Supplementary Appendix

Title: Human endogenous retrovirus-K (HERV-K) reverse transcriptase (RT) structure and biochemistry reveals remarkable similarities to HIV-1 RT and opportunities for HERV-K specific inhibition.

Authors: Eric T. Baldwin*, Matthias Götte, Egor P. Tchesnokov, Eddy Arnold, Margit Hagel, Charles Nichols, Pam Dossang, Marieke Lamers, Paul Wan, Stefan Steinbacher, and Donna L. Romero

*corresponding author, Eric T. Baldwin ROME Therapeutics, Boston MA, USA. ebaladin@rometx.com, 267-799-6585

This PDF file includes:

- Supplementary text
- Figure S1 to S15
- Table S1
Supplementary Information Text

Choice of DNA Substrate
The distance between HERV-K RNase H and polymerase active sites informed the range of optimal lengths of primer/template for use in our crystallization studies. Based upon the HIV-1 literature, best crystallization results have been obtained when oligos were used that spanned from the polymerase to the RNase H active sites. The sequence of the template strand was based upon the HIV-1 tRNA primer-binding site (U5-PBS from the HIV-1 reference genome, NC_001802.1) and incorporated into the final set of sequences that were used for crystallization (1). Biochemical assays were used to confirm that the sequences were good substrates for HERV-K RT. Based upon previous experience from HIV-1 RT crystallization work, we chose a template overhang of 2-4 bases.

Detailed Structural Comparison to HIV-1 RT
Beginning at the N-terminus of molA, the fingers subdomain provides important contacts to the dNTP and then (as in HIV-1) passes through the palm subdomain before returning to the fingers with secondary structural elements that recapitulate HIV-1 (Fig. 3B). Identical fingers subdomain residues between HIV and HERV-K are indicated in SI Appendix, Fig. S10. The HERV-K molA polypeptide chain returns to the palm subdomain and forms the conserved $\alpha_E$ helix and then on to the characteristic $\beta_9$ and $\beta_{10}$ and the catalytic YIDD motif. After the $\alpha_F$ helix HERV-K has a single-turn helix $\alpha_{F1}$ and the palm subdomain ends with a $\beta$-sheet similar to HIV-1 (SI Appendix, Fig. S8). The thumb subdomain of HERV-K (like HIV-1) includes three prominent $\alpha$ helices which grip the primer strand via the $\alpha_H$ helix (SI Appendix, Fig. S11D). There are a couple of important differences with HIV-1 RT of note as the molA chain passes through the connection subdomain. The initial $\beta$-sheet $\beta_{16}$, $\beta_{17}$, $\beta_{18}$ of HIV-1 is preserved but is followed by a short strand $\beta_{18a}$ that H-bonds to R$\beta_{4a}$ from the RNase H domain, creating a unique interaction between the connection subdomain and RNase H domain (Fig. 4a). In molA this new strand is followed by $\alpha_K$ and $\alpha_L$ as in HIV-1 (SI Appendix, Fig. S11C); whereas HERV-K presents an additional helix ($\alpha_{L1}$) similar to MoMLV RT (Fig. 4B, left) (main text ref 64, 49) (2, 3). The secondary structure of the RNase H domain is similar to that of HIV-1 (SI Appendix, Fig. S11D) with the exception of the Rs4a strand.

