0.6 nM and 19.5 ± 0.6 nM for Q258C-CL3 and M184C-CL3.1, respectively (Supplementary Figure S5). These values were very similar to the K_d of 18 nM that was previously reported for the free protein (34). They were also slightly different from those obtained using the fluorescent nucleotides, but this was expected when considering the differences in chemical structures of the two types of the dNTP analogs.

For efficient and specific binding of the incoming nucleotide, all of the elements of the polymerase active site need to be properly aligned. The divalent metal ions (Mg^{2+} or Ca^{2+}) at the active site coordinate the phosphate groups. Incoming nucleotide base stacking requires the correct orientation of the DNA primer and proper organization of the active site residue Arg72. Finally, base pairing of the incoming nucleotide requires appropriate orientation of the RNA template. The fact that the cross-linked complexes bound the incoming nucleotide indicates that all of the above requirements were met, and the architecture of the polymerase active site with regard to both protein conformation and substrate binding was not perturbed.

After establishing that the cross-linked complexes bind the incoming nucleotide, including its non-hydrolyzable analog, we tested whether its addition, which would result in stabilization of the primer end at the polymerase active site, affected RNase H cleavage. The results of the quenched-flow experiments showed that the presence of dCpNHpp did not affect the kinetics of RNase H cleavage within Q258C-CL3 and M184C-CL3.1 complexes (Figure 5C). Similarly, no effect on cleavage efficiency was observed when the reactions were performed for longer times (Supplementary Figure S6). Additionally, the pattern of cleavage remained unchanged, suggesting that multiple cuts resulted from the mobility of the RNase H domain rather than translocation of the substrate in the complex, which would be blocked by binding of the incoming nucleotide. These results confirm that displacement of the substrate from the polymerase active site is not required for RNase H cleavage and thus our computational and experimental results are in agreement.

DISCUSSION

Coordination of the polymerase and RNase H activities of RTs is a critical element of their mechanism of action and essential for the proper execution of reverse transcription. In particular, access of the substrate to the RNase H active site needs to be restricted. This is important for the execution of specialized and specific cleavage events, such as PPT generation (M. Figiel *et al*., in preparation). For HIV-1 RT, coordination of the two enzymatic activities relies on conformational changes in the protein-substrate complex (19), the detailed mechanism of which has been unclear. In this report, we used a chemical cross-linking approach and MD simulations to show that a complex of HIV-1 RT with RNA/DNA can adopt a conformation in which the substrate interacts with the active site of the polymerase and is simultaneously catalytically bound at the RNase H domain. An indication of such a simultaneous interaction, although not explicit, was provided by studies using RT-RNA/DNA complexes in which the substrate was stabilized at the polymerase active site through binding of the pyrophosphate analog inhibitor (35). However, a quantitative analysis of these data was not performed, and the possibility of transient rearrangement of this dynamic complex that allowed RNase H cleavage through substrate displacement from the polymerase active site could not be fully excluded. While the present report was under review, a study by Li *et al*. was published that described a pre-steady-state kinetic analysis of HIV-1 RT activity (36). The authors monitored concomitant DNA polymerization and RNase H cleavage by HIV-1 RT and used complex models of all possible transitions and reactions in this system together with global fitting to determine the corresponding kinetic parameters. They found that nucleotide incorporation and RNase H cuts occurred simultaneously and that both active sites interacted with the substrate in most HIV-1 RT–substrate complexes. Li *et al*. also found that the rate of polymerization was limited by pyrophosphate release (37), and RNA hydrolysis was limited by the requirement for an intact stretch of RNA downstream of the cleavage site. When these two elements were taken into account, the rates of polymerization and RNase H cuts during reverse transcription were comparable. Therefore, the single-turnover kinetic data are in full agreement with the simultaneous mode we describe herein. Our approach employing chemical tethering of the substrate at the polymerase active site provides a more direct, structure-based evidence of a simultaneous interaction between the substrate and the two active sites. Additionally, our MD simulations allowed us to characterize the conformation that the complex adopts in the simultaneous mode.

RNase H cuts have been shown to occur when polymerization of the DNA pauses (10–12). Pausing was previously presumed to be required to enable the substrate to disengage from the polymerase active site to allow the RNase H to perform cleavage. Our data suggest a different possibility. Polymerase pausing may give the complex sufficient time to undergo the conformational change that was observed in our MD simulations, which does not require substrate removal from the polymerase active site but still allows catalytic RNase H–substrate interactions. In fact, stronger interactions of the hybrid with the RNase H domain that resulted from this conformational change may further hinder substrate translocation and lengthen the pause of polymerization, leading to even tighter coupling of polymerase pausing and RNase H cuts. Therefore, our data are in agreement with coupling of polymerase pausing and RNase H cuts, and substrate removal from the polymerase active site does not need to be invoked to explain this relationship.

