Supplementary information

The main article describes the results of simulation of HIV-1 protease wildtype and resistant variants bound to lopinavir and ritonavir showing that cooperative interactions between mutations alter the level of water ingress to the binding site leading to differing resistance levels. In this document we provide details of the simulation and analysis methodology, figures showing the common backbone of the two inhibitors together with detailed binding free energy and water access data.

Supplementary Methods

Preparation and simulation set up was performed using the automated Binding Affinity Calculator (BAC) tool (full details of the tool and their simulation parameters employed are available in Sadiq et al. (*1*)). Models of the subtype B wildtype sequence of HIV-1 were constructed using the coordinates from PDB crystal structures 1MUI and 1HXW for LPV and RTV (*2, 3*) respectively. All mutations were inserted *in silico* using VMD (*4*), mutations being applied to both monomers of the protease homodimer. When it is necessary to distinguish residues in either monomer the residues of the first monomer in the PDB are labelled 1-99, those in the second monomer 100-199.

It is well established from the many crystal structures of ligand bound HIV-1 protease complexes that a water molecule mediates flap-ligand interactions (*5*). This water molecule is not resolved in the 1MUI crystal structure; consequently a water molecule was inserted in the correct tetrahedrally coordinated geometry between lopinavir and the protease flaps.

All systems were solvated in orthorhombic water boxes (using the TIP3P water model) with a minimum extension from the protein of 14 Å. Protein parameters were taken from the standard AMBER force field for bioorganic systems (ff03) (*6*). Drug parameters were produced using the general AMBER force field (GAFF) (*7*) following the procedure detailed in Sadiq et al (*1*). Before the production simulations reported here were run all systems were minimised and equilibrated for 2 ns equilibration using the protocol defined by the BAC. All simulations presented here were performed in the molecular dynamics package NAMD2 (*8*) in the NPT ensemble with a temperature of 300 K and a pressure of 1 bar, using a 2 fs time step. Langevin thermo- and barostats were used to couple temperature and pressure. For each system 50 simulations were performed. The only initial variation in each replica of the same sequence was the randomly seeded Maxwellian velocity distribution assigned to the atoms within the system. Individual LPV (WT/HM) and RTV (WT/HM) simulation systems were extended to 100, 50, 20, and 20 ns respectively, without leading to changes in water binding at the drug/catalytic residue interface. Backbone RMSD calculations of the ensemble simulations indicate a convergence to 1.6 Å after ~0.25 ns in each system, which changes minimally in extended simulations (due to loop motions).

RTV “swap” simulations were based on output frames from LPV trajectories. Crystal structure conformations of RTV were fitted to the structure of LPV based on 6 carbon and nitrogen atoms forming the diamine-alcohol backbone, before replacing LPV with RTV. Clashing water molecules were removed and the structure energy minimised for 2000 steps and simulations restarted from the new structure. LPV mutant reversion simulations (mutations from HM and AS, to WT) were generated from the same output frames and mutating back to WT using VMD, removing clashing water molecules.

Free energy analysis was conducted using the MMPBSA module of the AMBER 9 package (*9*). The MMPBSA computations were applied to configuration snapshots generated at a frequency of 100 snapshots/ns over the course of MD simulations. Further analyses and visualisation were performed using VMD. Free energies calculated by MMPBSA alone were used throughout the studies. Whilst

entropy can be taken into account through normal mode calculations (*10*), we previously found that predicted energies from MMPBSA alone correlated better with experimental data (*11*) due to noise introduced by NMODE calculations.

Hydrogen bonding frequency was calculated in VMD using a cut off distance of 3.5 Å and an angle of 30 °. The standard deviation of the bootstrapping means was found to be in all cases less than 0.01.

Catalytic Dyad Protonation State Assignment

Prior to a calculation of the binding free energies of the mutant complexes the protonation state of the catalytic dyad (D25/D250) when complexed to lopinavir and ritonavir was investigated using the MMPBSA methodology and normal mode analysis (again performed in AMBER 9, see (*9*) for full protocol details). These states were then used to assign the protonation states of each of the corresponding mutant-lopinavir complexes. Four systems, each consisting of a different catalytic-dyad protonation state were generated for the wildtype bound to each drug. These were denoted as the di-anionic (D-), diprotonated (D25,250), Asp 25 protonated (D25) and Asp 250 protonated (D250) states. Ensemble production simulations were performed up to 20 x 4 ns for each system (except for D-, for which 10 x 4 ns were performed) and results computed at the frequency mentioned above. The results suggest that the drugs are tightest bound in the D25 case for both lopinavir and ritonavir (see Supplementary Methods Table 1) and this was used for all subsequent simulations.

