Supporting Information

Drug Resistance Mutations Alter Dynamics of Inhibitor-Bound HIV-1 Protease

Yufeng Cai,[†, #] Wazo Myint,[‡, #] Janet L. Paulsen,[†] Celia A. Schiffer,[†,*] Rieko Ishima[‡,] and Nese Kurt Yilmaz[†,*]

[†] Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, United States

[‡] Department of Structural Biology, School of Medicine, University of Pittsburgh Biomedical Science Tower 3, 3501 Fifth Avenue, Pittsburgh 15260, United States

#-Authors contributed equally

*-Corresponding authors

NMR CT-CPMG R_2 dispersion Methods and Results

CT-CPMG R_2 relaxation dispersion data were analyzed in two steps: first, identify residues that undergo exchange, and second to calculate exchange parameters for a group of residues undergoing exchange.[1] In the first step, CT-CPMG R_2 relaxation dispersion data for each residue was fit using the following two models, i.e., with/without exchange in parallel, and the goodness of the fit was compared using partial F statistics as described below.

(I) No Exchange (Uniform R_2) Model: A uniform R_2 at all ν_{eff} values, assuming there is no significant exchange effect (equation (1)). Here, an intrinsic R_2 rate in the absence of chemical exchange (R_2^0) is optimized as a sole unknown parameter

$$R_2(\nu_{eff}) = R_2^0 \quad (1)$$

(II) Exchange (B-M) Model: A two-site exchange model that is expressed by Jen's solution to the Bloch-McConnell equations (see equation in reference[2]), in which a population of major state (p_a), a timescale of chemical exchange (τ_{ex}), a chemical shift difference between states undergoing chemical exchange ($\Delta\omega$), and R_2^0 are optimized.

The normalized χ^2/N values obtained by the fits to the Uniform R_2 and the B-M models were compared using partial F statistics. Residues with $F > 3.50$ ($\alpha = 0.01$) were identified as exhibiting significant conformational exchange. The F statistics used for selection of residues were verified by comparing with the qualitative measure of R_{ex}, derived from

$$R_{ex} = R_2(\nu_{eff} = 31.25 \text{ Hz}) - R_2(\nu_{eff} = 750 \text{ Hz}) \quad (2)$$

In the second step, groups of residues that showed $F > 3.50$ with ν_{ex} values within 1 standard deviation from the average within the group were selected for further analyzed by fitting the selected residues as a group with global p_a and τ_{ex} parameters. Once a group of residues is selected in either WT or Flap+, the same set of residues was subjected to the same analysis for consistence of comparison. The group fit was performed by a grid search, in which each grid point is defined by a pair of p_a and τ_{ex}, with minimization of χ^2 function for

optimization of local parameters R_2^0 and $\Delta\omega$ for individual residues in the group. Ranges of the grid were determined depending on the parameter ranges on the individual fits in the first step.

As an alternative to the B-M equation, group fitting was also performed by using the Luz-Meiboom equation (the L-M equation),[3] which is an approximate solution to the Bloch-McConnell equation:

$$R_2(\nu_{eff}) = R_2^0 + \Phi\tau_{ex}\left[1 - 4\cdot\nu_{eff}\tau_{ex}tanh\left(\frac{1}{4\cdot\nu_{eff}4\cdot\nu_{eff}\tau_{ex}}\right)\right] \quad (3)$$

with $\Phi = p_a(1-p_a)\cdot\Delta\omega^2$. The group fit using the L-M equation was performed with minimization of χ^2 function for optimization of τ_{ex} as a global parameter for the entire residues in the group, and Φ and R_2^0 as local parameters for individual residues in the group. For both B-M and for L-M equation fits, the uncertainties of optimized parameters were obtained from the confidence region that corresponds to the range with 1.0 higher χ^2/N than the minimum. [4]

