

## **Supplementary Information:**

### **High-resolution view of HIV-1 Reverse Transcriptase initiation complexes and inhibition by NNRTI drugs**

**Betty Ha<sup>1,2,7</sup>, Kevin P. Larsen<sup>1,3,5,7</sup>, Jingji Zhang<sup>1,7</sup>, Ziao Fu<sup>4,6</sup>, Elizabeth Montabana<sup>1</sup>, Lynnette N. Jackson<sup>1</sup>, Dong-Hua Chen<sup>1</sup>, Elisabetta Viani Puglisi<sup>1,\*</sup>**

<sup>1</sup> Department of Structural Biology, Stanford University School of Medicine, Stanford, CA. <sup>2</sup>Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA. <sup>3</sup>Program in Biophysics, Stanford University, Stanford, CA. <sup>4</sup>Department of Chemistry and Molecular Biophysics, Columbia University, New York, NY. <sup>5</sup>Present address: University of California Berkeley, Department of Molecular and Cell Biology, Berkeley, CA. <sup>6</sup>Present address: Laboratory of Molecular Neurobiology and Biophysics, The Rockefeller University, Howard Hughes Medical Institute, New York, NY. <sup>7</sup>These authors contributed equally to this work: Betty Ha, Kevin P. Larsen, Jingji Zhang. \*Correspondence: epuglisi@stanford.edu

## **Contents:**

**Supplementary Figure 1:** miniRTIC purification and activity validation.

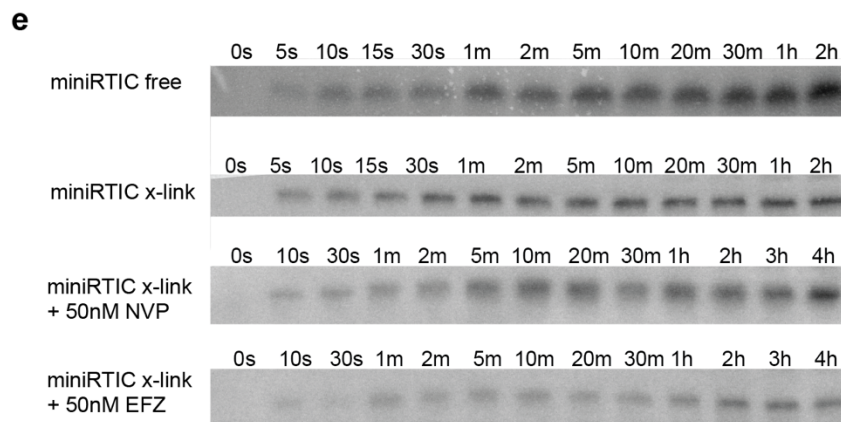
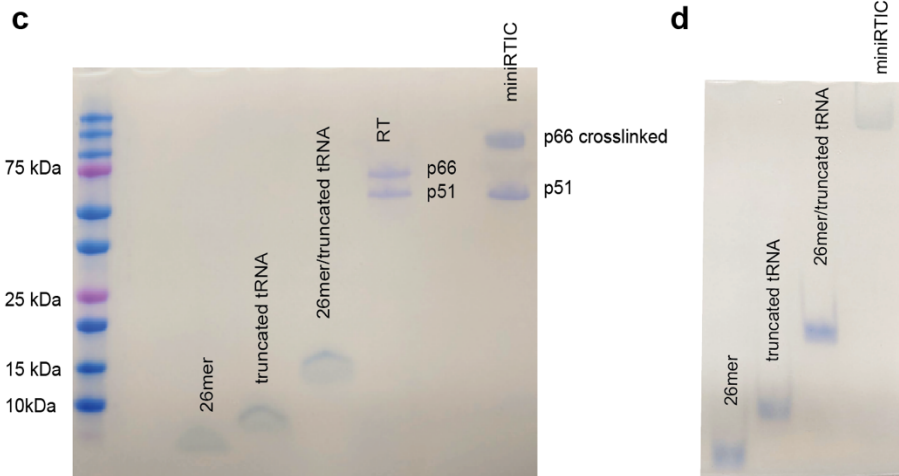
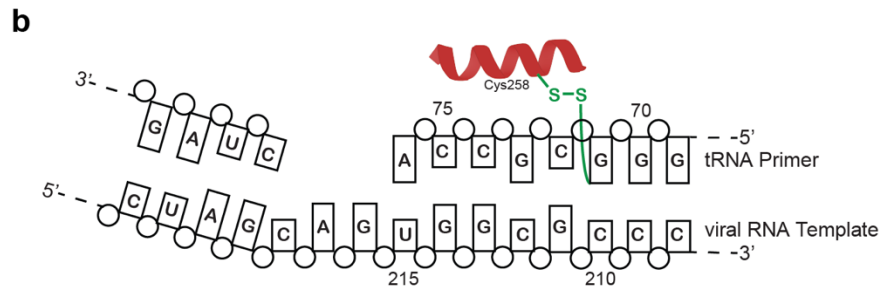
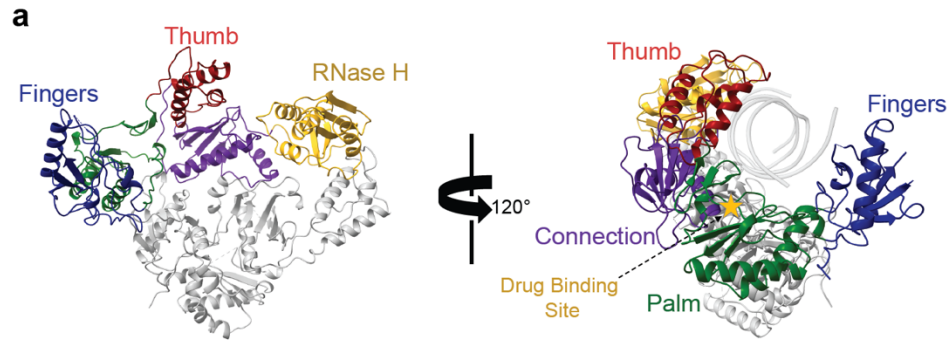
**Supplementary Figure 2:** Cryo-EM data processing and validation.

