# Elucidating the Origin of Resistance in HIV-1 Protease Using Atomistic Simulation

**David W. Wright[1], Benjamin A. Hall[1], Shantenu Jha[2] & Peter V. Coveney[1§]**

[1] Centre for Computational Science, Chemistry Department, University College London, London, UK

[2] Electrical and Computer Engineering, 627 CoRE Building, 94 Brett Road Busch Campus, Rutgers University Piscataway, NJ 08854, USA

[§] Corresponding Author: email: p.v.coveney@ucl.ac.uk

***Abstract: The emergence of drug resistance is a major challenge for the effective treatment of HIV. In this article we explore the application of atomistic molecular dynamics simulations to explore the structural origins of two separate resistance pathways found in patients treated with HIV-1 protease inhibitors.***

## Introduction

HIV-1 protease is one of the enzymes responsible for the replication of the HIV virus and as a result is the target of numerous HIV drugs (known as protease inhibitors). However the high mutation rate of HIV frequently leads to the emergence of drug resistant strains when under the selective pressure of inhibitory drugs. Not only can single point mutations have a significant impact on the ability of drugs to bind to the HIV-1 protease but it is increasingly becoming recognised that non-linear epistatic effects between mutations play an important role in determining the level of resistance [1, 2].

Ultimately, inhibitor resistance is caused by changes in the structure, dynamics and chemistry of the viral enzymes targeted by particular drugs. Using atomistic molecular dynamics simulations we aim to gain insight into the mechanistic causes of resistance and ultimately to be able to quantitatively predict the level of resistance provided by combinations of mutations. In this way we believe that simulations could potentially be used in future both to inform drug design and enhance patient specific treatment selection.

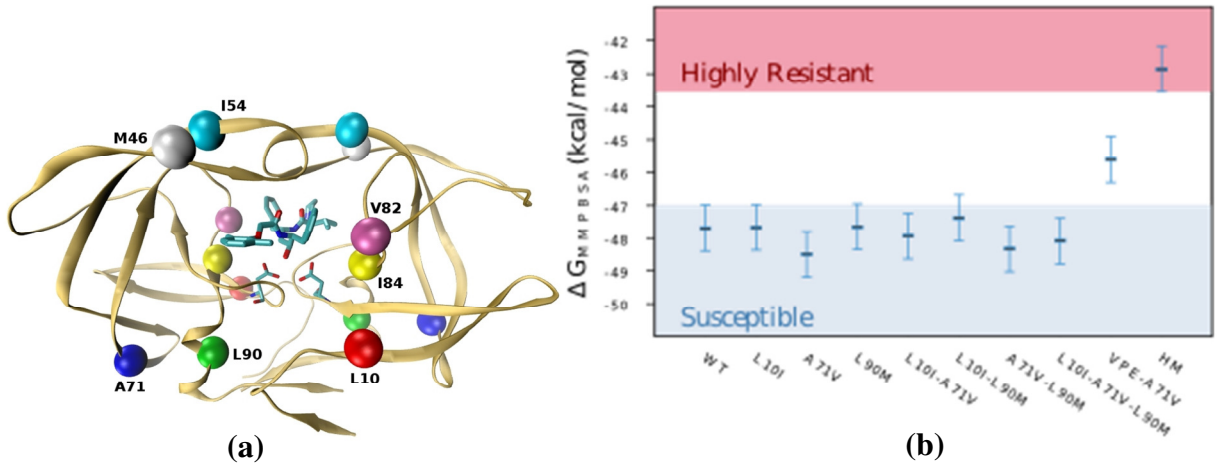
## Ensemble Sampling and Binding Affinity Calculations