Protonation State	ΔG_{MMPBSA}	$-T\,\Delta S_{NM}$	ΔG_{theor}
Lopinavir (LPV)			
D25	***-47.70***	***37.29***	***-10.41***
D125	-46.13	38.92	-7.21
D25/D125	-47.67	39.85	-7.82
D-	-32.15	38.23	6.08
Ritonavir (RTV)			
D25	**-57.48**	**37.06**	**-20.42**
D125	-56.14	38.64	-17.69
D25/D125	-53.72	38.33	-15.39
D-	-37.95	42.03	4.08

Supplementary Methods Table 1. Assessment of catalytic dyad protonation state for wildtype HIV-1 protease bound to the inhibitors lopinavir and ritonavir. All values are in kcal mol^{-1}. Data for lopinavir taken from previous work in our group (*1*). ΔG_{MMPBSA} = free energy calculated from MMPBSA, $-T\,\Delta S_{NM}$ = estimation of entropy (from normal mode analysis) multiplied by temperature, ΔG_{theor} = total estimated free energy of interaction

Computational Infrastructure

Simulations were performed on the Kraken machine at the National Institute for Computational Infrastructure (NICS) and the Legion cluster at University College London (UCL). MMPBSA analysis was performed on our local cluster and the Ranger machine at the Texas Advanced Computing Centre (TACC). The analysis performed on Ranger was facilitated by the use of BigJob functionality provided by the Simple API for Grid Application (SAGA) (*12*).

