pM to µM - Elucidating the Role of Distal Mutations in HIV-1 Protease in Conferring Drug Resistance

Mina Henes[+,1], Gordon J. Lockbaum[+,1], Klajdi Kosovrasti[1], Florian Leidner[1], Gily S. Nachum[1], Ellen A. Nalivaika[1], Sook-Kyung Lee[2], Ean Spielvogel[2], Shuntai Zhou[2], Ronald Swanstrom[2], Daniel N.A. Bolon[1], Nese Kurt Yilmaz[*,1], Celia A. Schiffer[*,1]

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

[2]Department of Biochemistry and Biophysics and the UNC Center for AIDS Research, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States

Abstract

Drug resistance continues to be a growing, global problem. The efficacy of small molecule inhibitors is threatened by pools of genetic diversity in all systems, including antibacterials, antifungals, cancer therapeutics, and antivirals. Resistant variants often include combinations of active site mutations, and distal "secondary" mutations which are thought to compensate for losses in enzymatic activity. HIV-1 protease is the ideal model system to investigate these combinations and underlying molecular mechanisms of resistance. Darunavir (DRV) binds WT HIV-1 protease with a potency of < 5 pM, but we have identified a protease variant that loses potency to DRV 150,000-fold, with 11 mutations in and outside the active site. To elucidate the roles of these mutations in DRV resistance, we used a multidisciplinary approach, combining enzymatic assays, crystallography, and molecular dynamics simulations. Analysis of protease variants with 1, 2, 4, 8, 9, 10 and 11 mutations showed that the primary active site mutations caused ~50-fold loss in potency (2 mutations), while distal mutations outside the active site further decreased DRV potency from 13 nM (8 mutations) to 0.76 µM (11 mutations). Crystal structures and simulations revealed that distal mutations induce subtle changes which are dynamically propagated through the protease. Our results reveal that changes remote from the active site directly and dramatically impact the potency of the inhibitor. Moreover, we find interdependent effects of mutations in conferring high levels of resistance. These mechanisms of resistance are likely applicable to many other quickly evolving drug targets, and the insights may have implications for the design of more robust inhibitors.

[*]Corresponding authors Celia A. Schiffer: Phone: +1 508 856 8008; Celia.Schiffer@umassmed.edu, Nese Kurt Yilmaz: Phone: +1 508 856 1867; Nese.KurtYilmaz@umassmed.edu.
[+]Joint first authors

Introduction

Drug resistance is a growing global problem [1,2]. The efficacy of current drugs is threatened by pools of genetic diversity in many therapeutic targets, including antibacterials, antifungals, cancer therapeutics, and antivirals[3]. Specifically, rapid evolution in viruses can cause emergence of variants that are not effectively inhibited by small molecule inhibitors. Human immunodeficiency virus (HIV) is capable of accumulating a large number of mutations in response to antiretroviral therapy[4, 5].

The viral enzymes coded by the HIV genome, including HIV-1 protease, have been targets of extensive drug development efforts. HIV-1 entry, nucleoside reverse transcriptase, non-nucleoside reverse transcriptase, integrase, and protease inhibitors are used in combination to curb the emergence of drug resistance. Resistance can develop quickly against a single agent, with mutations resulting from the high infidelity of the reverse transcriptase[6-8] and the high recombination rate of HIV[9, 10]. Mutations may also arise from patient non-compliance due to a number of factors, such as adverse side effects and high pill burden[11]. HIV-1 protease can tolerate a large number of mutations, with 45 out of 99 amino acid positions able to mutate while retaining adequate enzymatic activity to allow replication[12]. Viral variants from patients who failed therapy and in vitro selection experiments often contain complex combinations of multiple mutations. Especially with potent inhibitors, a single mutation is not sufficient and many mutations need to accumulate to confer appreciable levels of resistance. Under drug pressure, resistance and compensatory mutations accumulate in the gag and gag-pol viral polyproteins, resulting in often unpredictable and complex mechanisms of resistance [13, 14]. Given the prevalence and complexity of these mutations and underlying mechanisms, HIV-1 has served as a model system to identify, analyze, and characterize drug resistance.

Darunavir (DRV) is the latest and most potent FDA-approved HIV-1 protease inhibitor [15]. The potency of DRV is due to several factors, including a number of crucial hydrogen bonds with the protease backbone, conservation of the water-mediated hydrogen bonding network with the flaps, and extensive van der Waals (vdW) contacts with active site residues[16, 17]. Moreover, DRV fits well within the substrate envelope (the consensus volume occupied by the natural substrates[18]), which explains its low susceptibility to resistance[19-21]. DRV can effectively inhibit variants with common single primary resistance mutations such as I84V and I50V/A71V[22]. However, high levels of resistance are reached as mutations accumulate, both within and outside the active site of HIV-1 protease.

The combinations of mutations in HIV-1 protease can utilize complex mechanisms of resistance to reach significant levels of resistance. The primary mutations proximal to the active site are thought to directly confer drug resistance while distal mutations are considered compensatory and restore catalytic function[23]. However distal mutations can also weaken inhibitor binding, as has been reported not only for HIV-1 protease but other drug targets as well[24-27]. Molecular mechanisms such as reducing the coupled motions of the protease and inhibitor[28] and altering protein intra-molecular hydrogen bonds[29] have all been suggested in resistance mechanisms involving mutations distal to the inhibitor binding site. We have previously shown that cross-correlations between protein and inhibitor fluctuations

distinguish tight binders from weak binders in both HIV and HCV protease[29-31]. In addition to weakening inhibitor binding, mutations can alter the dynamic ensemble of the protease–inhibitor complex and the balance between inhibition and catalytic activity. However, changes to the structure and dynamic ensemble along a resistance pathway, molecular mechanisms of resistance due to combinations of mutations, and the role of distal mutations in conferring resistance versus restoring catalytic activity are not well understood.

Here we investigate highly mutated and highly DRV resistant variants of HIV-1 protease to understand the role of distal mutations and uncover underlying mechanisms of resistance. While DRV inhibits wild type HIV-1 protease with picomolar inhibition constant, a combination of 11 mutations (3 active site and 8 distal) confers near μM resistance to DRV. All of these mutations have been observed in patient isolates, although not in this specific combination[32, 33]. Our inhibition assays show that the active site mutations alone do not account for this severe loss in affinity to DRV. We generated protease variants with subsets of these 11 mutations (1, 2, 4, 8, and 10-mutant variants) and found increased resistance with accumulation of mutations. We also tease out the effect of two individual distal mutations on the viral evolution pathway from 8 to 10-mutant variants, through double-mutant cycle analysis of the two possible intermediate 9-mutant variants. Using high-resolution crystal structures, enzymatic and inhibition assays, and molecular dynamics simulations, our findings support the role of distal mutations in conferring drug resistance. In addition to restoring catalytic activity, distal mutations contribute to drug resistance through interdependent, dynamic processes propagated throughout the protease.

Results

Viral passaging under drug pressure selected for highly mutated variants

Viral passaging experiments were performed to elucidate drug resistance in HIV-1 protease under increasing DRV selective pressure, and identified 2 highly abundant variants with 8 and 10 mutations[34]. At the highest drug concentration, an additional mutation was selected, leading to a variant with 11 total mutations (11Mut) [Figure 1A,B]. Sequencing from early time points indicated that I84V (1Mut) was the first fixed mutation. A variant with both V82F and I84V (2Mut) was generated to investigate primary active site mutations. A midpoint variant between 2Mut and 8Mut was chosen to include 2 active site and 2 flap mutations (I84V, V82F, K45I, and M46I; 4Mut variant). Addition of 4 more mutations (I13V, G16E, V32I, and I33F) to 4Mut yielded the 8Mut variant that was highly abundant in the viral passaging at higher DRV concentrations. With increased DRV pressure, the next highly abundant variant included additional A71V and L76V mutations (10Mut). The possible 9-mutant intermediate variants en route to 10Mut from 8Mut were not observed in viral passaging. To understand the role of the additional A71V and L76V mutations individually, 9Mut-A71V and 9Mut-L76V variants were generated. Lastly, the addition of I54L mutation onto 10Mut created the 11Mut variant. For all these 8 variants [Figure 1A] and wild type protease, the catalytic activity, DRV inhibition constant, and high-resolution crystal structure were determined and MD simulations performed to elucidate the molecular mechanisms of resistance.

