HHS Public Access

Author manuscript

*J Chem Theory Comput*. Author manuscript; available in PMC 2018 November 14.

Published in final edited form as:

*J Chem Theory Comput*. 2017 November 14; 13(11): 5671–5682. doi:10.1021/acs.jctc.7b00601.

# Elucidating the Interdependence of Drug Resistance from Combinations of Mutations

**Debra A. Ragland**[1], **Troy W. Whitfield**[2,3], **Sook-Kyung Lee**[4], **Ronald Swanstrom**[4], **Konstantin B. Zeldovich**[3], **Nese Kurt-Yilmaz**[1], and **Celia A. Schiffer**[1]

[1]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605, USA

[2]Department of Medicine, University of Massachusetts Medical School, Worcester, MA 01605, USA

[3]Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, MA 01605, USA

[4]Department of Biochemistry and Biophysics, and the UNC Center for AIDS Research, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

## Abstract

HIV-1 protease is responsible for the cleavage of 12 non-homologous sites within the Gag and Gag-Pro-Pol polyproteins in the viral genome. Under the selective pressure of protease inhibition, the virus evolves mutations within (primary) and outside of (secondary) the active site allowing the protease to process substrates while simultaneously countering inhibition. The primary protease mutations impede inhibitor binding directly, while the secondary mutations are considered accessory mutations that compensate for a loss in fitness. However, the role of secondary mutations in conferring drug resistance remains a largely unresolved topic. We have shown previously that mutations distal to the active site are able to perturb binding of darunavir (DRV) via the protein's internal hydrogen-bonding network. In this study we show that mutations distal to the active site, regardless of context, can play an interdependent role in drug resistance. Applying eigenvalue decomposition to collections of hydrogen bonding and van der Waals interactions from a series of molecular dynamics simulations of 15 diverse HIV-1 protease variants, we identify sites in the protease where amino acid substitutions lead to perturbations in non-bonded interactions with DRV and/or the hydrogen-bonding network of the protease itself. While primary mutations are known to drive resistance in HIV-1 protease, these findings delineate the significant contributions of accessory mutations to resistance. Identifying the variable positions in the

Correspondence to: Celia A. Schiffer.

**Supporting Information**

The Supporting Information file includes table of $EC_{50}$ values for clinically-derived HIV-1 protease variants, the sequence identity matrix for all 15 HIV-1 protease variants in the panel, the root mean square deviation (RMSD) values for MD simulations, plots of the first two principal components for the variance of hydrogen bond occupancies and mean van der Waals energy, correlation matrix for dynamic hydrogen bond occupancies, the plots of density distribution of 15 variants along the first principal component for the 111 mean hydrogen bond occupancies in Figure 3A, and for the 64 mean van der Waals contact energies seen in Figure 5A. This information is available free of charge via the Internet at http://pubs.acs.org

protease that have the greatest impact on drug resistance may aid in future structure-based design of inhibitors.

## Introduction

Darunavir (DRV) is a highly potent protease inhibitor (PI) used in the treatment of patients infected with HIV-1. Unlike first generation PIs, DRV is able to withstand many mutations both within and outside of the protein's active site [1–2]. Due to a high barrier to resistance, single mutations do not individually cause significant DRV resistance, and substitutions responsible for cross-resistance to other PIs are still fairly susceptible to DRV inhibition [3]. Contributing factors to DRV's high genetic barrier to resistance include the tight binding affinity ($K_d = 4.5\times10^{-12}$ M) [4], extensive hydrogen bonding with several active site backbone atoms [5], favorable hydrophobic contacts within the active site and a good fit within the substrate envelope [6]. However, even with all these key attributes the protease is still able to develop complex mutational patterns that facilitate evasion of DRV inhibition. Previous studies [7] have demonstrated an interdependence among specific amino acid substitutions that together result in resistance.

In such complex mutational patterns, active site mutations physically alter inhibitor binding and are, therefore, readily identified. However, the role of mutations beyond the active site is more difficult to characterize. For instance, the DRV resistance-associated mutations I84V, I50V, V32I and I47V all lie in positions where the inhibitor atoms protrude beyond the substrate envelope at the active site [8]. In clinical trials of DRV, however, several non-active site (secondary or accessory) mutations are selected for at positions 11, 33, 54, 73, 76, 85, and 89 among others [9–10] , and the role of these mutations in DRV resistance is not understood. The widely accepted notion is that accessory mutations have the sole purpose of balancing the destabilizing effects of primary active site mutations [11]. Studying the mutational tolerance of the protease using an empirical scoring function has indicated that distal mutations can be beneficial not only via stabilizing monomeric folding and dimerization, but interactions with the substrate as well [12]. However, the direct role of accessory mutations in drug resistance has not been extensively probed [13–16] and even less well known are which specific variable positions outside the active site play a role in resistance [7]. Thus, while most mutations within the active site that arise to DRV are readily explained by the substrate envelope hypothesis, without a similar framework, evaluating the role of other mutations is not straightforward.

Previously, to gauge how secondary mutations away from the active site could play a role in protease inhibitor susceptibility, we examined several single mutations and one double mutant variant of HIV-1 protease[17]. These mutations included V32I, located at the periphery of the active site, and a combination of V32I/L33F. In addition, we examined the distal DRV resistance associated mutation L76V and the non-DRV resistance associated mutation L90M. A careful investigation of the crystal structures and molecular dynamics simulations of these variants bound to DRV showed that while these distal mutations alone do not drive significant levels of resistance, they were all able to perturb the network of hydrogen bonds within the protein, thereby propagating the effect to the protease active site causing slight

loss of affinity. This network model provided a general understanding as to how mutation of residues may communicate with one another and why some co-mutant relationships may be synergistic or redundant.

In this study we further investigate the role of mutations both near and distal from the active site in complex combinatorial backgrounds of a set of protease variants, and compare to the WT and single/double site mutants examined previously. Specifically, this study seeks to determine which variable positions in the protease are most relevant for resistance, given the complex sequence variations observed among heavily mutated variants. Several DRV-resistant protease variants were selected from viral passaging experiments as well as patient-derived sequences from the HIV Drug Resistance Database [18]. A series of molecular dynamics (MD) simulations were conducted on the panel of 15 protease variants (SF-2 and NL4-3 wild-types and 13 other variants) all bound to DRV. We find via analyses of root mean square fluctuations (RMSF) that sequence similarity alone may be indicative of backbone dynamics in the variants. In addition, mean hydrogen bond occupancies (both intra-protease and DRV–protease) were collected from these trajectories, along with mean per-residue DRV–protease van der Waals energies. For these two data sets, a combination of eigenvalue decomposition and statistical testing was used to identify mutations that can best explain the observed variance in physical properties across the protease panel. Through this analysis we find that alterations in hydrogen bonding network distinguish single and double mutants from the more complex variants. A71V and R41K, a known resistance-associated mutation and a polymorphic substitution respectively, also impact the hydrogen-bonding network. In addition, the primary mutation I84V in conjunction with the peripheral accessory mutation M46I and several other remote accessory mutations gives rise to alterations of van der Waals contacts with DRV. Thus, combining MD simulations of a diverse set of HIV-1 protease variants, both susceptible and resistant to inhibition, with unsupervised machine learning techniques yields mechanistic insights into how distal accessory mutations contribute to drug resistance.

## Results

To determine which variable positions specifically impact the structural and dynamic properties of protease–DRV binding, a combination of inhibitor-bound crystal structures and homology models were used as input for MD simulations for 15 variants of HIV-1 protease. Details of the models, nomenclature for the variants, and MD simulations are described in the Methods section. Using the resulting trajectories, root mean square fluctuations (RMSF), mean per-residue contact energies with DRV, and hydrogen bond occupancies throughout the protease were monitored for the panel of protease variants.

### Convergent evolution drives protease resistance in independent viral lineages

The diverse panel of 15 HIV-1 protease variants was chosen with a broad range of sequence substitutions containing single site mutants and more heavily mutated multi-drug resistant proteases (MDR-PRs). Both the SF-2 and NL4-3 wild-type (WT) proteases were used as controls for the variants in the panel. These two WT proteins share 95% identity, varying at positions 7, 14, 41, 63 and 64 (Figure 1A and S2). Both SF-2 and NL4-3 have high

susceptibilities to DRV, with single-digit pM $K_I$ values[17] at the limit of detection by enzymatic assays and an $EC_{50}$ of 4 nM in cell-based replicon assays (Table S1).

Of the highly mutated proteases in the panel, two were obtained from long-term viral passaging experiments ($DRV^r8$ and $DRV^r10$), conducted under DRV selective pressure. These two protease sequences differ from the NL4-3 WT by 8 and 10 amino acid substitutions, respectively (Figure 1A). The remaining MDR proteases were obtained from the HIV Drug Resistance Database [18–19]. The patient-derived MDR-PRs contain between 19 and 26 substitutions when compared to the SF-2 WT protease [7]. Taken together, the panel of 15 proteases has sequence variations at 50 of the 99 amino acid positions within each monomer (Figure 1A and B).