**Supplementary Figure 3:** Additional cryo-EM processing, local resolution estimates, and example density.

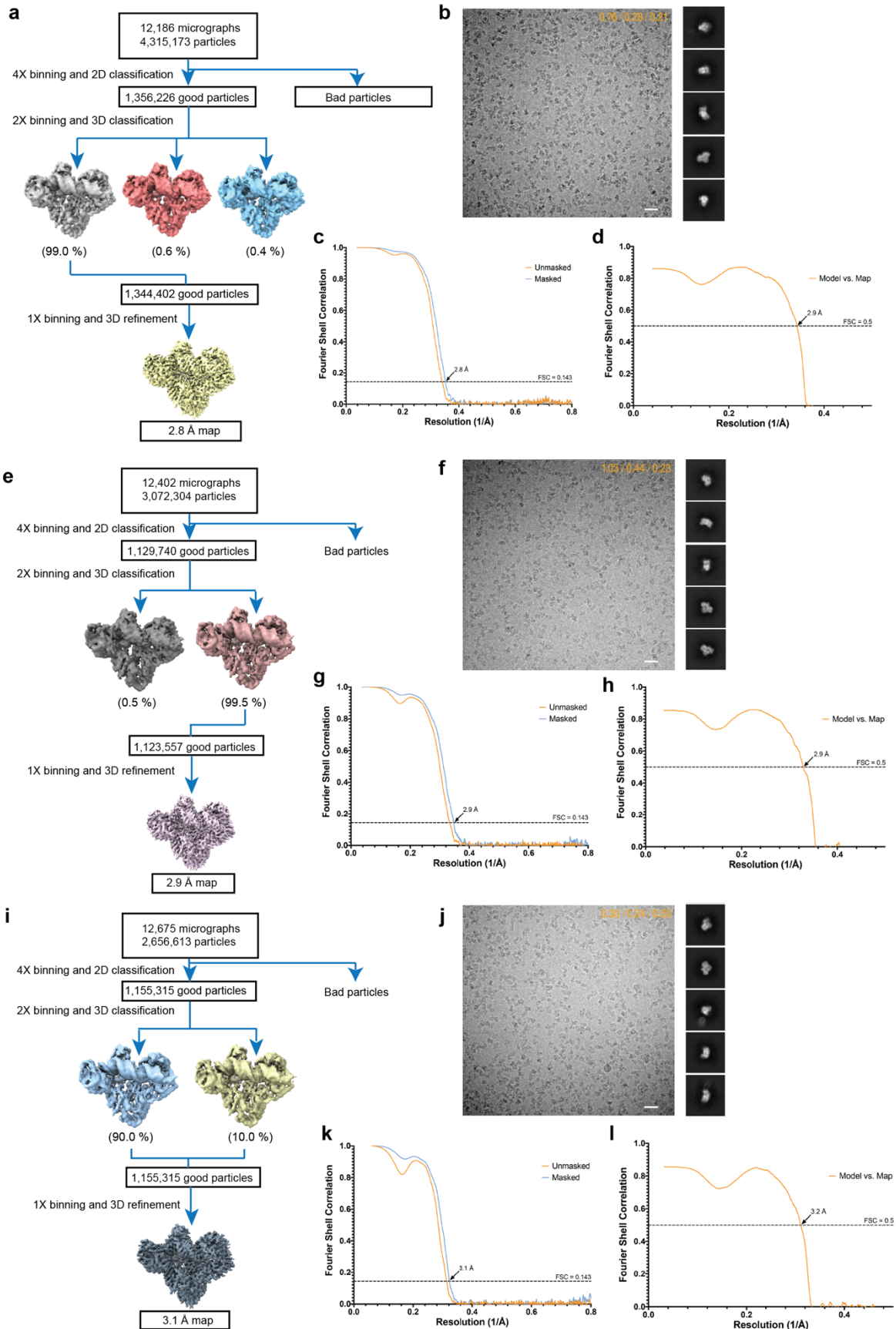
**Supplementary Figure 4:** RTIC cryo-EM density and modelling.

