Supplementary material: Relative domain orientation of the L289K HIV-1 reverse transcriptase monomer

Zhaoyong Xi[1,#], Tatiana V. Ilina[1, #], Michel Guerrero[1], Lixin Fan[2], Nicolas Sluis-Cremer[3], Yun-Xing Wang[4], Rieko Ishima [1,].*

[1]Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260, USA.

[2]Basic Science Program, Frederick National Laboratory for Cancer Research, SAXS Core Facility of the National Cancer Institute, Frederick, MD 21702, USA.

[3]Department of Medicine, Division of Infectious Diseases, University of Pittsburgh School of Medicine, Pittsburgh, PA, 15261, USA

[4]Protein-Nucleic Acid Interaction Section, Structural Biophysics Laboratory, National Cancer Institute, National Institutes of Health, Frederick, MD 21702, USA.

Corresponding Author

* Rieko Ishima, Email: ishima@pitt.edu

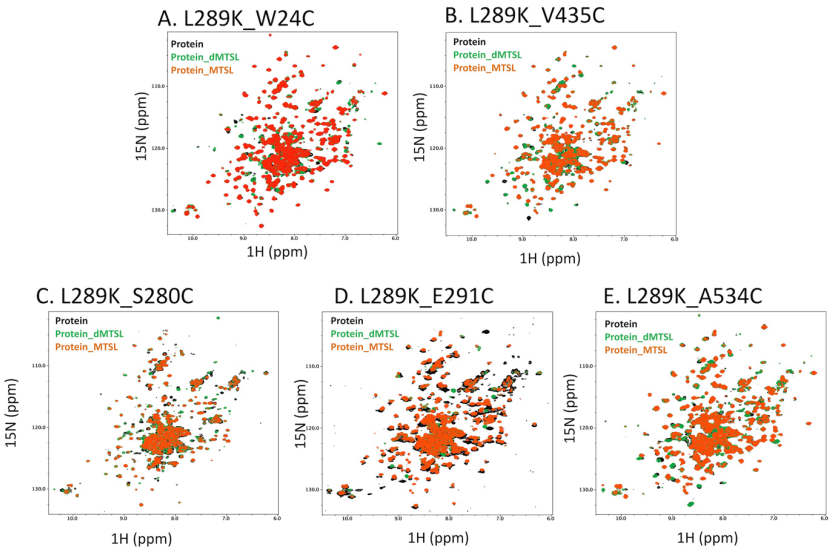

Figure S1. ^{1}H-^{15}N TROSY-HSQC spectra of $p66_{L289K}$ (black) and with MTSL labeling (red) or diamagnetic-MTSL labeling (green), for (A) $p66_{L289K/W24C}$, (B) $p66_{L289K/V435C}$, (C) $p66_{L289K/S280C}$, (D) $p66_{L289K/E291C}$ and (E) $p66_{L289K/A435C}$. MTSL or dMTSL was introduced to the Cys mutation site. To enable only one label in each protein, p66 containing C38V and C280S mutations was used in this study. In case when the spectra are compared with other NMR spectra, note that chemical shift changes occurred by TROSY selection, ~45 Hz, was not corrected in these spectra.

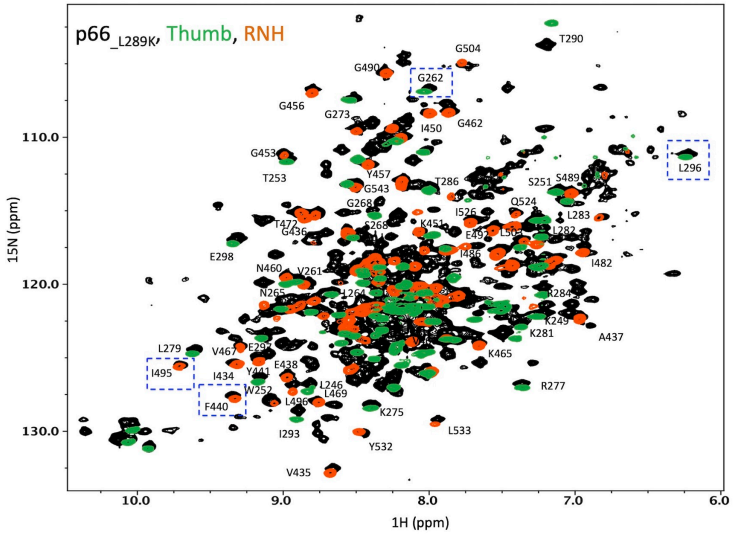

Figure S2. ^{1}H-^{15}N TROSY-HSQC spectra of $p66_{L289K}$ (black) overlaid with those of the isolated RNH domain (orange) and thumb domain (green) proteins. Qualitative signal assignments were obtained from the spectral overlay (1-3). Residues used for the PRE analysis are labeled. Blue dashed-rectangles indicate residues shown in Figure 2.

