Current Research in Structural Biology 2 (2020) 116–129

Contents lists available at ScienceDirect

Current Research in Structural Biology

journal homepage: www.journals.elsevier.com/current-research-in-structural-biology

Integrative structural biology studies of HIV-1 reverse transcriptase binding to a high-affinity DNA aptamer

Steve Tuske [a,1,2], Jie Zheng [b,1,3], Erik D. Olson [c], Francesc X. Ruiz [a], Bruce D. Pascal [b], Anthony Hoang [a], Joseph D. Bauman [a,2], Kalyan Das [a,3], Jeffrey J. DeStefano [d], Karin Musier-Forsyth [c], Patrick R. Griffin [b,**], Eddy Arnold [a,*]

[a] Center for Advanced Biotechnology and Medicine, And Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ, 08854, USA
[b] Department of Molecular Medicine, The Scripps Research Institute, Jupiter, FL, 33458, USA
[c] Department of Chemistry and Biochemistry, Center for RNA Biology, And Center for Retrovirus Research, The Ohio State University, Columbus, OH, 43210, USA
[d] Department of Cell Biology and Molecular Genetics, University of Maryland College Park, College Park, MD, 20740, USA

ARTICLE INFO

Keywords:
HDX-MS high-Affinity DNA aptamer
HIV-1 reverse transcriptase
Subunit-specific ^{15}N-labeling
SELEX
Integrative structural biology

ABSTRACT

The high-resolution crystal structure of HIV-1 reverse transcriptase (RT) bound to a 38-mer DNA hairpin aptamer with low pM affinity was previously described. The high-affinity binding aptamer contained 2′-O-methyl modifications and a seven base-pair GC-rich tract and the structure of the RT-aptamer complex revealed specific contacts between RT and the template strand of the aptamer. Similar to all crystal structures of RT bound to nucleic acid template-primers, the aptamer bound RT with a bend in the duplex DNA. To understand the structural basis for the ultra-high-affinity aptamer binding, an integrative structural biology approach was used. Hydrogen-deuterium exchange coupled to liquid chromatography-mass spectrometry (HDX-MS) was used to examine the structural dynamics of RT alone and in the presence of the DNA aptamer. RT was selectively labeled with ^{15}N to unambiguously identify peptides from each subunit. HDX of unliganded RT shows a mostly stable core. The p66 fingers and thumb subdomains, and the RNase H domain are relatively dynamic. HDX indicates that both the aptamer and a scrambled version significantly stabilize regions of RT that are dynamic in the absence of DNA. No substantial differences in RT dynamics are observed between aptamer and scrambled aptamer binding, despite a large difference in binding affinity. Small-angle X-ray scattering and circular dichroism spectroscopy were used to investigate the aptamer conformation in solution and revealed a pre-bent DNA that possesses both A- and B-form helical character. Both the 2′-O-methyl modifications and the GC tract appear to contribute to an energetically favorable conformation for binding to RT that contributes to the aptamer's ultra-high affinity for RT. The X-ray structure of RT with an RNA/DNA version of the aptamer at 2.8 Å resolution revealed a potential role of the hairpin positioning in affinity. Together, the data suggest that both the 2′-O-methyl modifications and the GC tract contribute to an energetically favorable conformation for high-affinity binding to RT.

Abbreviations: SELEX, systematic evolution of ligands by exponential enrichment; HDX-MS, hydrogen-deuterium exchange coupled to liquid chromatography mass spectrometry; HDX, hydrogen-deuterium exchange; HIV-1 RT, human immunodeficiency virus type-1 reverse transcriptase; p66, 66 kDa subunit of HIV-1 RT; p51, 51 kDa subunit of HIV-1 RT; WT aptamer, wild-type DNA aptamer; Δmethyl aptamer, DNA aptamer lacking 2′-O-methyl groups; RNA/DNA aptamer dU_3, RNA/DNA hybrid aptamer with three deoxyuridine residues in the aptamer loop region.

* Corresponding author.
** Corresponding author.
E-mail addresses: pgriffin@scripps.edu (P.R. Griffin), arnold@cabm.rutgers.edu (E. Arnold).

[1] The first two authors should be regarded as joint First Authors.
[2] Present addresses: Steve Tuske and Joseph Bauman: Amicus Therapeutics, Gene Therapy Center of Excellence, 3675 Market Street, Philadelphia, PA 19104.
[3] Jie Zheng: Shanghai Institute of Materia Medica, Shanghai, China; Kalyan Das: Rega Institute and Department of Microbiology and Immunology, KU Leuven, Belgium.

https://doi.org/10.1016/j.crstbi.2020.06.002
Received 11 March 2020; Received in revised form 3 June 2020; Accepted 19 June 2020

1. Introduction

HIV-1 reverse transcriptase (RT) is a highly dynamic enzyme that carries out multiple reactions during reverse transcription of single-stranded viral RNA to proviral double-stranded DNA (dsDNA). RT undergoes multiple conformational changes during its catalytic cycle upon binding nucleic acids, nucleotide substrates, nucleoside, and non-nucleoside drugs (Sarafianos et al., 1999, 2003, 2004, 2009; Boyer et al., 2004; Bauman et al., 2013; Das et al., 2012, 2014; Huang et al., 1998; Xavier Ruiz, 2020). Previously we showed that small molecule drug-like fragments can be used to detect allosteric changes in RT structure and can be used to identify new binding sites for structure-based drug design efforts (Bauman et al., 2008, 2013). Recently we solved the crystal structure of RT in the presence of a high-affinity (K_D = 14 pM) DNA 38-mer hairpin aptamer with two 2′-O-methyl modifications (WT aptamer, Fig. 1) developed using a modified version of the systematic evolution of ligands by exponential enrichment (SELEX) approach (Miller et al., 2015), and demonstrated that the complex is catalytically active *in crystallo* (Das et al., 2016). This structure represents the highest resolution crystal structure of an RT-nucleic acid complex solved to date, without chemical cross-linking or a monoclonal antibody Fab fragment for crystallization.

The 2.3 Å resolution structure of the RT-aptamer complex revealed specific contacts between RT and the 2′-O-methyl-modified template strand of the aptamer that may contribute to the very high affinity of the aptamer for RT. However, the unmethylated aptamer (Δmethyl aptamer), still had a relatively high affinity for RT (K_D = 163 pM). Thus, addition of the two 2′-O-methyl modifications only enhanced the K_D by ~10-fold indicating that other structural properties of the aptamer contributed to the three-orders-of-magnitude increase in affinity of RT for the aptamer compared to non-selected DNAs (~5 nM).

All crystal structures of RT bound to nucleic acid template-primers show the nucleic acid in a bent conformation near the polymerase active site 5–6 base pairs upstream from the primer terminus, with a conformation that is A-like near the active site and B-like towards the RNase H active site (Hu and Hughes, 2012). The bend, or kink, at the junction is formed by the sugar moieties (adopting a 3′-*endo* configuration) of the Watson-Crick base-paired template-primer near the RT polymerase site. When we reported the RT-aptamer crystal structure, we speculated that the seven consecutive G:C pairs near the RT polymerase active site could adopt a "pre-bent" conformation because crystal structures of G:C-rich DNA have been shown to favor A-form DNA (Ng et al., 2000). An aptamer in a "pre-bent" conformation could facilitate RT binding, as less energy would be required for RT to mold the nucleic acid to fit its nucleic acid-binding groove. We could also not completely rule out that an effect of the aptamer was to stabilize the conformation of RT, although no significant differences in the structure of the RT-aptamer complex compared to those of RT with unselected template-primers were found (Miller et al., 2015).

In an effort to further elucidate the mechanism of high-affinity binding of this aptamer to RT, to further explore the structural dynamics of RT, and to develop a general method for detecting protein allostery in the presence of nucleic acids, we employed hydrogen-deuterium exchange (HDX) coupled to liquid chromatography-mass spectrometry (LC-MS). A method involving subunit-specific labeling with ^{15}N was developed to distinguish the separate effects on the p66 and p51 subunits of RT (Goswami et al., 2015), and was recently applied to characterize the binding of an RNA aptamer to HIV-1 RT (Nguyen et al., 2020) Comparison of the HDX profiles of RT bound to WT aptamer vs. a scrambled aptamer (Fig. 1) revealed that binding of both DNAs lead to a similar degree of stabilization of RT dynamics. We employed small-angle X-ray scattering (SAXS) to determine the global shape of several aptamer DNA constructs (Fig. 1), revealing that the presence of the 2′-O-methyl modifications and the extended run of G:C base pairs gives rise to a "pre-bent" conformation of the aptamer in its unbound state. Circular dichroism (CD) spectroscopy confirmed that the aptamer displayed hybrid A- and B-form helical geometry, consistent with a "pre-bent" conformation. We also sought structures of RT complexed with the Δmethyl aptamer and with an RNA/DNA version of the aptamer (named RNA/DNA dU_3 aptamer), shown in Fig. 1. The RT-RNA/DNA dU_3 aptamer complex, has a similar structure to that adopted by the WT aptamer, but with a displaced and highly flexible hairpin loop region. Thus, coupled to its lower affinity for RT (K_D = 542 pM, 40-fold lower than for the DNA aptamer), the structure suggests that the aptamer

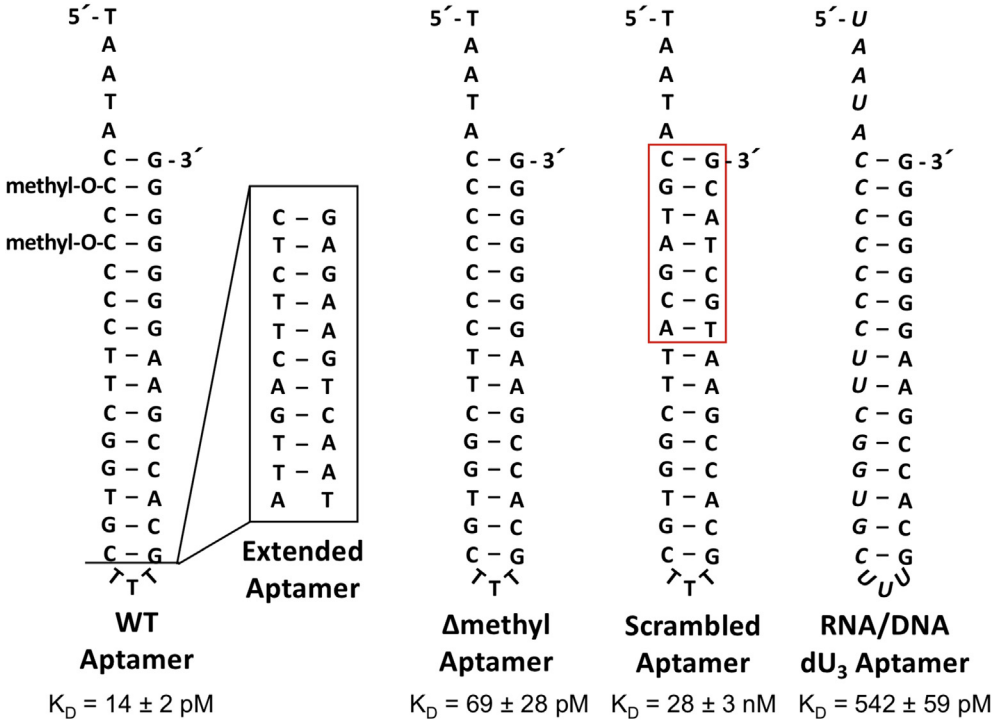

Fig. 1. Sequence and secondary structure of all nucleic acid constructs used in this study. The "methyl-O" signifies that the ribose sugar moiety of the indicated nucleotide is modified with a 2′-O-methyl group. The red box indicates the extended GC tract nucleotides that have been changed in the scrambled aptamer construct. The normal typeface letters correspond to DNA nucleotides, while the italic typeface letters correspond to RNA nucleotides. The binding affinities (K_D) of each aptamer for HIV-1 reverse transcriptase (RT) are reported with their standard deviations.

hairpin loop position and conformation may play an important role in attaining high affinity for RT.

