Supporting Information

Interdependence of Inhibitor Recognition in HIV-1 Protease

Janet L. Paulsen, Florian Leidner, Debra A. Ragland, Nese Kurt Yilmaz,

Celia A. Schiffer*

Department of Biochemistry and Molecular Pharmacology, University of Massachusetts

Medical School, Worcester, MA 01605, USA

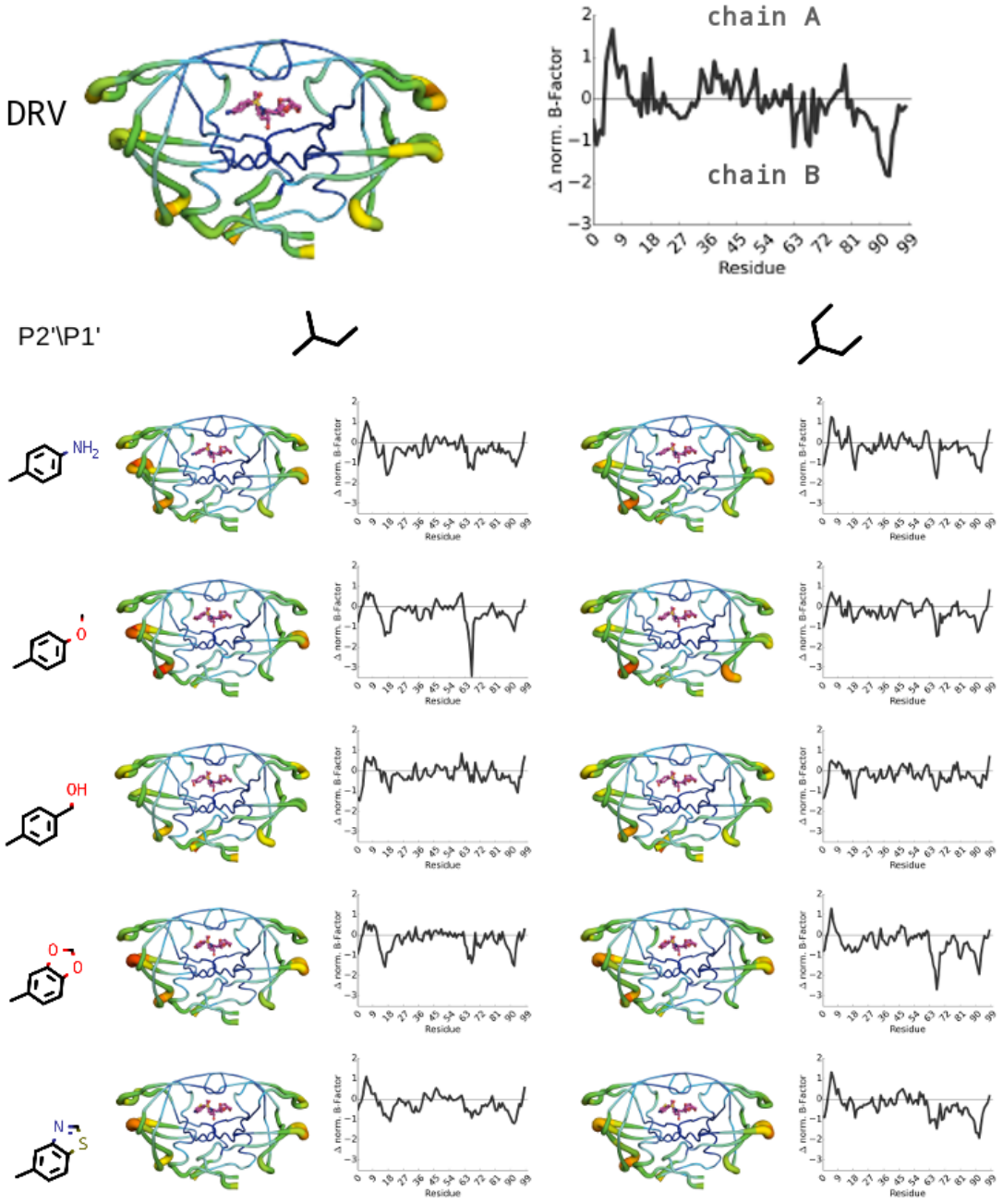

Figure S1. Comparison of the crystallographic B-factors across inhibitors and between protease monomers. Normalized crystallographic B-factors of the Cα atoms mapped onto the protease structure (left), and the difference in normalized B factors (ΔB-Factor) for the two monomers of HIV-1 protease, calculated as Chain A – Chain B (right).