For the apo form of HIV-1 protease, five residues in the flap region in the inhibitor-free WT and Flap+ were selected for a group fitting using the B-M equation (Figure S5). However, this fit yielded large uncertainties in τ_{ex} and p_a values (Table S1). On the other hand, fit using the L-M equation better fit the data and indicated significant differences between WT and Flap+ variants. In fitting using the L-M equation, not the populations but a composite parameter Φ defined by both p_a and $\delta\omega$ parameters is obtained. Assuming the chemical shift difference between the two populations $\delta\omega$ to be similar for the WT and the Flap+ HIV-1 PR, we can estimate the relative populations of the states undergoing exchange in the two protease variants. Although the two flap mutations G48V and I54V cause chemical shift changes between the WT and Flap+ variants, the above assumption was still tested as a first approximation because high F-values are observed throughout the entire region of the flap region, regardless of chemical shift changes due to mutations. The ratio of best fit Φ values for apo WT and Flap+ proteins for the same residues (residues 46, 48, 50, 51, and 52) used for group fitting showed that the WT values are on average 4 ± 2 times above the Flap+ values (Table S1). If the p_b of the Flap+ was estimated to be approximately 2% of the total population, the WT data would be approximately between 4-12% of the total population. This population shift is qualitatively consistent with

results obtained using the B-M equation. Overall, our current analyses suggest that Flap+ variant has a smaller population of the minor conformer than WT protease.

Table S1. The chemical exchange parameters for the WT and Flap+ HIV-1 proteases in apo form are shown.

Protein [a]	Fit using the B-M equation		Fit using the L-M equation	
	τ_{ex} (ms)	p_a [b]	τ_{ex} (ms)	$\Phi(s^{-2})$
WT	0.68 ± 0.13	0.992 ± n. d.	0.49 ± 0.025	6600 ± 3600
Flap+	1.3 ± 2.1	0.994 ± n. d.	0.89 ± 0.052	2400 ± 320

[a] For both fits, residues 46, 48, 50, 51, and 52 were included.

[b] Uncertainty for p_a was large and not determined.

Table S2. Ranksum analysis results for distances in Å between Cα atoms of residues in the flaps, 80s loop and the active site in WT and MT HIV-1 protease calculated over ten 100 ns trajectories. An h value of 1 signifies statistically significant difference between WT and MT ($p < 0.05$), z reflects the extent of this difference.

Atom Pair	p	h	z
b80b50	0.00	1	3.06
b25b50	0.16	0	1.40
a80b80	1.83E-04	1	3.74
a80b50	1.83E-04	1	3.74
a80a50	4.40E-04	1	3.52
a50b25	0.47	0	-0.72
a50a25	0.01	1	-2.46
a50b80	0.01	1	2.68
a50b50	0.47	0	-0.72
a25b50	0.05	0	-1.93