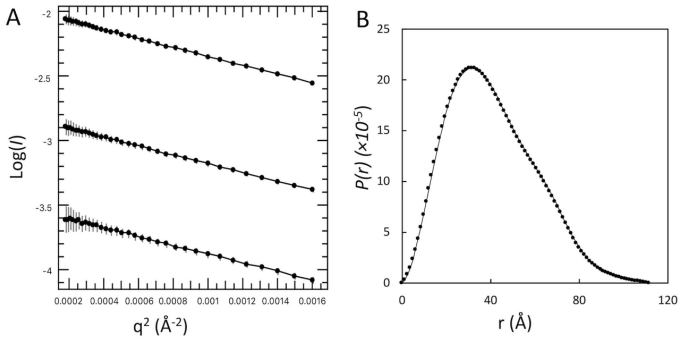

Figure S3. SAXS (A) Guinier plots for 1.0, 2.0 and 4.3 mg/mL and (B) the normalized pair-distance distribution function for $p66_{L289Ktr}$. In (A), the log of scattering intensity (I) is plotted against the scattering vector q. In (B), the pair-distance distribution, *P(r),* is plotted against distance, *r*. Detailed data for the radius of gyration and the maximum size of the protein are described in **Table 1**.

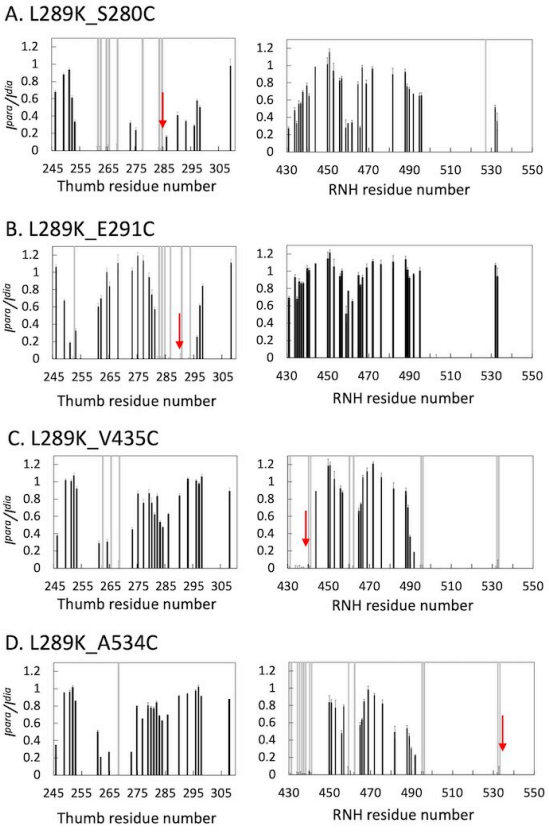

Figure S4. Plots of PRE, as I^{para}/I^{dia}, for residues in the thumb and RNH domains in $p66_{L289K}$. Black bars are those residues for which intensity ratios were obtained, while grey bars indicate residues with I^{para} of almost zero. Resonances were qualitatively assigned based on the spectral overlay in **Figure S2**. Residues that overlap with other resonances in $p66_{L289K}$, such as those in the mobile region, could not be assigned in this analysis. Red arrows indicate the sites to which MTSL-tag was attached.