Comparison of Crystallographic B-factors

Principal insights into protein dynamics can be gained from crystallographic B-factors [1] $B = \frac{8\pi^2\langle r^2\rangle}{3}$ where $\langle r^2\rangle = \langle u_x^2\rangle + \langle u_y^2\rangle + \langle u_z^2\rangle$ is the mean squared atomic displacement. The structure of HIV-1 protease in complex with the studied inhibitors had all been solved in the $P2_12_12_1$ space group, minimizing effects due to crystal packing. The resolution of the structures ranged from 1.20 Å (PDB ID: 1t3r) to 1.95 Å (PDB ID: 3o99)[2][3]. The B-factors of the Cα atoms were extracted from the PDB files and subsequently normalized by z-score normalization, $B_{zscore} = \frac{(B-\langle B\rangle)}{\sigma}$, where $\langle B\rangle$ is the arithmetic average and σ is the standard deviation of the Cα B-factors.

Darunavir and UMASS analogs form strong non-bonded interactions with the protease, restricting the atomic fluctuations at the active site. In the crystal structures of the co-complexes, active site residues had significantly lower B-factors compared to the residues distal from the active (Figure S1). Despite the asymmetric nature of the inhibitors, the B-factors of the active site residues showed little differences between the two monomers. Relatively large differences were observed distal from the active site: the N-terminal (residue 5-7) and C-terminal region had significant differences in dynamics. These residues are connected to the residues forming the S2 and S2' subsites through a network of intramolecular hydrogen bonding interactions. Two other regions displayed differences, and correspond to residues connecting the two distal beta-strands, residues 15 to18 and residue 63. These regions are connected to the active site through the hydrophobic core.

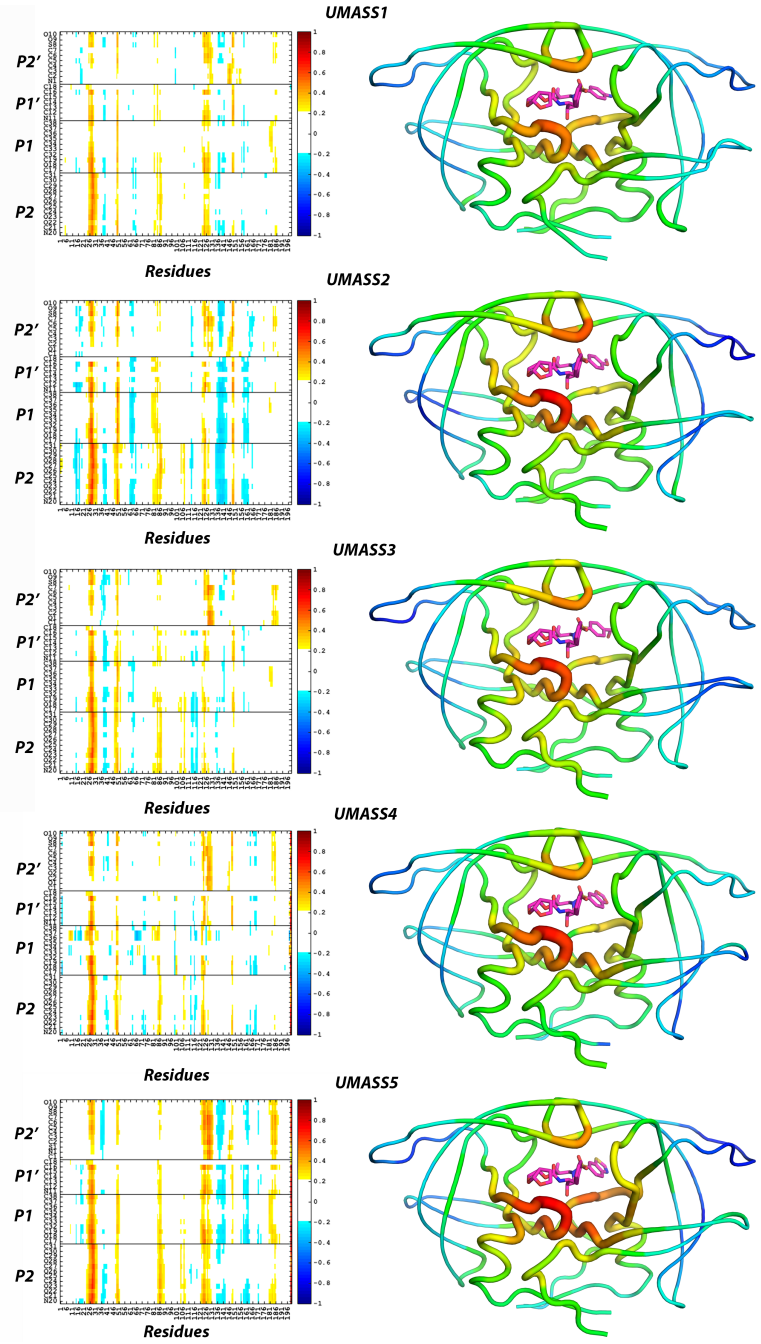

Figure S2. Cross-correlations in atomic fluctuations between UMASS1–5 inhibitor atoms and the C-alpha atoms of HIV-1 protease. Average cross-correlation intensities by residue determined from the matrix (left) are mapped onto the protease structure (right), colored blue to red and depicted thin to thick for increasing correlations.