**Supplementary Figure 5:** miniRTIC protein–RNA contact analysis.

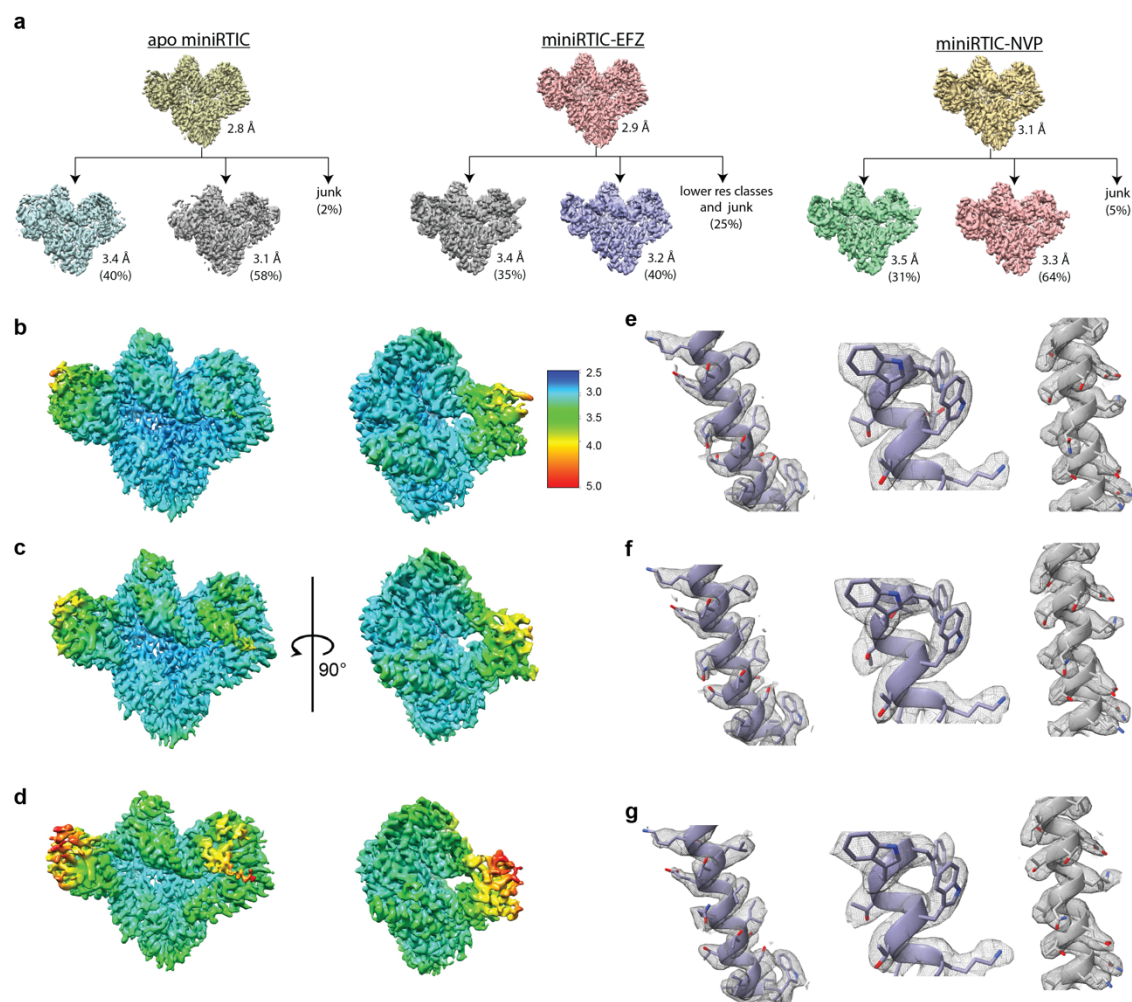
**Supplementary Figure 6:** NNRTI binding to the RTIC.