The activity of HIV RNase H can be influenced by multiple factors. Among these are mutations in regions that are distant from the active sites and the substrate binding cleft (e.g. (38)) or the addition of nucleocapsid protein (39). The sequence of the RNA/DNA substrate can also greatly impact RNase H cleavage that is important for PPT generation. The effect of many of these factors can be explained by considering the way in which they affect conformational changes in the RT–substrate complex that are required for RNase H activity.

Two types of retroviral RTs have been described based on the architecture of the active molecule: dimeric and

monomeric. In dimeric RTs, such as the HIV-1 enzyme, the RNase H domain is placed within the structure of the heterodimer, and its movements are restricted. In contrast, in monomeric RTs, such as the one from Moloney leukemia virus, polymerase and RNase H domains are tethered by a flexible linker, and the RNase H domain can move freely (40). Therefore, the regulation of access of RNase H to the substrate is different in dimeric and monomeric enzymes. For HIV-1 RT, the conformational changes in the RNA/DNA hybrid that are described herein are used to restrict the amount of RNase H activity. For monomeric enzymes, RNase H activity is limited because the mobile RNase H domain only occasionally interacts with the substrate and cleaves it. Despite this difference, the mechanism of the reverse transcription reaction is the same for both types of RTs. For example, the PPT sequences of HIV-1 and Moloney leukemia virus are very similar; in some cases, they can be used interchangeably by the two enzymes (41,42). Therefore, RTs from related retroviruses can use different mechanisms to modulate RNase H activity.

In summary, we devised a unique approach that combined MD with enzymatic studies of chemically cross-linked protein–nucleic acid complexes to demonstrate the existence of a conformation of the HIV-1 RT–substrate complex that has not been observed in crystal structures. It involves simultaneous interactions between the RNA/DNA and both polymerase and RNase H active sites. Such high energy states that cannot be captured in crystal structures but are relevant to the enzyme's mechanism likely also exist for other nucleic acid enzymes with multiple activities. Our work provides a conceptual and methodological approach for studies of such transient states.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank W. Yang for reading the manuscript; I. Ptasiewicz for technical assistance; R. Marquet, C. Isel and V. Vivet for help with DNA modification and cross-linking; and J. Trylska for access to the quenched-flow apparatus. J.S. also acknowledges support from Praemium Academiae.

FUNDING

Polish National Science Center [contract no. N N301 439738 to M.N.]; Howard Hughes Medical Institute [International Early Career Scientist grant to M.N.]; M.N. is a recipient of a Foundation for Polish Science 'Ideas for Poland' award; Czech Science Foundation [P208/12/1878 to J.S. and M.K.]; Ministry of Education, Youth and Sports of the Czech Republic [project LO1305 to J.S. and M.K.]; Research was performed using Centre for Preclinical Research and Technology (CePT) infrastructure [European Union POIG.02.02.00-14-024/08-00 project].
Conflict of interest statement. None declared.