Supplementary Figures

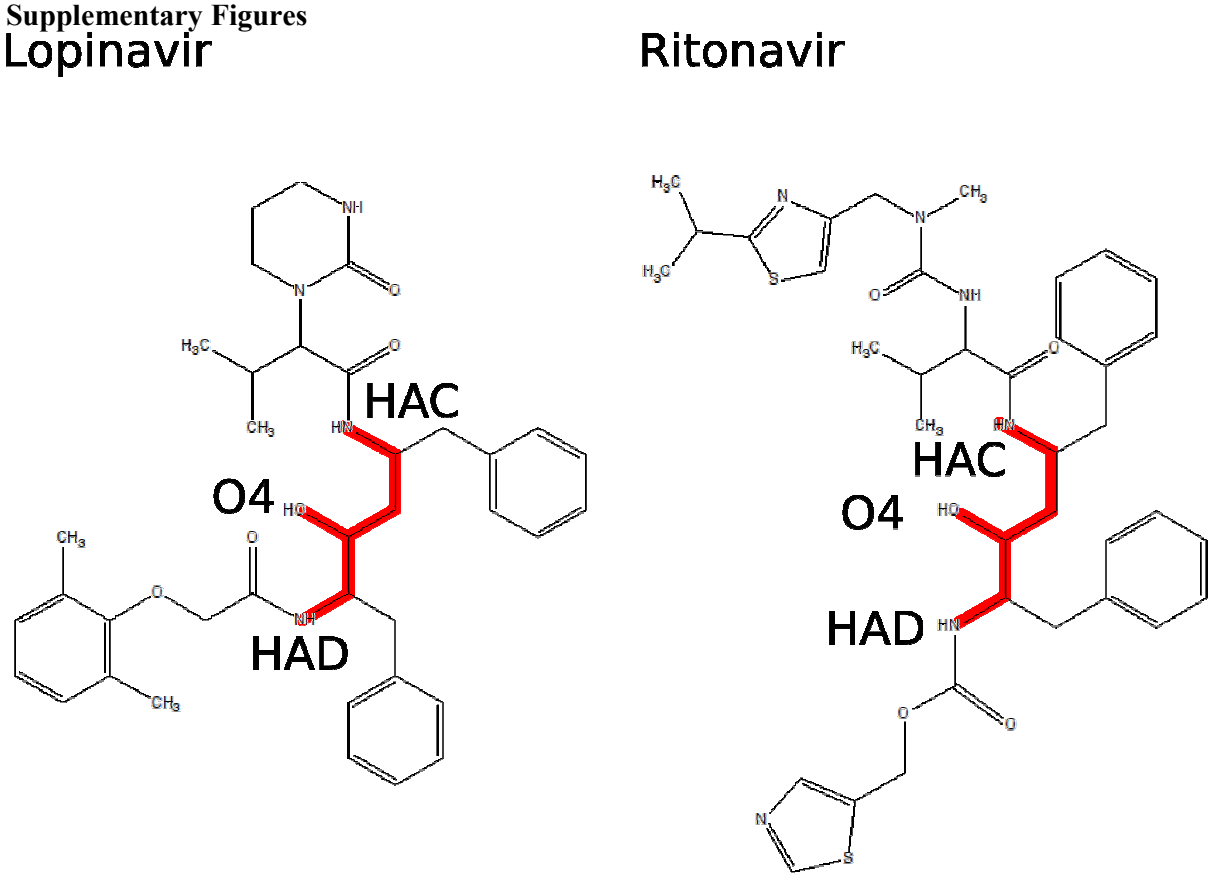

Figure S1. Lopinavir and ritonavir structures. Common diamine/alcohol backbone highlighted in red and specific water interaction sites on the backbone labelled with the names used in this article

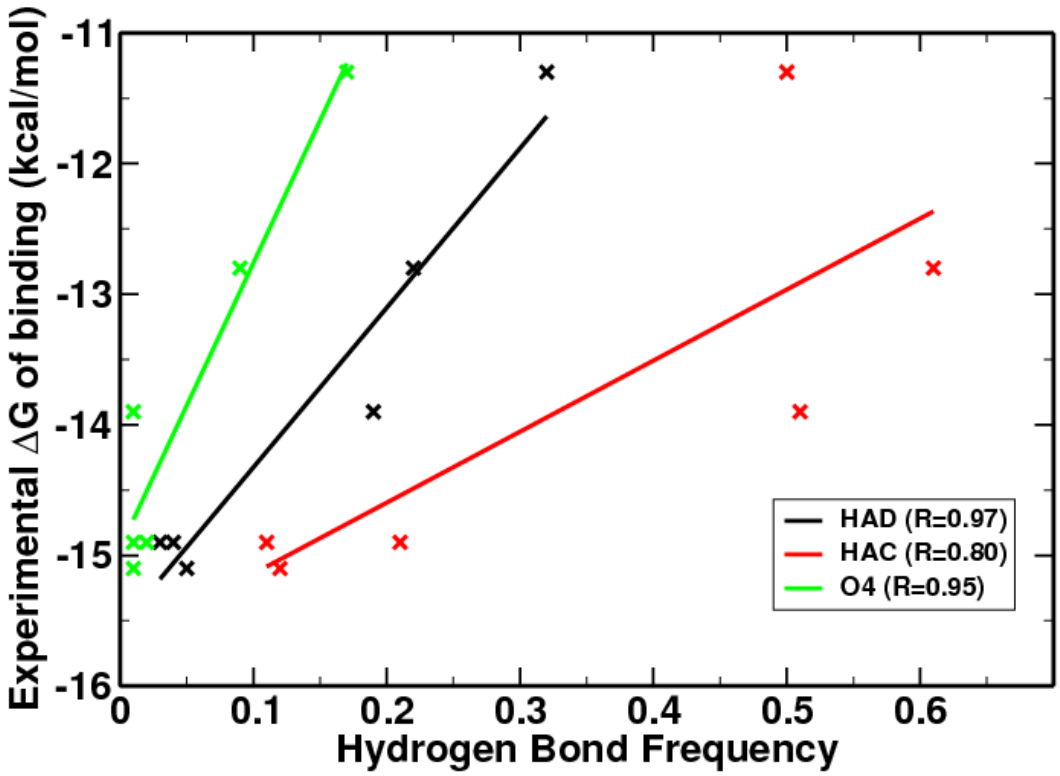

Figure S2. Correlation of hydrogen bond frequency with experimentally determined free energies of binding. Whilst water binding at all three defined sites shows correlation with the experimental results, the HAD site alone shows responsiveness across the complete range of experimentally determined results (i.e. O4 cannot discriminate between AS and WT/FL/DM mutations), and the highest correlation.