Although the viral population ancestry and temporal treatment history is not annotated for the patient-derived proteases in our simulation panel, they share common mutations with one another and also with the highly mutated variants derived from viral passaging experiments (Figures 1 and S1). Based on the dates when the samples were isolated and the high level of resistance to DRV inhibition (Table S1), there is a possibility that one patient-derived strain, $VEG_{23}$, was exposed to DRV treatment.

Considering the resistance to DRV inhibition in the $DRV^r8$ and $DRV^r10$ strains that is evident from viral passaging experiments, the measured resistance among the patient-derived strains (Table S1) and the shared sequence identities within the panel, cross-resistance to DRV among the patient isolates is apparent. These observations suggest that the phenotypes of these proteases have converged under the selective pressure of inhibition, perhaps driving similar mechanisms of resistance. A phylogenetic tree based on their sequences for the 15 proteases in the panel is shown in Figure 1C[20].

## Sequences of multi-drug resistant mutants correlate with changes in protease dynamics

The protease variants with available DRV bound structures were SF-2 (PDB: 1T3R), L76V, V32I and V32I/L33F [17, 21]. The NL4-3 wild type and remaining variants were modeled based on the DRV-bound wild-type SF-2 structure. The crystallographic water molecules, including the important bridge water between the inhibitor and the protease flaps, were preserved in each model. Three 100 ns replicate MD simulations with explicit solvent were performed and analyzed for each DRV complex.

The root mean square deviations (RMSD) reveal that the accumulation of mutations from single site to patient-derived variants leads to greater structural changes in order to reach thermal equilibrium starting from the modeled configuration (Figure S2). The changes in per-residue root mean square fluctuations (RMSF) about the mean appear to correlate well with protease lineage with the most pronounced changes in fluctuation seen in the flaps, flap hinge, and lower cantilever regions of the protease while the catalytic aspartate residues remain rigid across the simulations (Figure 2A). Intriguingly, hierarchical clustering of the per-residue RMSF profiles for the fifteen variants results similar groupings to those within the sequence-based phylogenetic tree (Figures 1C and 2B). This overlap of clustering suggests that sequence similarity alone may be a good predictor of similar backbone

dynamics as the regions of high variability contain residues found to be predictive of changes in dynamics (Figure 2C).

## Alterations in hydrogen bonding correlate with particular combinations of mutations

To characterize alterations in the hydrogen bonding observed among the variants and to determine which variable positions best explain alterations in hydrogen bond network among the proteases, a set of 143 hydrogen bonds (111 main chain and 32 side chain) were monitored over the simulations. This set included both intra-protease and protease–DRV hydrogen bonds. With an expanded panel of protease variants, relative to our earlier study [17], the use of algorithms for detecting patterns of altered occupancies and identifying specific mutations that may underlie these alterations becomes essential. A combination of principal component analysis (PCA), to detect alterations, followed by hypothesis testing based on amino acid substitution at specific sites in the protease, was employed. To begin, a 15×15 correlation matrix of mean occupancies for these 143 main and side chain hydrogen bonds was computed and used for PCA (see Methods for details). The first principal component (Figure S5), $u_1$, accounted for nearly all (89%, Figure S3) of the inter-variant variance in hydrogen bond occupancies. Comparing the ordering of protease variants along $u_1$, a striking similarity is observed with the phylogenetic ordering (Fig. 1C), suggesting that variations in overall hydrogen bond occupancies are dictated by lineage.

Focusing on the 111 main chain hydrogen bonds plus the 2 catalytic aspartate-DRV side chain bonds (excluding the other side-chain hydrogen bonds), the resistant variants tend to have higher values of the first principal component (Figure 3A) than the more susceptible variants, including the two wild-type strains. To infer which amino acid substitutions at specific positions in the protease account for the distribution of variants along $u_1$, and noting that the density, $\rho(u_1)$, of variants along this component is approximately bimodal (see Figure S6), hypothesis tests were performed using the Wilcoxon rank sum[22]. The hypothesis tests were conducted for pairs of distributions defined by the presence or absence of a specific mutation (e.g. I84V). The null hypothesis was no difference between the means of these two distributions. There is a lower bound on the p-value, defined by $\rho(u_1)$. Figure 3A and Table 3B summarize the results of these tests, identifying A71V as the single mutation that best accounts for the spread in hydrogen bonding patterns among the variants in our panel. Segregating the variants based on this mutation, we find that $DRV^r10$, $KY_{26}$, $SLK_{19}$, $VEG_{23}$ and $VSL_{23}$ all contain changes at position 71 (Figure 3A). To determine whether *pair-wise* substitutions of amino acids can be used to better recapitulate the bimodal distribution of variants with respect to $u_1$, additional statistical tests were performed: no pair or other combination of substitutions explains more of this variance than does the A71V/I mutation alone (Table 3C).

The variance in hydrogen bonding within this panel can be further explained by classifying the variants into those containing mutations at 10, 54, 71 and 41 simultaneously, and those that do not. Only the patient variants mentioned above contain this combination of mutations. While protease mutations at residue 41 are considered polymorphic, such mutations have been reported to play a role in resistance to protease inhibitors, including DRV [23]. Nonetheless, residue 41 is one of five residues whose changes distinguish the

variants in our panel in an NL4-3 background from those in the SF-2 background. This finding suggests that mutations at positions 10, 54, 71 and 41, being distal to the active site, may be relaying information about the global dynamics of the protein, consistent with our earlier hypothesis that mutations perturb the dynamic ensemble of the protease via the network of hydrogen bonds [17, 24]. The majority of variants in the panel contain very pronounced changes throughout the lower cantilever region of the protease, which includes residue 71 (Figure 4).

### Distal accessory mutations alter ligand–protease van der Waals interactions

Using the approach described above, amino acid substitutions that played key roles in altering the hydrophobic contacts of the protease with DRV were determined. The mean van der Waals (vdW) contact energies between the protease active site residues and DRV were calculated over the trajectories for each of the variants in the panel. Energies were collected for all 64 amino acids within the protease active site that have contacts with DRV during the simulations. Inspecting the distribution of the variants along the first principal component of the resulting correlation matrix, which accounts for 97.6% of the inter-variant variance of vdW energies (Figure S4), the variants segregate differently than when analyzing the internal hydrogen bonding. Interestingly, the variants $KY_{26}$ and $SLK_{19}$ segregate with both WT proteases and some single site variants as well (Figure 5A).

Following the same hypothesis testing approach, as explained above for mean hydrogen bond occupancies, I84V was determined to be the most predictive single site substitution for classifying the variants (Table 5B). Segregating the variants based solely on the presence of I84V captured most of the observed bimodality in $\rho(u_1)$ (Figure S7), such that variants $ATA_{21}$, $VEG_{23}$, $KY_{26}$, $DRV^r8$, $DRV^r10$ and, necessarily the I84V single mutant, were distinguished from the remainder of the protease panel (Figure 5A). The change from isoleucine to valine reduces the close packing of the isoleucine side chain and the P1’ phenylalanine-mimicking moiety of DRV (Figure 6D).

Among the candidate pairs of mutational sites, the pair of residues that is most predictive of the perturbations of the vdW contact energies was I84V and M46I (Table 5C). Segregating the variants based on this combination of mutations, $ATA_{21}$, $VEG_{23}$, $DRV^r8$ and $DRV^r10$ can be distinguished from the rest of the panel (Figure 5A). Furthermore, combination of substitutions at positions 13, 32, and 33 in addition to I84V were predictive of the distinguishing patterns within the vdW data. All four of the variants containing an amino acid substitution at positions 84 and 46 also contain mutations I13V, V32I and L33F. This finding suggests that there may be some coupling between mutations at positions 84, 46, 13, 32 and 33, resulting in the weakening of protease– DRV binding.

With the exception of primary resistance mutation I84V, these mutations are accessory mutations not located directly at the active site, which cause alterations in vdW contacts of other active site residues. In addition to I84, among the 64 active site residues that make vdW contacts with DRV, D30 and I50 in chain A and R8, D29, D30, G27, G48, and V82 in chain B, were perturbed the most by accessory mutations, as indicated by the departure from the mean for all residues across the 15 proteases (Figure 6A; see also Methods). These distinguishing variations in vdW contacts mostly impact generally immutable active site

residues, suggesting that preserving mostly hydrophobic contacts upon inhibitor binding is crucial for sustained targeting (Figure 6B). Mapping the difference (Δ*, Figure 6B) in departure from the mean values (σ*) onto the structure further details how vdW contacts of active site residues are impacted by distal mutations (Figure 6C).