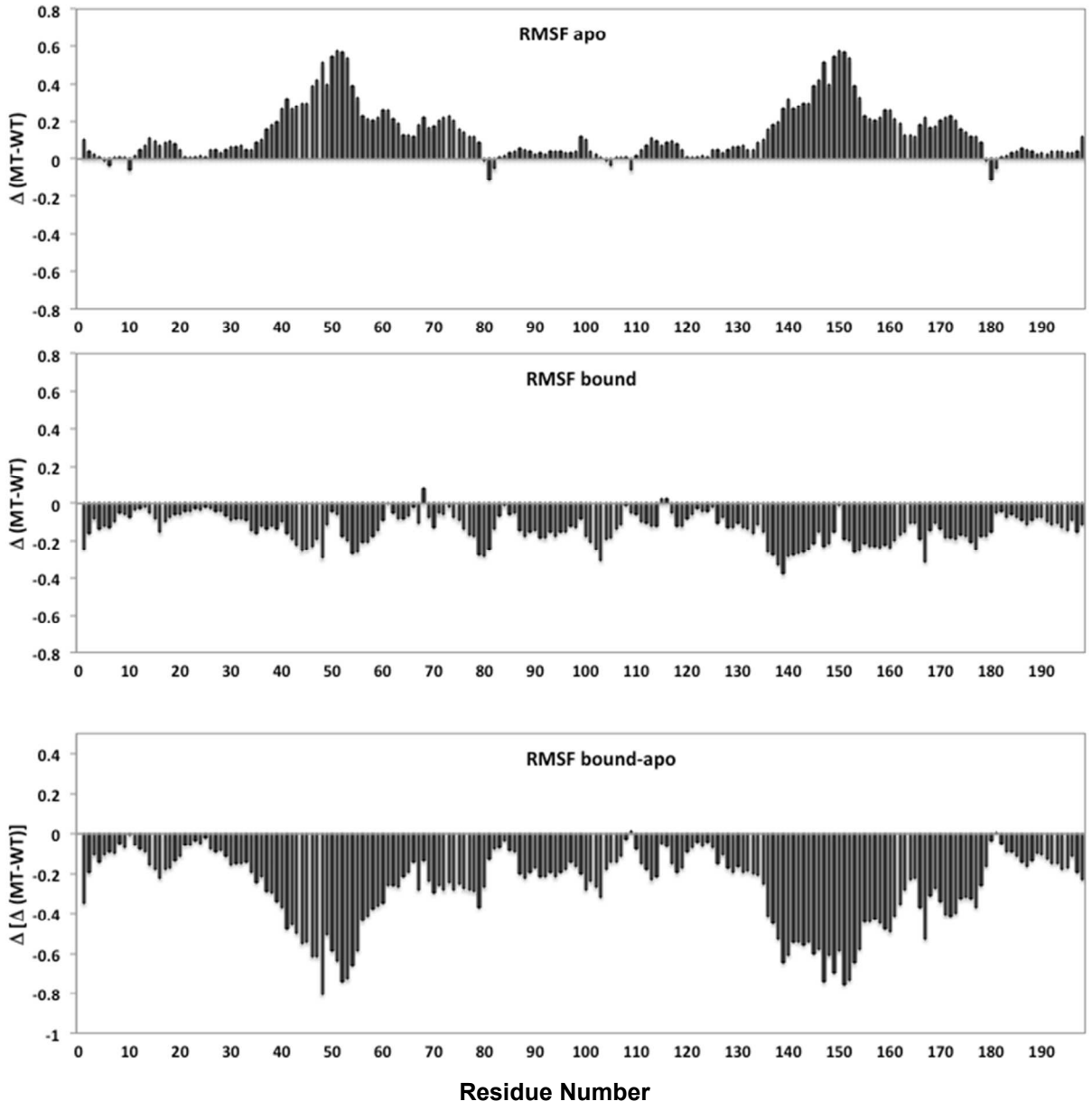

Figure S1. Difference in RMSF fluctuations of Cα atoms in Å during MD simulations between WT and MT protease in apo and bound forms. The predominantly positive and negative values, respectively, indicate that the mutant protease is more flexible in apo form and more rigid in inhibitor-bound form compared to the WT protease. The negative values in the double Δ plot shows that the flexibility loss due to inhibitor binding is more pronounced for MT compared to WT throughout the protease (0.27 ± 0.19 Å), and especially in the flap regions.