2. Materials and methods

Trifluoroacetic acid (TFA), D_2O, and hydrochloric acid (HCl, 37%) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Formic acid (FA, 98%) was acquired from Fluka (Buchs, Switzerland). Mobile phases were prepared using HPLC-grade water and acetonitrile from Fisher Scientific (Fair Lawn, NJ, USA). 100X MEM vitamins for MJ9 minimal media were purchased from Sigma-Aldrich.

2.1. ^{15}N-labeling and expression of HIV-1 reverse transcriptase (RT)

The 51 kDa subunit of RT (p51) was uniformly labeled with ^{15}N-isotope as described previously (Goswami et al., 2015). A single colony of *E. coli* BL-21 DE3 RIPL (Agilent Technologies CA, USA) harboring pCDF-Duet plasmid carrying the 6XHis-tagged p51 subunit of RT was inoculated into 0.5 mL of Luria-Bertani (LB) broth with 50 μg/mL streptomycin, 34 μg/mL chloramphenicol and incubated at 37 °C with vigorous shaking until the culture reached turbidity (~4–6 h). 20 μL of the starter culture was diluted into 20 mL of fresh MJ9 media with 50 μg/mL streptomycin and 34 μg/mL chloramphenicol with ^{15}N–NH_4SO_4 as the sole nitrogen source and incubated at 37 °C overnight with shaking. 20 mL of overnight culture was diluted into 1 L of MJ9 media with 50 μg/mL streptomycin and incubated at 37 °C with shaking until an OD600 of 0.6–0.8 was reached. Cultures were incubated at 4 °C for 1 h and p51 was expressed by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Induced cultures were incubated at 17 °C overnight with vigorous shaking. Cultures were harvested by centrifugation at 6000×*g* for 20 min at 4 °C. Cell paste was stored at −80 °C until further use. Unlabeled p66 was expressed identically to p51 except that LB broth was used instead of MJ9 minimal media.

2.2. Purification of p66 and p51 subunits

Approximately 11 g of cell paste was resuspended in 50 mL of lysis buffer containing 50 mM phosphate buffer, pH 8.0, 0.6 M NaCl, 0.1% Triton X-100, 5 mM imidazole, 1 mM 2-mercaptoethanol, one tablet of Roche cOmplete EDTA-free Protease Inhibitor Cocktail for each 50 mL of lysis buffer, and 1 mM phenylmethylsulfonyl fluoride (PMSF), added just before use. The resuspended cells were lysed by sonication (Misonix-3000) with a total processing time of 10 min and a 30 sec on/off pulse at an output of 7.0 (93 W) in an ice-water bath. The crude lysate was separated into soluble and insoluble fractions by centrifugation at 38,000×*g* for 45 min at 4 °C in a Sorvall RC 6+ centrifuge with a fixed-angle-rotor. The clarified supernatant was loaded onto a 2–4 mL nickel-NTA (Qiagen) column equilibrated with lysis buffer without PMSF. The column was washed with 20 column volumes of lysis buffer, then 5 column volumes of lysis buffer containing 1.5 M NaCl, and eluted with 10 mL of 50 mM phosphate buffer, pH 7.0, 200 mM NaCl, 1 mM 2-mercaptoethanol, and 250 mM imidazole. The p66 subunit, but not p51, was incubated overnight with a 1:25 mass:mass ratio of HRV-14 3C protease (produced in-house) to remove the hexahistidine tag from the N-terminus of p66. Eluate from the nickel-NTA column (10 mL) was concentrated to 0.5 mL using a Millipore Amicon ultracentrifugal concentrator (30 kDa MW cutoff), diluted to 10 mL with 50 mM diethanolamine (DEA) buffer, pH 8.9, and loaded onto a Mono Q 10/100-anion exchange column (GE Healthcare). The column was washed with 100 mL of 50 mM DEA buffer, pH 8.9 and eluted with a linear gradient from 0 to 24% buffer B: 50 mM DEA pH 8.9, 1 M NaCl. Elution was monitored at A_{280} nm and peak fractions were analyzed by SDS-PAGE. The p51 and p66 subunits eluted at ~120 mM NaCl. Fractions containing the purest sample judged by SDS-PAGE (>95%) were pooled and buffer exchanged into 10 mM Tris-HCl, pH 8.0, 75 mM NaCl using a centrifugal concentrator unit with a 30 kDa MW cutoff. RT subunits were quantitated using an extinction coefficient at A_{280} nm of 3.1 (mg/mL) $^{-1}$cm^{-1} for p51 and 3.5 (mg/mL)$^{-1}$cm^{-1} for p66.

2.3. Reconstitution and purification of ^{15}N-labeled HIV-1 RT

Reconstitution of unlabeled p66 and ^{15}N-labeled p51 into heterodimeric RT was carried out as described (Goswami et al., 2015) using a slight molar excess of p51:p66 (1.2:1). The two subunits were mixed and dialyzed against a buffer containing 50 mM Tris-HCl, pH 8.0, 25 mM NaCl, 1 mM EDTA, and 50% glycerol for five days at 4 °C with gentle stirring. Purification of heterolabeled RT was performed as described above using a nickel-NTA column followed by Mono Q anion-exchange chromatography. The remaining hexahistidine tag on ^{15}N-labeled p51 was cleaved using HRV14 3C protease after Ni-NTA purification, but before Mono Q column chromatography. RT was snap-frozen in liquid nitrogen and stored at −80 °C until use.

2.4. Intact protein analysis using MALDI-MS

The extent of ^{15}N labeling of p51 and the labeling of RT was measured by MALDI-TOF mass spectrometry by the Rutgers Core Mass Spectrometry Facility as described previously (Goswami et al., 2015). For this study the p51 subunit of reconstituted RT was found to be >99% ^{15}N-labeled (not shown).

2.5. RT-aptamer DNA binding experiments

Equilibrium dissociation constants (K_D) for the constructs shown in Fig. 1 were determined using nitrocellulose filter binding assays as described (Das et al., 2016). The buffer used contained 50 mM Tris-HCl, pH 8, 80 mM KCl, 1 mM DTT, 2 mM $MgCl_2$, and 0.1 mg/mL bovine serum albumin (BSA).

2.6. Hydrogen-deuterium exchange mass spectrometry (HDX-MS)

Solution-phase amide HDX was carried out with a fully automated system as described previously (Chalmers et al., 2006; Goswami et al., 2014). Briefly, 4 μl of 10 μM full-length RT alone or in the presence of aptamer or scrambled aptamer (Fig. 1) at a molar ratio of 1.1 and 1.5:1 DNA:RT, respectively, was diluted to 20 μl with D_2O and incubated at 4 °C for 10, 30, 60, 300, 900, or 3600 sec. Following on-exchange, unwanted forward- or back-exchange was minimized by reducing pH and temperature as follows: the protein was denatured by dilution to 50 μl with 0.1% (v/v) TFA in 5 M urea (held at 1 °C). Samples were then passed across an immobilized pepsin column (prepared in-house) at 50 μl min^{-1} (0.1% v/v TFA; 15 °C); the resulting peptides were trapped on a C8 trap cartridge (Hypersil Gold; Thermo Fisher Scientific). Peptic peptides were eluted with a linear gradient of 4% (w/v) to 40% (w/v) acetonitrile with 0.3% (w/v) formic acid over 5 min at 2 °C across a 1 × 50 mm C18 HPLC column (Hypersil Gold; Thermo Fisher Scientific) and electrosprayed directly into an Orbitrap mass spectrometer (Q-Exactive; Thermo Fisher Scientific). Peptide ion signals were confirmed if they had a MASCOT score of 20 or greater and had no ambiguous hits using a decoy (reverse) sequence in a separate experiment using a 60 min gradient. The intensity-weighted average m/z value (centroid) of each peptide's isotopic envelope was calculated with in-house developed software (Pascal et al., 2012) and corrected for back-exchange on an estimated 70% recovery and accounting for the known deuterium content of the on-exchange buffer. Identification of ^{15}N-labeled and unlabeled peptic peptides originating from RT was carried out using modified software described in (Goswami et al., 2015).

2.7. Aptamer DNA preparation for SAXS

The WT aptamer, extended aptamer, Δmethyl aptamer, and scrambled aptamer were purchased from Integrated DNA Technologies. DNAs

(300 μg) were refolded by heating at 90 °C for 3 min in the presence of 50 mM Tris-HCl, pH 8, 2 mM $MgCl_2$, 1 mM dithiothreitol (DTT), and 80 mM KCl, followed by incubation on ice for at least 30 min. The folded DNA was then purified via size exclusion chromatography (SEC) on a 24-mL Superdex 200 10/300 GL Increase column (GE Healthcare) in SEC running buffer containing 50 mM Tris-HCl (pH 7.4), 80 mM KCl, 1 mM DTT, 2 mM $MgCl_2$, and 3% glycerol (w/v) at a flow rate of 1 mL/min. Peak fractions containing the desired DNA were pooled and concentrated to 75–80 μl using an Amicon 0.5-mL 3 kDa MW cutoff spin concentrator (EMD Millipore). The SEC running buffer was used to serially dilute the DNAs, yielding three sample concentration ranges (2.0–2.6 μg/μl, 1.0–1.3 μg/μl, and 0.5–0.65 μg/μl). An aliquot of the SEC buffer used to purify the DNAs was reserved for use in SAXS for buffer subtraction. All submitted samples were run on 12% native polyacrylamide gels concomitant with SAXS data acquisition to verify their homogeneity.

2.8. SAXS data acquisition, analysis, and model building

Samples (30 μl) were shipped in 96-well plates (Axygen Scientific) at 4 °C to the 12.3.1 SIBYLS beamline at the Advanced Light Source (Lawrence Berkeley National Lab) (Putnam et al., 2007; Hura et al., 2009). Scattering data were acquired and buffer subtraction was performed by the SIBYLS beamline staff as described (Hura et al., 2009; Dyer et al., 2014). Subsequent SAXS data analysis was performed as previously described (Jones et al., 2014; Post et al., 2016). The best fitting D_{max} value was determined using the program GNOM (Konarev et al., 2006). This D_{max} was used to generate 20 *ab initio* envelopes using the DAMMIN program in jagged mode with no symmetry restraints imposed (Jones et al., 2014; Konarev et al., 2006), and the χ^2 fits and reproducibility (normalized spatial discrepancy, NSD) values of these envelopes were calculated. The 20 envelopes were averaged, packed with at least 2000 "dummy atoms", and used as the starting point for an additional 20 *ab initio* envelope calculations, generated in expert mode with no symmetry restraints imposed. These envelopes were averaged to generate the final envelope and their χ^2 fits and NSD values were determined. Overlay of envelopes was performed using the DAMSEL and DAMSUP programs (Konarev et al., 2006).

The program MONSA, which is an extended version of DAMMIN for multiphase bead modeling was used for simultaneous fitting of the aptamer and extended aptamer curves (Svergun, 1999). MONSA modeling was performed in general as described in (Rambo, 2015). Briefly, the aptamer and extended aptamer scattering data were truncated to 0.2 $Å^{-1}$ and Porod-Debye analysis was performed using the program Scatter to estimate the molecular volume of each DNA. The MONSA online web service (https://www.embl-hamburg.de/biosaxs/atsas-online/monsa.php) was used to generate 15 *ab initio* envelopes of the aptamer and extended aptamer DNAs using the following inputs: the D_{max} of the larger extended aptamer DNA was used to set the radius of the spherical search volume, the molecular volumes of the aptamer and extended aptamer DNAs were set, and it was indicated that there were two phases, with the aptamer and extended aptamer sharing a phase, while the extended aptamer contains an additional phase. These envelopes were averaged to generate the final envelope and their χ^2 fits and NSD values were determined. Overlay of envelopes was performed using the DAMSEL and DAMSUP programs.

2.9. CD analysis

The WT aptamer, Δmethyl aptamer, and scrambled aptamer constructs were refolded by heating 20 μM of DNA at 90 °C for 3 min in the presence of 50 mM Tris-HCl, pH 8, 2 mM $MgCl_2$, 1 mM DTT, and 80 mM KCl, followed by incubation on ice for at least 30 min. CD spectra of the folded DNAs were collected in a 1 mm quartz cuvette using a Jasco J-815 CD spectrometer. For each independent experiment (N = 2 for WT aptamer, and N = 3 for Δmethyl aptamer and scrambled aptamer), 5 CD spectra were collected, the buffer background was subtracted, and the final buffer-corrected spectra were averaged. The average spectra were smoothed using the smoothing macro in KaleidaGraph with a window size of 5.