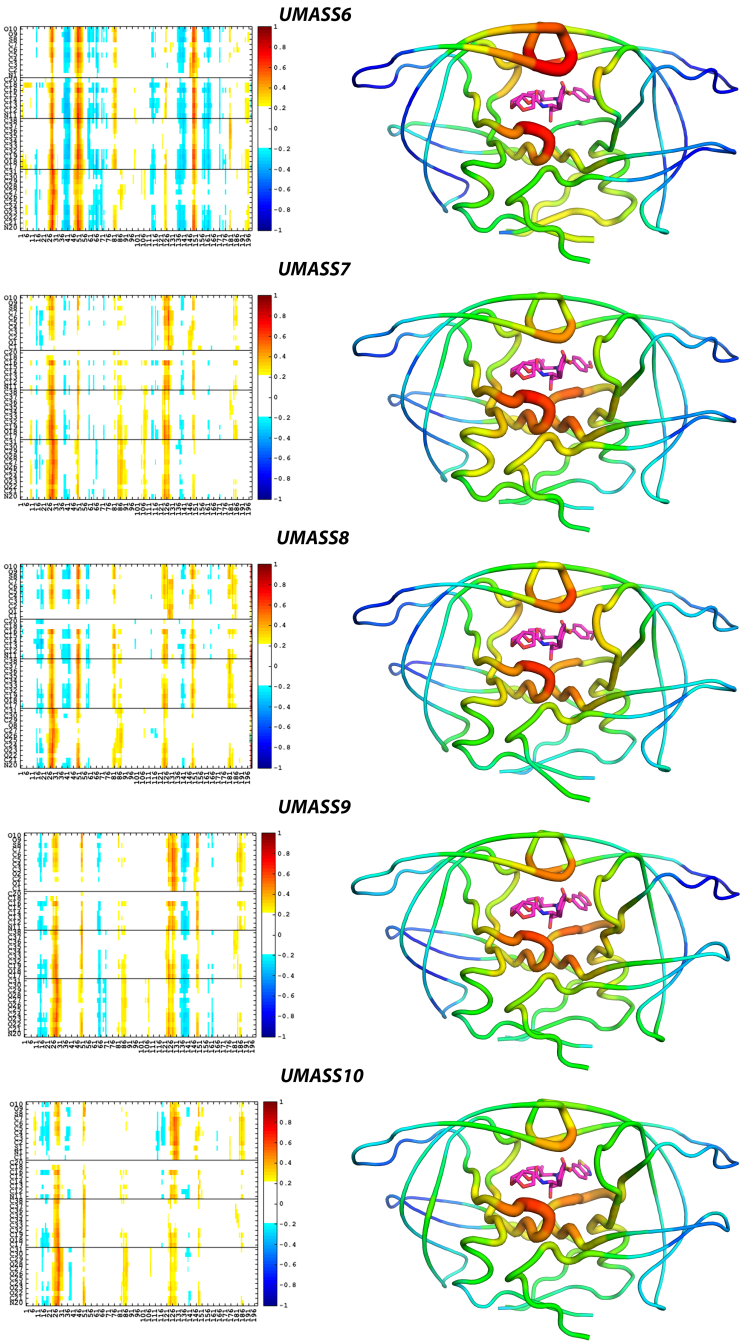

Figure S3. Cross-correlations in atomic fluctuations between UMASS6–10 inhibitor atoms and the C-alpha atoms of HIV-1 protease. Average cross-correlation intensities by residue determined from the matrix (left) are mapped onto the protease structure (right), colored blue to red and depicted thin to thick for increasing correlations.

Distance difference matrices show asymmetric inhibitors impact homodimeric protease monomers differentially

Distance difference matrices were generated using the C-alpha coordinate positions of each amino acid from the MD trajectories to determine the average distance between all C-alpha atom pairs. Distance difference matrix is useful for assessing subtle differences between structurally related proteins without superposition bias.