The fingers subdomain of molB of HERV-K (SI Appendix, Fig. S5B & Fig. 3C) is very similar to the p51 fingers of HIV-1 with the exception of a short terminal helix ($\alpha_{D1}$, SI Appendix, Fig. S12A). While the isolated molB palm superimposes well on the isolated p51 palm of HIV-1, in the context of the full molB/p51 subunit there is a significant rotation of the palm when compared to HIV-1. The molB chain continues with the palm subdomain with the expected secondary structure through $\alpha_F$. $\alpha_{F1}$ of HERV-K RT is also present in molB and is followed by a rearrangement of the molA $\beta$-sheet ($\beta_{12}$ thru $\beta_{14}$) to present $\beta_{11}$ and helix $\alpha_{F2}$, which is a prominent and unique feature of HERV-K molB (SI Appendix, Fig. S12B). The molB polypeptide chain continues into the thumb subdomain which contacts molA RNase H as it does in HIV-1 (SI Appendix, Fig. S12C). A couple of unique features of the connection domain of HERV-K RT molB are noteworthy. The molA $\beta_{18a}$ strand is an ordered coil and does not form an interaction with molB RNase H (which is not observed in this structure). However, a helix ($\alpha_{K0}$) precedes the dominant $\alpha_k$ helix of the connection domain and is similarly positioned to the equivalent helix of MoMLV RT (Fig. 4B, right). The $\alpha_{F1}$ helix of molA is retained in HERV-K molB so that two features of MoMLV RT connection are present in molB of HERV-K (SI Appendix, Fig. S12D). The HERV-K molB connection domain terminates with a unique helix ($\alpha_{L2}$) which makes hydrophobic interactions with the molB thumb.
subdomain (Fig. 4C). This terminal helical interaction with the thumb is very similar to the yeast Ty3 RT (main text ref 54) (4) where a helix from the RNase H domain of the p51-like subunit contacts the thumb (Fig. 4D). The structure-based sequence alignment of HIV-1 and HERV-K RT is shown in SI Appendix, Fig. S13.