2.10. RT-aptamer crystallization

HIV-1 RT plasmid RT152A allows for the co-expression of p66Δ555 and p51Δ428 and codes for mutations in p66:C280S, D498N; and in p51: C280S. HIV-1 RT152A was expressed, and purified following the previously published protocol (Bauman et al., 2008; Das et al., 2016). The RT-Δmethyl aptamer and RNA/DNA dU_3 aptamer were purchased from IDT and Dharmacon, respectively, and purified by polyacrylamide gel electrophoresis. Complexes of RT with the RNA/DNA dU_3 aptamer or Δmethyl aptamer were generated by mixing RT at a concentration of 85 μM (10 mg/mL) with a 1.2-fold excess (except for the, 2-fold excess). These complexes were directly used for crystallization or analyzed by SEC. The SEC-purified RT-aptamer peak was used for crystallization (after concentrating it to ~85 μM with Amicon Centrifugal Filter Units). SEC was done with a 24-mL Superdex 200 10/300 GL Increase column (GE Healthcare), but in buffer containing 50 mM HEPES pH 7.5, 100 mM NaCl at a flow rate of 1 mL/min.

The crystallization conditions screened, as reported previously (Das et al., 2016), were composed of 8–12% (w/v) PEG 8000, 25 mM Bis-tris Propane pH 6.8, 75 mM Bis-tris Propane pH 7.4, 50 mM ammonium sulfate, 5% (v/v) glycerol, and 5% (w/v) sucrose. A 1:1 ratio of protein-to-well solution was suspended in a 2 μl drop over a 1-mL well. Crystallization plates were incubated at least for 7–14 days at 4 °C. Crystals were obtained for the RT-Δmethyl aptamer complex set up by mixing RT and the DNA directly: no electron density was observed for the nucleic acid. RT- RNA/DNA dU_3: crystals were only obtained after SEC purification, yielding the X-ray structure reported herein. Crystal cryocooling was performed in two steps, 12.5% and 25% (v/v) glycerol plus well solution, and flash-cooled in liquid nitrogen.

2.11. Data collection, processing, and structure determination

Data were collected on flash-cooled crystals at 100 K on the 12-2 beamline at the Stanford Synchrotron Radiation Lightsource. A 270° dataset was collected at a wavelength of 0.98 Å with 1° oscillations. The data were integrated and scaled using HKL 2000 (Otwinowski and Minor, 1997). Phasing was performed by molecular replacement with the PDB 5D3G as model, i.e., RT-aptamer structure (subtracting the nucleic acid to avoid bias) using phenix. phaser (Adams et al., 2010). Prior to addition of the nucleic acid, two cycles of phenix. refine were done (Adams et al., 2010). The RNA/DNA aptamer dU_3 as built using the aptamer coordinates superimposed to the obtained mFo-DFc electron density map and adding the 2′-hydroxyl groups manually with Coot (Emsley and Cowtan, 2004). The RNA base-pairing restraints were generated with the 'PDB to 3D Restraints' server (http://rna.ucsc.edu/pdbrestraints/pdbtorestraints.html), and RCrane (RNA Constructed using RotAmeric NuclEotides) (Keating and Pyle, 2012) allowed identification and correction of backbone conformer errors. The refinement was completed with iterations of phenix. refine and manual rebuilding with Coot. As with the RT-aptamer crystal structure, a combination of NCS restraints and TLS refinement produced a final model with significantly improved agreement indices and stereochemical quality *vs.* other strategies of refinement. Data collection and refinement statistics are shown in Supplementary Table S3. Figures were generated using PyMol (DeLano, 2002). Structure analysis was performed using the web interface of 3DNA-DSSR (Hanson and Lu, 2017) and PyMol. The coordinates were deposited in the PDB under accession number 6BHJ. Note that, while both RT-RNA/DNA aptamer copies present in the asymmetric unit were analyzed, the results presented correspond to the copy with the best real space refinement Z-score (RSRZ) (see page 25 of the validation report file of PDB 6BHJ).

3. Results

3.1. Hydrogen-deuterium exchange of HIV-1 reverse transcriptase

HIV-1 RT is a well-studied enzyme and much is known about its biochemical function, three-dimensional structure, drug inhibition, and drug resistance from more than three decades of investigation (Das and Arnold, 2013a, 2013b). RT is known to be a dynamic protein that undergoes multiple conformational changes upon binding various substrates and inhibitors (Sarafianos et al., 1999; Bauman et al., 2013; Das et al., 2014; Huang et al., 1998; Rodgers et al., 1995; Hsiou et al., 1996) and is a good model for studying protein dynamics by HDX. Using our method of subunit-specific differential isotopic labeling, we undertook HDX on unliganded RT (Goswami et al., 2015). Tandem MS/MS was used to identify both ^{15}N-labeled (p51) and unlabeled (p66; natural isotopic distribution) peptides derived from each of the two subunits. Coverage was 96% and 85% for p66 and p51, respectively, with an overall coverage for RT of 91% as shown in Fig. 2. For the most part, our HDX results show a p51 that is similar to p66 in its dynamics despite its more compact shape observed in its many crystal structures, but consistent with the two subunits having mostly shared secondary structure (Kohlstaedt et al., 1992; Jacobo-Molina et al., 1993). When our HDX results were mapped onto the crystal structure of unliganded RT (Bauman et al., 2008) (PDB 3DLK), three regions of p66 known to play an important role in RT's function exhibit a high rate of exchange. Fig. 3 shows the tips of the fingers, thumb, and the RNase H domain form a crest comprised of relatively dynamic structural elements into which sits the more-ordered palm and connection subdomains. The results from HDX show that much of the RNase H domain is unprotected from exchange and highly dynamic in unliganded RT. This observation is consistent with previously reported work on HDX of RT by Wintrode and co-workers (Seckler et al., 2009) and with crystal structures where few crystal lattice contacts are made with the RNase H domain (Tuske et al., 2004) (PDB 1T05). Moreover, discrepancies in RT temperature factors and HDX can be explained by crystal contacts within the crystal lattice or differences in construct design in the RT used for these studies, which was engineered to give high-resolution crystal structures (Bauman et al., 2008).

3.2. HDX of HIV-1 reverse transcriptase bound to WT aptamer DNA

Previously, we used SELEX to develop aptamers that bind to RT with low-pM affinity (Fig. 1) (Das et al., 2016; Olimpo and DeStefano, 2010). Here, we used HDX to explore the effects of aptamer binding to RT on the exchange kinetics of regions found in the RT-aptamer crystal structure that directly contact DNA, and to see if these effects extended to other RT regions. HDX is a powerful method for exploring protein-ligand interactions in solution. While many studies have been reported that examine the effects of the binding of other proteins, small molecules, and inhibitors on protein dynamics (Gallagher and Hudgens, 2016), until recently, relatively few studies have been carried out to determine how

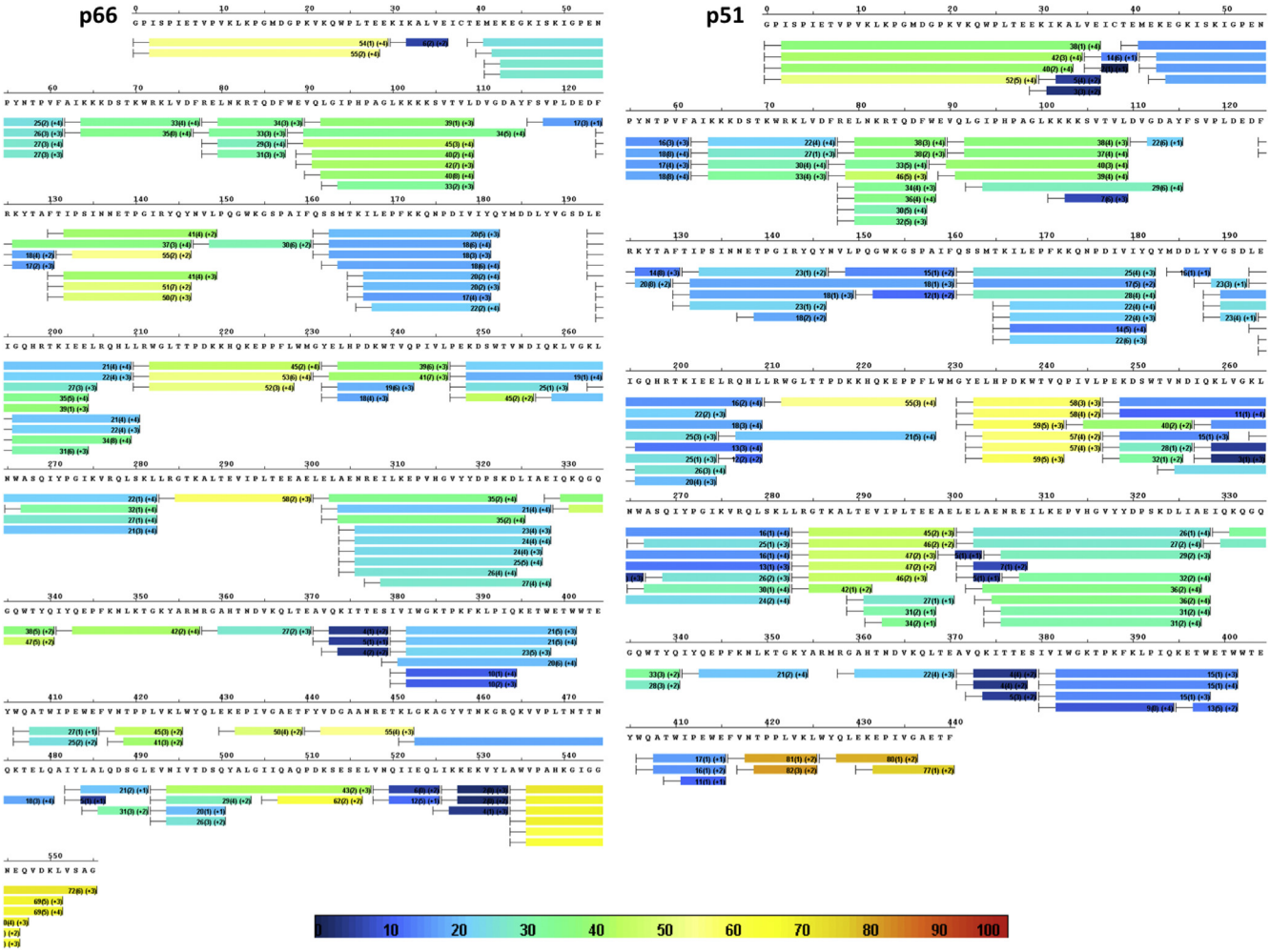

Fig. 2. Dynamics of unliganded reverse transcriptase probed by hydrogen/deuterium exchange. Percentages of hydrogen bonding activities are color-coded according to the color gradient key. Regions that have no sequence coverage and include proline residues that have no amide hydrogen exchange activity are blank. The HDX data from all overlapping peptides are consolidated to individual amino acid values using a residue averaging approach (Keppel and Weis, 2015).