DRV-bound protease was compared to protease bound to each DRV analog. The distance difference plots clustered into two groups. Compounds with a phenyl substituted moiety in the P2' position exhibited a similar pattern, (Figure S4A) for DRV versus UMASS1, where residues I15', G16', G17' and Q18' in the prime monomer (adjacent to the P2' pocket) showed the largest differences in distance between C-alpha pairs (see Figure S4B). Residues I15', G16', G17' and Q18' in protease structures bound to UMASS1, 2, 3, 6, 7 and 8 were observed to deviate as much as 3 Å and on average 1.3 Å when compared to protease bound to DRV. Differences with these same residues were observed in the opposite monomer (non-prime, monomer bound to the *bis*-THF moiety) but always to a lesser degree. An asymmetrical effect is observed in identical and symmetrical monomers in residues distal to the active site that is propagated from the asymmetrical inhibitor bound at the active site. Observations for DRV versus UMASS1 were similar for UMASS 2, 3, 6, 7 and 8 when compared to DRV (Figure S5.) The second group was UMASS 4, 5, 9 and 10 all possessing fused ring systems in the P2' position. When compared with DRV, on average deviations were much less than 1 Å (Figure S6). Congeneric pairs were also compared; however no significant features were observed (Figure S7).

In summary, phenyl substituted P2' compounds (UMASS1, 2, 3, 6, 7 and 8) compared to the fused ring substituents (UMASS 4, 5, 9 and 10) exhibited greater C-alpha distance differences compared to DRV in the prime monomer. The residues with the greatest deviation however, are not in direct contact with the P2' moiety but are distal to the active site (residues I15', G16', G17' and Q18'). The effect of the phenyl substitution in the P2' pocket is propagated differently to the prime and non-prime monomers of HIV-1 protease with the prime monomer always observed to deviate the greatest for all DRV analogs.

Materials and Methods: *Distance Difference Plots*

Distance difference plots were generated to reveal differences in protease structure when bound to DRV versus the DRV analogs. All C-alpha–C-alpha distances within each structure were calculated. Then differences between these distances in any two structures were calculated to determine where the largest differences occur. Calculations and plots were generated using custom python scripts developed in conjunction with Dmitry Lupyan (Schrödinger, Inc).

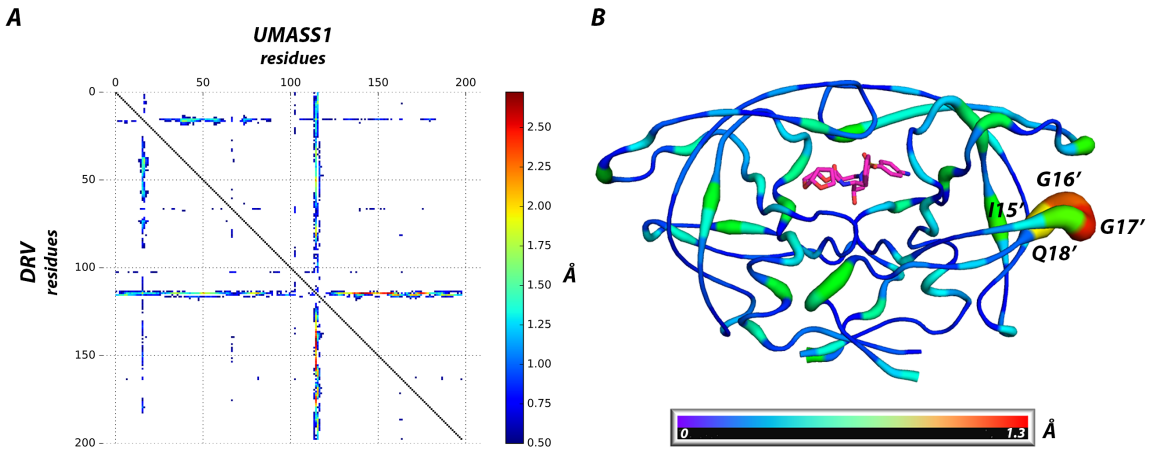

Figure S4. A) Distance difference plot of HIV-1 protease bound to DRV versus UMASS1. The absolute values of the C-alpha distance differences are plotted between DRV and UMASS1. B) Average distance differences between DRV and UMASS1 in panel A mapped onto HIV-1 protease structure. Residues with highest deviations are labeled.