## Discussion

The accumulation of mutations within a drug target allows the balance of substrate processing versus inhibitor binding to tip in favor of the former. The HIV-1 protease exhibits a high level of resiliency under selective pressure and resistant HIV-1 protease variants are sufficiently adapted to evade inhibition without a substantial growth penalty [25]. While associating active site mutations with weaker inhibitor binding may seem straightforward, the constellation of resistance-associated mutations that arise throughout the rest of the protein, as is the case with DRV, has been relegated to aiding in recovery of viral fitness [26–28]. In this study we examined the role of non-active site mutations in conferring drug resistance, by analyzing 15 protease variants that together contain substitutions at 50 of the 99 amino acid positions within the enzyme. Specifically, the effects of mutations on protease structure and dynamics have been investigated via hydrogen bonding, and vdW contacts with the inhibitor. Specific positions have been identified that account for the variance in these properties, including mutations at residues away from the active site.

We have shown previously that mutations compromise the hydrogen bonding and vdW contacts necessary for DRV binding to ensure that the mutations render the protein resistant while retaining its biological function [17]. In this study, we used a novel combination of MD simulations and unsupervised machine learning to characterize the variability in these functionally important quantities across a panel of 15 susceptible and resistant protease variants and identify specific mutations that can explain this variability. The specific mutations that were identified validate the significance of previously observed mutations occurring outside the active site. For example, we observed variants that had mutations at residue 71 to be a major contributor to the variance of the hydrogen bonds. The A71V mutation has been shown previously to be a key mutation in the re-stabilization of the enzyme in the presence of major mutations such as I50L/V, and has also been shown to propagate its effects from its position in the lower cantilever region of the protease to the active site via the hydrogen bond network of the protein [26–27, 29–33]. Similarly, other non-active site mutations that have been thought to merely contribute to overall protein stability in drug-resistant protease variants may have impacts propagating to the active site. Overall, our finding of mutations that are highly predictive of changes in the hydrogen bonding are all distal to the active site and mostly involve changes to larger hydrophobic residues. Consistent with our previous findings and others[17, 27, 33–34] this result suggests a ‘domino’ effect model, driven by mutations in distal residues whose impact propagate to the active site of the protein.

The presence of I84V mutation was the best predictor of alterations in vdW contacts with the inhibitor, in combination with M46I. The I84V mutation is common in protease inhibitor cross-resistance. The change from the bulkier beta-branched isoleucine to the smaller valine has been the Achilles heel of PI treatment since the I84V mutation was first observed in

saquinavir treatment [35]. M46I was thought to only be a compensatory mutation [25] until early MD studies found it to be a key modulator of flap dynamics [36]. The combination of I84V with M46I has been long studied as a major/minor co-mutant pair in the midst of other compensatory mutations able to drive resistance to early PIs [25, 37]. With the exceptions of V32I, which is at the periphery, the remaining mutations that were identified here as to underlie the observed variance in the vdW contacts lie outside of the active site and have previously been explored for their compensatory effects [7]. Figure 6 illustrates that the residues with the highest variability in vdW contact energies lie within the active site (G48, I50, I84, G27' and I84'), while the specific mutations that regulate this variability are either juxtaposed or distal to the active site.

The analysis presented here demonstrates how mutations outside of the active site impact DRV targeting. We find that mutations distal to the active site, whether they occur as single amino acid substitutions or as highly complex combinations, are able to perturb inhibitor binding through changes in certain key interactions between the enzyme and inhibitor. We also find overall backbone dynamics to be associated with sequence similarity, as expected, and to change with accumulating mutations and drug resistance. While primary mutations are known to drive resistance in HIV-1 protease, these findings delineate the significant contributions of accessory mutations to resistance. Identifying the variable positions that have the greatest impact on drug resistance may aid in future structure based designs for inhibitors, potentially for other quickly evolving targets as well which are susceptible to drug resistance.

## Methods

### Protease Panel & Nomenclature

The panel of 15 proteases used in this study consisted of the SF-2 and NL4-3 wild-type proteases along with 13 mutant variants. The WT proteases served as controls for the variants with respect to their subtype B backgrounds and laboratory origin. The L76V, V32I and V32I/L33F (PDB accession codes 3OY4, 4Q1X and 4Q1Y respectively) variants were taken from a previous study [17]. To this group of single and double mutants, the L33F single mutant was added. Another set of protease sequences were obtained from HIV-1 cell culture passaging studies in the presence of DRV. Briefly, *in vitro* selections were carried out with DRV using an initial mixture of 26 variants, each containing a single resistance mutation. The selections were carried out with increasing inhibitor concentrations between with final drug concentrations that were 1000-fold greater than the measured $IC_{50}$ in WT strains. Using the Primer ID-based paired-end MiSeq platform [38], mutations in the protease were analyzed based on RNA sequencing carried out at four time points (i.e. passage checkpoints) as inhibitor concentration was increased. The variants in this set include homology models of single site I93L and I84V variants complexed with DRV along with a variant containing eight mutations ($DRV^r8$) and a variant containing 10 mutations ($DRV^r10$) in an NL4-3 background. The two highly mutated variants, $DRV^r8$ and $DRV^r10$, were present at very high concentrations of DRV.

The remaining variants in the panel were selected from the patient-derived proteases in the HIV Drug Resistance Database [18–19]. This group of proteins contains 19–26 mutations

compared to the SF-2 WT protease[7]. Each patient variant is named based on which amino acid substitutions are unique to that variant and the number of mutations it contains compared to the SF-2 WT. For example, variant $KY_{26}$ is the only variant that contains substitutions H69K and C67Y and it has 26 mutations. All other variants are named for the mutations they contain (e.g. variant I84V only contains this mutation). The V32I+L33F double mutant is referred to as DM for "double mutant".

## Homology Modeling and MD Simulations

The SF-2 WT, L76V, V32I and V32I/L33F protease sequence variants had available crystal structures bound to DRV. The NL4-3 wild-type and remaining variants were all modeled based on the DRV bound structure (PDB ID: 1T3R). The crystallographic water molecules, including the important bridge water between the inhibitor and the protease flaps, were preserved in each model, as was DRV. Using the Prime Structure Prediction Wizard by Schrödinger (Release 2014-4, Schrödinger LLC, New York, NY [39–40]), each of the 11 variant sequences was used as a query to search for homologs via BLAST [41]. Both chains of PDB structure 1T3R were selected as templates to build the homodimer containing the appropriate variant sequence. Once the model was prepared, the structure was built retaining the ligand from the template structure. Water molecules from the template structure were added to the newly built variant structure and the side-chains of those residues that were mutated *in silico* were refined locally using Prime Refinement Tools followed by a complete refinement of the overall structure in the Protein Preparation Wizard. This utility processes the structures by assigning bond orders, adding hydrogen atoms, creating disulfide bonds, and filling in missing sidechains using Prime. Next, tautomerization states are optimized using Epik and hydrogen bond networks and protonation states were determined and optimized using PROPKA pH 7.0, with exhaustive sampling of water orientations and minimization of the hydrogen atom configurations of altered species. Finally, interaction energies of hydrogen atoms were minimized using the Impact Refinement Module and the OPLS2005 force field.

All MD simulations were performed using Desmond [42–45] with the OPLS2005 force field. Systems were prepared by solvating the structure in a cubic box that extended at least 10 Å beyond the nearest solute atom in all directions using the TIP3 water model [46]. Sodium chloride was added to the equivalent of 150 mM to simulate physiological conditions. The system was neutralized by adding counterions as needed ($Na^+$ or $Cl^-$).

The rigorous pre-equilibration model was employed as described elsewhere [47]. Briefly, a series of restrained minimization steps was performed to gradually relax the system. Initially all heavy solute atoms were restrained with a force constant of 1000 kcal $mol^{-1}$ $Å^{-2}$ for 10 steps of steepest decent followed by up to 2000 steps using the LBFG method to a convergence of 50 kcal $mol^{-1}$ $Å^{-2}$. Restraints were removed from side-chains using LBFG for 5000 step or until a convergence of 50 kcal $mol^{-1}$ $Å^{-2}$. The restraints on the backbone were gradually removed using the following decreasing force constants: 1000, 500, 250, 100, 50, 10, 1 and 0 50 kcal $mol^{-1}$ $Å^{-2}$ using the LBFG method to convergence of 50 kcal $mol^{-1}$ $Å^{-2}$.