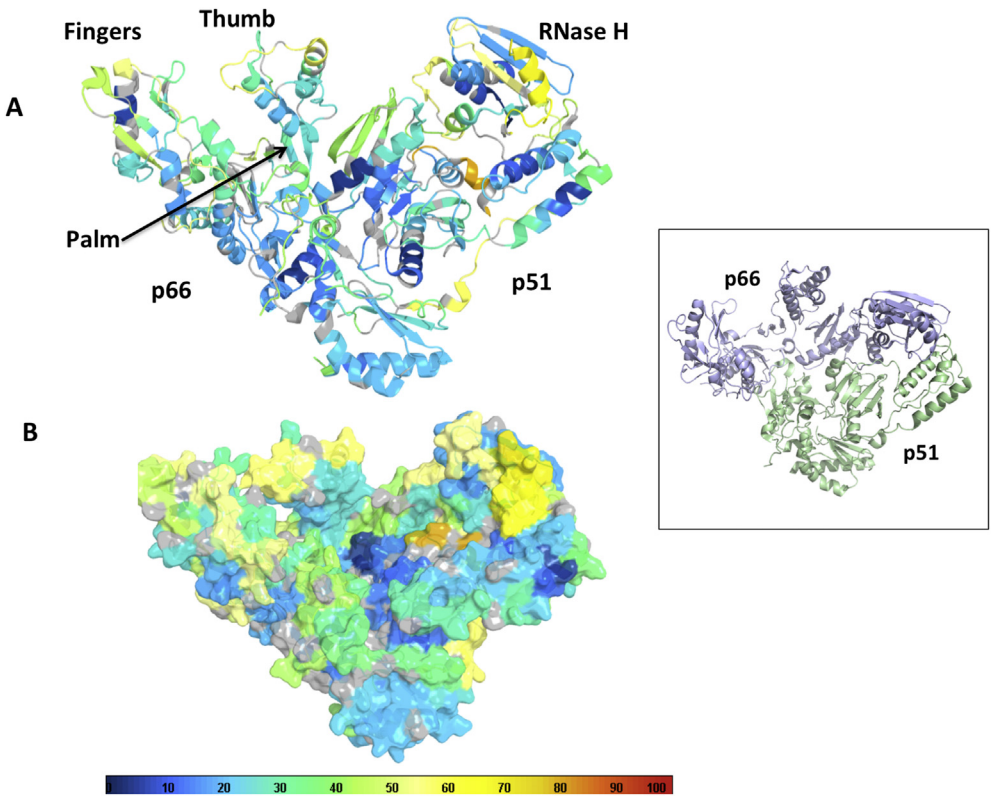

Fig. 3. HDX of unliganded RT mapped onto its crystal structure. HDX Workbench was used to map the HDX data to the RT-apo structure (3KLD). Extent of exchange from HêD is indicated by the legend at the bottom of the figure. (A) The HDX results from Fig. 2 displayed in ribbon representation in Pymol. (B) A Pymol surface representation showing regions of fast (orange and yellow), medium (green), and slow (blue) exchange. Gray indicates RT residues not identified during the HDX experiment. The two subunits of RT are shown inset; the p51 subunit is shown in green and the p66 subunit is blue.

nucleic acids affect HDX of their protein-binding partners (Morton et al., 2010; Bernardes and Batista, 2012; Zheng et al., 2015).

We wanted to develop a method using HDX to explore conformational changes in proteins upon binding cognate nucleic acids and to map protein-nucleic acid interactions. We found that using large excess molar concentrations of DNA to ensure >90% RT-DNA complex formation was technically challenging because the excess DNA tended to precipitate at the low pH and temperature used to quench the HDX reaction and fouled the LC-MS system. Moreover, relatively high concentrations of certain nucleic acids lead to aggregation of RT, complicating HDX analysis (data not shown). The high-affinity aptamer was ideal for our HDX studies because its tight binding ensures complete occupancy of RT at equimolar concentrations, as shown in previous work (DeStefano and Nair, 2008).

Fig. 4 shows the results of HDX on RT bound to the aptamer after subtracting the HDX of unliganded RT shown in Fig. 2. This differential HDX experiment is designed to quantify the net change in exchange of D→H in the presence of a ligand. The colored bars represent peptides identified during the experiment that had a net change in deuterium exchange in the presence of DNA. The net effect of aptamer DNA on the exchange kinetics of RT was a widespread stabilization of the protein, as indicated by a negative value of exchange (Fig. 4). No peptides isolated from either subunit of RT showed an increase in exchange when DNA was bound. When mapped onto the crystal structure of the RT-aptamer complex (PDB 5D3G) shown in Fig. 5, it is clear that the overall effect of the aptamer is to substantially rigidify the regions of RT that were the most dynamic and exchanged most rapidly in unliganded RT. The greatest changes are seen for the fingers, palm, thumb, and RNase H domains. The parts of RT colored green in Fig. 5 are those with modest protection, ≤10% exchange, and those colored blue show greater protection, 11–20% exchange. The protection from exchange afforded by binding the aptamer extends well beyond the regions harboring amino acid residues that make direct contact with the nucleic acid.

Table 1 lists the regions of RT protected from exchange in the presence of the aptamer and the associated protein-DNA interactions identified from the RT-aptamer crystal structure (5D3G). There is excellent agreement between interactions observed in the crystal structure of the RT-aptamer complex and the regions of RT protected from exchange when bound to the same aptamer in solution. Of the thirty amino acid residues found to directly contact the aptamer, 20 are located in the polymerase domain in the fingers, palm (including the primer grip), and thumb subdomains. Correspondingly, these regions show the most extensive and largest protection from exchange as shown in Fig. 5 and Table 1. Region 64–115 of p66 interacts with the template overhang of the WT aptamer, and along with region 147–188 forms a platform on which the template and primer terminal regions of the aptamer are situated. Region 212–231, which includes the primer grip, caps the regions of the fingers and palm and is strongly protected from exchange, as are residues 90–115. The portion of the thumb subdomain that fits into the minor groove of the aptamer also shows a dramatic reduction in exchange and includes helices H and I. Relatively fewer contacts between RT and the aptamer are made in the RNase H domain and connection subdomain, but a large reduction in exchange is observed for peptides of the RNase H domain, which is somewhat surprising since the aptamer is composed of 15 Watson-Crick base pairs and a loop of three thymidines that do not extend completely to the RNase H domain and only make four contacts with residues Y501, T473, N474, and Q475. No regions of the connection subdomain of p66 show net protection from exchange. This is partly due to fewer observed contacts and a concomitantly greater exposed surface area of the aptamer (Miller et al., 2015), and partly due to regions of RT not covered during the MS/MS peptide identification as described in Table 2. Four ionic contacts that were observed in the RT-aptamer crystal structure were not protected from exchange. Nor were surrounding residues (Table 2). Moreover, regions where no RT-DNA contacts are observed in the crystal structure are also strongly protected, such as residues spanning 536 to 555 in the RNase H domain at the C-terminus of p66 and residues in the p51 subunit for which no direct interactions are observed (132–160 and 408–425; Table 1). These observations can best be explained by invoking an allosteric effect of the aptamer on RT.

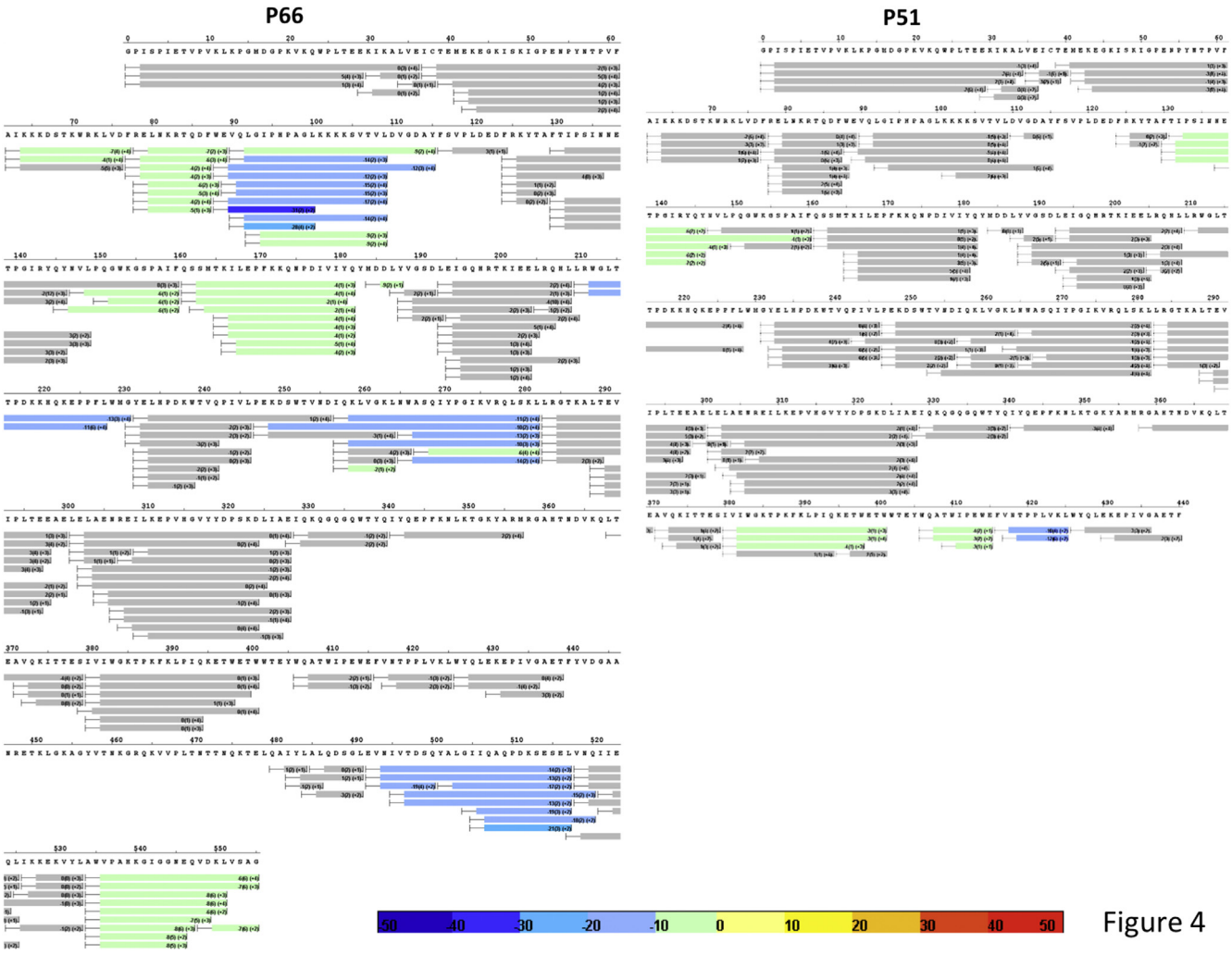

Fig. 4. Differential HDX of HIV-1 RT bound to the 38-mer aptamer. Bars represent peptides from the p66 (left) and p51 (right) subunits. Gray bars represent peptides where no net protection from exchange from D ê H was observed in the presence of the DNA aptamer. Colored bars represent differential protection from exchange. The color key at the bottom of the figure indicates the extent of protection. Each peptide is labeled with % exchange, standard deviation, and charge state (m/z).

3.3. HDX of HIV-1 reverse transcriptase bound to scrambled aptamer DNA

To determine whether the strong and widespread protection by the aptamer on the HDX of RT was due to its unusually high affinity, we constructed a 38-mer hairpin DNA with the 7-GC base pair sequence in the aptamer substituted with a random sequence (scrambled aptamer, Fig. 1). Since the scrambled aptamer had lower affinity for RT than the aptamer (28 nM vs.14 pM, respectively), we used a 1.5-fold molar excess of this DNA to ensure complete occupancy of RT without significantly disrupting the operation of the LC-MS. Fig. S1 and Table 3 show the peptides identified in the differential HDX experiment designed to map regions of RT that show a net change in exchange from D→H in the presence of the scrambled aptamer. As with the WT aptamer, no regions of RT exhibited increased exchange in the presence of the scrambled DNA. Surprisingly, we found little difference between the protection patterns observed for RT in the presence of the WT aptamer construct compared to that found in the presence of the scrambled construct, despite the much greater affinity of RT for the WT aptamer. For the scrambled aptamer we found one peptide in p66 (442–452), not identified by MS/MS for RT in the presence of the WT aptamer, which showed a decrease in exchange with the scrambled aptamer (Table 3). R448 of this peptide makes a direct interaction with the T-loop of the aptamer in the RT-aptamer crystal structure (Miller et al., 2015). Three regions of protection were found for RT in the presence of scrambled aptamer that were not protected from HDX for the RT-WT aptamer complex, although these peptides were identified by MS/MS during the HDX experiment for both complexes. Peptides 37–40 and 521–525 of p66 and 232–246 of p51 were protected from exchange by the scrambled aptamer (Table 3), but none of these regions is expected to make direct amino acid contact with DNA based on the RT-aptamer crystal structure (Miller et al., 2015).