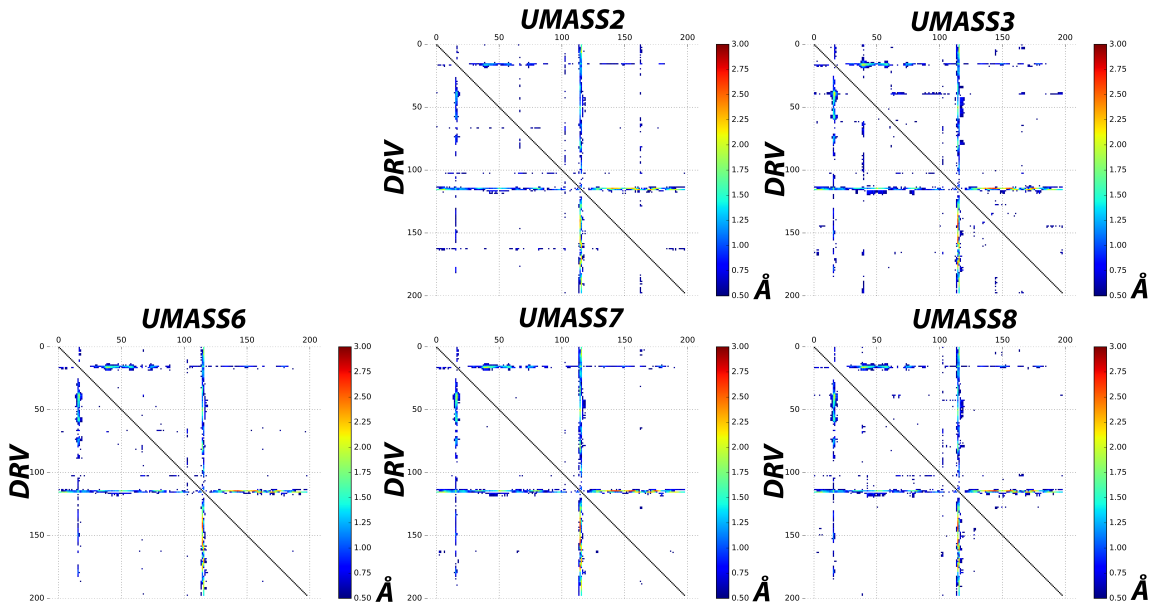

Figure S5. Distance difference plots of HIV-1 protease bound to DRV versus compounds that contain a substituted phenyl group at the P2' position (UMASS2, 3, 6, 7 and 8).

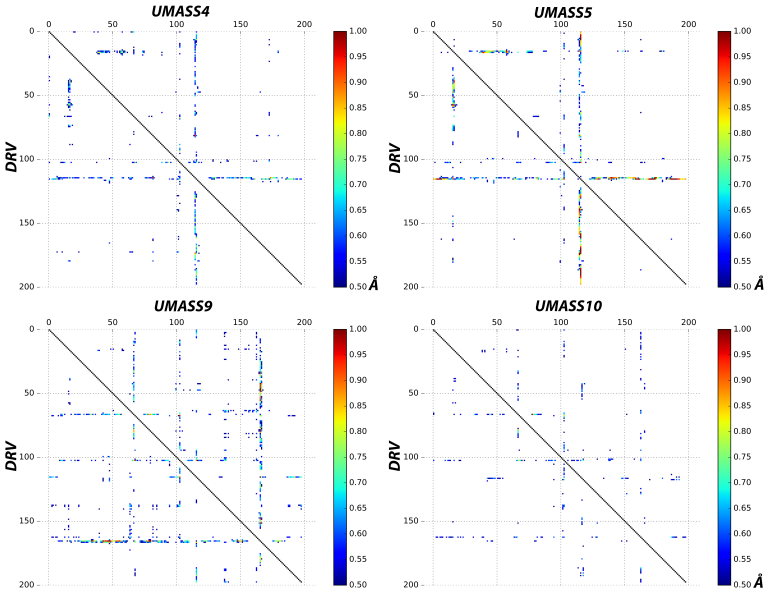

Figure S6. Distance difference plots of HIV-1 protease bound to DRV versus compounds that contain a fused ring moiety at the P2' position (UMASS4, 5, 9 and 10).