A series of short pre-production MD simulations were performed to equilibrate the system, starting with a 10 ps simulation in the NVT ensemble with 50 kcal $mol^{-1}$ $Å^{-2}$ harmonic restraints on solute heavy atoms using the Berendsen thermostat [48] at 10 K. A 1 fs time-step was used for bonded and short-range interactions (up to 9 Å) and a 3 fs time-step was used for long-range electrostatic interactions. A 10 ps MD simulation followed, using an NPT ensemble with a Berendsen thermostat followed, run at 10 K with a 2 ps time-step for bonded and short-range interactions and 6 fs for long-range electrostatics. Over 50 ps, the temperature of the system was increased to 300 K with restraints on heavy solute atoms followed by a 10 ps simulation where all harmonic restraints were removed. Production simulations were performed in the constant NPT ensemble using the Desmond implementation of the Martyna-Tobias-Klein (MTK) extended system [49]. Simulations were carried out with no harmonic restraints for 100 ns at 300 K and 1 bar. The cut off for non-bonded interactions was 9 Å; the smooth particle mesh Ewald (PME) method [50] was applied; the time-step was 2 fs for short-range interactions and 6 fs for long range interactions. All simulations were performed in triplicate, each with different random initial velocities for a total production time of 300 ns for each of the 15 simulated protease systems. The Simulation Event Analysis Tool within Maestro was used to determine the mean occupancies of 143 inter and intra-main chain and side chain hydrogen bonds, along with hydrogen bonds between the ligand and protease. In order to facilitate analysis, including computation of the mean protein-ligand van der Waals interaction energies, Visual Molecular Dynamics (VMD) version 1.9.2 [51] was used to translate the Desmond trajectories to PDB format.

## Evaluation of Hydrogen Bonding and van der Waals Interactions

For each hydrogen bond pair, the donor heavy atom along with its hydrogen and the acceptor atom were specified for calculation of hydrogen bonding occupancy. For each frame, only pairs that satisfied the hydrogen bond geometric criteria as set forth by Schrödinger were chosen: the distance between hydrogen atom and acceptor atom must be less than or equal to 2.5 Å, the angle between donor heavy atom and its hydrogen and the acceptor must be at least 120°, and the angle between the hydrogen and acceptor heavy atom must be at least 90°. The van der Waals contacts between the inhibitor and the protease were calculated using a simplified Lennard-Jones potential, following published protocols [52].

## Principal Components Analysis and Statistical Testing

The hydrogen bond and van der Waals observations were combined into matrices of dimension 15 x N, where N is the number of observations in each data set (for example, there were N = 64 protein–ligand van der Waals energies per variant). We made use of the correlation matrix in lieu of a covariance matrix so that any outlier data would not dominate the variance in the data set. Principal components can be defined in terms of an eigenvalue problem for the correlation matrix: $\boldsymbol{C}u_i = \lambda_i u_i$ where **C** is the correlation matrix for any two variables X and Y;

$$\boldsymbol{C} = \frac{\langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle}{\langle (X - \langle X \rangle) \rangle \langle (Y - \langle Y \rangle) \rangle}$$

This problem can be solved by diagonalization $\mathbf{C} = \mathbf{UC'U^{-1}}$ where the diagonal elements of **C'** are ordered components of the variance (i.e. the eigenvalues $\lambda_i$). This transformation preserves the trace of matrix **C** (Tr**C** = Tr**C'**). The proportion of total variance, Tr**C**, that is explained by eigenvector $u_i$ is defined as $\frac{\lambda_i}{Tr\mathbf{C}}$.

Eigenvalues and eigenvectors were calculated using an R interface to the LAPACK (Linear Algebra Package) library [53]. With the van der Waals and hydrogen bonding data, the explained variance was dominated by the first eigenvector (or principal component) and hypothesis testing was used to interpret the spread among the different protease variants. For each mean hydrogen bond occupancy or van der Waals contact, the following quantity was computed in order to measure how the within-class observations deviated from the global average for each of two classes of variants (e.g. those with or without the I84V mutation):

$$\sigma^* = \sum_{j=1}^{N_{class}} \sqrt{(E_j - \mu)^2}$$

where $N_{class}$ is the total number of variants in each class, $E_j$ is the van der Waals contact for variant $j$ and $\mu$ is the average across all 15 variants. This deviation allows us to identify important, or at least highly variable residue interactions. The difference between $\sigma^*$ for two classes, A and B, is defined as $\Delta^* = (\sigma^*_A - \sigma^*_B)$.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This research was supported by NIH P01 GM109767 and NIH F31 GM 111101. RS also receives support from P30 AI050410 and P30 CA16068. We would like to thank Janet Paulsen, postdoctoral associate, and Kristina Prachanronarong, MD/PhD graduate student, University of Massachusetts Medical School, for assistance with running and troubleshooting various MD simulations and post-simulation analyses. We thank Dr. Shuntai Zhou for analyzing the deep sequence data from the viral passaging experiments.

This work is dedicated in memory of the late Bridgett E. Ragland.

## References

1. Tenore SB, Ferreira PR. The Place of Protease Inhibitors in Antiretroviral Treatment. Braz. J. Infect. Dis. 2009; 13:371–374. [PubMed: 20428639]
2. Wang Y, Liu Z, Brunzelle JS, Kovari IA, Dewdney TG, Reiter SJ, Kovari LC. The Higher Barrier of Darunavir and Tipranavir Resistance for HIV-1 Protease. Biochem. Biophys. Res. Commun. 2011; 412:737–742. [PubMed: 21871444]
3. Bethell R, Scherer J, Witvrouw M, Paquet A, Coakley E, Hall D. Short Communication: Phenotypic Protease Inhibitor Resistance and Cross-Resistance in the Clinic from 2006 to 2008 and Mutational Prevalences in HIV from Patients with Discordant Tipranavir and Darunavir Susceptibility Phenotypes. AIDS Res. Hum. Retroviruses. 2012; 28:1019–1024. [PubMed: 22098079]
4. King NM, Prabu-Jeyabalan M, Nalivaika EA, Wigerinck P, de Bethune MP, Schiffer CA. Structural and Thermodynamic Basis for the Binding of TMC114, a Next-Generation Human Immunodeficiency Virus Type 1 Protease Inhibitor. J. Virol. 2004; 78:12012–12021. [PubMed: 15479840]

5. Ghosh AK, Chapsal BD, Weber IT, Mitsuya H. Design of HIV Protease Inhibitors Targeting Protein Backbone: An Effective Strategy for Combating Drug Resistance. Acc. Chem. Res. 2008; 41:78–86. [PubMed: 17722874]
6. Nalam MN, Ali A, Altman MD, Reddy GS, Chellappan S, Kairys V, Ozen A, Cao H, Gilson MK, Tidor B, Rana TM, Schiffer CA. Evaluating the Substrate-Envelope Hypothesis: Structural Analysis of Novel HIV-1 Protease Inhibitors Designed to Be Robust against Drug Resistance. J. Virol. 2010; 84:5368–5378. [PubMed: 20237088]
7. Varghese V, Mitsuya Y, Fessel WJ, Liu TF, Melikian GL, Katzenstein DA, Schiffer CA, Holmes SP, Shafer RW. Prototypical Recombinant Multi-Protease-Inhibitor-Resistant Infectious Molecular Clones of Human Immunodeficiency Virus Type 1. Antimicrob. Agents Chemother. 2013; 57:4290–4299. [PubMed: 23796938]
8. King NM, Prabu-Jeyabalan M, Nalivaika EA, Schiffer CA. Combating Susceptibility to Drug Resistance: Lessons from HIV-1 Protease. Chem. Biol. 2004; 11:1333–1338. [PubMed: 15489160]
9. de Meyer S, Vangeneugden T, van Baelen B, de Paepe E, van Marck H, Picchio G, Lefebvre E, de Bethune MP. Resistance Profile of Darunavir: Combined 24-Week Results from the Power Trials. AIDS Res. Hum. Retroviruses. 2008; 24:379–388. [PubMed: 18327986]
10. Prezista(R)[Package Insert]. Janssen Therapeutics, Division of Janssen Products, Lp; Titusville, Nj: 08560
11. Swanstrom R, Erona J. Human Immunodeficiency Virus Type-1 Protease Inhibitors: Therapeutic Successes and Failures, Suppression and Resistance. Pharmacol. Ther. 2000; 86:145–170. [PubMed: 10799712]
12. Humphris-Narayanan E, Akiva E, Varela R, S OC, Kortemme T. Prediction of Mutational Tolerance in HIV-1 Protease and Reverse Transcriptase Using Flexible Backbone Protein Design. PLoS Comput. Biol. 2012; 8:e1002639. [PubMed: 22927804]
13. Muzammil S, Ross P, Freire E. A Major Role for a Set of Non-Active Site Mutations in the Development of HIV-1 Protease Drug Resistance. Biochemistry. 2003; 42:631–638. [PubMed: 12534275]
14. Agniswamy J, Shen CH, Aniana A, Sayer JM, Louis JM, Weber IT. HIV-1 Protease with 20 Mutations Exhibits Extreme Resistance to Clinical Inhibitors through Coordinated Structural Rearrangements. Biochemistry. 2012; 51:2819–2828. [PubMed: 22404139]
15. Kozisek M, Lepsik M, Grantz Saskova K, Brynda J, Konvalinka J, Rezacova P. Thermodynamic and Structural Analysis of HIV Protease Resistance to Darunavir - Analysis of Heavily Mutated Patient-Derived HIV-1 Proteases. FEBS J. 2014; 281:1834–1847. [PubMed: 24785545]
16. Mahalingam B, Boross P, Wang YF, Louis JM, Fischer CC, Tozser J, Harrison RW, Weber IT. Combining Mutations in HIV-1 Protease to Understand Mechanisms of Resistance. Proteins. 2002; 48:107–116. [PubMed: 12012342]
17. Ragland DA, Nalivaika EA, Nalam MN, Prachanronarong KL, Cao H, Bandaranayake RM, Cai Y, Kurt-Yilmaz N, Schiffer CA. Drug Resistance Conferred by Mutations Outside the Active Site through Alterations in the Dynamic and Structural Ensemble of HIV-1 Protease. J. Am. Chem. Soc. 2014; 136:11956–11963. [PubMed: 25091085]
18. Shafer RW. Rationale and Uses of a Public HIV Drug-Resistance Database. J. Infect. Dis. 2006; (194 Suppl 1):S51–58. [PubMed: 16921473]
19. Rhee SY, Gonzales MJ, Kantor R, Betts BJ, Ravela J, Shafer RW. Human Immunodeficiency Virus Reverse Transcriptase and Protease Sequence Database. Nucleic Acids Res. 2003; 31:298–303. [PubMed: 12520007]
20. Tamura K, Stretcher G, Peterson D, Filipski A, Kumar S. Mega6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution. 2013; 30:2.
21. Surleraux DL, Tahri A, Verschueren WG, Pille GM, de Kock HA, Jonckers TH, Peeters A, De Meyer S, Azijn H, Pauwels R, de Bethune MP, King NM, Prabu-Jeyabalan M, Schiffer CA, Wigerinck PB. Discovery and Selection of TMC114, a Next Generation HIV-1 Protease Inhibitor. J. Med. Chem. 2005; 48:1813–1822. [PubMed: 15771427]
22. Hogg, RV., Tanis, Elliot A., Zimmerman, Dale L. Probability and Statistical Inference. 9. Pearson Higher Education Upper Saddle River; NJ: 2015.