The distribution of the extent and strength of the protections in the presence of these DNAs is instructive. The presence of the WT aptamer provides more protection to the fingers subdomain (p66) than does the presence of scrambled aptamer (compare Figs. 4 and 5 with Figs. S1, S2 and S3, and Table 3). Fingers region 79–115 is protected by both WT and scrambled aptamer, but p66 residues 64–77, which were identified in the MS/MS experiment for the scrambled aptamer, are not protected from HDX (see Table 3). This is consistent with interactions found in the RT-aptamer crystal structure that show specific effects of the 2′-O-methyl modifications of the aptamer interacting with Glu89 and a water-mediated contact with the backbone amide of Ile94 (Miller et al., 2015). The scrambled aptamer contains no 2′-O-methyl modifications and is not expected to make these interactions. The scrambled DNA, lacking these interactions, as well as the seven consecutive GC base pairs adjacent to the polymerase active site, may be free to slide within the nucleic acid binding cleft of RT, thereby weakening the interaction and the resulting protection by allowing for a greater exposure to solvent and exchange. Other differences with the scrambled aptamer are a somewhat

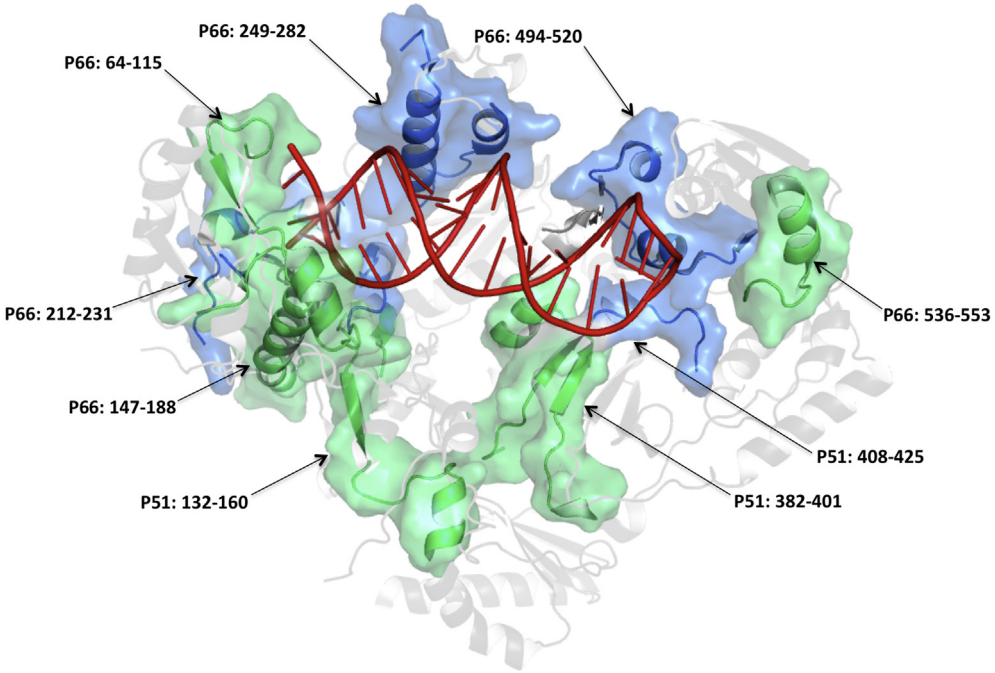

Fig. 5. Differential HDX of RT-aptamer – unliganded RT. Peptides shown in Fig. 4 and listed in Tables 1 and 3 mapped onto the 2.3 Å crystal structure (PBD 5D3G) to show regions protected from exchange in the presence of the aptamer DNA shown in red. Protected regions shown in green and blue retain 1–10% and 10%–20% deuterium compared to unliganded RT, respectively as shown in the key. RT regions with no net protection are shown in gray.

Table 1
Comparison of HDX protections and crystal structure contacts in the RT-aptamer complex.

Residues RT-aptamer	Contacts in RT-aptamer crystal structure	RT subdomain	Aptamer region
p66	p66		
64–115	V75, D76, R78, G89, Q91	Fingers/Palm	Template
	Y115	Palm	Primer
147–188	Q151, G152, Y183	Palm	Template
	Y183, D185	Palm	Primer
212–231	M230	Palm	Primer
249–282	N265, S280	Thumb	Template
	Q258, K259, G262, K263	Thumb	Primer
494–520	Y501	RNase H	Primer
536–553	None	RNase H	–
p51	p51		
132–160	None	Fingers	–
382–401	Q394, R395	Connection	Primer
408–425	None	Connection	–

Table 2
Crystal structure contacts not observed in HDX experiments with RT and hairpin DNAs.

Contact residues from crystal structure with no observed protection	K22 (p51) R284, K353, R356	p51 Fingers p66 Thumb	Primer
Contact residues from crystal structure missing from HDX experiment	A360, H361, T473, N474, Q475	p66 Connection RNase H	Template Template/ Loop

Table 3
RT peptides with differential protection in the presence of 38-mer hairpin DNA.

Peptides scrambled	Peptides aptamer	%Protection scrambled	%Protection aptamer
p66	p66		
37–40	***	1%–10%	0%
***	64–77	0%	1%–10%
79–89	79–89	1%–10%	1%–10%
90–115	90–115	10%–20%	10%–20%
149–160	147–160	1%–10%	1%–10%
163–183	163–182	1%–10%	1%–10%
184–187	186–188	10%–20%	1%–10%
212–231	212–231	1%–10%	10%–20%
249–264	249–282	1%–10%	10%–20%
265–282	265–282	1%–10%	10%–20%
442–452	–	1%–10%	–
494–520	494–520	10%–20%	10%–20%
521–525	***	1%–10%	0%
535–553	536–553	10%–20%	1%–10%
p51	p51		
126–160	132–160	1%–10%	1%–10%
232–246	***	1%–10%	0%
382–401	382–401	1%–10%	1%–10%
408–415	408–415	1%–10%	1%–10%
418–425	418–425	10%–20%	10%–20%

***Peptide present but not protected from HDX.
—Peptide missing from coverage.
Peptides marked 1%–10% protection are colored green in Fig. 4.
Peptides marked 10%–20% protection are colored blue in Fig. 4.
Peptides marked as 0% protection are colored gray in Fig. 4.

reduced strength of protection to the thumb of p66 and increased protection in the RNase H domain compared to the aptamer. However, these differences are small and the general pattern of protection of RT by these two DNA constructs is similar. Importantly, as with the aptamer, the scrambled DNA also results in protection from exchange in regions not in direct contact with DNA. Most notable, protection extends through the p51 subunit as shown in Table 3 and Figs. S1-S3.

3.4. SAXS analysis of aptamer DNA constructs

Previously we hypothesized that the high affinity with which the WT aptamer binds RT may be due, in part, to its ability to adopt a “pre-bent”

conformation (Miller et al., 2015). Given that HIV-1 RT imposes a bend characterized by a B- to A-form transition in the helical geometry of nucleic acids to which it binds (although dsRNA is likely to retain mostly A-form geometry when bound to RT), a pre-bent DNA would theoretically bind RT with more favorable energy as RT would not need to alter the conformation of the nucleic acid to fit its active site. To further test this hypothesis, we employed SAXS to investigate the structure of the unbound WT aptamer, an aptamer that lacks the 2′-O-methyl groups (Δmethyl aptamer), and the scrambled aptamer (Fig. 1). An extended aptamer containing a base-pair extension was also analyzed using SAXS to unambiguously determine the orientation of the WT aptamer SAXS envelope (Fig. 1, extended aptamer). All DNAs were re-folded using the same protocol used in the RT binding experiments and purified by SEC prior to SAXS data collection as described in Materials and Methods. Two independent SAXS datasets were obtained for each DNA construct. Analysis of the samples by native gel electrophoresis was performed concomitantly with SAXS data acquisition and confirmed that all aptamer constructs were present as a single homogeneous species in solution (data not shown).

Analysis of the Guinier and Kratky plots indicated that none of the samples were aggregated or experienced interparticle repulsions, and that all DNAs were well folded. The radius of gyration (R_g), which represents the mass distribution about a particle's center of gravity, was calculated from both the slope of the Guinier plot and the pairwise distance distribution (P(r)) function, a histogram of all inter-electron distances in the molecule (Figs. S4 and S5, and). For the WT aptamer, the R_g calculated by Guinier plot and P(r) function were found to be 19 Å and 20.1 Å, respectively. The R_g (Guinier) for both the Δmethyl and scrambled aptamers was determined to be 20.1 Å, while the R_g (P(r)) for the Δmethyl and scrambled was calculated to be 20.5 Å and 21.0 Å, respectively. The maximum electron-pair distance (D_{max}) as calculated from the P(r) distribution was found to be 73 Å, 78 Å, and 79 Å for the aptamer, Δmethyl, and scrambled constructs, respectively (Table S1). Taken together, analysis of the structural parameters derived from the SAXS data suggests that removal of the methyl groups from the aptamer and alteration of the GC tract sequence leads to a slight increase in the overall dimensions of the DNA. The extended aptamer was characterized by a Guinier-derived R_g of 27 Å and a P(r)-derived R_g of 30.7 Å, while the

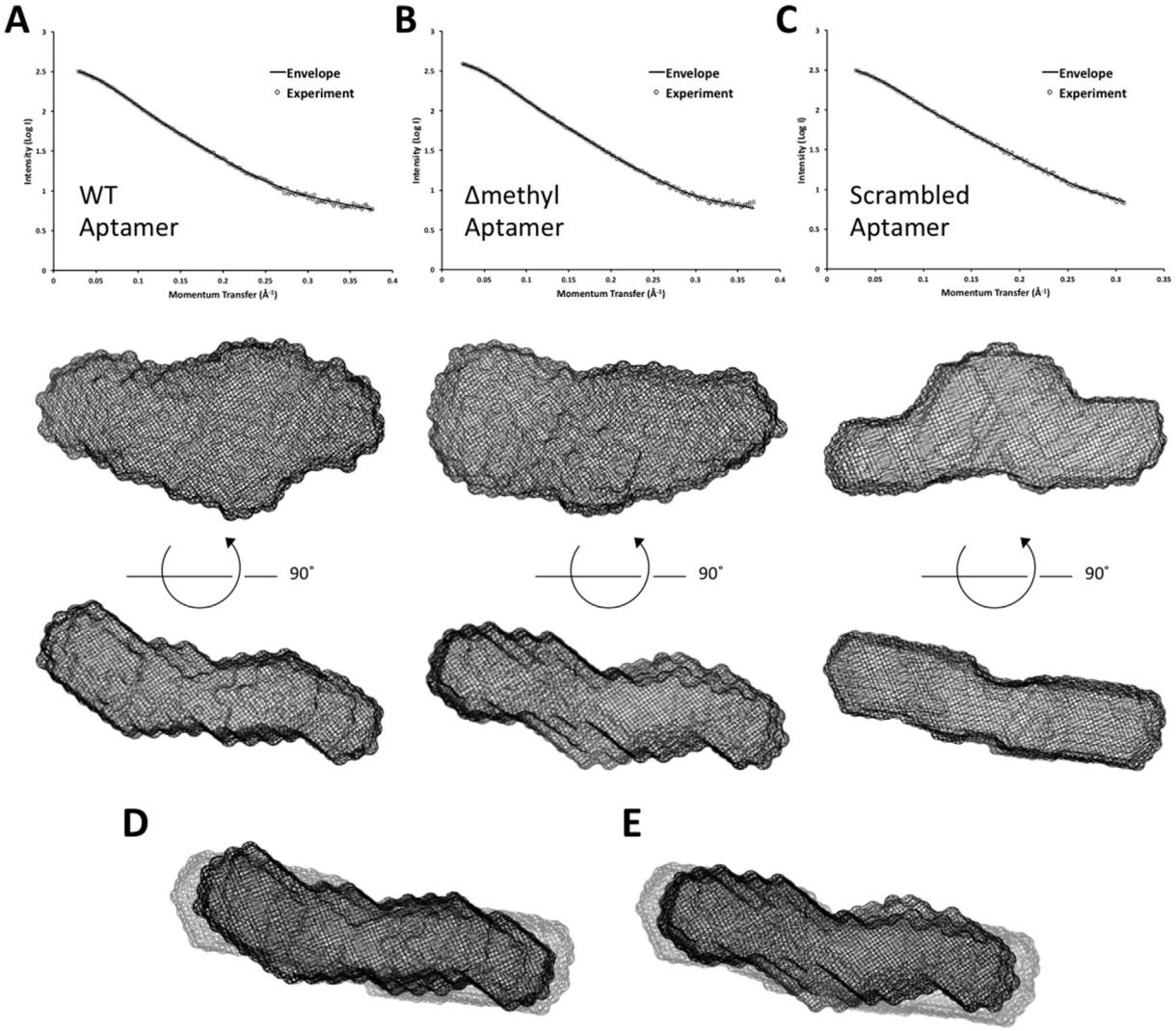

Fig. 6. SAXS data and *ab initio* models of the aptamer DNAs. Plot of intensity versus momentum transfer for the WT aptamer (A), Δmethyl aptamer (B), and scrambled aptamer (C) DNA constructs. The open circles indicate every fifth data point from the experimental SAXS curve and the black line represents the back-calculated scattering curve of the corresponding *ab initio* envelopes. Shown below each graph are two orientations of the *ab initio* envelopes of the respective aptamers. All aptamer envelopes are shown in the same orientation. (D) Comparison of the WT aptamer (black mesh) and scrambled aptamer (gray mesh) *ab initio* envelopes. (E) Comparison of the Δmethyl aptamer (black mesh) and scrambled aptamer (gray mesh) *ab initio* envelopes.