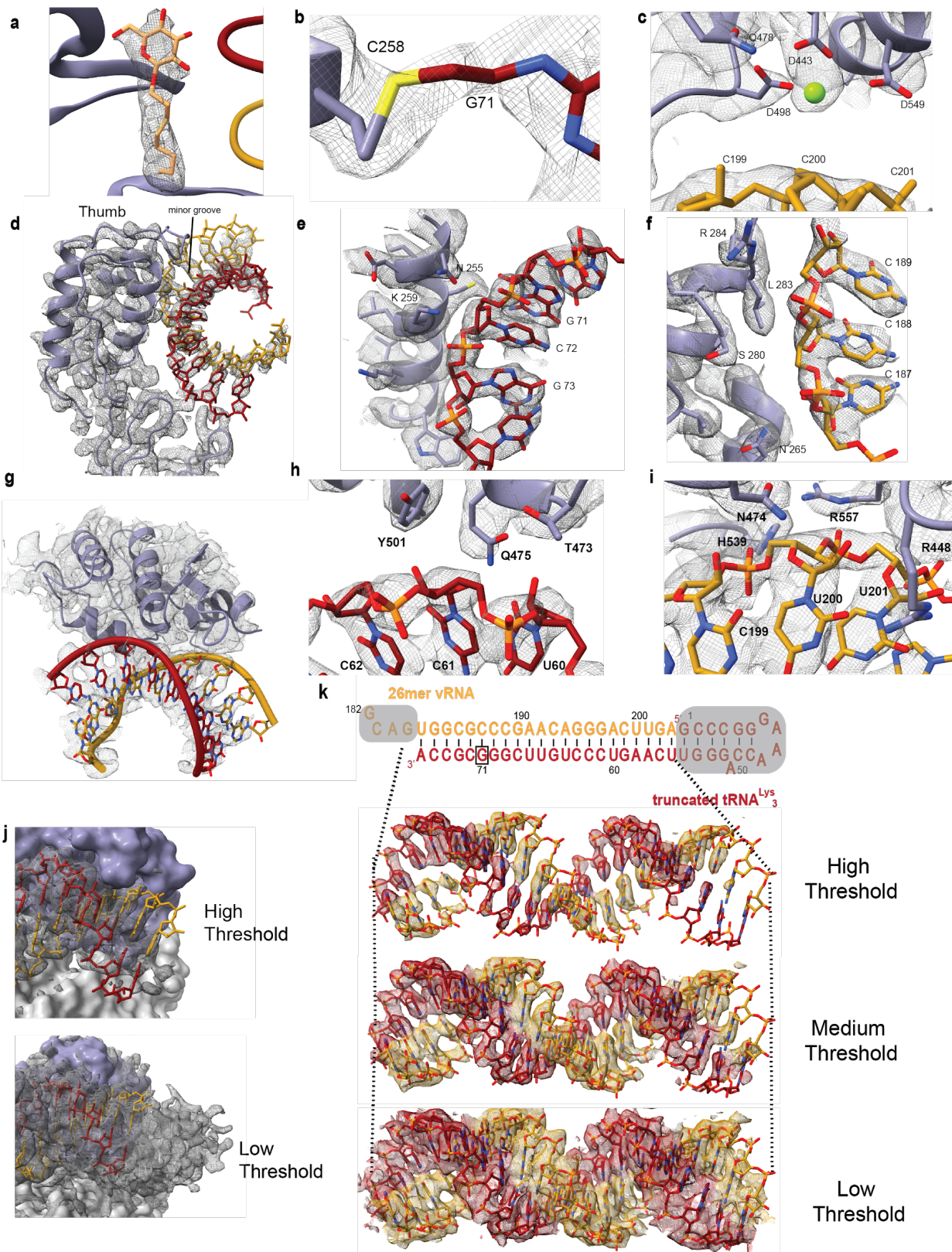
**Supplementary Fig. 1: miniRTIC purification and activity validation.** **(a)** RT domain architecture. The p66 subunit is coloured by domain (Fingers = blue, Palm = green, Thumb = red, Connection = purple, RNase H = yellow) and the p51 subunit is grey. The NNRTI binding pocket is marked with a star. **(b)** Crosslinking scheme of the miniRTIC (Thumb  $\alpha$ H=red, crosslink=green). The crosslinkable nt is placed at position 71. **(c)** Non-reducing SDS-PAGE of the vRNA (26mer), truncated tRNA, free vRNA–tRNA, reverse transcriptase (RT), and miniRTIC. The miniRTIC exhibits a shifted p66 subunit due to crosslinking with the vRNA–tRNA. **(d)** Native PAGE (20% polyacrylamide) of 26mer, truncated tRNA, vRNA–tRNA, and miniRTIC. The miniRTIC runs as a single band. Gel analysis in **c** and **d** was performed on all samples used in the manuscript ( $n > 5$ ) and was consistently reproducible. **(e)** Representative autoradiograph image of incorporation of dCTP over time in free RT+vRNA–tRNA (miniRTIC free), crosslinked miniRTIC (miniRTIC x-link), miniRTIC x-link in the presence of 50 nM nevirapine and miniRTIC x-link in the presence of 50 nM efavirenz. Results are quantified and plotted in Fig. 1d.