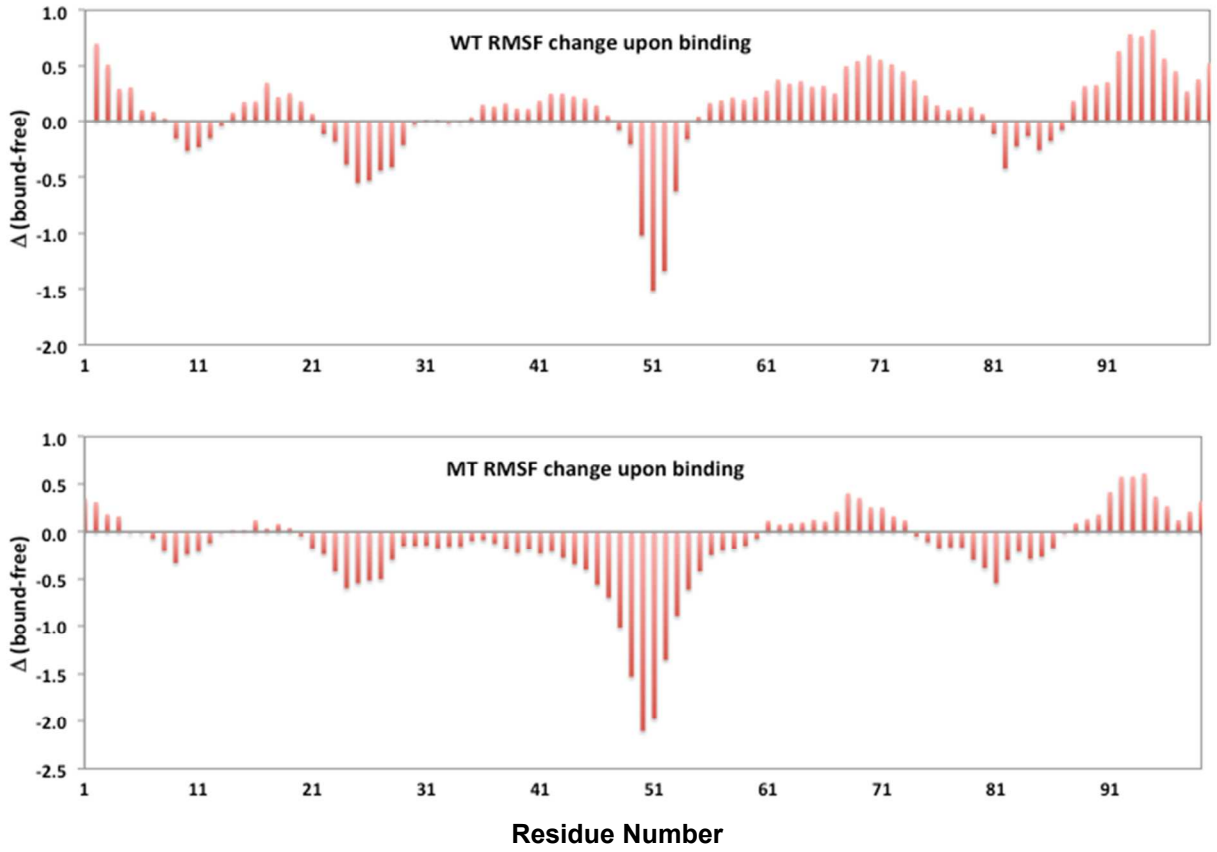

Figure S2. Change in RMSF fluctuations of Cα atoms in Å during MD simulations between bound and unliganded forms of protease. In both WT and MT protease, the active site region where the inhibitor binds loses considerable flexibility when inhibitor is bound (flap region around residue 50, catalytic loop including Asp25, and the 80s loop). This loss in flexibility due to inhibitor binding is more pronounced and widespread for the MT protease, especially in the flap regions.