D_{max} was calculated to be 123 Å (Table S1). These values are consistent with the overall greater length of the extended aptamer relative to the other constructs examined.

3.5. SAXS reveals that aptamer adopts a pre-bent conformation

Ab initio envelopes for the WT aptamer, Δmethyl aptamer, and scrambled aptamer constructs were generated as described in Materials and Methods. After two rounds of model generation, the average χ^2 fit of each model to the scattering data was very good, and the normalized spatial discrepancy (NSD) between the 20 models was within bounds of converging on a reproducible structure (Fig. 6 and Table S1). NSD is a measure of the overall variance between 3D shapes (here, the *ab initio* envelopes) and values < 1 are not generally considered to be significantly different. While *ab initio* models were made from both SAXS datasets collected for each aptamer construct, the models that had the best χ^2 fit to the data are shown in Fig. 6; Table S1 shows related statistics regarding the fits. The final averaged *ab initio* envelopes for the Δmethyl and scrambled aptamers were each aligned to the WT aptamer envelope using the DAMSEL and DAMSUP programs. The aligned envelopes are all shown in the same orientation in Fig. 6.

Comparison of the *ab initio* envelopes reveals a kink in the WT aptamer and Δmethyl aptamer, while little to no kink is evident in the scrambled aptamer (Fig. 6). The all-atom model derived from the aptamer crystallized with RT (see Materials and Methods) was overlaid onto the WT aptamer envelope, revealing that the A-form segment of the aptamer (red) fits predominantly in the upward kink region of the envelope, while the B-form segment of the aptamer (black) fills in the rest of the envelope (Fig. S6). Taken together, the envelope structures and similarity to the RT-aptamer crystal all-atom model suggest that the aptamer DNA adopts a "pre-bent" conformation in the apo state, and that the presence of the 2′-O-methyl groups and GC tract influence the extent of bending.

To unambiguously define the orientation of the aptamer SAXS envelope, SAXS analysis of an extended aptamer was carried out and MONSA analysis was carried out as described in the materials and methods. There was good agreement between the 15 MONSA models of both aptamers, and the χ^2 fit and NSD parameters for the models are listed in Table S2. The average of these MONSA models are shown in Fig. 7. The red arrows indicate the locations in the envelope where a distinct bend occurs on the same side of the structures. We interpret the side with the bend (left side of the figure) to correspond to the shared nucleotides of the two DNAs, while the right side of the red envelope (extended aptamer) corresponds to the extended hairpin region specific to that DNA. In conclusion, these data confirm the orientation of the envelopes and are consistent with an aptamer DNA sequence that maintains a pre-bent state in solution.

3.6. CD spectroscopy consistent with aptamer A-form helical character

CD spectroscopic analysis of the aptamer, Δmethyl aptamer, and scrambled aptamer DNA constructs was performed to measure and compare their helicity. The scrambled aptamer displays two positive peaks centered at ~220 nm and ~280 nm, and two negative peaks centered around ~255 nm and ~210 nm (Fig. S7). This pattern of peaks is consistent with a CD spectrum of purely B-form DNA (Kypr et al., 2009). The aptamer and Δmethyl aptamers both yielded highly similar CD spectra, with positive peaks around ~220 nm and ~270 nm, and a minor shoulder around ~280 nm; negative peaks were observed around ~210 nm and ~240 nm (Fig. S7). For the aptamer and Δmethyl aptamers, the shifted positive peak from ~280 to ~270 nm and shifted negative peak from ~255 to ~240 nm indicate a transition towards A-form DNA helical geometry (Kypr et al., 2009). The fact that the aptamer and Δmethyl aptamer CD spectra are not fully consistent with A-form DNA (i.e., positive peak ~265 nm and negative peak ~230 nm, see reference (Kypr et al., 2009)), suggests that these DNAs possess elements of both A- and B-form character. Thus, CD data support the conclusion the scrambled aptamer DNA adopts a predominantly B-form conformation, while the aptamer and Δmethyl aptamers possess characteristics of both A-form and B-form helices.

3.7. X-ray crystal structure of the RT-RNA/DNA dU_3 aptamer complex

The RT-RNA/DNA dU_3 aptamer structure (Fig. 1) was solved and refined to 2.8 Å resolution. Data collection and refinement statistics are shown in Table S2. The crystal asymmetric unit contains two copies of the RT-RNA/DNA dU_3 aptamer complex, with a large nucleic acid binding track along the length of the heterodimer (Fig. 8). Overall, it is similar to the previously reported RT-DNA aptamer structure (Fig. 9) (Miller et al., 2015). However, the K_D value for the RNA/DNA dU_3 aptamer for RT (542 pM ± 59, is 40-fold higher than that of the WT aptamer.

The RNA/DNA dU_3 aptamer shows electron density for the duplex region in both copies of the asymmetric unit for 35 of the 38 nucleotides (Fig. 8B). Conformational analysis of the RNA/DNA aptamer, using the 3DNA software package DSSR (see all output files in Supplementary Information) (Hanson and Lu, 2017), indicates that its conformation is similar to the DNA aptamer for only the first five base pairs of the duplex,

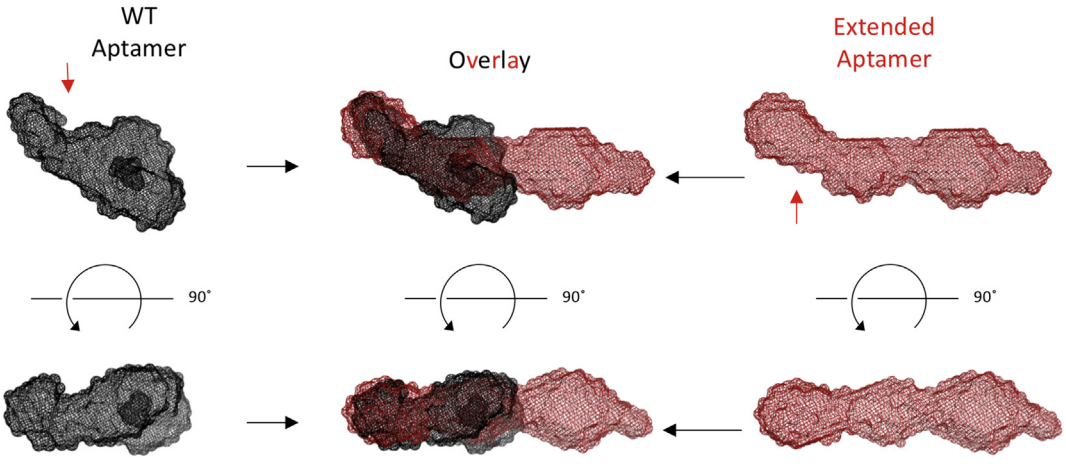

Fig. 7. Superposition of the SAXS envelopes generated in MONSA. MONSA-generated SAXS envelope of the WT aptamer DNA construct (black mesh) and the extended aptamer DNA construct (red mesh) are shown on the left and right, respectively. The MONSA-derived overlay of these two SAXS envelopes is shown in the middle, in the same orientations as the structures by themselves on the left and right. The red arrow indicates where the apparent bend is occurring in these envelopes. This overlay suggests that a pre-formed bend exists in these DNA constructs at the base of hairpin with the 5′overhang.

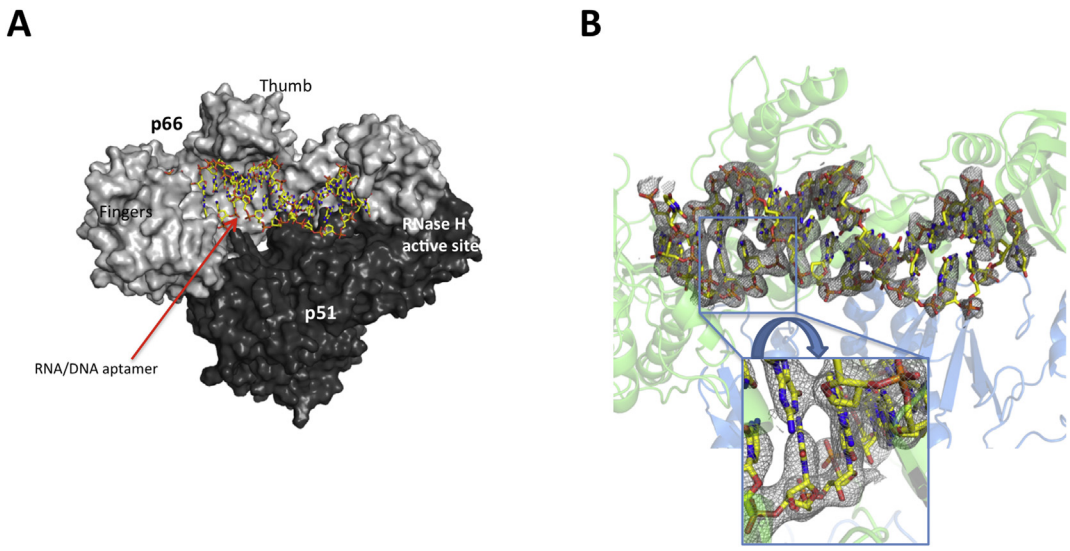

Fig. 8. Overview of HIV-1 RT-RNA/DNA aptamer dU_3 X-ray structure. (A) Overview of RT-RNA/DNA aptamer complex with HIV-1 RT p66/p51 heterodimer (light gray/dark gray) shown with the aptamer colored in yellow sticks. (B) mFo-DFc electron density map of the RNA/DNA aptamer at 2σ contour. Inset: representative electron density for a portion of the RNA template strand, including several 2′-OH groups.

which adopt an A-like conformation. The remaining base pairs of the DNA aptamer show B-like geometry while the RNA/DNA aptamer displays A-like conformation throughout the entire duplex region (they are not assigned to either A or B type, but they have C3′-endo puckering typical for an A-like conformation); nucleotides dC20, dG25, and dA26 have B-like conformation. This almost "all A-form" conformation can also be observed for the nucleic acid bound to RT in the highest resolution RT-RNA/DNA complex published to date (PDB ID 4PQU) (Das et al., 2014).

The main difference between the RT-RNA/DNA aptamer and the RT-aptamer structures is in the hairpin-loop region, between nucleotides 14 and 19 and in the vicinity of the RNase H primer grip (Fig. 9B and D). Differences in the phosphate-to-phosphate distances of nucleotides 17 and 16 range from 2.0 to 3.2 Å and deviations in the C1′-C1′ distances for these same nucleotides range from 2.6 to 4.3 Å, as shown in Fig. 9B.