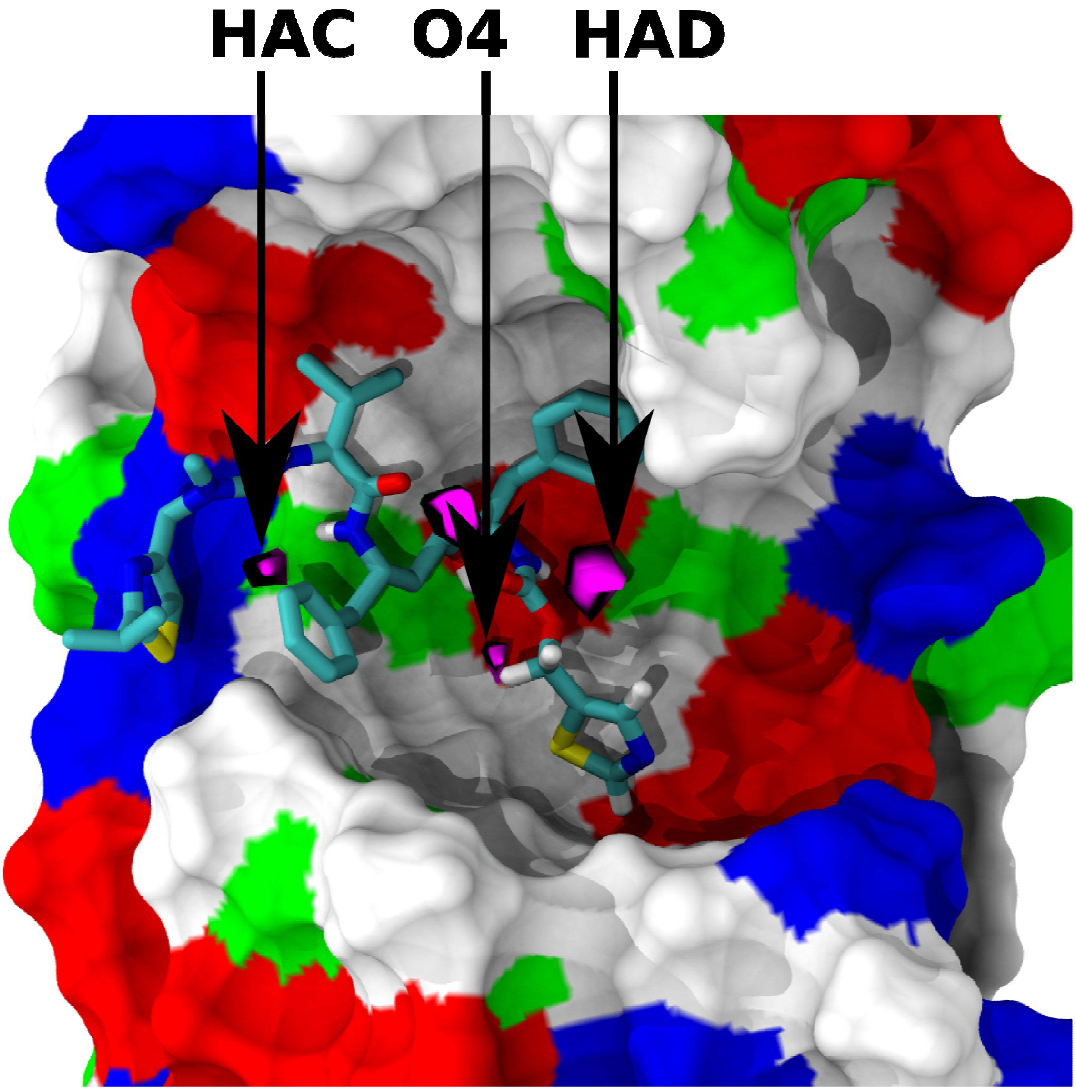

Figure S3. Images of the water binding environments in the HIV-PR active site. Non polar residues are white, polar green, basic blue, and acidic red, with water pockets indicated in magenta and drug molecule colored by atom type. Both HAD and HAC sites are adjacent to carbonyl oxygen atoms, whilst the HAD and O4 sites also interact with the catalytic aspartate residues in the core. Water entry occurs along this hydrophilic trench.