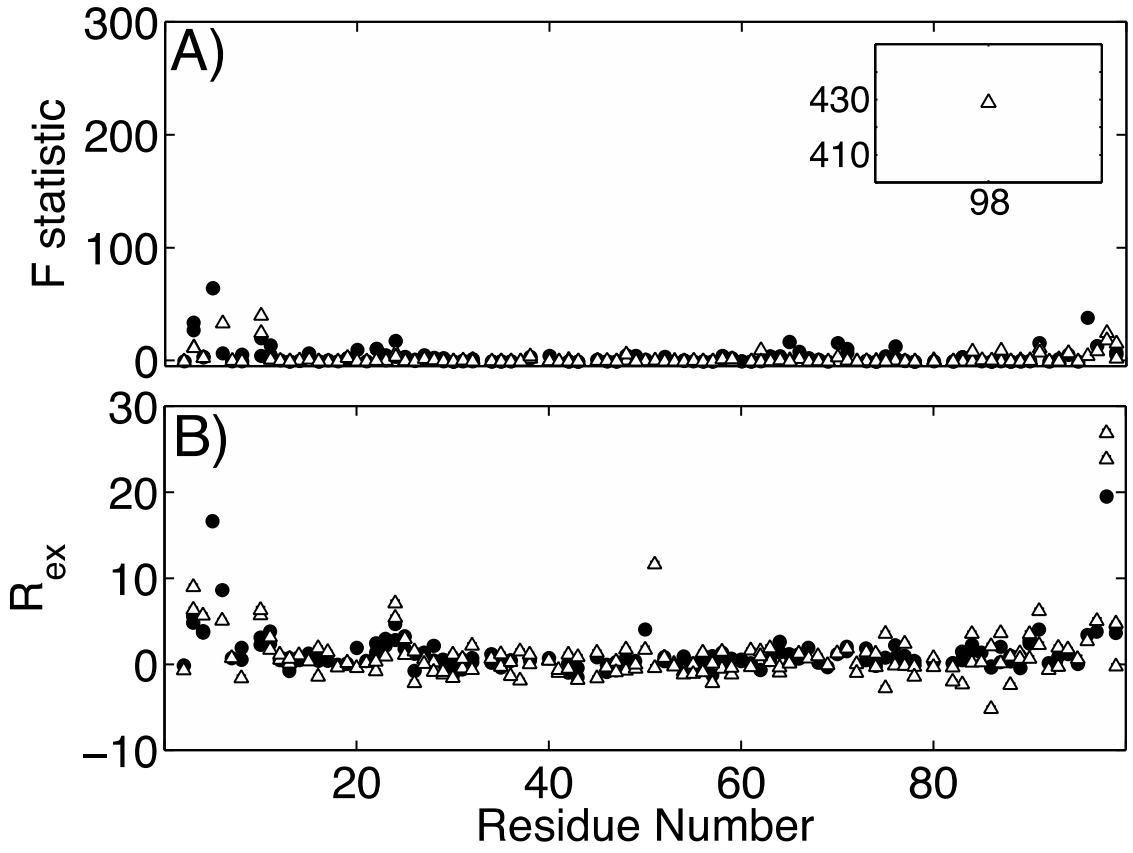

Figure S3. Conformational exchange due to motions in the ms-μs timescale in DRV-bound form. (A) The partial F statistic comparing the fits of individual residues and (B) Rex for WT (filled circle) and Flap+ (open triangle) HIV-1 PR bound to DRV. Two data points are displayed for some of the residues due to the loss of degeneracy between the two subunits on binding to asymmetric inhibitor DRV.

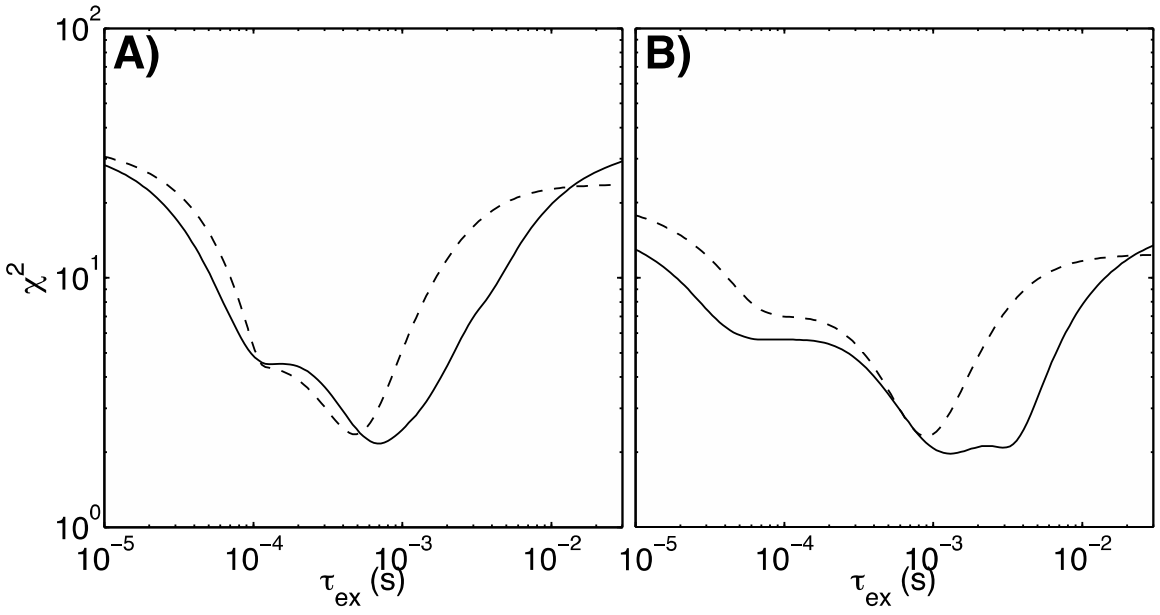

Figure S4. The best fit χ^2 values at specific values of τ_{ex} from group fitting to the Bloch-McConnell (solid line) equation and to Luz-Meiboom (dashed) equation for selected residues in the flap region (residues 43 to 58) for apo form (A) WT and (B) Flap+ HIV-1 PR. For fitting to the Bloch-McConnell equation, the p_a was kept at the lowest χ^2 based on grid search.