The key thermodynamic property governing the strength of binding of inhibitory drugs is the binding free energy, or binding affinity. In previous work, our group focused on a series of mutant proteases with experimentally known binding energies to the inhibitor lopinavir (LPV) we established a protocol which allows us to calculate binding affinities which are reproducible to within 0.7kcal/mol, achieving a correlation with the experimental values of 0.98 [3]. The systems included mutations in the dimer interface (L10I and L90M - DM), the active site (V82A and I84V - AS) and a flexible region which encloses the drugs known as the flaps (M46I and I54V - FL) compared to the hxb2 wildtype (WT). This work used the established MMPBSA methodology alongside normal mode conformational entropy estimates to calculate binding affinities from molecular dynamics trajectories. In order to gain accurate and converged results it was necessary to employ large-scale ensemble-based Molecular Dynamics (EnMD) simulations; for each system 50 replica systems were simulated for 6ns. This protocol is used here to investigate two further examples of protease inhibitor resistance.

## Virtual Patient Experiment

In collaboration with the ViroLab project (*www.virolab.org*) a patient derived protease sequence was identified which gave discordant resistance ratings for various protease inhibitors. The sequence contained the mutations L10I, I13V, K14K, I15V, K20T, L63P, A71IV, V77I, L90M, I93L. This sequence was deemed to be susceptible to LPV by HIVdb (*hivdb.stanford.edu*) but displayed intermediate resistance according to ANRS (*www.anrs.fr*) and Rega (*www.rega.kuleuven.be/cev/regadb*). All the drug ranking systems identified the mutations L10I,

A71IV and L90M alone as causing any resistance assigned to the sequence. By simulating not only the full patient sequences but also systems containing the constituent mutations, we gain insight into why resistance estimates vary and the interactions between the various mutations. We find that the three mutations identified do not alone (when inserted into a wildtype background) confer any changes in binding affinity; however, the full patient sequence binds much less strongly than wildtype. This resistance is seen to be accompanied by changes in both protein conformation and flexibility [4]. In particular we identify changes in the relative conformation of the two beta sheets that form the protease dimer interface which suggest an explanation of the relative frequency of different amino acids observed in patients at residue 71.




**(a)** **(b)**

**Figure 1**: (a) The three dimensional structure of HIV-1 protease bound to lopinavir showing the location of the resistance associates mutations simulated. (b) Binding affinities computed using MMPBSA for each simulated variant, showing only the full patient sequence (labelled VPE-A71V) exhibits any resistance.

**Quantized Water Entry**

Initial simulations of the inhibitor ritonavir (RTV) were not able to reproduce either absolute or relative affinities to protease variants (the same set of mutations as studied for LPV were investigated). This was a surprising result as LPV and RTV are closely related drugs. Further investigation showed that these simulations had become kinetically trapped preventing water entry to the active site seen in the LPV bound systems. By initiating simulation with RTV swapped into structures generated by previous LPV simulations our ability to distinguish mutants was recovered. The relative affinity results reproduce the experimentally observed super-additive effect of mutations (although changes between sequences are exaggerated compared to experiment and those

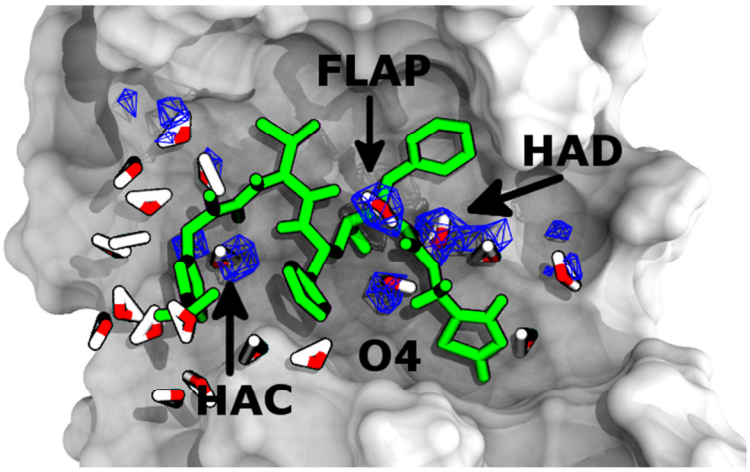



**Figure 2**: Water access to the HIV-1 protease active site. High occupancy water pockets rendered as blue isosurfaces overlaid onto a snapshot of the protein surface (white, excluding flaps) with a drug (green) and specific water molecules (red/white licorice).