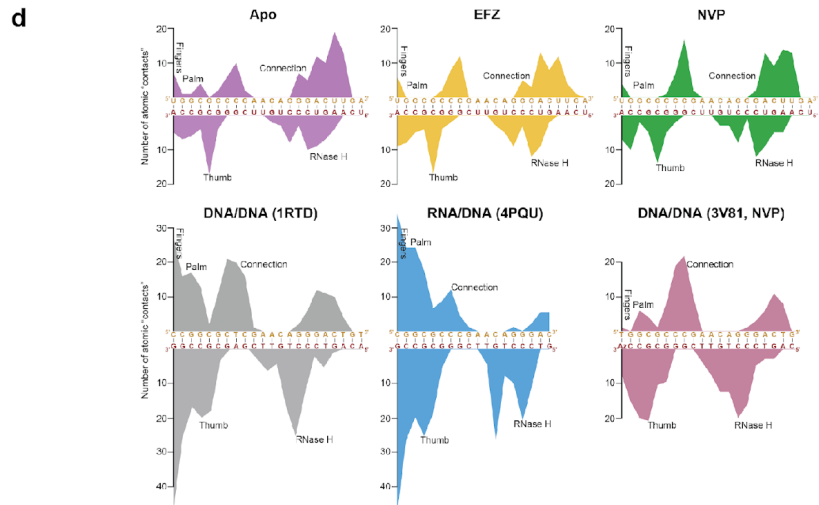
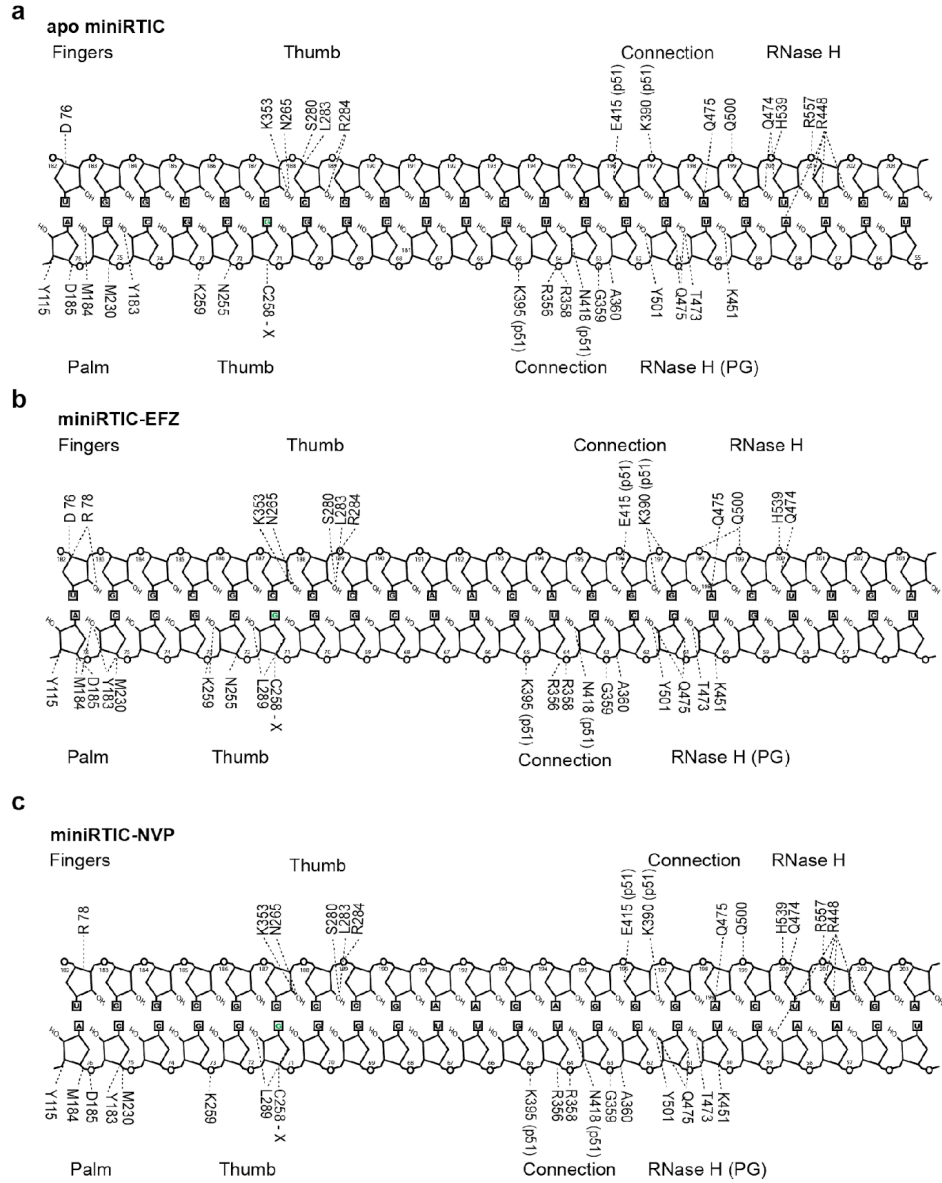
**Supplementary Fig. 2: Cryo-EM data processing and validation.** Cryo-EM processing tree of the **(a)** apo-miniRTIC, **(e)** miniRTIC-EFZ, and **(i)** miniRTIC-NVP. Representative micrographs and 2D averages of the **(b)** apo-miniRTIC, **(f)** miniRTIC-EFZ, and **(j)** miniRTIC-NVP. Scale bar is 20 nm (white, lower right corner). Gold standard Fourier shell correlation (FSC) curve of the **(c)** apo-miniRTIC, **(g)** miniRTIC-EFZ, and **(k)** miniRTIC-NVP. The global resolution is determined by an FSC cut-off value of 0.143 (dashed line and arrow). Map-vs-model FSC curves of the **(d)** apo-miniRTIC, **(h)** miniRTIC-EFZ, and **(l)** miniRTIC-NVP. The FSC cut-off value of 0.5 is indicated by a dashed line.