23. De Meyer S, Azijn H, Surleraux D, Jochmans D, Tahri A, Pauwels R, Wigerinck P, de Bethune MP. Tmc114, a Novel Human Immunodeficiency Virus Type 1 Protease Inhibitor Active against Protease Inhibitor-Resistant Viruses, Including a Broad Range of Clinical Isolates. Antimicrob. Agents Chemother. 2005; 49:2314–2321. [PubMed: 15917527]
24. Sinha N, Smith-Gill SJ. Protein Structure to Function Via Dynamics. Protein Pept Lett. 2002; 9:367–377. [PubMed: 12370024]
25. Condra JH, Schleif WA, Blahy OM, Gabryelski LJ, Graham DJ, Quintero JC, Rhodes A, Robbins HL, Roth E, Shivaprakash M, et al. In Vivo Emergence of HIV-1 Variants Resistant to Multiple Protease Inhibitors. Nature. 1995; 374:569–571. [PubMed: 7700387]
26. Nijhuis M, Schuurman R, de Jong D, Erickson J, Gustchina E, Albert J, Schipper P, Gulnik S, Boucher CA. Increased Fitness of Drug Resistant HIV-1 Protease as a Result of Acquisition of Compensatory Mutations During Suboptimal Therapy. AIDS. 1999; 13:2349–2359. [PubMed: 10597776]
27. Erickson JW. The Not-So-Great Escape. Nat. Struct. Biol. 1995; 2:523–529. [PubMed: 7664118]
28. Hoffman NG, Schiffer CA, Swanstrom R. Covariation of Amino Acid Positions in HIV-1 Protease. Virology. 2003; 314:536–548. [PubMed: 14554082]
29. Skalova T, Dohnalek J, Duskova J, Petrokova H, Hradilek M, Soucek M, Konvalinka J, Hasek J. HIV-1 Protease Mutations and Inhibitor Modifications Monitored on a Series of Complexes. Structural Basis for the Effect of the A71v Mutation on the Active Site. J. Med. Chem. 2006; 49:5777–5784. [PubMed: 16970402]
30. Meher BR, Wang Y. Interaction of I50/V Mutant and I50l/A71V Double Mutant HIV-Protease with Inhibitor Tmc114 (Darunavir): Molecular Dynamics Simulation and Binding Free Energy Studies. J. Phys. Chem. B. 2012; 116:1884–1900. [PubMed: 22239286]
31. Mittal S, Bandaranayake RM, King NM, Prabu-Jeyabalan M, Nalam MN, Nalivaika EA, Yilmaz NK, Schiffer CA. Structural and Thermodynamic Basis of Amprenavir/Darunavir and Atazanavir Resistance in HIV-1 Protease with Mutations at Residue 50. J. Virol. 2013; 87:4176–4184. [PubMed: 23365446]
32. Logsdon BC, Vickrey JF, Martin P, Proteasa G, Koepke JI, Terlecky SR, Wawrzak Z, Winters MA, Merigan TC, Kovari LC. Crystal Structures of a Multidrug-Resistant Human Immunodeficiency Virus Type 1 Protease Reveal an Expanded Active-Site Cavity. J. Virol. 2004; 78:3123–3132. [PubMed: 14990731]
33. Chang MW, Torbett BE. Accessory Mutations Maintain Stability in Drug-Resistant HIV-1 Protease. J. Mol. Biol. 2011; 410:756–760. [PubMed: 21762813]
34. Appadurai R, Senapati S. Dynamical Network of HIV-1 Protease Mutants Reveals the Mechanism of Drug Resistance and Unhindered Activity. Biochemistry. 2016; 55:1529–1540. [PubMed: 26892689]
35. Gulnik SV, Suvorov LI, Liu B, Yu B, Anderson B, Mitsuya H, Erickson JW. Kinetic Characterization and Cross-Resistance Patterns of HIV-1 Protease Mutants Selected under Drug Pressure. Biochemistry. 1995; 34:9282–9287. [PubMed: 7626598]
36. Collins JR, Burt SK, Erickson JW. Flap Opening in HIV-1 Protease Simulated by ‘Activated’ Molecular Dynamics. Nat. Struct. Biol. 1995; 2:334–338. [PubMed: 7796268]
37. Weber IT, Harrison RW. Molecular Mechanics Analysis of Drug-Resistant Mutants of HIV Protease. Protein Eng. 1999; 12:469–474. [PubMed: 10388843]
38. Jabara CB, Jones CD, Roach J, Anderson JA, Swanstrom R. Accurate Sampling and Deep Sequencing of the HIV-1 Protease Gene Using a Primer ID. Proc. Natl. Acad. Sci. U. S. A. 2011; 108:20166–20171. [PubMed: 22135472]
39. Jacobson MP, Friesner RA, Xiang Z, Honig B. On the Role of the Crystal Environment in Determining Protein Side-Chain Conformations. J. Mol. Biol. 2002; 320:597–608. [PubMed: 12096912]
40. Jacobson MP, Pincus DL, Rapp CS, Day TJ, Honig B, Shaw DE, Friesner RA. A Hierarchical Approach to All-Atom Protein Loop Prediction. Proteins. 2004; 55:351–367. [PubMed: 15048827]
41. Altschul SF, Gish W Fau - Miller W, Miller W Fau - Myers EW, Myers Ew Fau - Lipman DJ, Lipman DJ. Basic Local Alignment Search Tool.