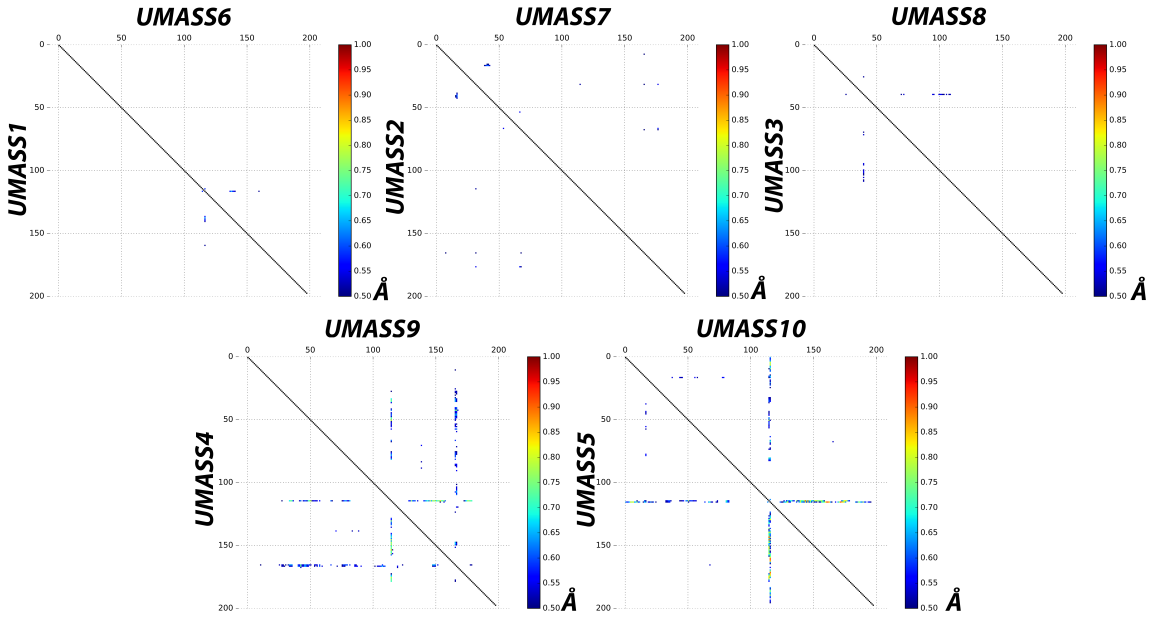

Figure S7. Distance difference plots of HIV-1 protease bound to congeneric pairs that only differ at the P1' position (UMASS 1-5, methyl butyl and UMASS6-10, ethyl butyl).

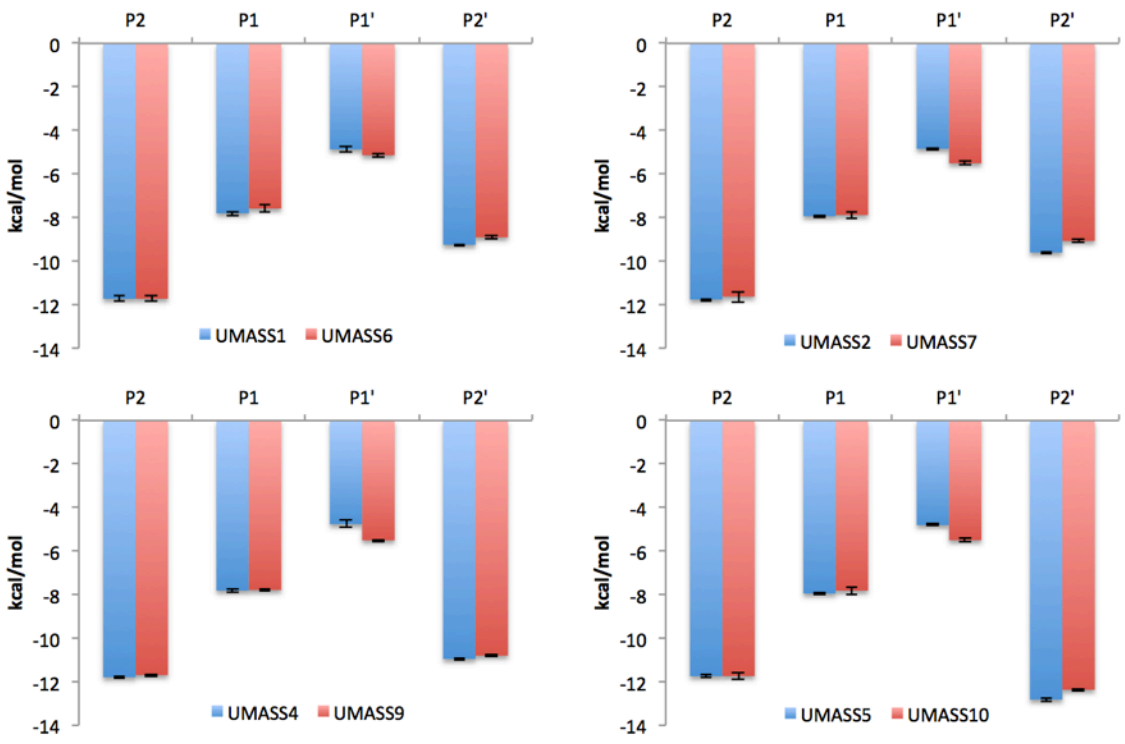

Figure S8. Intermolecular van der Waals contact energies by subsite between DRV analogs and HIV-1 protease for the congeneric pairs that differ only at the P1' position.

Interdependency of S1' and S2' subsites mediated by P2' moiety

Additional comparisons of the DRV, UMASS1 and UMASS were made as this group of congeners was the only one to show a dependency between the P2' moiety and interactions at the S1' subsite. DRV, UMASS1 and UMASS 6, each has an aniline group at the P2' position but differs in carbon length at P1' moiety. A similar pattern and frequency of hydrogen bonding existed for DRV and UMASS6 (Figure S9 & Table S1). However, UMASS1 had a hydrogen bond between the P2' moiety and the backbone nitrogen of D30 rather than the backbone oxygen as in DRV and UMASS6. This hydrogen bond between the P2' moiety of UMASS1 and the D30' nitrogen was observed at a lower frequency (19%) than the corresponding frequency of the hydrogen bond between the P2' moiety of DRV and UMASS6 and the D30' oxygen (71% and 68% respectively). UMASS1 compensated for this loss by a dramatic increase in the frequency of the hydrogen bond between the P1' moiety and terminal oxygen of catalytic D25 (78%) over the frequency that this same hydrogen bond was observed for DRV and UMASS6 (50% and 34% respectively).