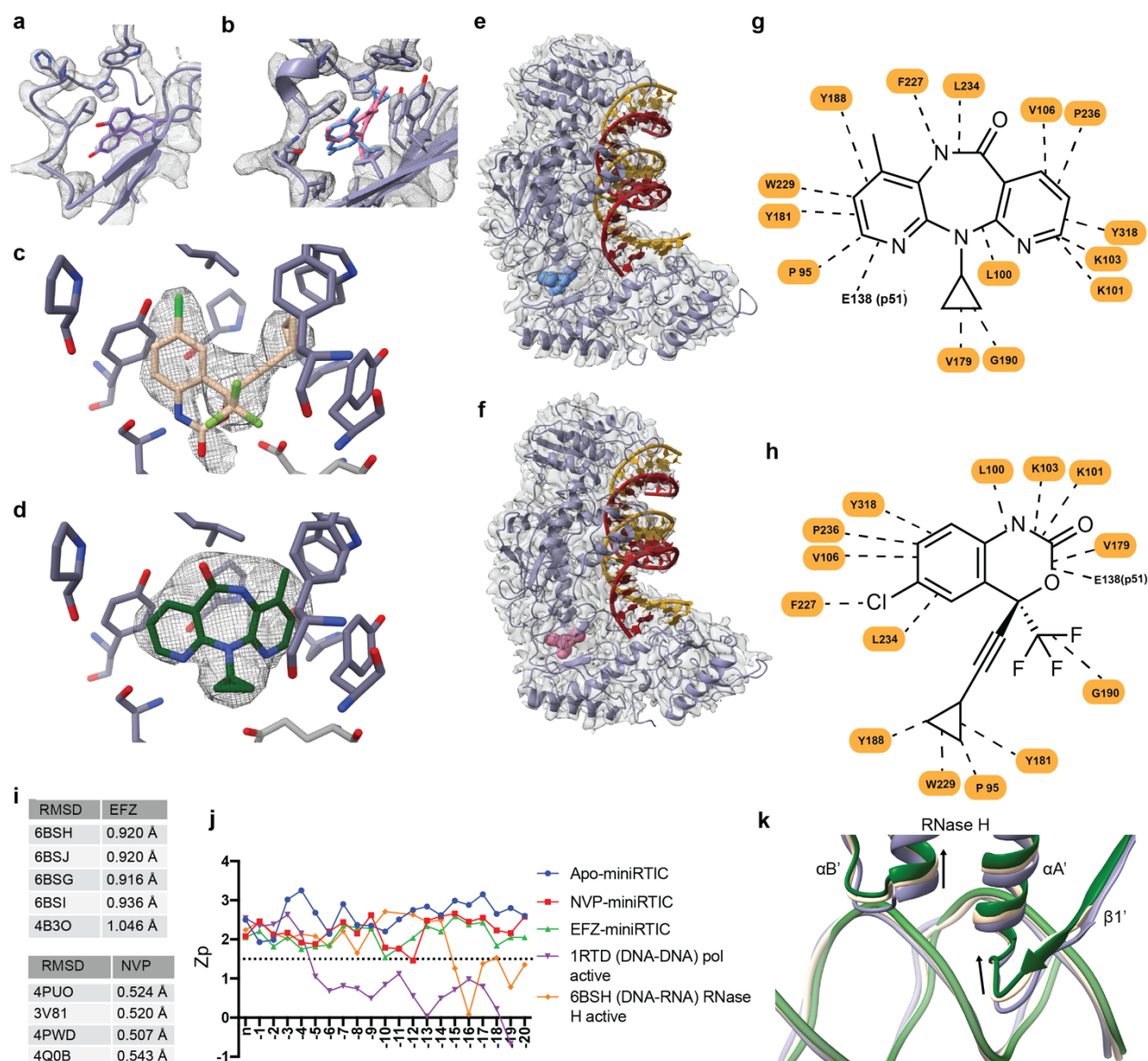
**Supplementary Fig. 3: Additional cryo-EM processing, local resolution estimates, and example density.** (a) Representative examples of attempts to further classify the final particle sets of all three data sets. All attempts to identify additional amounts of underlying heterogeneity did not demonstrably improve the final reconstructions in terms of resolution and quality of the density. The (b) apo-miniRTIC 2.8 Å, (c) miniRTIC-EFZ 2.9 Å and (d) miniRTIC-NVP 3.1 Å maps coloured according to local resolution estimated by Relion. Density and model fit for p66 subunit residues 364-383 (left), 395-403 (middle) and for p51 subunit residues 365-382 (right) for (e) apo-miniRTIC, (f) miniRTIC-EFZ and (g) miniRTIC-NVP.