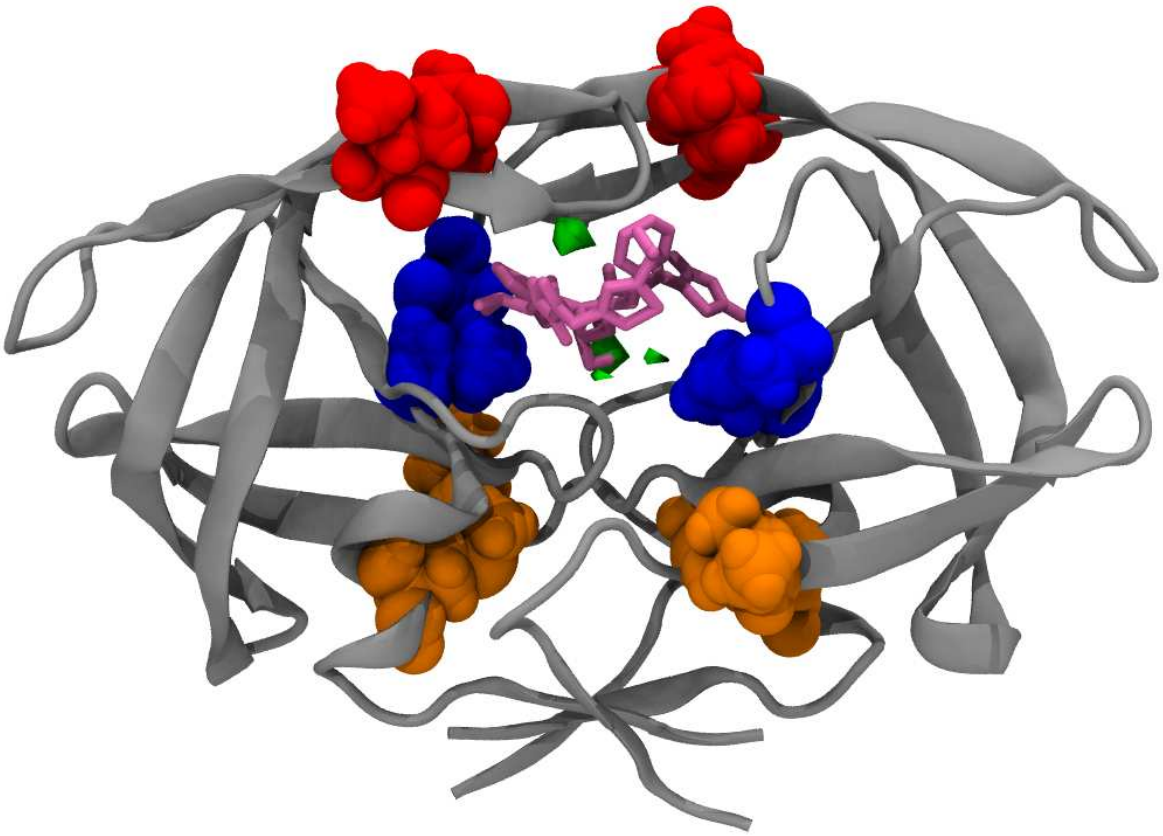

Figure S4. Mutations examined in this study, overlaid with protein backbone (grey cartoon) water binding pockets (green) and drug molecule pink. Mutations, rendered as spacefill models are coloured by type: AS blue, FL red, and DM orange.