Resistance mutations are selected to maintain catalytic activity

To determine the effects of mutation accumulation on protease catalytic activity, an established enzyme assay for cleaving the natural MA/CA site was used to measure the Michaelis-Menten constant (K_M) and turnover rate (k_{cat}) [**Table S1**]. The ratio of the two measured values was used to calculate the catalytic efficiency (k_{cat}/K_M) for all protease variants. Wild type protease had a K_M of 71.4 μM and k_{cat} of 1282.7 s^{-1}, yielding a catalytic efficiency of 17.1 $μM^{-1}s^{-1}$. 1Mut, 2Mut, and 4Mut variants had WT-like K_M values. However, k_{cat} of these variants were 2–7 fold lower than WT protease, resulting in decreased catalytic efficiency. 2Mut variant had a 5-fold decrease in k_{cat}/K_M compared to WT (k_{cat}/K_M = 3.4 ± 0.2 $μM^{-1}s^{-1}$).

Relative to WT protease, K_M values of the highly mutated 8Mut, 9Mut-A71V, 9Mut-L76V, and 10Mut variants increased 1.4 to 2.4-fold. The k_{cat} value of the 8Mut and 10Mut decreased 2–3 fold, and the k_{cat}/K_M 3–5-fold relative to WT. The most striking loss was in the turnover rates of the 9Mut variants, with 2 orders of magnitude slowing in the cleavage rate (k_{cat} = 16.9 ± 0.3 and 17.7 ± 0.3, respectively for 9Mut-A71V and 9Mut-L76V). Combined with the increase in K_M, these variants were extremely inefficient with k_{cat}/K_M values 171-fold lower relative to WT, representing a severe enzymatic penalty for these intermediate variants that were undetectable in viral passaging. Strikingly, co-occurrence of the A71V and L76V mutations in the 10Mut variant restored catalytic efficiency (k_{cat}/K_M = 5.3 ± 0.1 $μM^{-1}s^{-1}$). The 11Mut exhibited restored, WT-like K_M, but was 13-fold less catalytically efficient compared to WT protease (k_{cat}/K_M = 1.3 ± 0.4 $μM^{-1}s^{-1}$).

This enzymatic assay uses a single cleavage site for a relatively short amount of time. To examine how the highly mutated variants process multiple cleavage sites in full length Gag polyprotein (p55) over a longer time scale, a gel cleavage assay was utilized[35,36] [Figure S1]. Wild-type protease was extremely efficient cleaving 100% of the p55 polyprotein after 90 minutes. All the mutants were able to cleave the initial p2-NC cleavage site at varying rates. The intermediate 9Mut variants were much slower but were still able to process the multiple cleavage sites in the polyprotein. Overall, all protease variants were able to process substrates, with the highly mutated and abundant 10Mut and 11Mut variants having restored catalytic activity.

Accumulation of mutations progressively increases DRV resistance

The enzyme inhibition constant, K_i, of DRV was measured against WT and the selected protease variants using an established assay with an optimized fast-cleaving substrate[37]. Wild type protease is highly susceptible to DRV with a K_i in single digit pM range below the assay limit of reliable detection (K_i < 0.005 nM), as we previously reported[38]. Against 1Mut, 2Mut, and 4Mut, DRV maintained picomolar inhibition, but with K_i increasing progressively 6–80 fold relative to WT [Figure 1C]. The 8Mut variant was significantly more resistant to DRV, with a 2,560-fold increase in resistance compared to WT (K_i = 12.8 nM). Addition of A71V mutation (9Mut-A71V) resulted in a ~2-fold further increase in resistance (K_i = 23.2 nM). In contrast, addition of L76V (9Mut-L76V) caused an order of magnitude increase in DRV resistance (K_i = 172.7 nM), indicating that the distal L76V mutation contributes a significant level of resistance. The 10Mut variant harboring both

distal mutations exhibited 9Mut-L76V-like inhibition (K_i = 156.4 nM). Addition of I54L in 11Mut resulted in dramatic DRV resistance, with a 152,000-fold decrease in inhibition compared to WT (K_i = 759.2 nM). Active site mutations alone do not account for the high levels of resistance seen in the highly mutated variants, as only with the addition of distal mutations the protease was able to exhibit nanomolar-level inhibition. The relationship between DRV inhibition and catalytic efficiency is shown in Figure S2.

Double mutant cycle reveals the interdependency of A71V and L76V

Only the 8Mut to 10Mut variants were observed in viral passaging, thus the addition of A71V and L76V mutations individually were examined as well as the double-mutant cycle analysis [Figure 1D] to elucidate the interdependency of these two distal mutations which cause 12-fold increase in resistance. The change in free energy of inhibitor binding (estimated from $\Delta G = RTln[K_i]$)[39] between the reference variant 8Mut and 10Mut (ΔG_{10Mut}) was calculated. This was then compared to the sum of individual changes for 9Mut-A71V and 9Mut-L76V relative to 8Mut ($\Delta G_{9Mut-A71V} + \Delta G_{9Mut-L76v}$). When the two mutations are independent, the two values are equal ($\Delta G_{10Mut} = \Delta G_{9Mut-A71V} + \Delta G_{9Mut-L76V}$). The sum of free energy changes for the 9Mut intermediates were more than that for 10Mut [Figure 1D]. $\Delta G_{9Mut-A71V}$ was 0.35 ± 0.07 kcal/mol and $\Delta G_{9Mut-L76V}$ was 1.55 ± 0.10 kcal/mol, summing up to 1.90 ± 0.12 kcal/mol; however, ΔG_{10Mut} was 1.45 ± 0.05 kcal/mol. Thus the double mutant cycle analysis shows that A71V and L76V are interdependent[39] in conferring DRV resistance.

Crystal structures show resistance mutations reduce vdW contacts with DRV

To elucidate the changes in inhibitor binding and molecular mechanisms of resistance, we determined cocrystal structures of DRV bound to the wild type HIV-1 protease of the NL4-3 strain, as well as all eight mutated variants at high resolution (1.9–2.2 Å). Structure of DRV bound to WT protease and the I84V variant were previously determined[22] and the 2Mut crystallized in the same space group ($P2_12_12_1$), which contained one homodimer in the asymmetric unit and one inhibitor bound in a single orientation, allowing direct comparison. The variants containing four or more mutations had a different preferred crystallographic lattice and were solved in the $P6_1$ space group. For reliable and direct comparison with these variants, WT protease structure was also determined in the hexagonal space group. These structures were also solved containing one homodimer in the asymmetric unit with one inhibitor bound in a single orientation [Table S2]. Structures of protease variants were overall similar to that of WT protease, with 0.3–0.4 Å overall RMSD [Figure S3]. Following established convention, throughout this analysis the B chain that contacts the aniline moiety of DRV will be denoted the prime (') chain, while the A chain will remain non-prime. The crystal structures determined provided insights into changes in the overall structure and protein–inhibitor interactions due to the accumulating resistance mutations.

Resistance mutations selected under DRV selective pressure are expected to directly or indirectly perturb inhibitor binding and result in reduced potency. To elucidate alterations in inhibitor packing at the active site, we calculated vdW contacts of DRV with the protease residues from the crystal structures [Table S3]. Compared to WT protease, the common multidrug resistance mutation I84V resulted in a loss of 1.2 and 1.8 kcal/mol in vdW

contacts of residue 84 and 84', respectively [Figure S4A, Table S3]. Similarly, all remaining variants, which also contain this mutation, had 1–2.5 kcal/mol losses in vdW contacts at these two residues. In the presence of single I84V mutation, the location of the other active site mutation, V82, had a compensatory gain in vdW contacts in chain B, which was progressively alleviated as the mutations were added. Overall, V82F exhibits a minor gain in vdW contacts (0.2 kcal/mol on average), but the larger nonpolar side chain significantly shifts Arg8' away from DRV resulting in 0.6–1.1 kcal/mol loss in vdW contacts. Addition of flap mutations K45I and M46I caused reduced vdW contacts of I50 in 4Mut variant, and a net loss in total vdW contacts which correlate with reduced potency [Figure S4B, Table S3]. The highly resistant 8Mut variant structure continued the trend, exhibiting a net reduction in vdW contacts. Notably, although the V32I mutation results in a larger amino acid at the edge of the active site, vdW contacts with DRV were reduced 0.4–0.6 and 0.5–1.0 kcal/mol due to the orientation of the additional methyl group away from the inhibitor [Figure S4A, Table S3]. The combination of I13V, L33F and V32I exacerbated losses of vdW contacts at residues I84V and I84V' by 0.2 and 0.7 kcal/mol, respectively. Each mutation, both proximal and distal, either directly or indirectly reduced vdW contacts with the inhibitor, thereby destabilizing inhibitor binding.