**Supplementary Fig. 4: RTIC cryo-EM density and modelling.** Colouring for all panels: p66 = purple, p51 = gray, vRNA = gold, tRNA = red. **(a)**  $\beta$ -OG partially inserted into the NNRTI binding pocket. **(b)** Density for crosslink between C258 and modified dG71 of the tRNA primer. **(c)** RNase H active site with single bound  $Mg^{2+}$  cation. vRNA template strand (gold) is in close proximity, but not engaged. **(d)** Thumb interaction with minor groove of vRNA–tRNA. **(e)** Closeup of thumb domain  $\alpha$ -H helix interaction with tRNA backbone. **(f)** Closeup of thumb domain  $\alpha$ -H and  $\alpha$ -I with vRNA backbone. **(g)** RNase H domain interaction with minor groove of vRNA–tRNA. **(h)** RNase H primer grip residues interaction with tRNA backbone **(i)** RNase H residues proximal to active site interaction with vRNA backbone. **(j)** High threshold (top) and low threshold (bottom) of density (mesh) for peripheral stem, bulged A and tetraloop density not modeled in. Density for this region is weak and could not be modeled. **(k)** Secondary structure of bimolecular RNA complex, with 26 nucleotides of vRNA (gold) annealed to 39 nucleotides of truncated tRNA<sup>Lys</sup><sub>3</sub> (red). Grey boxes on vRNA template overhang, tRNA stem, and tetraloop cap indicate RNA that was not modeled. Dotted lines correlate with modeled RNA nucleotides built into 2.8 Å EM map density. Map density is shown in high (top), medium, and low contour thresholds (bottom). Colouring of the RNA is the same as depicted in the secondary structure.