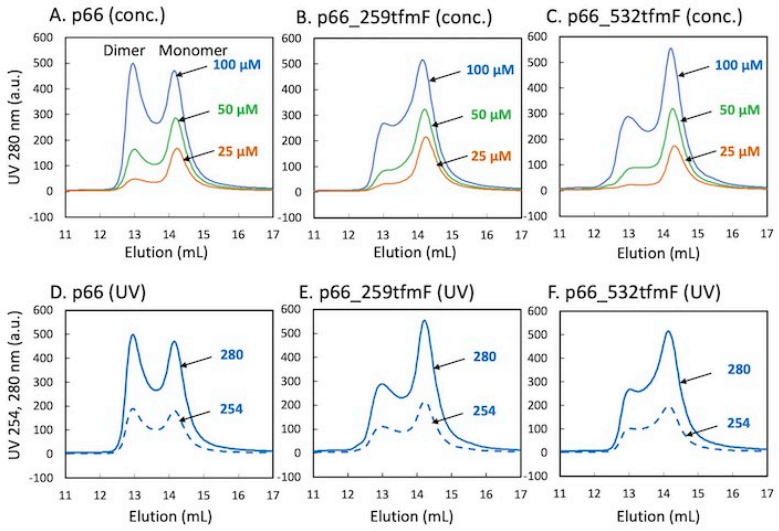

Figure S5. SEC elution profiles of (A) p66 and (B, C) 4-trifluoromethyl phenylalanine (tfmF) p66 with the label at (B) residue 259 or (C) residue 532 site, at protein concentrations, 25 µM (orange), 50 µM (green) and 100 µM (blue), and (D-F) comparison of detection at UV 254 nm and UV 280 nm for the 100 µM data. SEC was run using a 24-ml analytical Superdex 200 Increase 10/300 GL column (GE Healthcare), equilibrated with a 25 mM Bis-tris buffer, pH 7.0, containing 100 mM NaCl with 0.02% sodium azide, at a flow rate of 0.5 ml/min. In each experiment, a 60 µL sample was injected, and protein elution was monitored by UV absorbance at 254 and 280 nm. Empirically, ×10 fold dilution in the column is expected from those of the injected protein concentrations. Panels D-F are shown to verify that the purified proteins were not contaminated by nucleic acids.

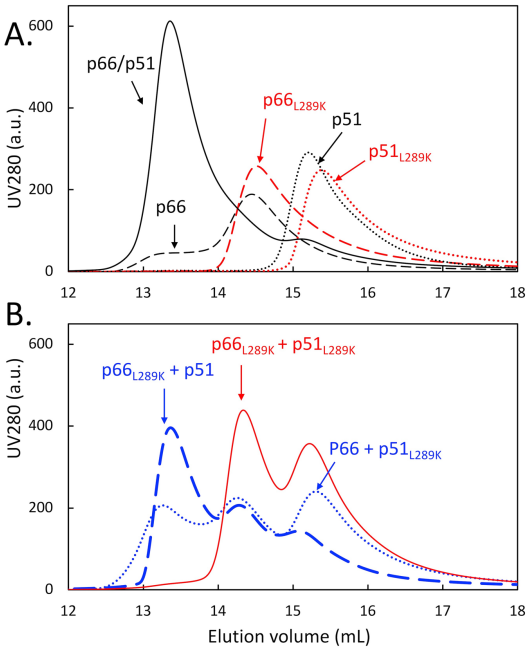

Figure S6. SEC elution profiles of (A) p66/p51 (black solid line), p66 (black dashed line), p51 (black dotted line), $p66_{L289K}$ (red dashed line) and $p51_{L289K}$ (red dotted line), and (B) of a mixture of $p66_{L289K}$ and $p51_{L289K}$ (red solid line), $p66_{L289K}$ and p51 (blue dashed line) and p66 and $p51_{L289K}$ (blue dotted line). Proteins were prepared in a 25 mM Bis-tris buffer, pH 7.0, containing 100 mM NaCl with 0.02% sodium azide, at 30 µM concentration, mixed. 80µL volume was injected each time.