REFERENCES

1. Abbink,T.E. and Berkhout,B. (2008) HIV-1 reverse transcription initiation: a potential target for novel antivirals? *Virus Res.*, **134**, 4–18.
2. Basu,V.P., Song,M., Gao,L., Rigby,S.T., Hanson,M.N. and Bambara,R.A. (2008) Strand transfer events during HIV-1 reverse transcription. *Virus Res.*, **134**, 19–38.
3. Cote,M.L. and Roth,M.J. (2008) Murine leukemia virus reverse transcriptase: structural comparison with HIV-1 reverse transcriptase. *Virus Res.*, **134**, 186–202.
4. Le Grice,S.F.J. and Nowotny,M. (2014), *Nucleic Acid Polymerases*. Springer-Verlag Berlin Heidelberg, NY, Vol. **30**, pp. 189–214.
5. Gopalakrishnan,V., Peliska,J.A. and Benkovic,S.J. (1992) Human immunodeficiency virus type 1 reverse transcriptase: spatial and temporal relationship between the polymerase and RNase H activities. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 10763–10767.
6. Furfine,E.S. and Reardon,J.E. (1991) Reverse transcriptase. RNase H from the human immunodeficiency virus. Relationship of the DNA polymerase and RNA hydrolysis activities. *J Biol. Chem.*, **266**, 406–412.
7. Gotte,M., Maier,G., Gross,H.J. and Heumann,H. (1998) Localization of the active site of HIV-1 reverse transcriptase-associated RNase H domain on a DNA template using site-specific generated hydroxyl radicals. *J. Biol. Chem.*, **273**, 10139–10146.
8. Kati,W.M., Johnson,K.A., Jerva,L.F. and Anderson,K.S. (1992) Mechanism and fidelity of HIV reverse transcriptase. *J. Biol. Chem.*, **267**, 25988–25997.
9. DeStefano,J.J., Buiser,R.G., Mallaber,L.M., Myers,T.W., Bambara,R.A. and Fay,P.J. (1991) Polymerization and RNase H activities of the reverse transcriptases from avian myeloblastosis, human immunodeficiency, and Moloney murine leukemia viruses are functionally uncoupled. *J. Biol. Chem.*, **266**, 7423–7431.
10. Driscoll,M.D., Golinelli,M.P. and Hughes,S.H. (2001) In vitro analysis of human immunodeficiency virus type 1 minus-strand strong-stop DNA synthesis and genomic RNA processing. *J. Virol.*, **75**, 672–686.
11. Purohit,V., Balakrishnan,M., Kim,B. and Bambara,R.A. (2005) Evidence that HIV-1 reverse transcriptase employs the DNA 3′ end-directed primary/secondary RNase H cleavage mechanism during synthesis and strand transfer. *J. Biol. Chem.*, **280**, 40534–40543.
12. Purohit,V., Roques,B.P., Kim,B. and Bambara,R.A. (2007) Mechanisms that prevent template inactivation by HIV-1 reverse transcriptase RNase H cleavages. *J. Biol. Chem.*, **282**, 12598–12609.
13. Schultz,S.J., Zhang,M.H. and Champoux,J.J. (2004) Recognition of internal cleavage sites by retroviral RNases H. *J. Mol. Biol.*, **344**, 635–652.
14. Wisniewski,M., Balakrishnan,M., Palaniappan,C., Fay,P.J. and Bambara,R.A. (2000) Unique progressive cleavage mechanism of HIV reverse transcriptase RNase H. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 11978–11983.
15. Wisniewski,M., Balakrishnan,M., Palaniappan,C., Fay,P.J. and Bambara,R.A. (2000) The sequential mechanism of HIV reverse transcriptase RNase H. *J. Biol. Chem.*, **275**, 37664–37671.
16. Hostomsky,Z., Hostomska,Z., Hudson,G.O., Moomaw,E.W. and Nodes,B.R. (1991) Reconstitution in vitro of RNase H activity by using purified N-terminal and C-terminal domains of human immunodeficiency virus type 1 reverse transcriptase. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 1148–1152.
17. Smith,J.S., Gritsman,K. and Roth,M.J. (1994) Contributions of DNA polymerase subdomains to the RNase H activity of human immunodeficiency virus type 1 reverse transcriptase. *J. Virol.*, **68**, 5721–5729.
18. Huang,H.F., Chopra,R., Verdine,G.L. and Harrison,S.C. (1998) Structure of a covalently trapped catalytic complex of HIV-I reverse transcriptase: implications for drug resistance. *Science*, **282**, 1669–1675.
19. Lapkouski,M., Tian,L., Miller,J.T., Le Grice,S.F.J. and Yang,W. (2013) Complexes of HIV-1 RT, NNRTI and RNA/DNA hybrid reveal a structure compatible with RNA degradation. *Nat. Struct. Mol. Biol.*, **20**, 230–236.