Although the 9Mut variants were not observed in viral passaging, determining the structures of these complexes highlighted the interdependent structural changes of the mutations at residue 71 and 76 connecting the 8Mut and 10Mut variants. The A71V mutation, located on the distal "70's loop", is a well-known compensatory mutation that restores catalytic activity of protease variants containing I50V mutation. The mean differences in internal C_α distances [Figure S5], also known as distance difference matrices, were calculated to quantify the structural alterations. Comparing the 8Mut and 9Mut-A71V structures, the A71V mutation pushes the flexible "70s loop" outward up to 0.61 and 0.63 Å for residues 67 and 68', respectively. This outward shift can perturb the beta strand (residues 69-76) that directly interacts with the neighboring beta strand extending up to the flaps (residues 53-65). Although there are very few changes in the internal C_α distances of 8Mut and 9Mut-L76V, when comparing 8Mut and 10Mut the changes in distances reflect an additivity of the differences between the 8Mut versus the two 9Mut structures [Figure S5].

Comparing the complexes of 10Mut and 11Mut helps explain the 5-fold reduction in potency. The longer side chain resulting from the I54L mutation pushed against Pro79 creating a mean difference of 0.36 Å at residues 54 and 54' and 0.20 Å at residues 79 and 79' [Figure S5]. This widening at the active site may destabilize flap closing, and most notably resulted in reduced vdW contacts at residues 49 and 50 of 0.9 and 0.8 kcal/mol, respectively. 11Mut–DRV cocrystal structure had the lowest total intermolecular vdW contacts, correlating with the highest DRV resistance.

Resistance Mutations Perturb Packing in the Hydrophobic Core

Many of the mutations observed were changes from one hydrophobic side chain to another, located in the hydrophobic core of the protein, that considerably change the packing between hydrophobic side chains. K45I results in a side chain rotamer that increases hydrophobic packing with residues I47, V56, L76 at the base of the flaps [Figure S4B].

While I84V may only bridge the gap between V82F and V32I [Figure S4A], V32I is involved in both the L76V cluster (also including K45I, I47, and V56) [Figure S4C] and the I54L cluster (also including I47, I50, and V56) [Figure S4D]. Mutations that perturb the packing of these clusters may also affect "hydrophobic sliding" and thus conformational dynamics of the protease[40, 41].

Resistance Mutations Increase Protease Fluctuations

Despite 5 orders of magnitude increase in the inhibition constant (K_i) of the 11Mut variant relative to WT protease, DRV displayed only minor conformational differences in the cocrystal structures, and MD simulations were performed to better interrogate these structures. Starting from the cocrystal structures of the DRV complexes with WT protease and the resistant variants, three replicates of fully hydrated 100 ns MD simulations were performed for each complex. All simulations reached convergence [Figure S6]. The protein backbone dynamics were compared via root-mean-square fluctuations (RMSF) of C_α atoms [Figure S7]. In WT–DRV complex, the entire backbone was stable, including the flaps (50s region) that close upon the bound inhibitor. In I84V variant, the flaps are slightly more mobile but the active site remains stable. Relative to I84V, the 2Mut (I84V/V82F) protease has reduced flaps fluctuations [Figure S8]. The addition of flap mutations K45I and M46I in the 4Mut resulted in higher fluctuations at chain A flap and residues 79-82. Chain B flap and residues 79'-82' had WT-like fluctuations, remaining stable throughout the MD simulations. Overall, the backbone flexibility of the highly mutated 8Mut, 9Mut-A71V, 9Mut-L76V and 10Mut variants were similar to WT complex, except increased fluctuations at residue 65' in 10Mut. 11Mut variant, bearing flap mutation I54L, displayed extremely large fluctuations at both flaps (residues 42-55 and 45'-52'), increasing more than 0.6 and 0.5 Å in chain A and B, respectively, relative to WT protease. In addition, 11Mut protease had increased fluctuations at residues 79-82 and 78'-80' as well as residues 26-28 in the active site. Thus, the addition of I54L mutation resulted in considerable increase in protease backbone fluctuations at the active site and flaps.

To compare the dynamic ensemble of DRV-bound variants, mean internal C_α distances from MD simulations were compared, similar to the comparison of crystal structures [Figure 2]. In this analysis, the internal distances between all C_α carbons were measured for every frame of the trajectory and an average was taken. Comparing 8Mut to 9Mut-A71V, changes were mostly localized to the 70s beta-sheets and residues 15–20 in both chains, with small changes at the flap elbows [Figure 2A]. Comparing 8Mut to 9Mut-L76V, changes were located in both flaps and the flap elbows. The 80s loop (residues 78 to 82) also had slight changes. In addition, the B chain 70s beta-sheet was affected while the A chain beta-sheet showed very little change [Figure 2B]. As was observed with the crystal structures, alterations in the two 9-Mut variants were additive when 10Mut was compared to 8Mut. The exception was that changes at residues 78-82 were attenuated when A71V and L76V were simultaneously present [Figure 2C]. To complete the cycle, the two 9Muts were compared with 10Mut. Similar to the effects noted earlier, A71V induced changes at the 70s beta-sheets and residues 15–20, while L76V distally affected the flaps and flap elbows [Figure 2D-E]. These alterations in the dynamic ensemble were consistent with the comparison of crystal structures [Figure S5]. However, alterations in the 11Mut variant were accentuated in

the conformational dynamics in the MD simulations, with large structural changes at both flaps and the active site [Figure 2F]. Compared to other variants, the flaps of 11Mut were more open as indicated by the increasing distance between the C_α atoms of I50-V84' and I50'-V84 [Figure S9]. Thus, I54V mutation in 11Mut resulted in destabilization and increased fluctuations of the protease, particularly in the flaps.

Inhibitor fluctuations increase and interactions are destabilized due to resistance mutations

In the cocrystal structures, conformation of DRV was highly similar with the most notable alterations occurring at the P1' moiety [Figure S10]. To investigate how the dynamics of DRV may vary, the RMSF of every DRV heavy atom from the MD simulations was calculated and grouped according to the four DRV moieties [Figure 3]. In complex with WT protease DRV was stable, with RMSF below 0.75 Å except for the P1' moiety. In 1Mut, 2Mut and 4Mut variants fluctuations of DRV moieties were similar to those in WT complex, with few minor exceptions at the P1' and P1 moieties (data not shown). When bound to all the highly mutated variants, DRV had increased flexibility, especially at the P1 and P2 moieties [Figure 3]. In 10Mut complex, the P2 bis-THF moiety had high RMSF compared to WT and all other variants. This moiety makes several hydrogen bonds to protease active site residues and contributes significantly to the high potency of DRV. In complex with the 11Mut variant, the entirety of DRV experienced significant flexibility with the P1 moiety showing the greatest RMSF. Overall, resistance mutations in the protease variants resulted in increased flexibility and higher fluctuations of DRV, suggesting the inhibitor binding and intermolecular interactions are destabilized.

Underlying the high potency of DRV against WT HIV-1 protease is an extensive hydrogen bonding network and strong vdW interactions with active site residues. In highly mutated variants these interactions may be perturbed, as suggested by increased DRV fluctuations. The vdW packing around DRV in complex with 8Mut, 9Mut-A71V, 9Mut-L76V, and 10Mut decreased slightly around the P2 bis-THF moiety compared to WT [Figure 3B]. In 11Mut variant, packing around DRV was significantly reduced at the P1 phenyl, P1' isobutyl, and P2' aniline moieties. In the cocrystal structure bound to WT protease, DRV makes a number of hydrogen bonds with the protease backbone and a water-mediated hydrogen-bonding network with the flaps [Figure 4A]. All these hydrogen bonds were stable and highly conserved throughout the MD simulations [Figure 4B]. In the highly mutated variants, frequency of hydrogen bonds during the MD simulations decreased relative to WT complex, with a severe loss in the hydrogen bond between D25' and the central hydroxyl of DRV. In 8Mut and 9Mut-A71V variants, DRV hydrogen bonds with P2 bis-THF and flaps destabilized 8–17%, and frequency of the hydrogen bond between P2' aniline and backbone nitrogen of D30' decreased 18%. In 9Mut-L76V and 10Mut variants there was a further decrease in the frequency of these hydrogen bonds. In 11Mut variant, water-mediated hydrogen bonding network with the flaps was severely impacted, where two hydrogen bonds were completely lost and the other two maintained only 25% during the MD simulations. The hydrogen bond between D25 and the central hydroxyl was reduced to 62% where it was previously maintained at 83–99% for WT and all other resistant variants. Additionally, the P2' aniline hydrogen bond with D30' was significantly reduced to only 9%, an 85% decrease

compared to WT. Thus, the impact of distal mutations in these highly mutated and resistant variants propagated to the active site to destabilize inhibitor vdW and hydrogen bonding interactions with the protease.