**Supplementary Fig. 5: miniRTIC protein–RNA contact analysis.** Contact map of (a) apo-miniRTIC, (b) miniRTIC–EFZ and (c) miniRTIC–NVP. The top strand is the vRNA template and the bottom strand is the tRNA primer. All residues, unless indicated, are located on the p66 subunit. (d) RT–nucleic acid contact landscape. Contacts defined as atom distances within 5 Å. Raw atom contact counts for each nucleotide on the template (top) and primer (bottom) strands are plotted on the y-axis. Apo-miniRTIC=purple, miniRTIC–EFZ=yellow, miniRTIC–NVP=green, RT–dsDNA=gray, RT–RNA–DNA=blue, RT–dsDNA–NVP=pink.



**Supplementary Fig. 6: NNRTI binding to the RTIC.** (a) View of apo miniRTIC binding pocket. (b) View of miniRTIC–NVP NNRTI binding pocket with EFZ and NVP aligned on binding pocket. Both drugs make similar interactions with RT within the binding pocket. (c) Front on view of EFZ (beige) and interacting residues in p66 subunit (purple) and p51 subunit (grey). (d) Front on view of NVP (green) and interacting residues in p66 subunit (purple) and p51 subunit (grey). (e) 2.9 Å map and model of miniRTIC–EFZ. EFZ is coloured in blue. (f) 3.1 Å map and model of miniRTIC–NVP. NVP is coloured in pink. (g) NVP contact map with RT residues. (h) EFZ contact map with RT residues. (i) Root mean square deviation (r.m.s.d.) of efavirenz all-atom positions bound to RTIC for structures of efavirenz bound to RT-nucleic acid structures listed by PDB ID (top); r.m.s.d. of nevirapine all-atom positions bound to RTIC nevirapine bound to RT-nucleic acid structures listed by PDB ID (bottom); r.m.s.d. values calculated after alignment on palm subdomain backbone residues. (j) Helical analysis using Zp data from 3DNA. Zp parameter on y-axis and base pair position on x-axis. Base pairs above the dotted line at Zp = 1.5 generally categorized

as A-like and below the line as B-like. **(k)** Compared to apo miniRTIC (purple), upon drug binding the RNase H domain shifts by  $\sim 2.0$  Å in the miniRTIC–EFZ (beige) miniRTIC–NVP (green).