The comparison of the asymmetric dimer interfaces of HIV-1 and HERV-K (SI Appendix, Fig. S13) reveal a very similar pattern of contact points without an apparent enrichment in strict conservation. The interface is defined by important contacts between the molA and molB polymerase domains, between the molA and molB connection subdomains, and the molB thumb subdomain with the RNase H domain (SI Appendix, Fig. S14B and C). While there are abundant conserved interactions at the polymerase (SI Appendix, Fig. S15A) and RNase H domain interfaces (SI Appendix, Fig. S15D), far fewer identical interactions are found in the connection subdomain interface (SI Appendix, Fig. S15C). Interesting differences in the polymerase domain include HERV-K RT molB W34 and HIV-1 RT p66 W88, which are functionally equivalent. In both cases, a tryptophan side chain spans the asymmetric dimer interface; but in HIV-1 the side chain projects from p66, and in HERV-K the tryptophan is contributed by molB (SI Appendix, Fig. S15B). The asymmetric dimer interface residues contributed by molA and molB from the connection subdomain are similar in HIV-1 and HERV-K. Both demonstrate similar positioning of molA/p66 αK and molB/p51 αL. However, HERV-K includes an additional helix: αL1 from molA. The interfacial salt bridge between p51 E396 and p66 R358 in HIV-1 is not retained in HERV-K. Rather, the equivalent glutamate (HERV-K molB E407) makes a salt bridge with R410 also from molB. The P39/P25 molB/p51 conserved residue from the interface makes very similar hydrophobic interactions with αK from molA/p66. There is a conserved glycine (G394/G384) that defines the end of αK and is directly above P39/P25 in the asymmetric dimer interface (SI Appendix, Fig. S15C). The important interfacial residue W401 in HIV-1 is not retained in HERV-K (main text ref 58) (5). The contacts between the molB thumb subdomain and the RNase H domain are largely conserved between HIV-1 and HERV-K RT (SI Appendix, Fig. S15D). Thus, the structural conservation of this asymmetric dimer interface, despite the extensive sequence drift and variation in secondary structure, suggests that there is strong selective pressure to maintain the functional asymmetry of this class II retroviral RT.
Fig. S1. Foscarnet-dependent inhibition of the DNA polymerization activity of HIV-1 and HERV-K RTs. (A) Chemical structure of foscarnet (PFA). (B) Denaturing PAGE migration patterns of the products of DNA polymerization activity of purified HIV-1 and HERV-K RTs on a DNA/RNA primer/template as shown on top of the panel. The numbering on top of the RNA template indicates the size (nucleotide, nt) of the primer (19-nt) and of the full template-length product (40-nt). 5′-end of the primer is radiolabeled with $^{32}$P to facilitate the detection of products. DNA polymerization activity was monitored at 37°C for 10 minutes in the presence of dNTP cocktail, MgCl$_2$ and the indicated concentrations of PFA. Lane m illustrates the migration pattern of the 5′-end radiolabeled 19-nt long primer. Asterisk indicates the site for PFA inhibition which occurs in the context of HIV-1 RT only.
**Fig. S2.** Schematic representation of the location of the RNase H cleavage with respect to the HIV-1 RT DNA polymerase active site and the size of the products of the RNase H activity. (A) RNA template and DNA primer hybrid used in RNase H activity assays. Letters P and N refer to the HIV-1 RT primer- and nucleotide-binding sites, respectively. When the 3' end of the primer occupies the P-site, the HIV-RT/primer-template complex assumes a post-translocation (post) conformation. When the 3' end of the primer occupies the N-site, the HIV-RT/primer-template complex assumes a pre-translocation (pre) conformation. Pre- and post-translocation conformations exist in equilibrium which is illustrated by vertical arrows. The numbering on top of the RNA template indicates the size in nucleotides (nt) of the RNA template (37-nt), the first nucleotide at the 3' end of the template which is radiolabeled, and the size of the products of the RNase H activity. 19nt on top of the horizontal bracket illustrates the distance between the DNA polymerase and the RNase H active sites. The products of the RNase H activity differ by one nucleotide, depending on the translocation conformation of the HIV-RT/primer-template complex. (B) Same notation as in panel A except that the 3' end of the RNA template is fluorescently labeled with fluorescein dye (Fl).