Protease–DRV cross-correlations of fluctuations are lost with increasing resistance

Cross-correlation of collective motions between a ligand and target protein can distinguish tight and weak binders, as we previously reported[28, 30, 42]. For HIV-1 protease, tight binding inhibitors have greater correlation of fluctuations with proximal protease residues, namely the active site and flaps. Contrary to tight binders, weak binding inhibitors lose correlations with these residues corresponding to disrupted intermolecular interactions[30]. Cross-correlations between the fluctuations of DRV and protease residues were calculated from the MD simulation trajectories. As expected, DRV fluctuations were highly correlated with catalytic residues especially at chain A, flaps tips (49-51/49'-51'), and residues 26-31 adjacent to the catalytic D25 [Figure 5A]. DRV also had moderate correlations with residues 81–87, while the flap elbows and 70's beta-sheets were anti-correlated. In 8Mut, DRV maintained correlation with residues 49'-50' but decoupled from chain A flap tip and the active site residues [Figure S11]. While 9Mut-A71V restored the lost correlations of DRV with the catalytic residues, the flap tips became further decoupled. In 9Mut-L76V and 10Mut, DRV fluctuations were further decoupled from the flap tips, but in both variants DRV somewhat maintained correlation with catalytic residues [Figure 5]. In 11Mut variant, positive correlations with DRV were almost completely lost, and the inhibitor established anti-correlations with various regions of the protein (chain B flap, S2' subsite, chain B 70's beta-sheet and 80's loop). Disruptions in the correlations between DRV and 11Mut fluctuations are consistent with increased resistance and weaker intermolecular interactions.

Distal mutations impact the dynamic ensemble to confer resistance

The existing paradigm reasons that distal mutations play a compensatory role in HIV-1 protease, acting to restore catalytic efficiency lost by the accumulation of mutations. However, while certain distal mutations indeed play a compensatory role, they also confer resistance as demonstrated above. We reported previously that distal mutations may propagate their effects to the active site via altering the conformational dynamics of the inhibitor–protease complex, via hydrophobic sliding[40, 41] and the internal hydrogen bonding network[29]. Here we used Jensen-Shannon divergence (JSD) to analyze perturbations in the dynamic ensemble in response to mutations, where protein backbone and side chain dihedral angles were compared between variants during MD simulations. We focused on the impact of 3 additional non-active site mutations, from 8Mut to 11Mut, which results in almost 60-fold increase in resistance.

Consistent with the C_α distance difference analysis, A71V caused only local perturbations in the dynamic ensemble in the 70's beta-sheets [Figure 6A]. L76V also perturbed the dynamics of the 70's loop (chain B) but also had distal effects on the dihedrals of residues in both flaps [Figure 6B]. In combination the effects of A71V and L76V were mostly additive, but together they also perturbed the side chain dihedrals of catalytic D25/25' residues (Φ of D25 and χ^2 of D25') [Figure 6C]. However, the alterations in D25/D25' were subtle and did not significantly change the hydrogen bonds with DRV [Figure 4]. Comparison of 9Mut

variants with 10Mut confirmed the earlier observation that addition of L76V mostly impacted the flaps [Figure 6D], while A71V had local effects on the proximal 70's beta-sheets [Figure 6E]. Lastly, addition of I54L in 11Mut caused considerable changes to the dynamic ensemble compared to 10Mut, both proximal and distal. Most noticeably, side chain dihedrals of catalytic residues were impacted (Ψ and χ^2 of D25 and χ^2 of D25'), which is consistent with decreasing hydrogen bonding between the catalytic residues and the central hydroxyl of DRV [Figure 4]. In addition to changes at the active site, the dynamic ensemble of the flaps and 80's loop of chain B were affected. These results indicate that distal resistance mutations propagate their effects throughout the protein structure, which can be captured as alterations to the dynamic ensemble of protein dihedral angles.

Discussion

Drug resistance occurs when the target evolves and gains mutations to thwart inhibition while still maintaining biological function. We investigated a highly evolved and mutated variant of HIV-1 protease that accumulated 11 mutations to confer high-level resistance to DRV. We systematically dissected the mutations in this variant by generating 8 different subset variants, including separating active site mutations from distal mutations and leveraging a double mutant cycle analysis [Figure 1A]. Integrating activity and inhibition assays, high-resolution crystal structures, and molecular dynamics simulations, we determined how active site and distal mutations contribute to drug resistance. Our results show that resistance progressively increased as the mutations were added, while catalytic activity was maintained. Mutations were selected both inside and outside the active site where the inhibitor binds. We found that the distal mutations play a pivotal role in conferring drug resistance, as proximal mutations alone cannot explain the observed near μM resistance. Double mutant cycle analysis revealed that two such distal mutations are interdependent, with A71V restoring catalytic activity while L76V decreasing inhibition. Our findings challenge the convention that distal mutations are merely compensatory and have secondary or minimal contribution to resistance. On the contrary, we demonstrated that distal mutations act interdependently through complex molecular mechanisms to confer high levels of resistance.

In addition to avoiding inhibition, the mutated protease needs to continue processing viral substrates to allow viral replication. The highly mutated variants (8Mut, 10Mut, 11Mut) detected from viral passaging had somewhat compromised enzymatic activity relative to WT protease, although still allowed for robust viral replication. Unlike the detectable variants, both 9Mut-A71V and 9Mut-L76V variants had severely compromised catalytic activity. The protease must evolve through one of the 9Mut intermediates en route from 8Mut to 10Mut variant, but the virus cannot efficiently replicate to populate detectable levels. A71V is a well-known compensatory mutation that is often observed together with I50V and was reported to restore catalytic activity lost due to this primary mutation[43, 44]. The 8Mut variant does not contain I50V, and when added to the 8Mut variant A71V actually obliterated catalytic activity (9Mut-A71V). In the presence of the L76V mutation though the catalytic activity was restored (10Mut). Thus, the effect of A71V on catalytic activity is context dependent, and may be beneficial or detrimental depending on the background mutation(s).

While resistance due to primary active site mutations are rather straightforward to explain especially using the substrate envelope, how combinations of mutations including distal mutations act interdependently to confer resistance requires further investigations. The crystal structures showed consistent losses in protein-inhibitor vdW contacts, differences in internal distances, and subtle alterations in the binding conformation of DRV. However, given the high levels of resistance and orders of magnitude change in inhibition constants, we expected further and amplified impact on the conformational dynamics of the protease–inhibitor complex. MD simulations highlighted the extent of unfavorable protease–DRV interactions with increasing mutation accumulation. Highly mutated variants resulted in increased protein and inhibitor fluctuations and, most importantly, reduced cross-correlated motions. We previously found that the loss in correlated inhibitor–target fluctuations correlated with reduced inhibitor potency for HCV NS3/4A protease[30] and analogous inhibitors binding to WT HIV-1 protease28. We find here that the same loss exists for highly mutated and resistant variants of HIV-1 protease. In addition, the hydrogen bonds and vdW interactions that contribute to DRV's pM potency in WT HIV protease were greatly diminished. Alterations in the dynamics of the protease–inhibitor complex are crucial in the molecular mechanisms of resistance, and thus the molecular dynamics need to be investigated in addition to static crystal structures structural to elucidate how the effect of distal mutations can be propagated to cause alterations in inhibitor binding.

Here we develop an alternative approach to previously explored methods to elucidate how distal mutations propagate dynamic changes throughout the protease. Expanding upon hydrophobic sliding[40, 41] and internal hydrogen bonding networks[29], quantitative comparison of protein dihedral angles can be used to examine dynamic structural changes. As dihedral angles from MD simulations comprise highly similar probability distributions, a rigorous statistical analysis involving Jensen-Shannon divergence implementation is needed to determine distinct changes. This approach previously quantified differences in long-range dynamics of proteins with and without a ligand bound[45]. Dihedral angle comparison illustrated the propagation of dynamic effects due to distal mutations throughout the protein, including impact of A71V/L76V on the catalytic residues. In the most mutated and most resistant variant 11Mut, addition of I54L mutation conferred many additional changes to the active site and flaps while also propagating changes throughout the protease structure. This analysis demonstrated how mutations cause dynamic changes that propagate to distal regions of the protein, and thus how considering local changes to the static structure is not sufficient to elucidate molecular mechanisms of drug resistance.