The hydrogen bonding patterns and frequencies were similar for DRV and UMASS6, as both of these ligands have symmetrical functional groups in the P1' subsite. Whereas UMASS1, which has an asymmetric functional group in the P1', differed in pattern and frequency in the P1 and P2' subsites suggesting that the symmetry of the hydrophobic group at P1' altered the hydrogen bonding patterns in the P2' pocket. However this dependency appears to be mediated by the functional group in the P2' position as the asymmetry-induced affect was not observed for other congeneric pairs but only for the compounds with the aniline group at the P2' position.

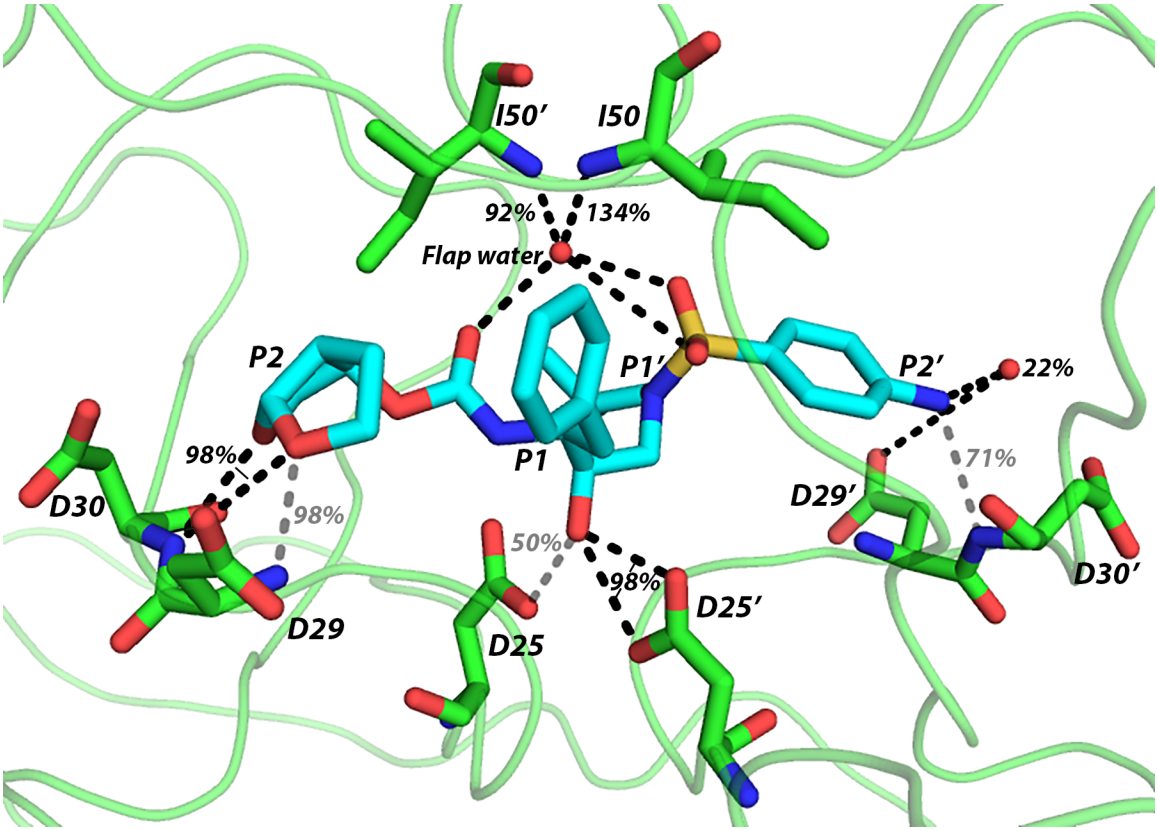

Figure S9. Hydrogen bonding pattern and frequencies between HIV-1 protease and DRV, averaged over 300 ns MD simulations. Where more than one hydrogen bond interact with a given residue (black dash lines) the frequencies were summed. Sums are listed in Supplementary Table 1 and 2.

Table S1. Direct hydrogen bonds and frequencies between HIV-1 protease residues and DRV analogs over 300 ns MD simulations.