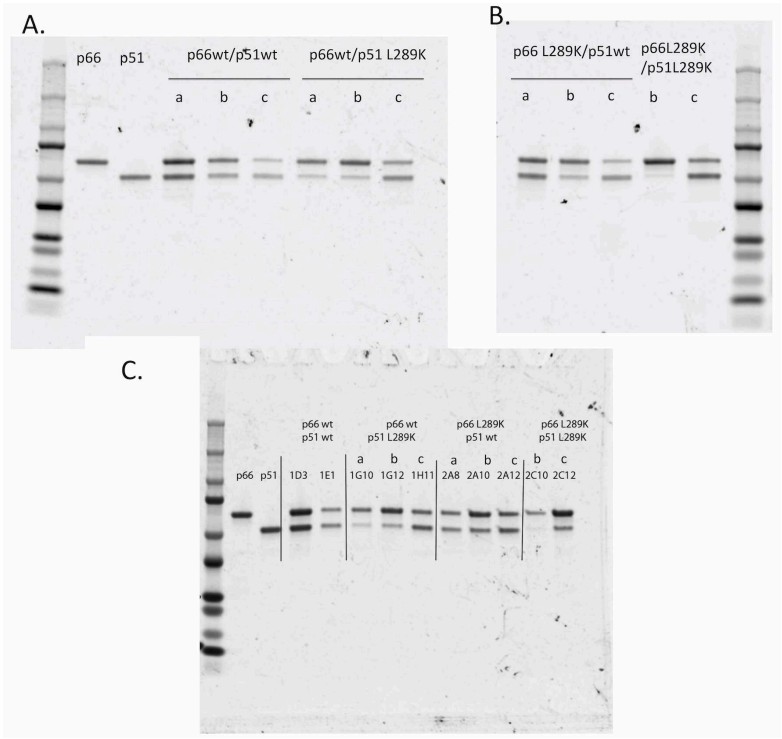

Figure S7. (A, B) the entire SDS gel shown in Figure 5C, indicating that there are no smaller degradation bands, and (C) repeated data.

Table S1. Molecular masses of the elution peaks at position ***a*** in the SEC in Figure 5.

Sample	Detected	Theoretical	Protein
p66 WT + p51 WT	65537.36	65538.27	p66 WT
	52595.40	52596.60	p51 WT
p66 WT + p51 L289K	65536.64	65538.27	p66 WT
	63631.96	*63634.10*	Cleavage after denaturation [a]
	63291.67	*63293.64*	Cleavage by formic acid [b]
	52610.00	52611.61	p51 L289K
p66 L289K + p51 WT	65552.49	65553.29	p66 L289K
	63307.52	*63308.65*	Cleavage by formic acid [b]
	52594.98	52596.60	p51 WT

a. We observed protease-cleaved fraction in the mass spec data. Since we do not see this band in the SDS gel, we conclude that the fragmentation happened in the column applied for the mass spectrometry.
b. Chemical cleavage by formic acid for denaturation to acquire mass spec data.

Table S2. Selected mass spectrometry results to identify the samples prepared for this study.

proteins for SAXS and PRE			
Date_#	Proteins	Measured (Da)	Theoretical (Da)
190423_1	p66 (N-terminal His, cleaved)	64906.8027	64905.64
190522_1	p66(1-556) (N-terminal His, cleaved)	64411.6646	64409.98

proteins for NMR - with C-terminal His to avoid truncation by TAG for tfmF labeling			
Date_#	Proteins	Measured (Da)	Theoretical (Da)
210618_1	p66 (C-terminal his)	65539.3935	65538.27
210521_1	p66_259_tfmF	65628.0941	65625.27
180327_2	p66_532_tfmF	65590.9873	65590.27
210819_1	p66_L289K	65551.6278	65553.29
201106_1	p66_L289K_259_tfmF	65641.3709	65640.29
180403_1	p66_L289K_532_tfmF	65604.3588	65605.29

p51 proteins for NMR - with N-terminal Strep to separately select p51 from p66.			
210728_1	p51	52595.3582	52596.6
210726_1	p51_L289K	52610.7329	52611.61

References

1. Pari K, Mueller GA, DeRose EF, Kirby TW, London RE. (2003) Solution structure of the RNase H domain of the HIV-1 reverse transcriptase in the presence of magnesium. Biochemistry. 42(3):639-50.
2. Sharaf NG, Poliner E, Slack RL, Christen MT, Byeon IJ, Parniak MA, et al. (2014) The p66 immature precursor of HIV-1 reverse transcriptase. Proteins. 82(10):2343-52.
3. Sharaf NG, Brereton AE, Byeon IL, Andrew Karplus P, Gronenborn AM. (2016) NMR structure of the HIV-1 reverse transcriptase thumb subdomain. J Biomol NMR.