**Fig. S3.** Determination of the size of products of RNase H activity of HIV-1 or HERV-K RTs. The denaturing PAGE migration patterns of the products of RNase H activity of purified HIV-1 and HERV-K RTs on a DNA/RNA primer/template labeled with either $^{32}\text{P}$ or fluorescein dye at the 3'-end of the RNA template. RNase H activity was monitored at 37°C for 1 minute in the presence of MgCl$_2$ and heparin trap. Lane “m” illustrates the migration pattern of the 3'-end radiolabeled 37-nt long template. The numbering on the left indicates the nucleotide-length of the hydrolytic products of RNA template degradation. Heparin traps RT molecules that are not in complex with DNA/RNA hybrid, which prevents reassociation of the free enzyme with the DNA/RNA substrate. “Control trap” lane shows migration pattern of the products of the RNase H activity when heparin is added to the reaction mixture prior to the initiation of catalysis with MgCl$_2$. “Plus trap” lanes show migration pattern of the products of the RNase H activity when heparin is added to the reaction mixture simultaneously with MgCl$_2$. “post” and “pre” indicate the products of RNase H activity associated with post- and pre-translocation conformation of the RT in complex with DNA/RNA hybrid. The red and purple rectangles delineated with dotted lines isolate the relative distribution of pre- and post-translation conformations at equilibrium in absence and presence of PFA, respectively. PFA promotes the pre-translocation conformation of HIV-1 and HERV-K RTs in complex with DNA/RNA hybrids.
Fig. S4. The effect of PFA on the pre- and post-translocation conformations of HIV-1 and HERV-K RTs in complex with DNA/RNA primer/template. (A) The denaturing PAGE migration pattern of the products of RNase H activity of purified HIV-1 and HERV-K RTs on a DNA/RNA primer/template labeled with the fluorescein dye at the 3'-end of the RNA template. RNaseH activity was monitored at 37°C for 1 minute in the presence increasing concentrations of PFA after a simultaneous addition of MgCl₂ and heparin trap. Lane “m” illustrates the migration pattern of the 3'-end fluorescein-labeled 36-nt long template. Lane “c” illustrates the migration pattern of the products of the RNase H activity initiated with MgCl₂ after the addition of the heparin trap. “post” and “pre” indicate the products of RNase H activity associated with post- and pre-translocation conformation of the RT in complex with DNA/RNA hybrid. The purple rectangles delineated with dotted lines isolate the relative distribution of pre- and post-translocation conformations at the highest concentration of PFA. PFA promotes pre-translocation conformation of HIV-1 and HERV-K RTs in complex with DNA/RNA hybrids. (B) Graphical representation of the data shown in panel A. A PFA-dependent interconversion of the post- to pre-translocation conformation indicates that the conformations exist in equilibrium.
Fig. S5. Overview of the HERV-K RT Structure. (A) Stereoview of the ternary complex of HIV-1 RT (5D3G.PDB) superimposed upon the HERV-K RT ternary complex. The RMSD was 3.69 Å for 617 Cα atoms. The HERV-K RT molecule A (molA) subunit in green and the molecule B (molB) subunit of the asymmetric dimer in blue. The dsDNA is shown in orange cartoon. The p66 subunit of HIV-1 is in gray and the p51 subunit in bronze. The hairpin dsDNA of 5D3G is shown in tan cartoon. The RNase H domain of HERV-K molB would project behind the current view and is disordered in the HERV-K RT structure. (B) SDS-PAGE gel of HERV-K ternary complex crystals shows that a homodimer is present and that there is no processing of the molB subunit to release the disordered RNase H domain.
| FINGERS | PALM | THUMB | CONNECTION | RNase H |
|---------|------|-------|------------|---------|
| H K42-K58 αA | H P108-M110 αB1 | H L264-L281 αH | β Q337-F341 β16 | β T464-S470 βR1 |
| β I61-P63 β2 | β L117-D121 β6 | H T285-L295 αI | β T348-Q353 β17 | β G473-T478 βR2 |
| β V72-K77 β3 | H B167-K185 αE | H P309-B324 αJ | β D356-B361 β18 | β R483-K486 βR3 |
| β G80-T86 β4 | β Y190-Y194 β9 | β F371-L373 β18a | H A492-D506 RaA |
| H R89-L94 αB | β D197-A201 β10 | H Y374-C393 αK | H I511-S515 βR4 |
| H C125-T128 αC | H S228-K230 αF1 | H K399-V402 β19 | H A518-V527 RaB |
| H E133-K138 αD | β F237-Y239 β10 | H K406-N415 αL | H A530-I532 βR4a |
| β A140-A143 β7 | β M242-I244 β13 | H G417-A424 αL1 | H D538-K553 RaD |
| β T153-W157 β8 | β I249-Q252 β14 | β I429-D431 β20 | β F558-H562 βR5 |