Subsite	Residue	Atom	UMASS1	UMASS2	UMASS3	UMASS4	UMASS5
P2	D29	N[a]	98%	98%	97%	97%	97%
	D30	N	82%	82%	82%	78%	81%
P1	D25	OD2	78%	78%	72%	64%	77%
P1'	D25'	OD1/2[a]	97%	97%	97%	99%	98%
P2'	D29'	N			21%		
	D30'	OD1/2[a]			68%		
	D30'	O					
	D30'	N	19%	19%	27%	54%	74%

Subsite	Residue	Atom	UMASS6	UMASS7	UMASS8	UMASS9	UMASS10
P2	D29	N[a]	97%	97%	97%	97%	98%
	D30	N	76%	83%	85%	75%	79%
P1	D25	OD2	34%	62%	28%	73%	59%
P1'	D25'	OD1/2[a]	97%	97%	96%	99%	98%
P2'	D29'	N			12%		
	D30'	OD1/2[a]			63%		
	D30'	O	68%				
	D30'	N		14%	29%	50%	69%

[a] *Sum of two or more hydrogen bonds between inhibitor and residue atom.*

Table S2. Water-mediated hydrogen bonds between HIV-1 protease residues and DRV analogs over 300 ns MD simulations.

Subsite	Residue	Atom	UMASS1	UMASS2	UMASS3	UMASS4	UMASS5
P1	I50	N[a]	117%	141%	135%	128%	131%
P1'	I50'	N[a]	84%	78%	67%	84%	83%
P2'	D29'	OD1/2[a]	16%		32%		

Subsite	Residue	Atom	UMASS6	UMASS7	UMASS8	UMASS9	UMASS10
P1	I50	N[a]	128%	135%	125%	111%	117%
P1'	I50'	N[a]	109%	63%	48%	101%	86%
P2'	D29'	OD1/2[a]	16%		49%		

[a] *Sum of two or more hydrogen bonds between inhibitor and residue atom.*

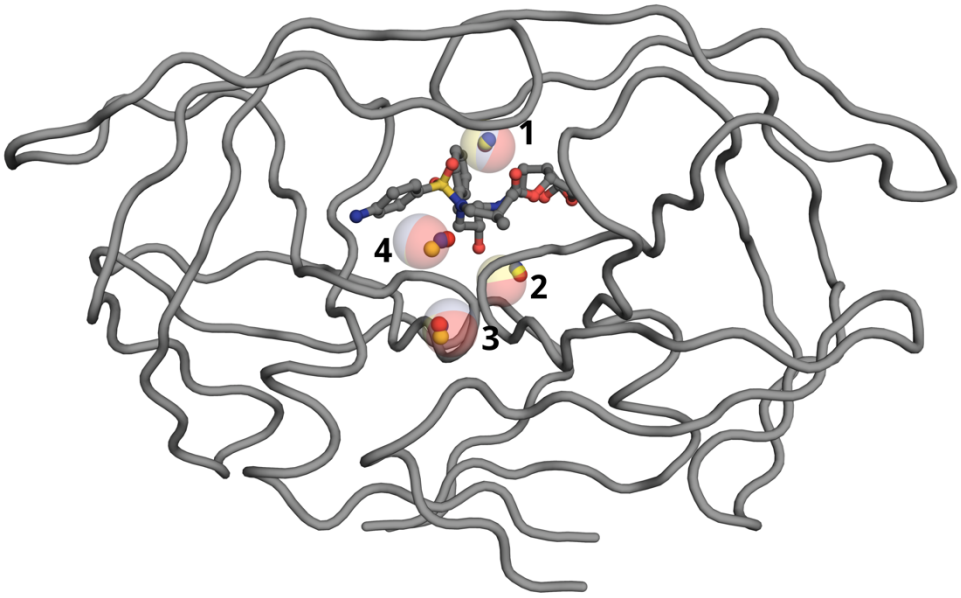

Figure S10. High occupancy water sites are conserved in the crystal structures. High occupancy water sites, which are conserved across all 11 (DRV and the 10 UMASS analogs) inhibitor–protease complexes, are depicted as solid spheres on the HIV-1 protease–DRV complex structure (PDB ID: 1T3R). The location of the nearest crystallographic water molecules are shown as transparent spheres. Color scheme and numbering adopted from Figure 7.

Supporting Information References:

[1] Trueblood, K. N., et al. "Atomic dispacement parameter nomenclature. Report of a subcommittee on atomic displacement parameter nomenclature." *Acta Crystallographica Section A: Foundations of Crystallography* 52.5 (1996): 770-781.

[2] Nalam, Madhavi NL, et al. "Substrate envelope-designed potent HIV-1 protease inhibitors to avoid drug resistance." *Chemistry & Biology* 20.9 (2013): 1116-1124.

[3] Surleraux, Dominique LNG, et al. "Discovery and selection of TMC114, a next generation HIV-1 protease inhibitor." *Journal of Medicinal Chemistry* 48.6 (2005): 1813-1822.