20. Das,K., Martinez,S.E., Bandwar,R.P. and Arnold,E. (2014) Structures of HIV-1 RT-RNA/DNA ternary complexes with dATP and nevirapine reveal conformational flexibility of RNA/DNA: insights into requirements for RNase H cleavage. *Nucleic Acids Res.*, **42**, 8125–8137.
21. Cornell,W.D., Cieplak,P., Bayly,C.I., Gould,I.R., Merz,K.M., Ferguson,D.M., Spellmeyer,D.C., Fox,T., Caldwell,J.W. and Kollman,P.A. (1995) A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J. Am. Chem. Soc.*, **117**, 5179–5197.
22. Perez,A., Marchan,I., Svozil,D., Sponer,J., Cheatham,T.E., Laughton,C.A. and Orozco,M. (2007) Refinenement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys. J.*, **92**, 3817–3829.
23. Zgarbova,M., Otyepka,M., Sponer,J., Mladek,A., Banas,P., Cheatham,T.E. and Jurecka,P. (2011) Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J. Chem. Theory Comput.*, **7**, 2886–2902.
24. Zgarbova,M., Luque,F.J., Sponer,J., Cheatham,T.E., Otyepka,M. and Jurecka,P. (2013) Toward improved description of DNA backbone: revisiting epsilon and zeta torsion force field parameters. *J. Chem. Theory Comput.*, **9**, 2339–2354.
25. Hornak,V., Abel,R., Okur,A., Strockbine,B., Roitberg,A. and Simmerling,C. (2006) Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins*, **65**, 712–725.
26. Maier,J.A., Martinez,C., Kasavajhala,K., Wickstrom,L., Hauser,K.E. and Simmerling,C. (2015) ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.*, **11**, 3696–3713.
27. Himmel,D.M., Myshakina,N.S., Ilina,T., Van Ry,A., Ho,W.C., Parniak,M.A. and Arnold,E. Structure of a dihydroxycoumarin active-site inhibitor in complex with the RNase H domain of HIV-1 reverse transcriptase and structure-activity analysis of inhibitor analogs. *J. Mol. Biol.*, **426**, 2617–2631.
28. Krepl,M., Havrila,M., Stadlbauer,P., Banas,P., Otyepka,M., Pasulka,J., Stefl,R. and Sponer,J. (2015) Can we execute stable microsecond-scale atomistic simulations of protein-RNA complexes? *J. Chem. Theory Comput.*, **11**, 1220–1243.
29. Nowotny,M., Gaidamakov,S.A., Crouch,R.J. and Yang,W. (2005) Crystal structures of RNase H bound to an RNA/DNA hybrid: Substrate specificity and metal-dependent catalysis. *Cell*, **121**, 1005–1016.
30. Nowotny,M., Gaidamakov,S.A., Ghirlando,R., Cerritelli,S.M., Crouch,R.J. and Yang,W. (2007) Structure of human RNase H1 complexed with an RNA/DNA hybrid: Insight into HIV reverse transcription (vol 28, pg 264, 2007). *Mol. Cell*, **28**, 513–513.
31. Dash,C., Yi-Brunozzi,H.Y. and Le Grice,S.F. (2004) Two modes of HIV-1 polypurine tract cleavage are affected by introducing locked nucleic acid analogs into the (-) DNA template. *J. Biol. Chem.*, **279**, 37095–37102.
32. Huang,H., Harrison,S.C. and Verdine,G.L. (2000) Trapping of a catalytic HIV reverse transcriptase*template:primer complex through a disulfide bond. *Chem. Biol.*, **7**, 355–364.
33. Banerjee,A., Santos,W.L. and Verdine,G.L. (2006) Structure of a DNA glycosylase searching for lesions. *Science*, **311**, 1153–1157.
34. Rittinger,K., Divita,G. and Goody,R.S. (1995) Human-immunodeficiency-virus reverse-transcriptase substrate-induced conformational-changes and the mechanism of inhibition by non-nucleoside inhibitors. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 8046–8049.
35. Beilhartz,G.L. and Gotte,M. (2010) HIV-1 ribonuclease H: structure, catalytic mechanism and inhibitors. *Viruses*, **2**, 900–926.
36. Li,A., Li,J. and Johnson,K.A. (2016) HIV-1 Reverse transcriptase polymerase and RNase H active sites work simultaneously and independently. *J. Biol. Chem.*, **291**, 26566–26585.
37. Li,A., Gong,S. and Johnson,K.A. (2016) Rate-limiting pyrophosphate release by HIV reverse transcriptase improves fidelity. *J. Biol. Chem.*,**291**, 26554–26565.
38. Sevilya,Z., Loya,S., Adir,N. and Hizi,A. (2003) The ribonuclease H activity of the reverse transcriptases of human immunodeficiency viruses type 1 and type 2 is modulated by residue 294 of the small subunit. *Nucleic Acids Res.*, **31**, 1481–1487.
39. Roda,R.H., Balakrishnan,M., Hanson,M.N., Wohrl,B.M., Le Grice,S.F.J., Roques,B.P., Gorelick,R.J. and Bambara,R.A. (2003) Role of the reverse transcriptase, nucleocapsid protein, and template structure in the two-step transfer mechanism in retroviral recombination. *J. Biol. Chem.*, **278**, 31536–31546.
40. Nowak,E., Potrzebowski,W., Konarev,P.V., Rausch,J.W., Bona,M.K., Svergun,D.I., Bujnicki,J.M., Le Grice,S.F. and Nowotny,M. (2013) Structural analysis of monomeric retroviral reverse transcriptase in complex with an RNA/DNA hybrid. *Nucleic Acids Res.*, **41**, 3874–3887.
41. Luo,G.X. and Taylor,J. (1990) Template switching by reverse transcriptase during DNA synthesis. *J. Virol.*, **64**, 4321–4328.
42. Pullen,K.A., Rattray,A.J. and Champoux,J.J. (1993) The sequence features important for plus strand priming by human immunodeficiency virus type 1 reverse transcriptase. *J. Biol. Chem.*, **268**, 6221–6227.