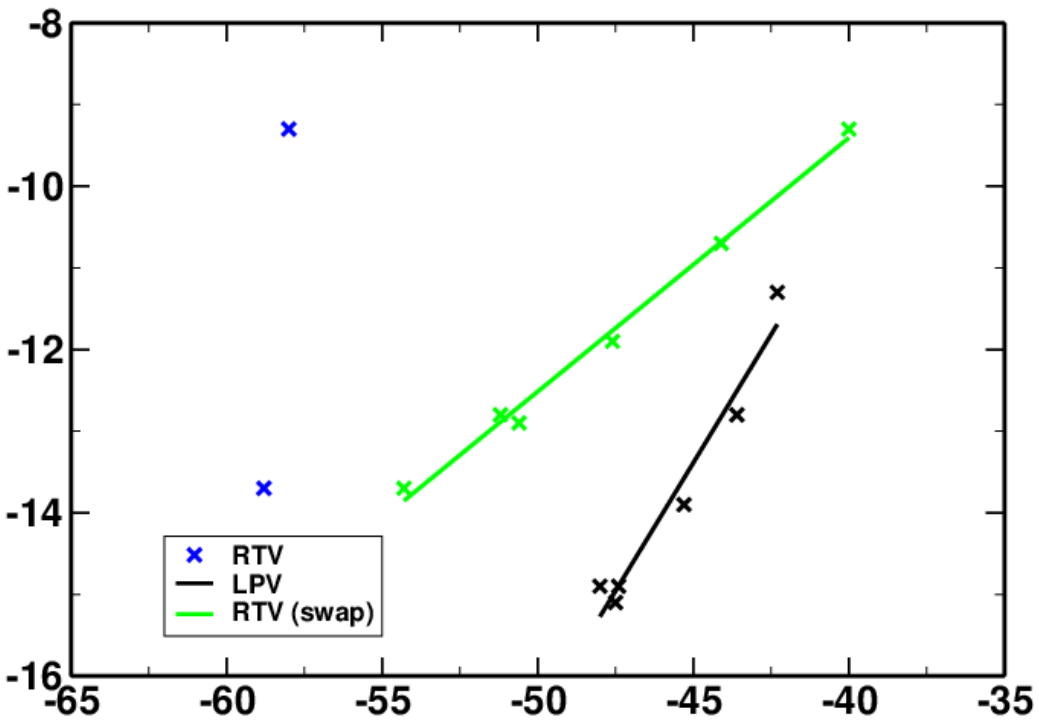

Figure S5. Correlation of MMPBSA calculated energy with experimental energies. Whilst both LPV and RTV (swap) results show clear correlation with experiment, the absolute values are incorrect preventing cross drug comparison. In contrast, RTV results for WT and HM show that discrimination between these mutations outside of the swap protocol is not possible.

Supplementary Tables

	HAD	HAC	O4	O1	Exp. ΔG kcal/ mol
LPV					
WT	0.05	0.12	0.01	0.47	-15.1
FL	0.04	0.11	0.02	0.45	-14.9
DM	0.03	0.21	0.01	0.44	-14.9
AS	0.19	0.51	0.01	0.43	-13.9
QM	0.22	0.61	0.09	0.45	-12.8
HM	0.32	0.5	0.17	0.52	-11.3
LPV Reversions to WT					
From AS	0.08	0.40	0.19	0.58	-15.1
From HM	0.01	0.43	0.04	0.51	-15.1
RTV					
WT	0.01	0	0	NA	-13.7
HM	0.02	0	0	NA	-9.3
RTV Swap					
WT	0.03	0.02	0.01	NA	-13.7
FL	0.04	0.01	0.04	NA	-12.8
DM	0.01	0.02	0.02	NA	-12.4
AS	0.16	0.08	0.05	NA	-11.9
QM	0.25	0.15	0.06	NA	-10.7
HM	0.44	0.24	0.19	NA	-9.3

Table S1. Fractional frequency of water hydrogen bonding at different sites and experimentally determined energy of binding (Exp. ΔG kcal/ mol, Ohtaka et al, 2003). O1 site is a bulk solvent exposed oxygen atom included to show lack of sensitivity outside the active site. Color coded to highlight high, medium and low binding affinities (green, yellow, red).