Fig. S6. Secondary structural elements of HERV-K RT for molA (top) and molB (bottom).
Fig. S7. Comparison of protein-nucleic acid interactions between HIV-1 (5TXL.PDB) and HERV-K RT (colored by subdomain). Template strand in yellow and primer strand in orange carbons. The N-site dTTP is shown with bronze carbons and the Mg$^{2+}$ ion is a tan sphere. The p66 subunit of STXL HIV-1 RT is shown with light gray carbons and dsDNA from that complex is omitted for clarity. (A) Stereoview of the N-site. The conserved interactions with triphosphate and Mg$^{2+}$ ion include K77, R83, Q162, and K230 from the HERV-K polymerase domain. R83 also makes an interaction with the face of the N-site nucleotide base. (B) Stereoview of primer strand interactions. I195 from the YIDD loop sits beneath the deoxyribose ring of the P-site. I195 is a methionine in HIV-1 RT. Above this loop, L240 of the primer grip is positioned for interaction with the P-1 site deoxyribose and W276 from the base of the thumb stacks against the deoxyribose of the P-2 site. (C) Stereoview of template strand interactions. Several conserved hydrophobic amino acids (W38, F73, I75, and L85) are indicated in this stereoview of the single-stranded overhang on the DNA template. These conserved residues could be involved in the advancement of the template strand. (D) Stereoview of the RNase H active site. HERV-K RT RNase H domain is colored orange. The tan cartoon is the dsDNA from X105. The light gray molecule is the RNase H domain from HIV-1 RT (SD3G.PDB) with hairpin dsDNA omitted for clarity. The conserved active site residues are labeled.
Fig. S8. A potential allosteric site in molB involves C171. This residue is exposed to the surface at the bottom of a deep pocket. In the HERV-K RT structure we see three 1,4-dioxane molecules and one ethylene glycol molecule bound, which suggests that small molecules tend to bind at this location.
Fig. S9. The asymmetric unit of these HERV-K RT crystals (X105) contain two distinct complexes. (A) The two complexes present in the asymmetric unit of crystal X105. The tan shading indicates the base pair in the P-site and the blue shading represents the N-site. In the top sequence the ddTTP has reacted with the primer and now occupies the P-site. There is no 3'-OH on the primer strand and the N-site is occupied by dCTP. In the lower sequence the primer terminates with a G which has the 3'-OH and dTTP occupies the N-site. (B) Ternary complex 1 in the asymmetric unit of the P31 crystal form (X105). The primer strand is terminated by a dideoxythymidine nucleotide which lacks a 3'-OH in the P-site. No positive or negative difference density is observed in the map, which indicates a good fit. Likewise, the N-site is occupied by a dCTP molecule with good density for the 3'-OH. The phosphates of the dCTP (orange atoms) wrap around the metal ion, which has strong density. (C) Ternary complex 2 in the asymmetric unit of the P31 crystal form (X105). The primer strand is terminated by a dGMP residue in the P-site and the 3'-OH is clearly visible in the map. The N-site is occupied by a dTTP with density for the 3'-OH. There is a slight suppression of this density which may be expected since the crystal was grown in the presence of both ddTTP and dTTP.
Fig. S10. Detailed comparison of the fingers subdomain of HERV-K and the equivalent subdomain from HIV-1 RT (5D3G.PDB). The HERV-K residues are labeled. (A) Stereoview of the superposition 5D3G HIV-1 RT p66 (gray) onto the isolated fingers subdomain (blue) of HERV-K RT molA. RMSD of 1.085 Å for 84 Cα atoms. The Mg\(^{2+}\) ion is shown as a green sphere and the dTTP is shown with bronze carbons. (B) Stereoview of the superposition of 5XTL.PDB HIV-1 RT p66 ternary complex (gray) onto the isolated fingers subdomain (blue) of HERV-K RT molA. RMSD of 0.927 Å for 87 Cα atoms. The tips of the fingers of 5XTL grasp the Mg\(^{2+}\) ion in the N-site more tightly than 5D3G. 5XTL best reflects these interactions when compared to HERV-K X105.
Fig. S11 Detailed comparison of each subdomain of molA HERV-K and the equivalent subdomain from p66 HIV-1 RT (5D3G.PDB). The HERV-K residues are labeled. (A) Stereoview of the superposition of 5D3G HIV-1 RT p66 (gray) onto the isolated palm subdomain (red) of HERV-K RT molA. RMSD 0.827 Å for 68 Cα atoms. (B) Stereoview of the superposition of 5D3G HIV-1 RT p66 (gray) onto the isolated thumb subdomain (green) of HERV-K RT molA. RMSD 3.425 Å for 57 Cα atoms. (C) Stereoview of the superposition of 5D3G HIV-1 RT isolated p66 connection domain (gray) onto the isolated Connection domain (yellow) of HERV-K RT molA. RMSD 0.824 Å for 36 Cα atoms. (D) Stereoview of the superposition of 5D3G HIV-1 RT p66 subunit (gray) onto the isolated RNase H domain (orange) of HERV-K RT molA. RMSD 1.225 Å for 78 Cα atoms.