Our work provides evidence that distal mutations do indeed contribute to drug resistance through complex dynamic processes, significantly altering protease dynamics and perturbing inhibitor binding. In other HIV-1 protease variants and, perhaps more significantly, other drug targets where distal mutations are observed in clinically relevant resistant variants, the possibility that distal mutations directly confer resistance cannot be ignored. This also has the implication that drug design needs to consider both structural and dynamic changes to the protein–inhibitor complex to be able to effectively target these variants. The increasing threat of drug resistance in many clinically relevant systems highlights the need to characterize resistance mutations, including distal mutations and combinations of mutations that interdependently confer resistance through complex dynamic resistance mechanisms.

Methods

Protein Preparation

Protease gene construction.—Protease gene construction was carried out as previously described[30, 46]. In short, the WT and 10Mut protease variant genes were purchased on a pET11a plasmid with codon optimization for protein expression in *E. coli* (Genewiz). The remaining variants were constructed using QuikChange site-directed mutagenesis (Genewiz). A Q7K mutation was included to prevent autoproteolysis[47].

Protein expression and purification.—The expression, isolation, and purification of WT and mutant HIV-1 proteases used for all assays and crystallization were carried out as previously described[46, 48]. Briefly, the gene encoding the desired HIV protease was subcloned into the heat-inducible pXC35 expression vector (ATCC) and transformed into *E. coli* TAP-106 cells. Cells grown in 6 L of Terrific Broth were lysed with a cell disruptor twice, and the protein was purified from inclusion bodies[49]. The inclusion body centrifugation pellet was dissolved in 50% acetic acid followed by another round of centrifugation at 19K rpm for 30 minutes to remove impurities. Size exclusion chromatography was carried out on a 2.1-L Sephadex G-75 superfine (Sigma Chemical) column equilibrated with 50% acetic acid to separate high molecular weight proteins from the desired protease. Pure fractions of HIV protease were refolded into a 10-fold dilution of refolding buffer [0.05 M sodium acetate at pH 5.5, 5% ethylene glycol, 10% glycerol, and 5 mM DTT]. Folded protein was concentrated down to 0.5–3 mg/mL and stored. The stored protease was used in K_M and K_i assays. For crystallography, an additional purification step was performed with a Pharmacia Superdex 75 FPLC column equilibrated with refolding buffer. Protease fractions purified from the size exclusion column was concentrated to 1–2 mg/mL using an Amicon Ultra-15 10-kDa device (Millipore) for crystallization.

Enzyme Kinetics and Inhibition

K_m Assay.—K_m values were determined as previously described[22, 37, 50]. Briefly, a 10-amino acid substrate containing the natural MA/CA cut site with an EDANS/DABCYL FRET pair was dissolved in 8% DMSO at 40 nM and 6% DMSO at 30 nM. The 30 nM substrate was 4/5 serially diluted from 30 nM to 6 nM. HIV protease was diluted to 120 nM and, using a PerkinElmer Envision plate reader, and 5 μL were added to the 96-well plate to obtain a final concentration of 10 nM. Fluorescence was observed with an excitation at 340 nm and emission at 492 nm and monitored for 200 counts. A FRET inner filter effect correction was applied as previously described[51]. Data corrected for the inner filter effect was analyzed with Prism7.

K_i Assay.—Enzyme inhibition constants (K_i values) were determined as previously described[22, 37, 50]. Briefly, in a 96-well plate, DRV was 3/5 serially diluted from 2000 nM for 8Mut, 1000 nM for 9Mut-A71V, 5000 nM for 9Mut-L76V, 5000 nM for 10Mut, or 10,000 nM for 11Mut. All samples were incubated with 5 nM protein for 1 hour. A 10-amino acid substrate containing an optimized protease cut site[37], purchased from Bachem, with an EDANS/DABCYL FRET pair was dissolved in 4% DMSO at 120 mM. Using a PerkinElmer Envision plate reader, 5 μL of the 120 mM substrate were added to the 96-well

plate to a final concentration of 10 mM. Fluorescence was observed with an excitation at 340 nm and emission at 492 nm and monitored for 200 counts. Data was analyzed with Prism7.

Crystal Structure Determination and Analysis

Crystallography.—Discovery of the condition producing hexagonal cocrystals of DRV bound to a highly mutated variant (8Mut) without seeding was achieved using the Protein Complex Suite Screen (Qiagen), in well G5, consisting of 1 M ammonium sulfate and 1 M potassium chloride and 0.1 M HEPES buffer at pH 7 with a protease concentration of 1.2 mg/mL and 5-fold molar excess of DRV. With previously generated orthorhombic and hexagonal cocrystals in hand, reproducible cocrystals of DRV bound to all variants were achieved as previously described[22]. Briefly, all cocrystals were grown at room temperature by hanging drop vapor diffusion method in 24-well VDX hanging-drop trays (Hampton Research) with protease concentrations between 1.0 to 2.4 mg/mL with 3, 5, or 10-fold molar excess of DRV. Crystallization drops were set with the reservoir solution consisting of 18–26% (w/v) ammonium sulfate and 0.1 M bis-Tris-methane buffer at pH 5.5 set with 2 mL of well solution and 1 mL protein and microseeded with a cat whisker. Crystal morphology and space group was entirely dependent on the microseeds. Diffraction quality crystals were obtained within 1 week. As data was collected at 100 K, cryogenic conditions contained the precipitant solution supplemented with 25% glycerol. For direct structural analysis without possible crystal lattice contact bias, structure of WT protease bound to DRV was also generated and determined in the hexagonal space group.

Data Collection and Structure Solution.—Diffraction data were collected and solved as previously described[22, 30, 46]. Diffraction quality crystals were flash frozen under a cryostream when mounting the crystal at our in-house Rigaku Saturn944 X-ray system. The data for 10Mut cocrystal structure was collected at the Advanced Photon Source at the Argonne National Laboratory, beamline 19-ID. All cocrystal diffraction intensities were indexed, integrated, and scaled using HKL3000[52]. Structures were solved using molecular replacement with PHASER[53]. Model building and refinement were performed using Coot[54] and Phenix[55]. During refinement, all crystals utilized optimized stereochemical weights and non-crystallographic symmetry operators. Hexagonal crystals grew as pseudo-merohedral twins and were solved with a twin law applied (h,-h-k,-l). Ligands were designed in Maestro and the output sdf file was used in the Phenix program eLBOW[56] to generate the cif file containing atomic positions and constraints necessary for ligand refinement. Iterative rounds of crystallographic refinement were carried out until convergence was achieved. To limit bias throughout the refinement process, 5% of the data were reserved for the free R-value calculation[57]. MolProbity[58] was applied to evaluate the final structures before deposition in the PDB. Structure analysis, superposition, and figure generation were performed using PyMOL[59]. X-ray data collection and crystallographic refinement statistics are presented in the Supporting Information [Table S2].

Internal Distance Analysis of Crystal Structures.—Distance-difference matrices were generated as previously described[60] to reveal structural changes between inhibitor–protease pairs.

Intermolecular vdW Contact Analysis of Crystal Structures.—To calculate the intermolecular van der Waals (vdW) interaction energies the crystal structures were prepared using the Schrodinger Protein Preparation Wizard[61]. Hydrogen atoms were added, protonation states determined and the structures were minimized. The protease active site was monoprotonated at D25. Subsequently, forcefield parameters were assigned using the OPLS2005 force field[62]. Interaction energies between the inhibitor and protease were estimated using a simplified Lennard-Jones potential, as previously described in detail[63]. Briefly, the vdW energy was calculated for pairwise interactions depending on the types of atoms interacting and the distance in-between. For each protease residue, the change in vdW interactions relative to a WT complex in the same space group was also calculated for each mutant structure.

Molecular Dynamics Simulations and Analysis

System Preparation.—High-resolution crystal structures were prepared using the Protein Preparation Wizard from Maestro within the Schrodinger Suite[61] as previously described[64]. Briefly, cocrystallized phosphates were removed, missing atoms were added using Prime[65], and PROPKA[66, 67] was used to determine the protonation state of side chains at pH 7.0. Lastly, the structure was minimized to a convergence criterion of 0.3 Å using Impref[68].

Molecular Dynamics Simulations.—The prepared systems were placed in a cubic TIP3P water box measuring 12 Å on each side. MD simulations were carried out as previously described[64] using Desmond software suite within Schrodinger[69]. Briefly, chloride ions were used to neutralize the system and 0.15 M salt were added using sodium and chloride ions. The OPLS3 force field was used to parameterize the ligand and protein. Prior to starting the 100 ns MD simulations, the solvated system was minimized using the stepwise procedure described previously[64]. Triplicates of 100 ns simulations for WT, I84V, 2Mut, 4Mut, 8Mut, 9Mut-A71V, 9 Mut L76V, 10Mut, and 11Mut each with a randomized velocity were started using the protocol previously developed[64]. The root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) of Cα atoms were calculated by utilizing the simulation interactions diagram from Maestro within the Schrodinger Suite.