Complex	Sim. ΔG_{MMPBSA}	Sim. $\Delta\Delta G_{MMPBSA}$	Exp. ΔG	Exp. $\Delta\Delta G$
LPV				
WT	-47.5	0	-15.1	0
FL	-48	-0.5	-14.9	0.2
DM	-47.4	0.1	-14.9	0.2
AS	-45.3	2.2	-13.9	1.2
QM	-43.6	3.9	-12.8	2.3
HM	-42.3	5.2	-11.3	3.7
RTV				
WT-RTV	-58.8	0	-13.7	0
HM-RTV	-58	0.8	-9.3	4.4
RTV Swap				
WT	-54.3	0	-13.7	0
FL	-51.2	3.2	-12.8	0.9
DM	-50.6	3.7	-12.9	0.8
AS	-47.6	6.7	-11.9	1.8
QM	-44.1	10.2	-10.7	3
HM	-40	14.3	-9.3	4.4

Table S2. Energy of binding predicted by MMPBSA calculations, experimentally determined energy of binding (Exp. ΔG, Ohtaka et al, 2003) and changes in free energy of binding on mutation (Exp. $\Delta\Delta G$), all values in kcal mol^{-1}. Color coded to highlight high, medium and low binding affinities (green, yellow, red).

	HAD	HAC	O4	O1	Exp. ΔG kcal/ mol
LPV					
WT	0.17	0.22	0.2	1	-15.1
FL	0.24	0.22	0.22	1	-14.9
DM	0.22	0.38	0.24	1	-14.9
AS	0.35	0.73	0.21	1	-13.9
QM	0.43	0.96	0.33	1	-12.8
HM	0.49	0.84	0.42	1	-11.3
RTV					
WT	0.14	0.10	0.00	NA	-13.7
HM	0.17	0.02	0.04	NA	-9.3
RTV Swap					
WT	0.06	0.13	0.03	NA	-13.7
FL	0.10	0.10	0.07	NA	-12.8
DM	0.11	0.33	0.09	NA	-12.4
AS	0.37	0.43	0.28	NA	-11.9
QM	0.49	0.56	0.28	NA	-10.7
HM	0.70	0.62	0.53	NA	-9.3

Table S3. Fraction of simulations showing water binding and experimentally determined energy of binding (Exp. ΔG kcal/ mol, Ohtaka et al, 2003). O1 site is a bulk solvent exposed oxygen atom included to show lack of sensitivity outside the active site. Color coded to highlight high, medium and low binding affinities (green, yellow, red).

References

1. Sadiq, S. K., Wan, S., and Coveney, P. V. (2007) Insights into a mutation-assisted lateral drug escape mechanism from the HIV-1 protease active site., *Biochemistry 46*, 14865-14877.
2. Kempf, D. J., Marsh, K. C., Denissen, J. F., McDonald, E., Vasavanonda, S., Flentge, C. A., Green, B. E., Fino, L., Park, C. H., and Kong, X. P. (1995) ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans., *Proc Natl Acad Sci USA 92*, 2484-2488.
3. Stoll, V., Qin, W., Stewart, K. D., Jakob, C., Park, C., Walter, K., Simmer, R. L., Helfrich, R., Bussiere, D., Kao, J., Kempf, D., Sham, H. L., and Norbeck, D. W. (2002) X-ray crystallographic structure of ABT-378 (lopinavir) bound to HIV-1 protease., *Bioorganic \& Medicinal Chemistry 10*, 2803-2806.
4. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD - Visual molecular dynamics., *J. Molec. Graph. 14*, 33-38.
5. Wlodawer, A., and Vondrasek, J. (1998) Inhibitors of HIV-1 protease: a major success of structure-assisted drug design., *Annual review of biophysics and biomolecular structure 27*, 249-284.
6. Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations., *Journal of Computational Chemistry 24*, 1999-2012.
7. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field., *Journal of Computational Chemistry 25*, 1157-1174.
8. Phillips, J., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD, *J. Comp. Chem. 26*, 1781-1802.
9. Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs, *Journal of Computational Chemistry 26*, 1668-1688.
10. Singh, N., and Warshel, A. (2010) Absolute binding free energy calculations: on the accuracy of computational scoring of protein-ligand interactions., *Proteins: Structure, Function, and Bioinformatics 78*, 1705-1723.
11. Wright, D. W., and Coveney, P. V. (2011) Resolution of Discordant HIV-1 Protease Resistance Rankings Using Molecular Dynamics Simulations, *Journal of Chemical Information and Modeling 51*, 2636-2649.
12. Luckow, A., Lacinski, L., and Jha, S. 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, In *2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing*, pp 135-144, IEEE.