Fig. S12 Detailed comparison of each subdomain of molB HERV-K and the equivalent subdomain from p51 HIV-1 RT (5D3G.PDB). The HERV-K residues are labeled. (A) Stereoview of the superposition of 5D3G HIV-1 RT p51 subunit (gray) onto the isolated fingers subdomain (periwinkle) of HERV-K RT molB. RMSD 0.973 Å for 78 Cα atoms. (B) Stereoview of the superposition of 5D3G HIV-1 RT p51 subunit (gray) onto the isolated palm subdomain (pink) of HERV-K RT molB. RMSD 3.901 Å for 57 Cα atoms. (C) Stereoview of the superposition of 5D3G HIV-1 RT p51 subunit (gray) onto the isolated thumb subdomain (light green) of HERV-K RT molB. RMSD 3.374 Å for 57 Cα atoms. (D) Stereoview of the superposition of 5D3G HIV-1 RT p51 isolated connection subdomain (gray) onto the isolated connection subdomain (mustard) of HERV-K RT molB. RMSD 1.159 Å for 35 Cα atoms.
**Fig. S13.** Detailed analysis of the asymmetric dimer interface HIV-1 RT vs. HERV-K RT. (A) The structure-based sequence alignment of HERV-K RT used in the present work and HIV-1 RT BH10 from the 5D3G.PDB structure. Colored bars indicate the various subdomains of RT: fingers=blue, palm=red, thumb=green, connection=yellow, and RNase H=orange. Identical residues are shown in red print. Gray shading indicates that the residue contributes to the asymmetric dimer interface from molA/p66 and the blue shading indicates that the residue contributes to the asymmetric dimer interface from molB/p51. Conserved residues in the interface are indicated by bold red lettering.
**Fig. S14.** Surface renderings of the interface between molA and molB of HERV-K RT. (A) The asymmetric dimer in surface rendering (Maestro). MolA is shown in green and molB in light blue. (B) MolB residues (blue) that interact with molA (gray) are shown. The orange cartoon indicates the dsDNA. Polymerase domain interactions are shown on the left, connection subdomain interactions in the center and thumb subdomain and RNase H domain interactions on the right. (C) MolA residues (green) that interact with molB (gray) are shown. Polymerase domain interactions are shown on the left, connection subdomain interactions in the center and thumb subdomain and RNase H domain interactions on the right.
Fig. S15. Stereoview of the dimer interface with comparison to HIV-1 RT. HERV-K molA is shown in green, HERV-K molB in blue. HIV-1 RT p66 is gray (5D3G) and HIV-1 RT p51 is bronze. Only residues in which at least one atom contribute to the interface are shown. Those in stick are conserved between HIV-1 RT and HERV-K RT. (A) Polymerase domain interface. The palm subdomain interface includes the conserved I146NNXXP motif (molB/p51) that interacts with αK from the molA connection. Also, contacts from the fingers molA Q102 and palm molA αE are present. (B) Polymerase domain interface. HERV-K W34 molB and HIV-1 RT W88 p66 are functionally equivalent. A tryptophan side chain spans the asymmetric dimer interface. In one case the side chain projects from the HIV-1 p66 subunit and in the other (HERV-K) the tryptophan is contributed by molB. (C) Connection subdomain interface. The asymmetric dimer interface residues contributed by molA and molB from the connection subdomain are similar between HIV-1 and HERV-K. Both demonstrate similar positioning of αK from molA/p66 and αL molB/p51. However, HERV-K includes an additional helix α1 from molA. The interfacial salt bridge between E396 p51 and R358 p66 in HIV-1 is not retained in HERV-K. Rather, the equivalent glutamate E407 HERV-K molB makes a hydrogen bond with R410 also from molB. The P39/P25 molB/p51 conserved residue from the interfaces make very similar hydrophobic interaction with αK from molA/p66. There is a conserved glycine (G394/G384) which defines the end of αK and is directly above P39/P25 at the asymmetric dimer interface. (D) Thumb-RNase H interface. The interactions between the thumb of molB/p51 and the RNase H domain of molA/p66. This alignment is asymmetric dimer on
asymmetric dimer. One can see that there is some rotation and translation between the molecules because of the global alignment. Nonetheless, a large number of identical residues are found at this conserved interface contributed from the molB/p51 thumb and RNase H. In addition to HERV-K RT residue labeling, the conserved LRG motif from the HIV-1 RT p51 thumb residues L283 and R284 are labeled.
TABLE S1. Data Collection and Refinement Statistics