Intermolecular vdW Contact Analysis of MD Trajectories.—Intermolecular vdW interactions were calculated using a previously published protocol utilizing Lennard-Jones potential[46, 70]. The force field is not optimized for a sulfur violating the octet rule such as that found in the P2' moiety of DRV. As such, the packing around the P2' sulfur was calculated by averaging the vdW packing of the four adjacent atoms.

Dynamic Structural Analysis

Cross-Correlations of Protease/DRV Fluctuations.—The cross-correlation coefficients between the protease C$_\alpha$ atoms and DRV heavy atoms were calculated using a previously published protocol[28]. Briefly, the atom fluctuations were determined for each MD simulation. The cross-correlation between atom pairs, such as atoms i and j, was determined according to

$$C_{i,j} = \frac{\langle \Delta R_i \cdot \Delta R_j \rangle}{\sqrt{\langle \Delta R_i^2 \rangle \langle \Delta R_j^2 \rangle}}$$

where R_i and R_j are the positional changes of atom i and j, respectively, and the angle brackets denote an ensemble average. The cross-correlation coefficient ranges from –1 to 1, where –1 represents complete anticorrelation between the atom pair, zero represents no correlation, and 1 represents complete correlation. The cross-correlation values were calculated using an in-house script and mapped onto the protease structure by replacing the B-factors in the PDB coordinate file, using PyMOL[59].

Distance-Difference Matrices.—The distances between all C_α atom pairs for a given structure were calculated as a *198×198* matrix for all frames in the trajectory. The distance was calculated for each C_α atom pair, giving the mean distance over the trajectory. The distance difference matrix was produced by subtracting each mean distance in the reference structure from the corresponding mean distance in the given structure. The overall difference for each residue was then calculated by taking the average of the absolute values of all the 198 distance differences involving that residue. PyMOL[59] was used for visualization of distance differences where the protease backbone was represented as a cartoon-putty with increasing thickness and warmer color for increasing deviation.

Jensen-Shannon Divergence Analysis of Dihedral Angles.—Over the MD simulations, all the Φ, Ψ, χ dihedral angles of protease residues were measured. These dihedral angle measurements were used as input for the MutInf software package[45]. Utilizing this package, the Jensen-Shannon divergence (JSD), which analyzes the difference between two probability distributions, was calculated. An α, of 0.05 was used as a threshold to filter out non-statistically significant differences in distributions, where a p-value > α was considered non-statistically significant. In such cases, the JSD was set to zero. For visualization purposes, using PyMOL[59], if a residue had multiple dihedral angles with a calculated JSD, the greatest value was selected.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

[1]. O'Neill J Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations. http://amr-review.org/sites/default/files/AMR%20Review%20Paper%20-%20Tackling%20a%20crisis%20for%20the%20health%20and%20wealth%20of%20nations_1.pdf (accessed April 12, 2019).

[2]. Editors, P. M. (2016) Antimicrobial Resistance: Is the World UNprepared?, PLoS Med 13, e1002130. [PubMed: 27618631]

[3]. (2018) World Health Organization. Fact Sheet: Antimicrobial Resistance.

[4]. Galvani AP (2005) The role of mutation accumulation in HIV progression, Proc Biol Sci 272, 1851–1858. [PubMed: 16096099]

[5]. Maldarelli F, Kearney M, Palmer S, Stephens R, Mican J, Polis MA, Davey RT, Kovacs J, Shao W, Rock-Kress D, Metcalf JA, Rehm C, Greer SE, Lucey DL, Danley K, Alter H, Mellors JW, and Coffin JM (2013) HIV populations are large and accumulate high genetic diversity in a nonlinear fashion, J Virol 87, 10313–10323. [PubMed: 23678164]

[6]. Patel PH, and Preston BD (1994) Marked infidelity of human immunodeficiency virus type 1 reverse transcriptase at RNA and DNA template ends, Proc Natl Acad Sci U S A 91, 549–553. [PubMed: 7507249]

[7]. Abram ME, Ferris AL, Das K, Quinones O, Shao W, Tuske S, Alvord WG, Arnold E, and Hughes SH (2014) Mutations in HIV-1 reverse transcriptase affect the errors made in a single cycle of viral replication, J Virol 88, 7589–7601. [PubMed: 24760888]

[8]. Abram ME, Ferris AL, Shao W, Alvord WG, and Hughes SH (2010) Nature, position, and frequency of mutations made in a single cycle of HIV-1 replication, J Virol 84, 9864–9878. [PubMed: 20660205]

[9]. Smyth RP, Schlub TE, Grimm AJ, Waugh C, Ellenberg P, Chopra A, Mallal S, Cromer D, Mak J, and Davenport MP (2014) Identifying recombination hot spots in the HIV-1 genome, J Virol 88, 2891–2902. [PubMed: 24371048]

[10]. Onafuwa-Nuga A, and Telesnitsky A (2009) The remarkable frequency of human immunodeficiency virus type 1 genetic recombination, Microbiol Mol Biol Rev 73, 451–480, Table of Contents. [PubMed: 19721086]

[11]. Yasuda JM, Miller C, Currier JS, Forthal DN, Kemper CA, Beall GN, Tilles JG, Capparelli EV, McCutchan JA, Haubrich RH, and California Collaborative Treatment, G. (2004) The correlation between plasma concentrations of protease inhibitors, medication adherence and virological outcome in HIV-infected patients, Antivir Ther 9, 753–761. [PubMed: 15535413]

[12]. Wu TD, Schiffer CA, Gonzales MJ, Taylor J, Kantor R, Chou S, Israelski D, Zolopa AR, Fessel WJ, and Shafer RW (2003) Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments, J Virol 77, 4836–4847. [PubMed: 12663790]

[13]. Clavel F, and Mammano F (2010) Role of Gag in HIV Resistance to Protease Inhibitors, Viruses 2, 1411–1426. [PubMed: 21994687]

[14]. Dam E, Quercia R, Glass B, Descamps D, Launay O, Duval X, Krausslich HG, Hance AJ, Clavel F, and Group AS (2009) Gag mutations strongly contribute to HIV-1 resistance to protease inhibitors in highly drug-experienced patients besides compensating for fitness loss, PLoS Pathog 5, e1000345. [PubMed: 19300491]

[15]. 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, and Wigerinck PB (2005) Discovery and selection of TMC114, a next generation HIV-1 protease inhibitor, J Med Chem 48, 1813–1822. [PubMed: 15771427]

[16]. Nalam MN, Peeters A, Jonckers TH, Dierynck I, and Schiffer CA (2007) Crystal structure of lysine sulfonamide inhibitor reveals the displacement of the conserved flap water molecule in human immunodeficiency virus type 1 protease, J Virol 81, 9512–9518. [PubMed: 17596316]

[17]. Cai Y, and Schiffer CA (2010) Decomposing the energetic impact of drug resistant mutations in HIV-1 protease on binding DRV, J Chem Theory Comput 6, 1358–1368. [PubMed: 20543885]

[18]. Prabu-Jeyabalan M, Nalivaika E, and Schiffer CA (2002) Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes, Structure 10, 369–381. [PubMed: 12005435]

[19]. Ali A, Bandaranayake RM, Cai Y, King NM, Kolli M, Mittal S, Murzycki JF, Nalam MN, Nalivaika EA, Ozen A, Prabu-Jeyabalan MM, Thayer K, and Schiffer CA (2010) Molecular Basis for Drug Resistance in HIV-1 Protease, Viruses 2, 2509–2535. [PubMed: 21994628]

[20]. Lv Z, Chu Y, and Wang Y (2015) HIV protease inhibitors: a review of molecular selectivity and toxicity, HIV AIDS (Auckl) 7, 95–104. [PubMed: 25897264]

[21]. Nalam MN, Ali A, Reddy GS, Cao H, Anjum SG, Altman MD, Yilmaz NK, Tidor B, Rana TM, and Schiffer CA (2013) Substrate envelope-designed potent HIV-1 protease inhibitors to avoid drug resistance, Chem Biol 20, 1116–1124. [PubMed: 24012370]