seen in the LPV simulations). This suggested that a common resistance mechanism was at work in the cases of both drugs. Our previous study had indicated that active site water entry was altered by resistance causing mutations. Detailed investigation of active site water binding allowed us to identify a series of water binding sites whose frequency of occupation correlated strongly with resistance level. In particular these were found neighboring the two secondary amine groups of the drug (labeled HAD and HAC, Figure 2), as well as deep penetrating waters interacting with the central hydroxyl group and catalytic aspartic acid residues (labeled O4, Figure 2). The correlation between hydrogen bond formation between drug and water at the HAD site is shown in Figure3. We propose that small alterations of the packing of the protease flaps around the drug in the DM and FL systems are insufficient to allow water access to the drug. In contrast, AS, QM and HM mutations allow differential access to the active site by opening the flaps to differing degrees, allowing individual water molecules to access different water binding sites. The efficiency of sampling enabled by the EnMD approach is required to identify the individual water binding sites due to the stochastic nature of water entry; individual simulations even in the most resistant systems do not uniformly display water entry, which could prevent observation in a single simulation.

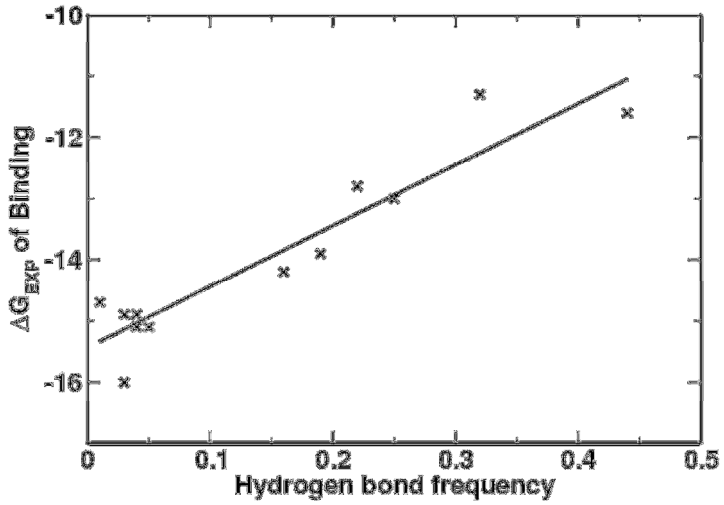



**Figure 3**: Correlation of HAD hydrogen bond frequency with experimentally determined[1] free energy of binding.

**Acknowledgements**

This work has been supported by the EU FP6 ViroLab and the EU FP7 CHAIN, VPH-SHARE and P-Medicine projects. Access to US TeraGrid was made available by the National Science Foundation , and access to to the EU DEISA network by the Virtual Physiological Human Virtual Community managed by the VPH Network of Excellence (*www.vph-noe.eu*). We are grateful to EPSRC and CoMPLEX for funding the Ph.D. studentship of D.W.W.

**References**

1. Ohtaka, H., A. Schön, and E. Freire, *Multidrug resistance to HIV-1 protease inhibition requires cooperative coupling between distal mutations.* Biochemistry, 2003. **42**: p. 13659-66.
2. Zhang, J., et al., *Detecting and understanding combinatorial mutation patterns responsible for HIV drug resistance.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**: p. 1321-6.
3. Sadiq, S.K., et al., *Accurate ensemble molecular dynamics binding free energy ranking of multidrug-resistant HIV-1 proteases.* Journal of chemical information and modeling, 2010. **50**: p. 890-905.
4. Wright, D.W. and P.V. Coveney, *Resolution of Discordant HIV-1 Protease Resistance Rankings Using Molecular Dynamics Simulations.* Journal of chemical information and modeling, 2011. **51**: p. 2636-49.