In terms of flexibility, which can be assessed by analyzing B-factors, the RNA/DNA aptamer is much more mobile than the DNA aptamer, highlighted by the hairpin region (Fig. 9C). Given the difference in resolution between the two structures (2.8 Å vs. 2.3 Å), we plotted the normalized B-factors (residue B-factor normalized per aptamer mean B-factor, calculated with the Phenix structure comparison GUI, https://www.phenix-online.org/version_docs/1.8-1069/structure_comparison.html) for both aptamers (Fig. S8). The most significant differences are in: i) nucleotide 16 (the first hairpin-loop nucleotide) and nucleotides 19–22 (primer nucleotides in the vicinity of the hairpin), with higher values for the RNA/DNA aptamer; and ii) nucleotide 18 (the last of the hairpin-loop region), with lower values for the RNA/DNA aptamer. Thus, the relative distribution of B-factors after normalization supports the notion that the relative mobilities of the aptamers in the hairpin region are different.

4. Discussion

We have used HDX to probe the structural dynamics of RT in the presence of two different DNAs (WT aptamer and scrambled aptamer) in a template-primer configuration. The 2′-O-methyl-modified aptamer, which also contains a 7-base pair GC tract, was previously used to solve a high-resolution crystal structure of an RT-aptamer complex that revealed detailed information on the contacts made with RT and gave some insight into the elements of this DNA that provided high affinity, especially the contributions of the 2′-O-methyl modifications located on the template strand of the aptamer (Miller et al., 2015). However, these modifications accounted for only an approximately 10-fold improvement in binding affinity over the unmodified aptamer. Additional mutation of the GC tract to a scrambled sequence (scrambled aptamer) drastically reduces the binding affinity further (~400-fold). Surprisingly, there was very little difference between the exchange protection observed with the aptamer and the scrambled aptamer. Thus, no large conformational changes occurred that could explain the increased affinity in terms of structural contributions from RT.

Another surprising finding from this work is the extent to which either DNA induced long-range changes in the structural dynamics of RT far from sites of direct interaction with DNA. This was especially notable for changes in the p51 subunit. In Fig. S9, an extensive network of interactions among regions with altered exchange kinetics can be observed that traverses almost the entire surface of RT. Only a gap in protected regions between the p66 thumb and RNase H domain breaks a nearly continuous network of interconnected regions protected from D→H exchange in RT. Example regions showing protection that are not in the direct vicinity of bound aptamer include p51 residues 132–169 (fingers) and p66 residues 535–553 (RNase H). These observations are consistent with our understanding that RT is a dynamic enzyme that undergoes large conformational changes associated with binding dNTP substrate and NNRTI inhibitors (Xavier Ruiz, 2020). Our current HDX study shows how RT "tectonics" extends throughout both subunits. A recent study on the sortilin receptor using HDX to probe receptor-ligand interactions showed that interaction of sortilin with its protein ligand progranulin stabilized a membrane-proximal region (10CC) distant from the PGRN binding site which the authors described as allosteric stabilization (Trabjerg et al., 2019).

Given the striking similarity with which the aptamer and scrambled DNAs stabilize RT dynamics, we ruled out the possibility that the high affinity of the aptamer was due to additional stabilization that this DNA imparts to RT *after* binding. Since all nucleic acid template-primers bound to RT in crystal structures show a bent conformation (Hu and Hughes, 2012), it is likely that the straight-to-bent transition represents an energy barrier that the nucleic acid must overcome to bind RT. The SAXS envelope of the aptamer construct revealed that it adopts a pre-bent conformation in solution; thus, it possesses a lower energetic barrier to binding RT, most likely leading to the DNA's extremely high binding affinity. The SAXS envelopes of the mutant aptamers (Δmethyl and scrambled) show that removal of the GC tract leads to the greatest

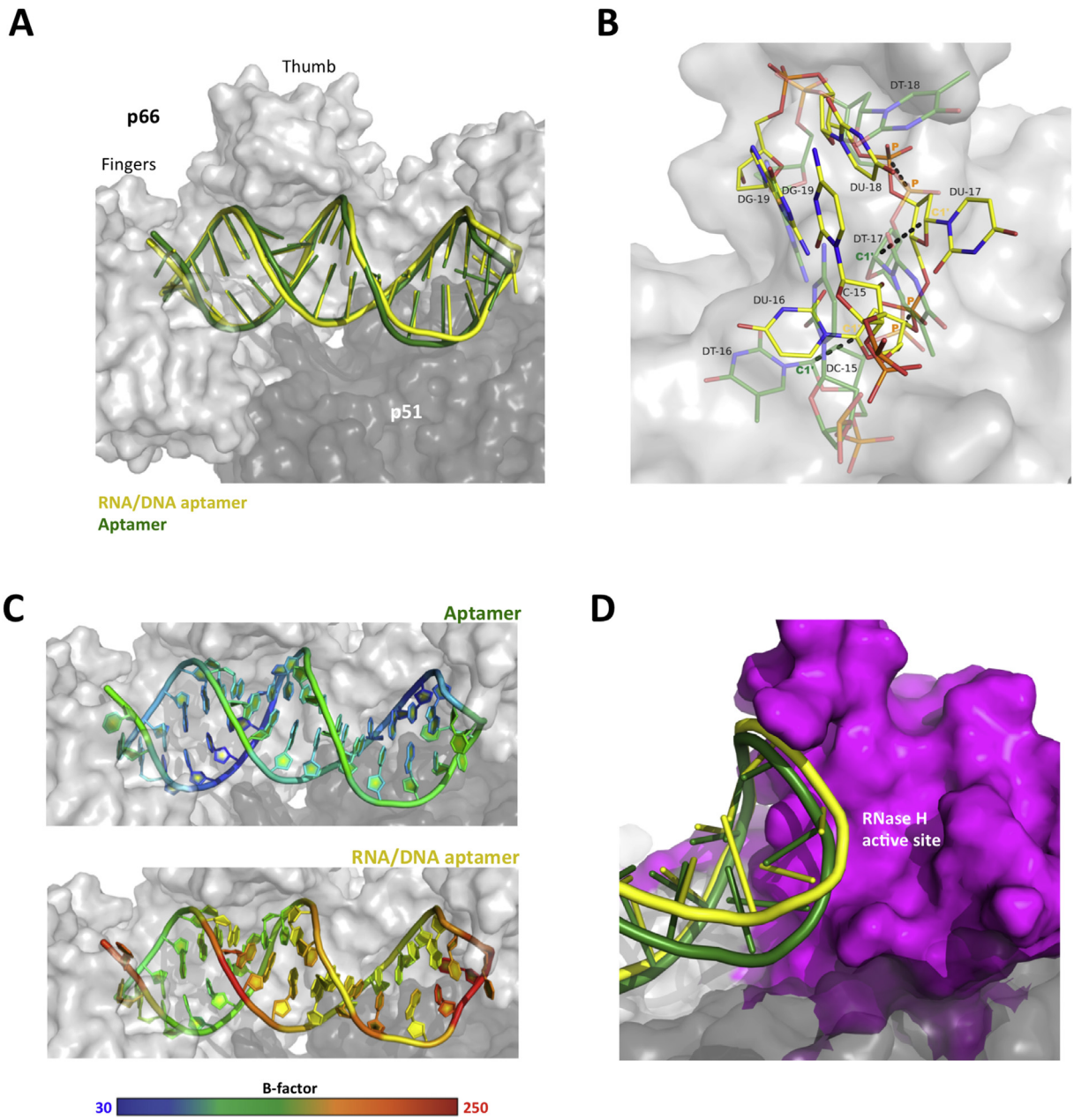

Fig. 9. Comparison of RT-DNA aptamer and RT-RNA/DNA aptamer X-ray structures. (A) Superimposition of both structures, with HIV-1 RT p66/p51 heterodimer (light gray/dark gray) and the aptamers (in nucleic acid cartoon) colored as indicated. (B) Zoom of the former superimposition displaying nucleotides 15 to 19 of both aptamers (same coloring); displacement between nucleotides 16 and 17 of both aptamers displayed through dashed lines between their P atoms and C1′ atoms (atoms labeled, distances in the main text). (C) Region of the nucleic acid binding cleft, with RT-DNA aptamer (top) and RT-RNA/DNA aptamer (down) colored according to the average B-factor per nucleotide (in $Å^2$, see legend on the bottom). (D) Model of the hairpin loop region near the RNase H active site, with RT represented in surface mode, highlighting the RNase H active site colored in magenta.

reduction in the pronounced kink in the envelope. This is consistent with the observation that extended GC tracts in DNA can induce a transition from B-to-A form geometry (Ng et al., 2000). It is also known that 2′-O-methyl-modified nucleotides in the context of DNA can induce A-form sugar pucker and geometry (Lubini et al., 1994). The CD spectra show that both the WT aptamer and Δmethyl aptamer possess features of both A- and B-form helical geometry, which would be expected from a "pre-bent" aptamer, whereas the scrambled aptamer adopts a purely B-form conformation, as expected for a canonical double-stranded DNA helix. Overall, the SAXS and CD data help to explain the RT binding data and suggest that the enhanced aptamer affinity for RT is due to a pre-bent conformation of the apo aptamer.

Additionally, the X-ray structure of the RT- RNA/DNA aptamer complex has allowed us to isolate the role of the hairpin in binding. The RNA/DNA aptamer affinity drops 40-fold relative to the WT aptamer, most likely stemming from the distortion and displacement of the hairpin loop in the former. The B-factor comparison between the nucleotides in both structures also points to a more flexible hairpin of the RNA/DNA aptamer. Notably, dT16 in the aptamer provides intra-loop interactions (presumably lost in the RNA/DNA aptamer, Fig. 9), whose strength has been observed to increase dramatically for small loops (Kuznetsov et al., 2008). This might partly account for the difference in affinity for both aptamers. Additionally, as the number of RT-nucleic acid interactions in the two structures are very similar (data not shown), the difference in the

conformation of both nucleic acids after the polymerase active site might also be involved, most likely due to different hairpin positioning. The improved shape complementarity for the DNA aptamer hairpin *vs.* the RNA/DNA aptamer hairpin (Fig. 9) could also contribute to stronger recognition of the former (Rohs et al., 2010) and might explain the surprisingly increased protection found in the RNase H domain despite few contacts as shown with HDX of the RT-aptamer complex. Notably, this property of the DNA aptamer has yielded a stable RT-nucleic acid N-site structure with the pyrophosphate mimic foscarnet, without the need for chemical cross-linking or an antibody Fab fragment, as the hairpin impedes the nucleic acid translocation *in crystallo* because of the abrogation of its sliding (Miller et al., 2015).

Using an integrative structural biology approach, we have been able to dissect the unique properties of the DNA aptamer, which possesses a pre-formed bend, partial A-form geometry, and a hairpin loop that tighten and stabilize its binding to HIV-1 RT.

Funding

This work was supported by National Institutes of Health (NIH) grants U54 AI150472 (to EA, PG, and KM-F), R37 MERIT Award AI027960 (to EA), and R01 AI150493 (to KM-F). EDO was supported byNIH grants T32 GM008512 and F31 AI120868. The SAXS measurements were performed at the Advanced Light Source (ALS), a national user facility operated by Lawrence Berkeley National Laboratory. The ALS is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy (DOE) under Contract No. DE-AC02-05CH11231, and through the Integrated Diffraction Analysis Technologies (IDAT) program, supported by the DOE Office of Biological and Environmental Research. Additional support comes from the NIH project MINOS (R01 GM105404). We are grateful for beamtime access from the Stanford Synchrotron Research Laboratory.

Authors contribution

S.T., J.Z., P.R.G., and E.A. conceived the project; S.T. and A.H. engineered and produced protein samples; J.Z. carried out HDX-MS experiments with support from B.D.P. on software and data curation; E.D.O. carried out SAXS and CD studies; F.X.R. carried out X-ray crystallography; S.T., J.Z., E.D.O. F.X.R, K.M.-F. and E.A. analyzed data; J.J.D., K.M.-F., P.R.G., and E.A. provided supervision and financial support; S.T. and E.A. wrote and edited the manuscript with contributions from all authors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We acknowledge the staff at the Stanford Synchrotron Radiation Lightsource (SSRL), and also Dr. Matthew Miller, Dr. Sergio Martinez, and Jerry Joe Harrison for assistance and helpful discussions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crstbi.2020.06.002.

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66, 213–221.