|                                | X105 HERV-K RT/dsDNA/dTTP/dCTP | X144 HERV-K RT/dsDNA/Hg-dCTP |
|--------------------------------|--------------------------------|-----------------------------|
| **PDB code**                   | 7SR6                           | N/A                         |
| **Data Collection & Scaling**  |                                |                             |
| wavelength (Å)                 | 1.49993                        | 1.00966                     |
| number of merged data sets     | 1                              | 1                           |
| resolution range (Å)           | 88.9-2.62 (2.71-2.62)          | 64.44-3.01 (3.09-3.01)      |
| space group                    | P3₁                            | P6₁                         |
| unit cell (Å)                  | 177.697, 177.697, 117.414      | 177.95, 177.95, 117.54      |
| angle (deg)                    | 90, 90, 120                    | 90, 90, 120                 |
| unique reflections             | 114382 (5724)                  | 41878 (3095)                |
| redundancy                     | 10.7 (10.4)                    | 21.2 (21.2)                 |
| completeness (%)               | 91.8 (49.4)                    | 99.5 (100)                  |
| mean I/σ(I)                    | 17.6 (2.0)                     | 14.7 (0.7)                  |
| Wilson B-factor (Å²)           | 66.06                          | 99.4                        |
| Rmerge                         | 0.08 (1.08)                    |                             |
| Rmeas                          | 0.08 (1.14)                    |                             |
| Rpim                           | 0.03 (0.35)                    |                             |
| CC1/2                          | 1.0 (0.78)                     |                             |
| **Refinement & Model Analysis**|                                |                             |
| resolution range (Å)           | 88.9-2.62 (2.69-2.62)          |                             |
| completeness (%)               | 91.8 (46.0)                    |                             |
| Rwork                          | 0.164 (0.205)                  |                             |
| Rfree                          | 0.221 (0.297)                  |                             |
| number of atoms                | 18221                          |                             |
| macromolecules                 | 17594                          |                             |
| ligands                        | 343                            |                             |
| water molecules                | 284                            |                             |
| protein residues               | 2441                           |                             |
| rms bonds (Å)                  | 0.007                          |                             |
| rms angles (deg)               | 1.626                          |                             |
| Ramachandran favored (%)       | 96.53                          |                             |
| Ramachandran allowed (%)       | 3.42                           |                             |
| Ramachandran outliers (%)      | 0.05                           |                             |
| rotamer outliers (%)           | 3.78                           |                             |
| clashscore                     | 12.59                          |                             |
| average B-factor (Å²)          | 71.86                          |                             |
| macromolecules                 | 72.86                          |                             |
| ligands                        | 55.57                          |                             |
| water molecules                | 52.60                          |                             |
| Number of TLS groups           | 4                              |                             |

*Statistics for the highest-resolution shell are shown in parentheses.*
1. Kang SM & Morrow CD (1999) Genetic analysis of a unique human immunodeficiency virus type 1 (HIV-1) with a primer binding site complementary to tRNAMet supports a role for U5-PBS stem-loop RNA structures in initiation of HIV-1 reverse transcription. *Journal of virology* 73(3):1818-1827.

2. Das D & Georgiadis MM (2004) The crystal structure of the monomeric reverse transcriptase from Moloney murine leukemia virus. *Structure* 12(5):819-829.

3. Nowak E, et al. (2013) Structural analysis of monomeric retroviral reverse transcriptase in complex with an RNA/DNA hybrid. *Nucleic Acids Res* 41(6):3874-3887.

4. Nowak E, et al. (2014) Ty3 reverse transcriptase complexed with an RNA-DNA hybrid shows structural and functional asymmetry. *Nat Struct Mol Biol* 21(4):389-396.

5. Tachedjian G, Aronson HE, & Goff SP (2000) Analysis of mutations and suppressors affecting interactions between the subunits of the HIV type 1 reverse transcriptase. *Proc Natl Acad Sci U S A* 97(12):6334-6339.