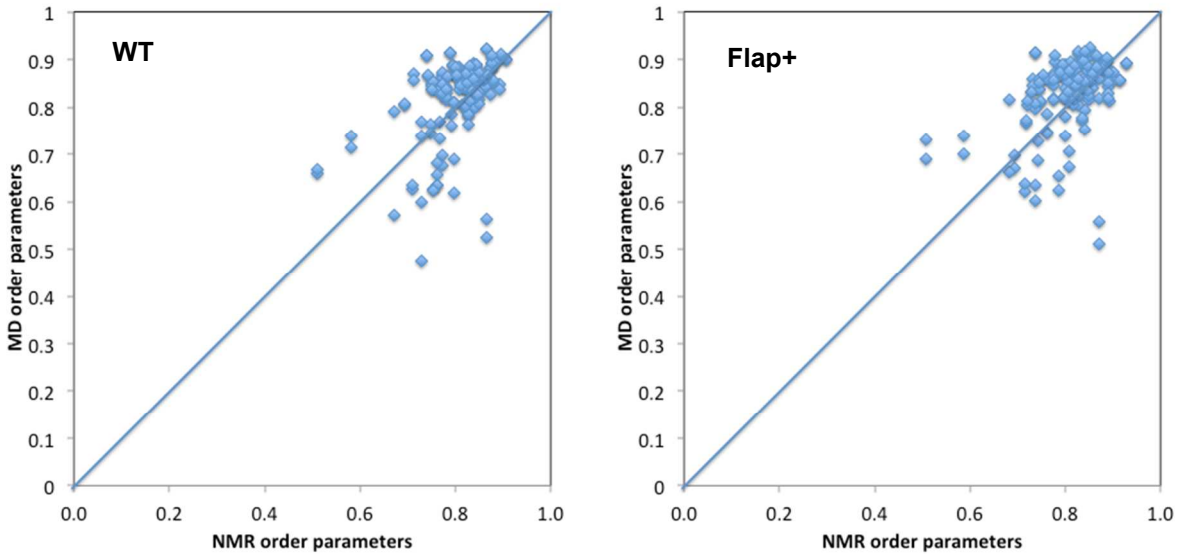

Figure S5. HIV-1 protease order parameters derived from MD simulation trajectories using a 1 ns time window plotted against those from NMR experiments. A reasonable agreement and a good match of the average order parameter value are reflected by the distribution of data around the diagonal.

References

1. Jee, J.; Ishima, R.; Gronenborn, A. M., Characterization of specific protein association by 15N CPMG relaxation dispersion NMR: the GB1(A34F) monomer-dimer equilibrium. *J Phys Chem B* **2008,** *112* (19), 6008-12.
2. Jen, J., Chemical Exchange and Nmr-T2 Relaxation. *Advances in Molecular Relaxation Processes* **1974,** *6* (2), 171-183.
3. Luz, Z.; Meiboom, S., Nuclear Magnetic Resonance Study of the Protolysis of Trimethylammonium Ion in Aqueous Solution—Order of the Reaction with Respect to Solvent. *The Journal of Chemical Physics* **1963,** *39* (2), 366-370.
4. Bevington, P. R.; Robinson, D. K., *Data reduction and error analysis for the physical sciences*. 2nd ed.; McGraw-Hill: New York, 1992; p xvii, 328 p.
5. Ishima, R.; Louis, J. M., A diverse view of protein dynamics from NMR studies of HIV-1 protease flaps. *Proteins* **2008,** *70* (4), 1408-15.