42. Shivakumar D, Williams J, Wu Y, Damm W, Shelley J, Sherman W. Prediction of Absolute Solvation Free Energies Using Molecular Dynamics Free Energy Perturbation and the Opls Force Field. J. Chem. Theory Comput. 2010; 6:1509–1519. [PubMed: 26615687]
43. Bowers KJ, Chow E, Xu H, Dror RO, Eastwood MP, Gregersen BA, Klepeis JL, Kolossvary I, Moraes MA, Sacerdoti FD, Salmon JK, Shan Y, Shaw DE. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. Proceedings of the 2006 ACM/IEEE conference on Supercomputing. 2006:84.
44. 2015-4, S. R., Desmond Molecular Dynamics System. D.E. Shaw Research, Maestro-Desmond Interoperability Tools. 2015
45. Guo Z, Mohanty U, Noehre J, Sawyer TK, Sherman W, Krilov G. Probing the Alpha-Helical Structural Stability of Stapled P53 Peptides: Molecular Dynamics Simulations and Analysis. Chem. Biol. Drug Des. 2010; 75:348–359. [PubMed: 20331649]
46. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. The Journal of Chemical Physics. 1983; 79:926–935.
47. Ozen A, Sherman W, Schiffer CA. Improving the Resistance Profile of Hepatitis C Ns3/4a Inhibitors: Dynamic Substrate Envelope Guided Design. J. Chem. Theory Comput. 2013; 9:5693–5705. [PubMed: 24587770]
48. Berendsen HJC, Postma JPM, Gunsteren WFv, DiNola A, Haak JR. Molecular Dynamics with Coupling to an External Bath. The Journal of Chemical Physics. 1984; 81:3684–3690.
49. Martyna GJ, Tobias DJ, Klein ML. Constant Pressure Molecular Dynamics Algorithms. The Journal of Chemical Physics. 1994; 101:4177–4189.
50. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A Smooth Particle Mesh Ewald Method. The Journal of Chemical Physics. 1995; 103:8577–8593.
51. Humphrey W, Dalke A, Schulten K. Vmd: Visual Molecular Dynamics. J. Mol. Graph. 1996; 14:33–38. 27–38. [PubMed: 8744570]
52. Ozen A, Haliloglu T, Schiffer CA. Dynamics of Preferential Substrate Recognition in HIV-1 Protease: Redefining the Substrate Envelope. J. Mol. Biol. 2011; 410:726–744. [PubMed: 21762811]
53. Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., Sorensen, D. {LAPACK} Users' Guide. 3. Society for Industrial and Applied Mathematics; Philadelphia, PA: 1999.

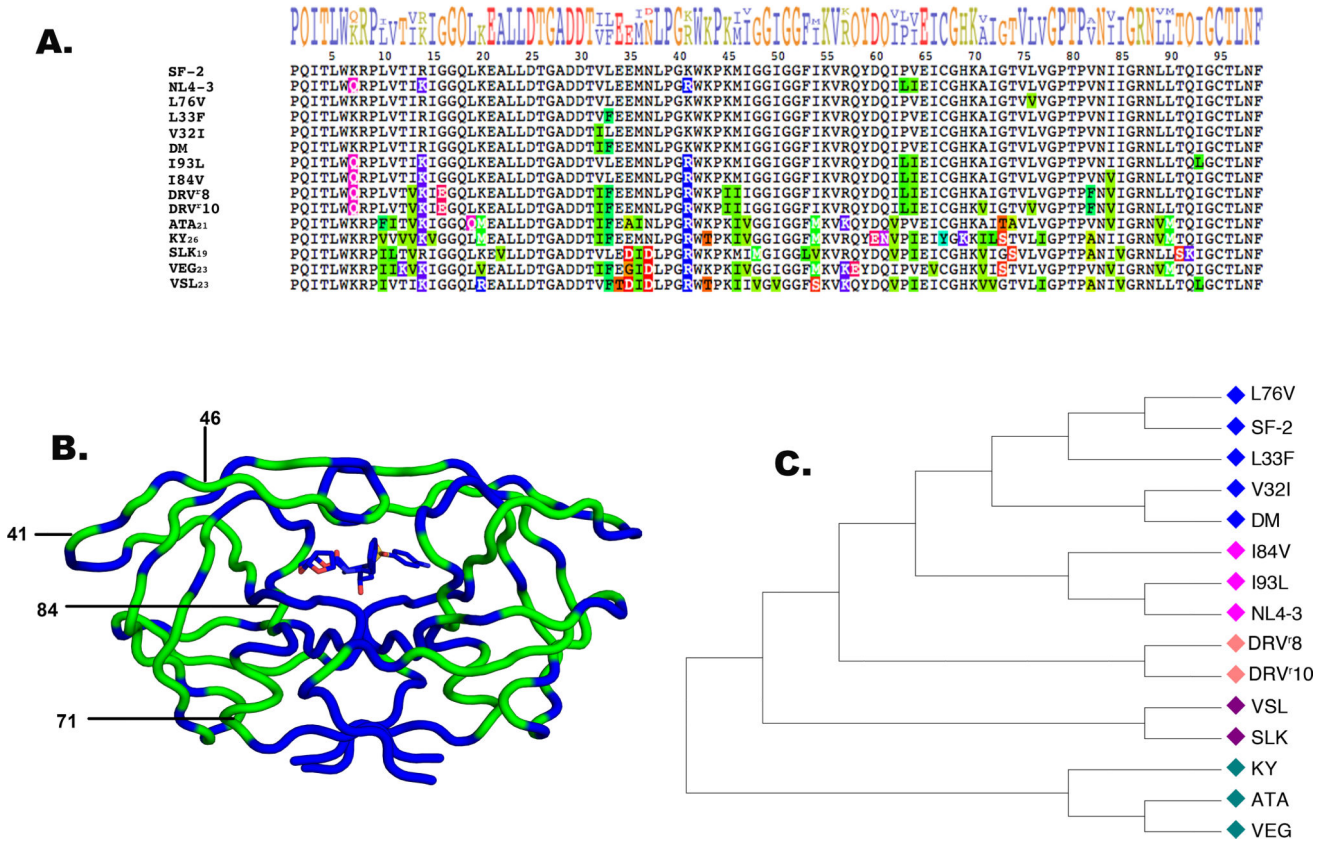



**Figure 1. Proteases derived from multiple ancestors converge evolutionarily to drive resistance**

**A**. Sequence alignment of all proteases included in the panel. Mutations are highlighted as compared to the SF-2 WT protease (PDB accession code 1T3R. PDB code 2HB4 denotes NL4-3 WT). **B**. Protease structure mapped in blue with all 50 sequence substitutions within the panel highlighted in green, protease inhibitor DRV is shown in the active site. **C**. Neighbor-joining phylogenetic tree of all 15 protease variants. Colored annotations denote sequence similarity among variants.

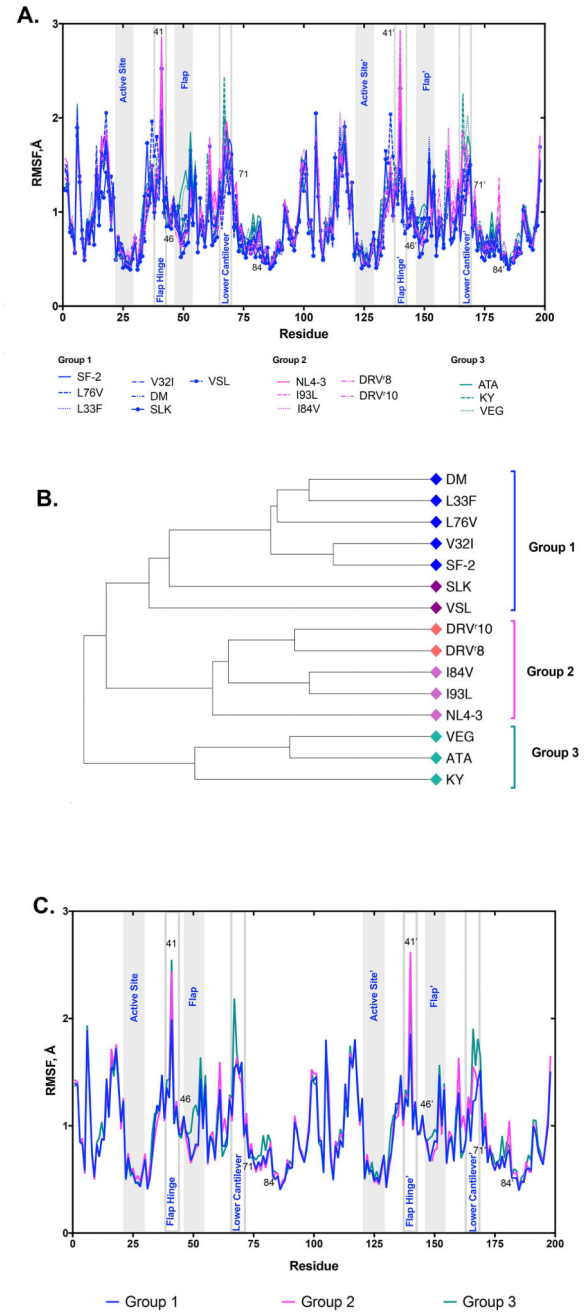
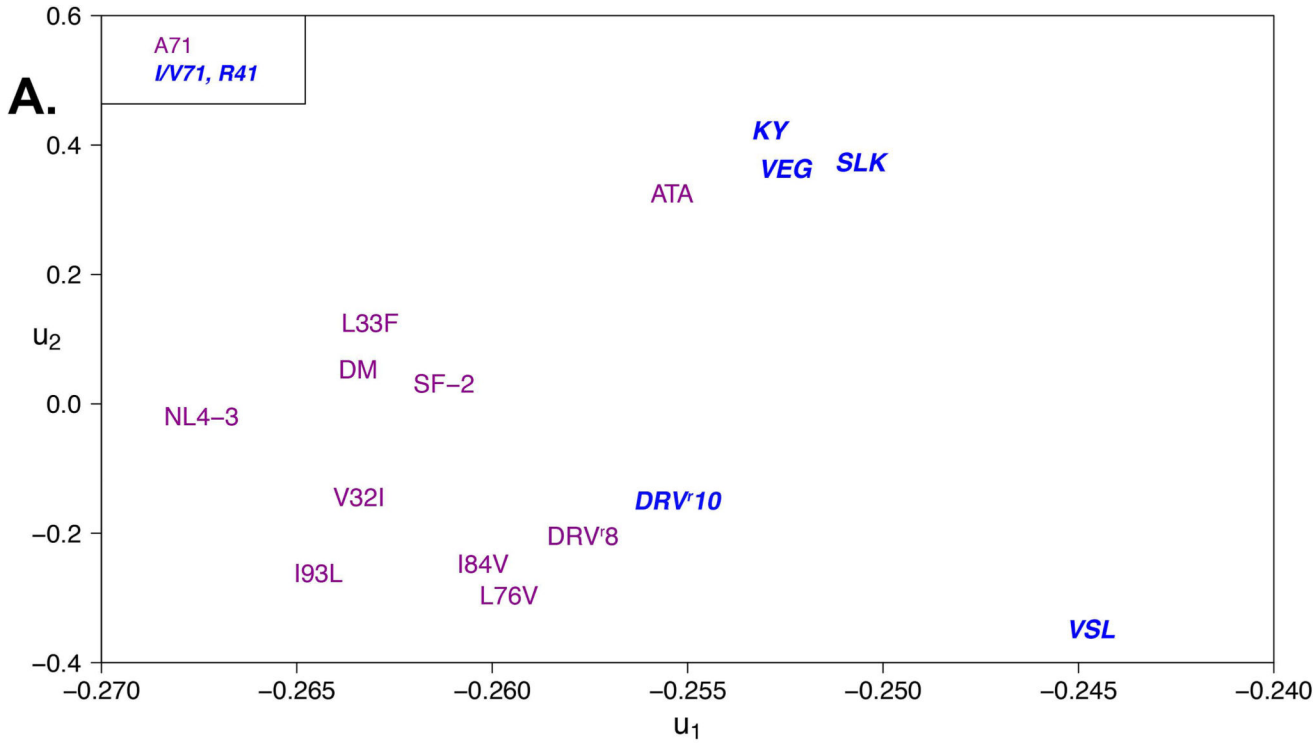