[22]. Lockbaum GJ, Leidner F, Rusere LN, Henes M, Kosovrasti K, Nachum GS, Nalivaika EA, Ali A, Yilmaz NK, and Schiffer CA (2019) Structural Adaptation of Darunavir Analogues against Primary Mutations in HIV-1 Protease, ACS Infect Dis 5, 316–325. [PubMed: 30543749]

[23]. Wensing AM, van Maarseveen NM, and Nijhuis M (2010) Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance, Antiviral Res 85, 59–74. [PubMed: 19853627]

[24]. Muzammil S, Ross P, and Freire E (2003) A major role for a set of non-active site mutations in the development of HIV-1 protease drug resistance, Biochemistry 42, 631–638. [PubMed: 12534275]

[25]. Ohtaka H, Schon A, and Freire E (2003) Multidrug resistance to HIV-1 protease inhibition requires cooperative coupling between distal mutations, Biochemistry 42, 13659–13666. [PubMed: 14622012]

[26]. Clemente JC, Moose RE, Hemrajani R, Whitford LR, Govindasamy L, Reutzel R, McKenna R, Agbandje-McKenna M, Goodenow MM, and Dunn BM (2004) Comparing the accumulation of active- and nonactive-site mutations in the HIV-1 protease, Biochemistry 43, 12141–12151. [PubMed: 15379553]

[27]. Prachanronarong KL, Canale AS, Liu P, Somasundaran M, Hou S, Poh Y-P, Han T, Zhu Q, Renzette N, Zeldovich KB, Kowalik TF, Kurt-Yilmaz N, Jensen JD, Bolon DNA, Marasco WA, Finberg RW, Schiffer CA, and Wang JP (2019) Mutations in Influenza A Virus Neuraminidase and Hemagglutinin Confer Resistance against a Broadly Neutralizing Hemagglutinin Stem Antibody, Journal of Virology 93, e01639–01618. [PubMed: 30381484]

[28]. Paulsen JL, Leidner F, Ragland DA, Kurt Yilmaz N, and Schiffer CA (2017) Interdependence of Inhibitor Recognition in HIV-1 Protease, J Chem Theory Comput 13, 2300–2309. [PubMed: 28358514]

[29]. Ragland DA, Nalivaika EA, Nalam MN, Prachanronarong KL, Cao H, Bandaranayake RM, Cai Y, Kurt-Yilmaz N, and Schiffer CA (2014) 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 136, 11956–11963. [PubMed: 25091085]

[30]. Soumana DI, Kurt Yilmaz N, Ali A, Prachanronarong KL, and Schiffer CA (2016) Molecular and Dynamic Mechanism Underlying Drug Resistance in Genotype 3 Hepatitis C NS3/4A Protease, J Am Chem Soc 138, 11850–11859. [PubMed: 27512818]

[31]. Matthew AN, Leidner F, Newton A, Petropoulos CJ, Huang W, Ali A, KurtYilmaz N, and Schiffer CA (2018) Molecular Mechanism of Resistance in a Clinically Significant Double-Mutant Variant of HCV NS3/4A Protease, Structure 26, 1360–1372 e1365. [PubMed: 30146168]

[32]. Rhee SY, Gonzales MJ, Kantor R, Betts BJ, Ravela J, and Shafer RW (2003) Human immunodeficiency virus reverse transcriptase and protease sequence database, Nucleic Acids Res 31, 298–303. [PubMed: 12520007]

[33]. Shafer RW (2006) Rationale and uses of a public HIV drug-resistance database, J Infect Dis 194 Suppl 1, S51–58. [PubMed: 16921473]

[34]. Ragland DA, Whitfield TW, Lee SK, Swanstrom R, Zeldovich KB, Kurt-Yilmaz N, and Schiffer CA (2017) Elucidating the Interdependence of Drug Resistance from Combinations of Mutations, J Chem Theory Comput 13, 5671–5682. [PubMed: 28915040]

[35]. Lee SK, Potempa M, Kolli M, Ozen A, Schiffer CA, and Swanstrom R (2012) Context surrounding processing sites is crucial in determining cleavage rate of a subset of processing sites

in HIV-1 Gag and Gag-Pro-Pol polyprotein precursors by viral protease, J Biol Chem 287, 13279–13290. [PubMed: 22334652]

[36]. Deshmukh L, Louis JM, Ghirlando R, and Clore GM (2016) Transient HIV-1 Gag-protease interactions revealed by paramagnetic NMR suggest origins of compensatory drug resistance mutations, Proc Natl Acad Sci U S A 113, 12456–12461. [PubMed: 27791180]

[37]. Windsor IW, and Raines RT (2015) Fluorogenic Assay for Inhibitors of HIV-1 Protease with Sub-picomolar Affinity, Sci Rep 5, 11286. [PubMed: 26261098]

[38]. King NM, Prabu-Jeyabalan M, Nalivaika EA, Wigerinck P, de Bethune MP, and Schiffer CA (2004) Structural and thermodynamic basis for the binding of TMC114, a next-generation human immunodeficiency virus type 1 protease inhibitor, J Virol 78, 12012–12021. [PubMed: 15479840]

[39]. Horovitz A (1996) Double-mutant cycles: a powerful tool for analyzing protein structure and function, Fold Des 1, R121–126. [PubMed: 9080186]

[40]. Goldfarb NE, Ohanessian M, Biswas S, McGee TD Jr., Mahon BP, Ostrov DA, Garcia J, Tang Y, McKenna R, Roitberg A, and Dunn BM (2015) Defective hydrophobic sliding mechanism and active site expansion in HIV-1 protease drug resistant variant Gly48Thr/Leu89Met: mechanisms for the loss of saquinavir binding potency, Biochemistry 54, 422–433. [PubMed: 25513833]

[41]. Foulkes-Murzycki JE, Scott WR, and Schiffer CA (2007) Hydrophobic sliding: a possible mechanism for drug resistance in human immunodeficiency virus type 1 protease, Structure 15, 225–233. [PubMed: 17292840]

[42]. Ozen A, Sherman W, and Schiffer CA (2013) Improving the Resistance Profile of Hepatitis C NS3/4A Inhibitors: Dynamic Substrate Envelope Guided Design, J Chem Theory Comput 9, 5693–5705. [PubMed: 24587770]

[43]. Nijhuis M, Schuurman R, de Jong D, Erickson J, Gustchina E, Albert J, Schipper P, Gulnik S, and Boucher CA (1999) Increased fitness of drug resistant HIV-1 protease as a result of acquisition of compensatory mutations during suboptimal therapy, AIDS 13, 2349–2359. [PubMed: 10597776]

[44]. Meher BR, and Wang Y (2012) Interaction of I50V mutant and I50L/A71V double mutant HIV-protease with inhibitor TMC114 (darunavir): molecular dynamics simulation and binding free energy studies, J Phys Chem B 116, 1884–1900. [PubMed: 22239286]

[45]. McClendon CL, Hua L, Barreiro A, and Jacobson MP (2012) Comparing Conformational Ensembles Using the Kullback-Leibler Divergence Expansion, J Chem Theory Comput 8, 2115–2126. [PubMed: 23316121]

[46]. Ozen A, Lin KH, Kurt Yilmaz N, and Schiffer CA (2014) Structural basis and distal effects of Gag substrate coevolution in drug resistance to HIV-1 protease, Proc Natl Acad Sci U S A 111, 15993–15998. [PubMed: 25355911]

[47]. Rose JR, Salto R, and Craik CS (1993) Regulation of autoproteolysis of the HIV-1 and HIV-2 proteases with engineered amino acid substitutions, J Biol Chem 268, 11939–11945. [PubMed: 8505318]

[48]. King NM, Melnick L, Prabu-Jeyabalan M, Nalivaika EA, Yang SS, Gao Y, Nie X, Zepp C, Heefner DL, and Schiffer CA (2002) Lack of synergy for inhibitors targeting a multi-drug-resistant HIV-1 protease, Protein Sci 11, 418–429. [PubMed: 11790852]

[49]. Hui JO, Tomasselli AG, Reardon IM, Lull JM, Brunner DP, Tomich C-SC, and Heinrikson RL (1993) Large scale purification and refolding of HIV-1 protease fromEscherichia coli inclusion bodies, Journal of protein chemistry 12, 323–327. [PubMed: 8397790]

[50]. Matayoshi ED, Wang GT, Krafft GA, and Erickson J (1990) Novel fluorogenic substrates for assaying retroviral proteases by resonance energy transfer, Science 247, 954–958. [PubMed: 2106161]

[51]. Liu Y, Kati W, Chen CM, Tripathi R, Molla A, and Kohlbrenner W (1999) Use of a fluorescence plate reader for measuring kinetic parameters with inner filter effect correction, Anal Biochem 267, 331–335. [PubMed: 10036138]

[52]. Otwinowski Z, and Minor W (1997) [20] Processing of X-ray diffraction data collected in oscillation mode, Methods Enzymol 276, 307–326.