Bauman, J.D., Das, K., Ho, W.C., Baweja, M., Himmel, D.M., Clark Jr., A.D., Oren, D.A., Boyer, P.L., Hughes, S.H., Shatkin, A.J., et al., 2008. Crystal engineering of HIV-1 reverse transcriptase for structure-based drug design. Nucleic Acids Res. 36, 5083–5092.

Bauman, J.D., Patel, D., Dharia, C., Fromer, M.W., Ahmed, S., Frenkel, Y., Vijayan, R.S., Eck, J.T., Ho, W.C., Das, K., et al., 2013. Detecting allosteric sites of HIV-1 reverse transcriptase by X-ray crystallographic fragment screening. J. Med. Chem. 56, 2738–2746.

Bernardes, A., Batista, F.A., 2012. de Oliveira Neto M, Figueira AC, Webb P, Saidemberg D, Palma MS, Polikarpov I: low-resolution molecular models reveal the oligomeric state of the PPAR and the conformational organization of its domains in solution. PloS One 7, e31852.

Boyer, P.L., Imamichi, T., Sarafianos, S.G., Arnold, E., Hughes, S.H., 2004. Effects of the Delta67 complex of mutations in human immunodeficiency virus type 1 reverse transcriptase on nucleoside analog excision. J. Virol. 78, 9987–9997.

Chalmers, M.J., Busby, S.A., Pascal, B.D., He, Y., Hendrickson, C.L., Marshall, A.G., Griffin, P.R., 2006. Probing protein ligand interactions by automated hydrogen/deuterium exchange mass spectrometry. Anal. Chem. 78, 1005–1014.

Das, K., Arnold, E., 2013. HIV-1 reverse transcriptase and antiviral drug resistance. Part 1. Curr Opin Virol 3, 111–118.

Das, K., Arnold, E., 2013. HIV-1 reverse transcriptase and antiviral drug resistance. Part 2. Curr Opin Virol 3, 119–128.

Das, K., Martinez, S.E., Bauman, J.D., Arnold, E., 2012. HIV-1 reverse transcriptase complex with DNA and nevirapine reveals non-nucleoside inhibition mechanism. Nat. Struct. Mol. Biol. 19, 253–259.

Das, K., Martinez, S.E., Bandwar, R.P., 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. 10.

Das, K., Balzarini, J., Miller, M.T., Maguire, A.R., DeStefano, J.J., Arnold, E., 2016. Conformational states of HIV-1 reverse transcriptase for nucleotide incorporation vs pyrophosphorolysis-binding of foscarnet. ACS Chem. Biol. 11, 2158–2164.

DeLano, W.L., 2002. The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA.

DeStefano, J.J., Nair, G.R., 2008. Novel aptamer inhibitors of human immunodeficiency virus reverse transcriptase. Oligonucleotides 18, 133–144.

Dyer, K.N., Hammel, M., Rambo, R.P., Tsutakawa, S.E., Rodic, I., Classen, S., Tainer, J.A., Hura, G.L., 2014. High-throughput SAXS for the characterization of biomolecules in solution: a practical approach. Methods Mol. Biol. 1091, 245–258.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60, 2126–2132.

Gallagher, E.S., Hudgens, J.W., 2016. Mapping protein-ligand interactions with proteolytic fragmentation, hydrogen/deuterium exchange-mass spectrometry. Methods Enzymol. 566, 357–404.

Goswami, D., Callaway, C., Pascal, B.D., Kumar, R., Edwards, D.P., Griffin, P.R., 2014. Influence of domain interactions on conformational mobility of the progesterone receptor detected by hydrogen/deuterium exchange mass spectrometry. Structure 22, 961–973.

Goswami, D., Tuske, S., Pascal, B.D., Bauman, J.D., Patel, D., Arnold, E., Griffin, P.R., 2015. Differential isotopic enrichment to facilitate characterization of asymmetric multimeric proteins using hydrogen/deuterium exchange mass spectrometry. Anal. Chem. 87, 4015–4022.

Hanson, R.M., Lu, X.J., 2017. DSSR-enhanced visualization of nucleic acid structures in Jmol. Nucleic Acids Res. 10.

Hsiou, Y., Ding, J., Das, K., Clark Jr., A.D., Hughes, S.H., Arnold, E., 1996. Structure of unliganded HIV-1 reverse transcriptase at 2.7 A resolution: implications of conformational changes for polymerization and inhibition mechanisms. Structure 4, 853–860.

Hu, W.S., Hughes, S.H., 2012. HIV-1 reverse transcription. Cold Spring Harb Perspect Med 2.

Huang, H., Chopra, R., Verdine, G.L., Harrison, S.C., 1998. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science 282, 1669–1675.

Hura, G.L., Menon, A.L., Hammel, M., Rambo, R.P., Poole 2nd, F.L., Tsutakawa, S.E., Jenney Jr., F.E., Classen, S., Frankel, K.A., Hopkins, R.C., et al., 2009. Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). Nat. Methods 6, 606–612.

Jacobo-Molina, A., Ding, J., Nanni, R.G., Clark Jr., A.D., Lu, X., Tantillo, C., Williams, R.L., Kamer, G., Ferris, A.L., Clark, P., et al., 1993. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. Proc. Natl. Acad. Sci. U. S. A. 90, 6320–6324.

Jones, C.P., Cantara, W.A., Olson, E.D., Musier-Forsyth, K., 2014. Small-angle X-ray scattering-derived structure of the HIV-1 5′ UTR reveals 3D tRNA mimicry. Proc. Natl. Acad. Sci. U. S. A. 10.

Keating, K.S., Pyle, A.M., 2012. RCrane: semi-automated RNA model building. Acta Crystallogr D Biol Crystallogr 68, 985–995.

Keppel, T.R., Weis, D.D., 2015. Mapping residual structure in intrinsically disordered proteins at residue resolution using millisecond hydrogen/deuterium exchange and residue averaging. J. Am. Soc. Mass Spectrom. 26, 547–554.

Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A., Steitz, T.A., 1992. Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science 256, 1783–1790.

Konarev, P.V., Petoukhov, M.V., Volkov, V.V., Svergun, D.I., 2006. ATSAS 2.1, a program package for small-angle scattering data analysis. J. Appl. Crystallogr. 39, 277–286.

Kuznetsov, S.V., Ren, C.C., Woodson, S.A., Ansari, A., 2008. Loop dependence of the stability and dynamics of nucleic acid hairpins. Nucleic Acids Res. 36, 1098–1112.

Kypr, J., Kejnovska, I., Renciuk, D., Vorlickova, M., 2009. Circular dichroism and conformational polymorphism of DNA. Nucleic Acids Res. 37, 1713–1725.
Lubini, P., Zurcher, W., Egli, M., 1994. Stabilizing effects of the RNA 2′-substituent: crystal structure of an oligodeoxynucleotide duplex containing 2′-O-methylated adenosines. Chem. Biol. 1, 39–45.
Miller, M.T., Tuske, S., Das, K., DeStefano, J.J., Arnold, E., 2015. Structure of HIV-1 reverse transcriptase bound to a novel 38-mer hairpin template-primer DNA aptamer. Protein Sci. 10.
Morton, V.L., Burkitt, W., O'Connor, G., Stonehouse, N.J., Stockley, P.G., Ashcroft, A.E., 2010. RNA-induced conformational changes in a viral coat protein studied by hydrogen/deuterium exchange mass spectrometry. Phys. Chem. Chem. Phys. 12, 13468–13475.
Ng, H.L., Kopka, M.L., Dickerson, R.E., 2000. The structure of a stable intermediate in the A <–> B DNA helix transition. Proc. Natl. Acad. Sci. U. S. A. 97, 2035–2039.
Nguyen, P.D.M., Zheng, J., Gremminger, T.J., Qiu, L., Zhang, D., Tuske, S., Lang, M.J., Griffin, P.R., Arnold, E., Chen, S.-J., Zou, X., Heng, X., Burke, D.H., 2020. Binding interface and impact on protease cleavage for an RNA aptamer to HIV-1 reverse transcriptase. Nucleic Acids Res. 48 (5), 2709–2722. https://doi.org/10.1093/nar/gkz1224.
Olimpo, J.T., DeStefano, J.J., 2010. Duplex structural differences and not 2′-hydroxyls explain the more stable binding of HIV-reverse transcriptase to RNA-DNA versus DNA-DNA. Nucleic Acids Res. 38, 4426–4435.
Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326.
Pascal, B.D., Willis, S., Lauer, J.L., Landgraf, R.R., West, G.M., Marciano, D., Novick, S., Goswami, D., Chalmers, M.J., Griffin, P.R., 2012. HDX workbench: software for the analysis of H/D exchange MS data. J. Am. Soc. Mass Spectrom. 23, 1512–1521.
Post, K., Olson, E.D., Naufer, M.N., Gorelick, R.J., Rouzina, I., Williams, M.C., Musier-Forsyth, K., Levin, J.G., 2016. Mechanistic differences between HIV-1 and SIV nucleocapsid proteins and cross-species HIV-1 genomic RNA recognition. Retrovirology 13, 89.
Putnam, C.D., Hammel, M., Hura, G.L., Tainer, J.A., 2007. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Q. Rev. Biophys. 40, 191–285.
Rambo, R.P., 2015. Resolving individual components in protein-RNA complexes using small-angle X-ray scattering experiments. Methods Enzymol. 558, 363–390.
Rodgers, D.W., Gamblin, S.J., Harris, B.A., Ray, S., Culp, J.S., Hellmig, B., Woolf, D.J., Debouck, C., Harrison, S.C., 1995. The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. Proc. Natl. Acad. Sci. U. S. A. 92, 1222–1226.
Rohs, R., Jin, X., West, S.M., Joshi, R., Honig, B., Mann, R.S., 2010. Origins of specificity in protein-DNA recognition. Annu. Rev. Biochem. 79, 233–269.
Sarafianos, S.G., Das, K., Ding, J., Boyer, P.L., Hughes, S.H., Arnold, E., 1999. Touching the heart of HIV-1 drug resistance: the fingers close down on the dNTP at the polymerase active site. Chem. Biol. 6, R137–R146.
Sarafianos, S.G., Clark Jr., A.D., Tuske, S., Squire, C.J., Das, K., Sheng, D., Ilankumaran, P., Ramesha, A.R., Kroth, H., Sayer, J.M., et al., 2003. Trapping HIV-1 reverse transcriptase before and after translocation on DNA. J. Biol. Chem. 278, 16280–16288.
Sarafianos, S.G., Das, K., Hughes, S.H., Arnold, E., 2004. Taking aim at a moving target: designing drugs to inhibit drug-resistant HIV-1 reverse transcriptases. Curr. Opin. Struct. Biol. 14, 716–730.
Sarafianos, S.G., Marchand, B., Das, K., Himmel, D.M., Parniak, M.A., Hughes, S.H., Arnold, E., 2009. Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. J. Mol. Biol. 385, 693–713.
Seckler, J.M., Howard, K.J., Barkley, M.D., Wintrode, P.L., 2009. Solution structural dynamics of HIV-1 reverse transcriptase heterodimer. Biochemistry 48, 7646–7655.
Svergun, D.I., 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys. J. 76, 2879–2886.
Trabjerg, E., Abu-Asad, N., Wan, Z., Kartberg, F., Christensen, S., Rand, K.D., 2019. Investigating the conformational response of the sortilin receptor upon binding endogenous peptide- and protein ligands by HDX-MS. Structure 27, 1103–1113.
Tuske, S., Sarafianos, S.G., Clark Jr., A.D., Ding, J., Naeger, L.K., White, K.L., Miller, M.D., Gibbs, C.S., Boyer, P.L., Clark, P., et al., 2004. Structures of HIV-1 RT-DNA complexes before and after incorporation of the anti-AIDS drug tenofovir. Nat. Struct. Mol. Biol. (11), 469–474. Epub 2004 Apr 2025.
Xavier Ruiz, F., 2020. Arnold E: evolving understanding of HIV-1 reverse transcriptase structure, function, inhibition, and resistance. Curr. Opin. Struct. Biol. 61, 113–123.
Zheng, J., Yong, H.Y., Panutdaporn, N., Liu, C., Tang, K., Luo, D., 2015. High-resolution HDX-MS reveals distinct mechanisms of RNA recognition and activation by RIG-I and MDA5. Nucleic Acids Res. 43, 1216–1230.