**Figure 2. Protease dynamics may be indicative of resistance**

**A**. Per-residue RMSF values for all 15 protease variants plotted in three groups as defined by the clustering of the hierarchical dendrogram. **B**. Hierarchichal clustering dendrogram of per-residue RMSF values for all 15 variants. Colored annotations are similar to Figure 1C denoting similarity in per-residue RMSF among variants. **C**. Averages of variants as grouped in A. Group colors are the same as noted in A and B.




**B.**

| Residue Position | p–value |
|---|---|
| 71 | 0.000666 |
| 10 | 0.00133 |
| 54 | 0.00133 |
| 37 | 0.00439 |
| 35 | 0.00586 |

**C.**

| Residue Position 1 | Residue Position 2 | p–value |
|---|---|---|
| 41 | 71 | 0.000666 |
| 10 | 41 | 0.00133 |
| 10 | 54 | 0.00133 |
| 41 | 54 | 0.00133 |
| 10 | 71 | 0.00146 |

**Figure 3. Variations in mean hydrogen bond occupancies delineate susceptibility of variants**

**A**. Fifteen sequence variants projected onto the first two principal components for the correlation matrix of 111 dynamic hydrogen bond occupancies. Variants are partitioned into two groups, those variants bearing a substitution at position 71 (bold, blue) and those that do not (purple). Variants that contain a substitution at 71 also contain a substitution at position 41. **B**. List of top five single positions found to most likely underlie alterations in hydrogen bond occupancy patterns. **C**. List of top five position pairs found to most likely underlie hydrogen bond occupancy changes.

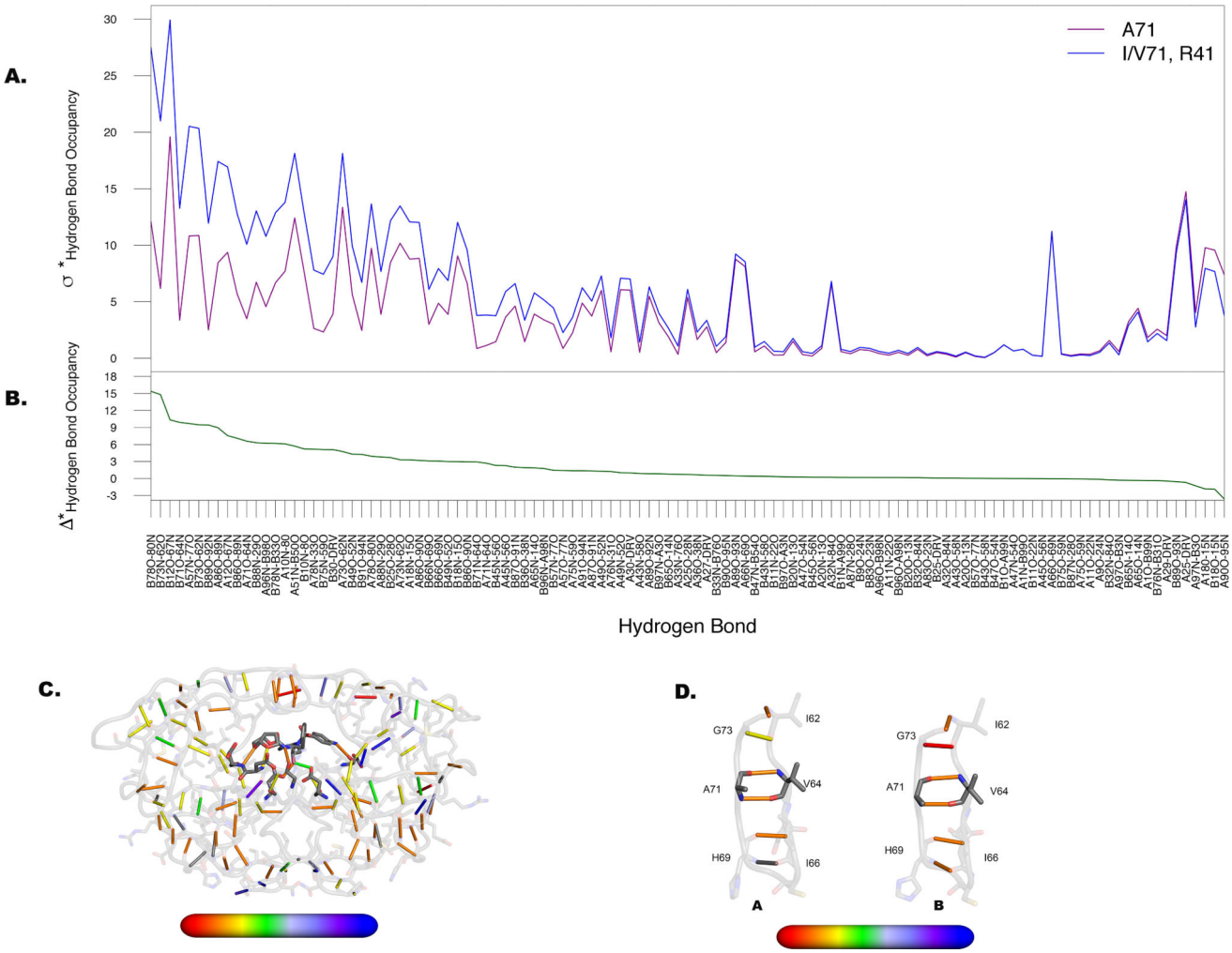
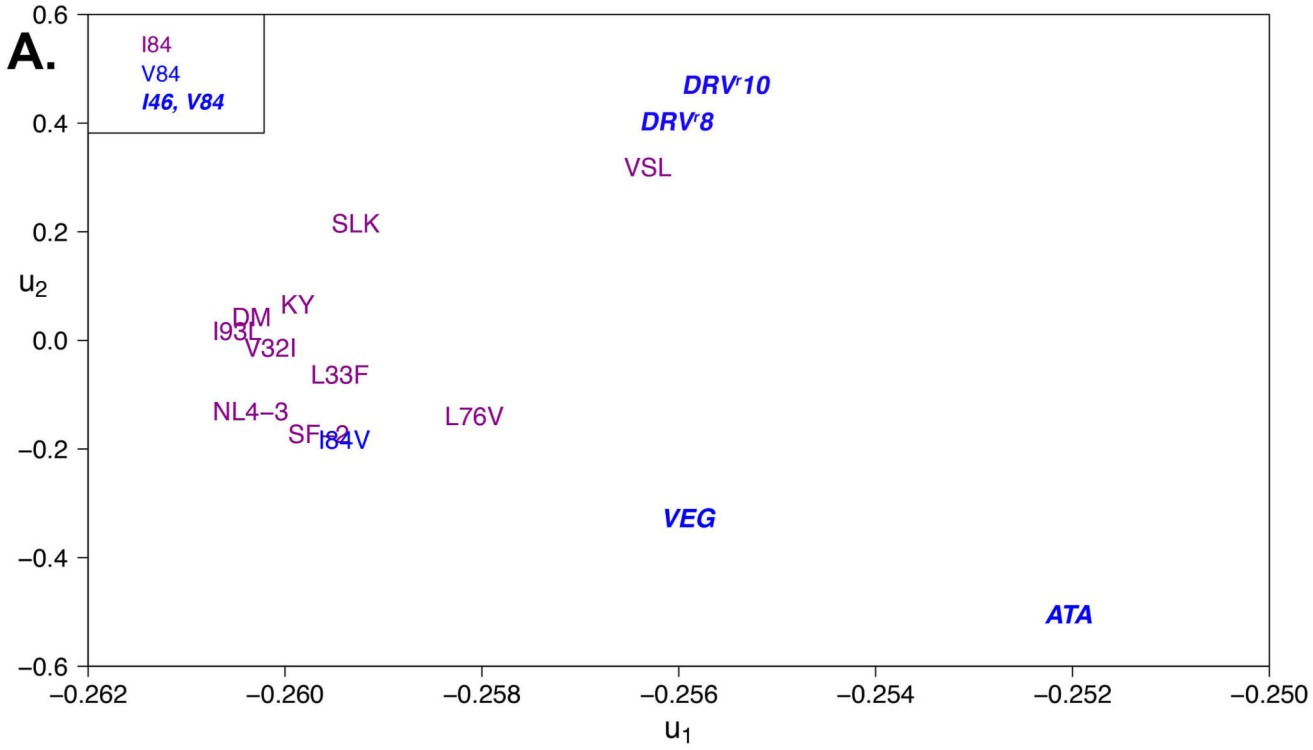



**Figure 4. Dominant alterations in hydrogen bond occupancies emanate from lower cantilever region of the protease**

**A**. Departure from the mean (σ*) as calculated for the two groups in Figure 3A. Maximal separation between those variants lacking substitutions at position 71 (purple) and those containing substitutions at 71 (blue) occurs predominantly at hydrogen bonds formed with residues within the 70s $\beta$-strand. **B**. Difference between two lines in A (Δ*) with values plotted onto the structure **C** from red (high variability) to blue (low variability). **D**. Hydrogen bonds formed between residues surrounding A71 as labeled, for both monomers of the protease. Hydrogen bonds in this region have higher alterations in chain B than in chain A.




**B.**

| Residue Position | p–value |
|---|---|
| 84 | 0.0046 |
| 46 | 0.00759 |
| 13 | 0.0175 |
| 57 | 0.0307 |
| 35 | 0.0395 |

**C.**

| Residue Position 1 | Residue Position 2 | p–value |
|---|---|---|
| 13 | 84 | 0.00146 |
| 32 | 84 | 0.00146 |
| 33 | 84 | 0.00146 |
| 46 | 84 | 0.00146 |
| 14 | 84 | 0.00466 |

**Figure 5. Ligand-protease van der Waals contact energies reveal energetic similarities between single site accessory RAMs and clinically-derived variants**

**A**. Fifteen sequence variants projected onto the first two principal components for the correlation matrix of 64 mean protease–DRV van der Waals contacts. Variants are partitioned into two groups, those bearing substitutions at positions 84 and/or 46 (blue) and those that do not (purple), similar to Figure 3A. **B–C** List of top single positions and paired positions most likely underlying changes in van der Waals contact energies.

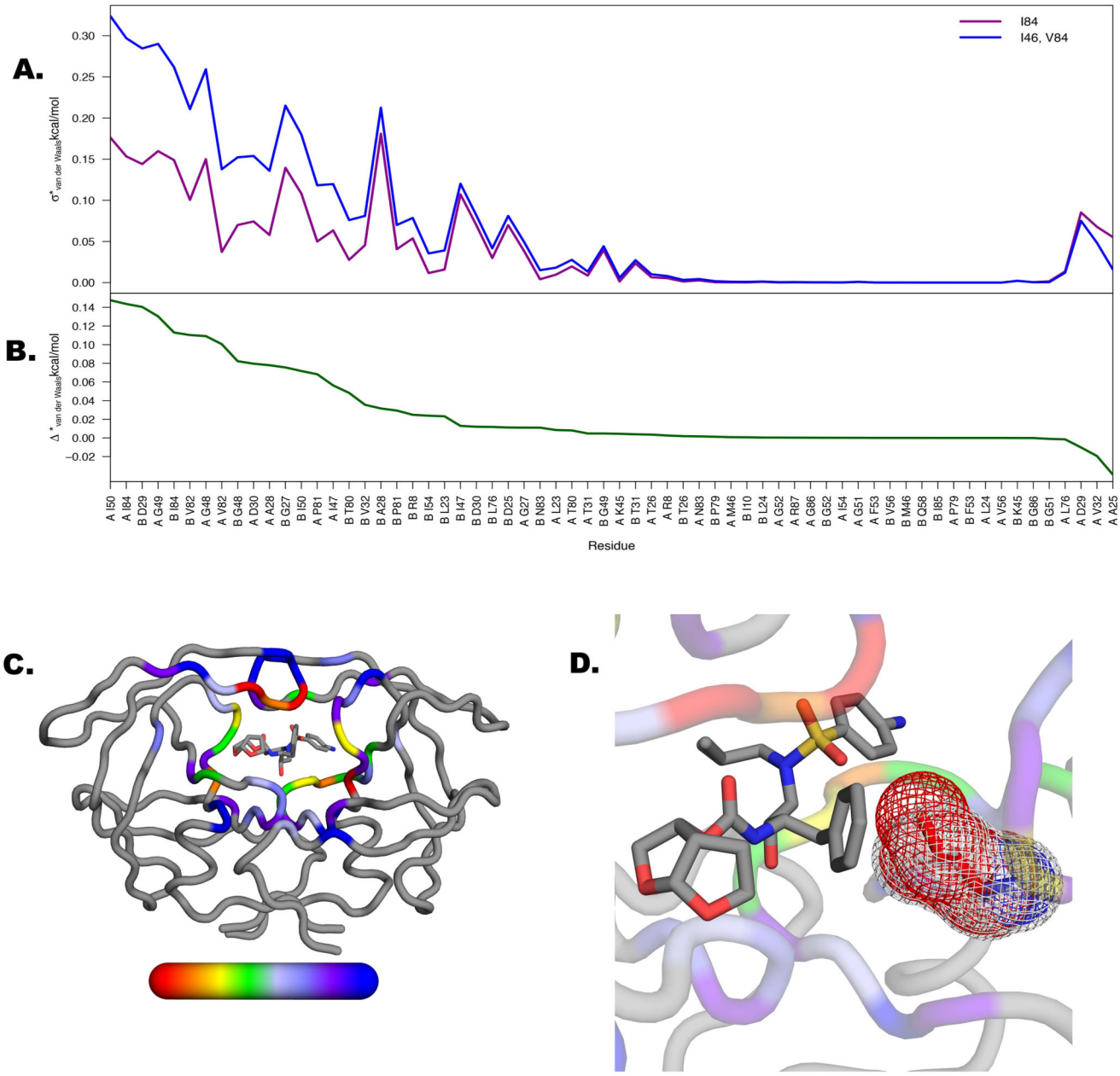



**Figure 6. Mutations cause perturbations of some non-mutable active site residue van der Waals contacts**

**A**. Departure from the mean (σ*) as calculated for two groups in Figure 5A. Maximal separation between variants not bearing substitutions at positions 84 and 46 (purple) and those containing substitutions at positions 84 and 46 (blue) occurs predominantly at residues within the active site. **B**. Difference between two lines in A (Δ*) with values plotted onto the structure **C** from red (high variability) to blue (variability). **D**. Residue I84 has the most perturbed van der Waals contact energy likely due to its juxtaposition with the P1' moiety of DRV, which is alleviated with the change to V84.