[53]. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, and Read RJ (2007) Phaser crystallographic software, J Appl Crystallogr 40, 658–674. [PubMed: 19461840]

[54]. Emsley P, and Cowtan K (2004) Coot: model-building tools for molecular graphics, Acta Crystallogr D Biol Crystallogr 60, 2126–2132. [PubMed: 15572765]

[55]. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, and Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr D Biol Crystallogr 66, 213–221. [PubMed: 20124702]

[56]. Moriarty NW, Grosse-Kunstleve RW, and Adams PD (2009) electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation, Acta Crystallogr D Biol Crystallogr 65, 1074–1080. [PubMed: 19770504]

[57]. Brunger AT (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures, Nature 355, 472–475. [PubMed: 18481394]

[58]. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB 3rd, Snoeyink J, Richardson JS, and Richardson DC (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids, Nucleic Acids Res 35, W375–383. [PubMed: 17452350]

[59]. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

[60]. Prabu-Jeyabalan M, Nalivaika EA, Romano K, and Schiffer CA (2006) Mechanism of substrate recognition by drug-resistant human immunodeficiency virus type 1 protease variants revealed by a novel structural intermediate, J Virol 80, 3607–3616. [PubMed: 16537628]

[61]. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, and Sherman W (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments, J Comput Aided Mol Des 27, 221–234. [PubMed: 23579614]

[62]. Jorgensen WL, Maxwell DS, and Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, Journal of the American Chemical Society 118, 11225–11236.

[63]. Nalam MN, Ali A, Altman MD, Reddy GS, Chellappan S, Kairys V, Ozen A, Cao H, Gilson MK, Tidor B, Rana TM, and Schiffer CA (2010) Evaluating the substrate-envelope hypothesis: structural analysis of novel HIV-1 protease inhibitors designed to be robust against drug resistance, J Virol 84, 5368–5378. [PubMed: 20237088]

[64]. Leidner F, Kurt Yilmaz N, Paulsen J, Muller YA, and Schiffer CA (2018) Hydration Structure and Dynamics of Inhibitor-Bound HIV-1 Protease, J Chem Theory Comput 14, 2784–2796. [PubMed: 29570286]

[65]. Jacobson MP, Friesner RA, Xiang Z, and Honig B (2002) On the role of the crystal environment in determining protein side-chain conformations, J Mol Biol 320, 597–608. [PubMed: 12096912]

[66]. Olsson MH, Sondergaard CR, Rostkowski M, and Jensen JH (2011) PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions, J Chem Theory Comput 7, 525–537. [PubMed: 26596171]

[67]. Sondergaard CR, Olsson MH, Rostkowski M, and Jensen JH (2011) Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values, J Chem Theory Comput 7, 2284–2295. [PubMed: 26606496]

[68]. Banks JL, Beard HS, Cao Y, Cho AE, Damm W, Farid R, Felts AK, Halgren TA, Mainz DT, Maple JR, Murphy R, Philipp DM, Repasky MP, Zhang LY, Berne BJ, Friesner RA, Gallicchio E, and Levy RM (2005) Integrated Modeling Program, Applied Chemical Theory (IMPACT), J Comput Chem 26, 1752–1780. [PubMed: 16211539]

[69]. Bowers KJ, Chow DE, Xu H, Dror RO, Eastwood MP, Gregersen BA, Klepeis JL, Kolossvary I, Moraes MA, Sacerdoti FD, Salmon JK, Shan Y, and Shaw DE (2006) Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters, In SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, pp 43–43.

[70]. Ozen A, Haliloglu T, and Schiffer CA (2011) Dynamics of preferential substrate recognition in HIV-1 protease: redefining the substrate envelope, J Mol Biol 410, 726–744. [PubMed: 21762811]
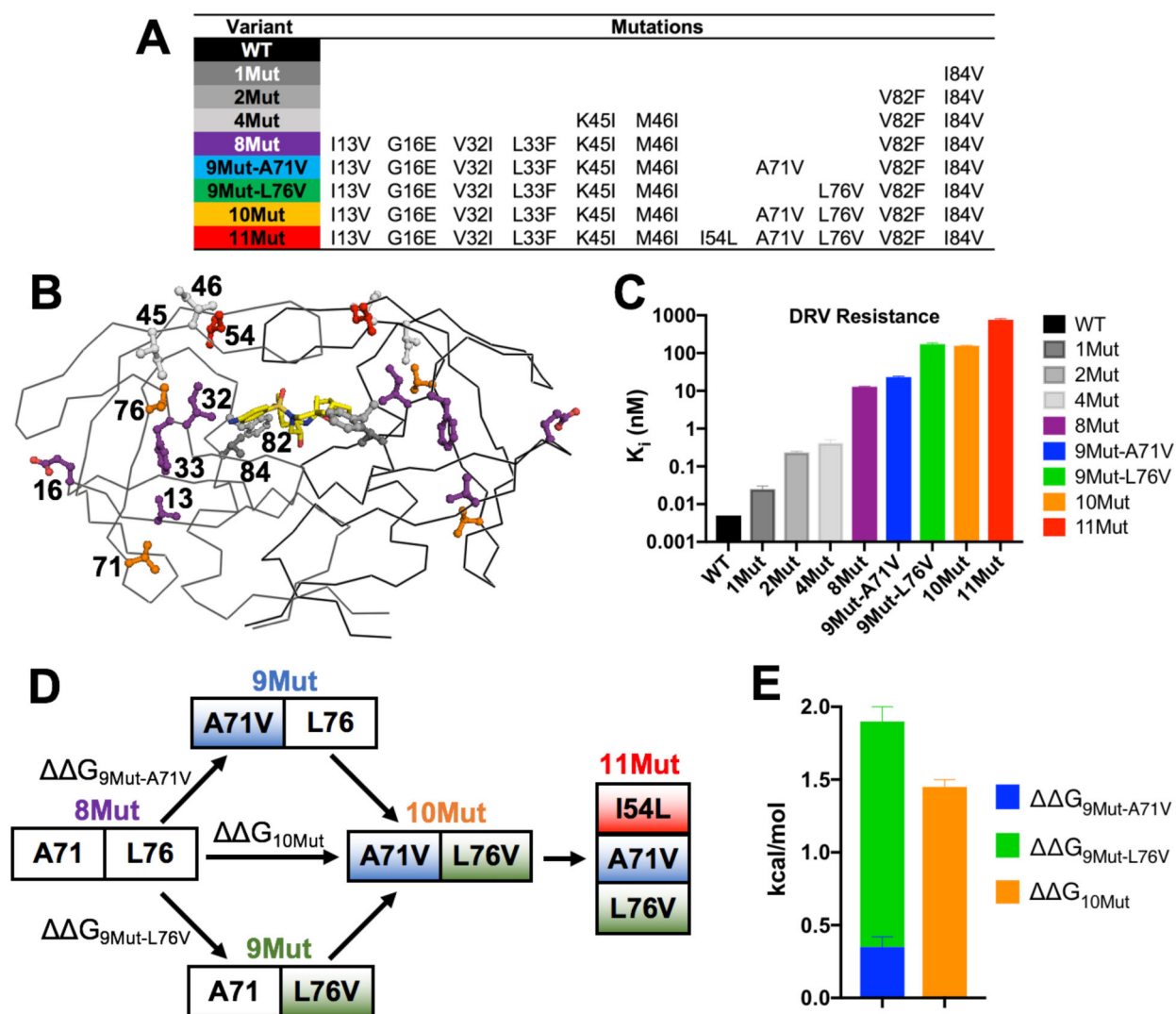
Figure 1.
A-B) Location of resistance mutations in the variants displayed on HIV-1 protease structure, with side chain of residues mutated shown as ball and stick and colored to match the variant added to in panel B. DRV depicted as yellow sticks. Mutations V32I, V82F, and I84V are classified as proximal while all other mutations are distal. **C)** Inhibition constant, Ki, of each variant against DRV. Resistance increases as mutations accumulate. **D)** Double mutant cycle of distal mutations A71V and L76V going from 8Mut to 10Mut variant with 12-fold higher resistance. The additional I54L mutation (11Mut) further increases resistance 5-fold. **E)** Mutations are interdependent ($\Delta\Delta G_{\text{9Mut-A71V}} + \Delta\Delta G_{\text{9Mut-L76V}} \neq \Delta\Delta G_{\text{10Mut}}$).
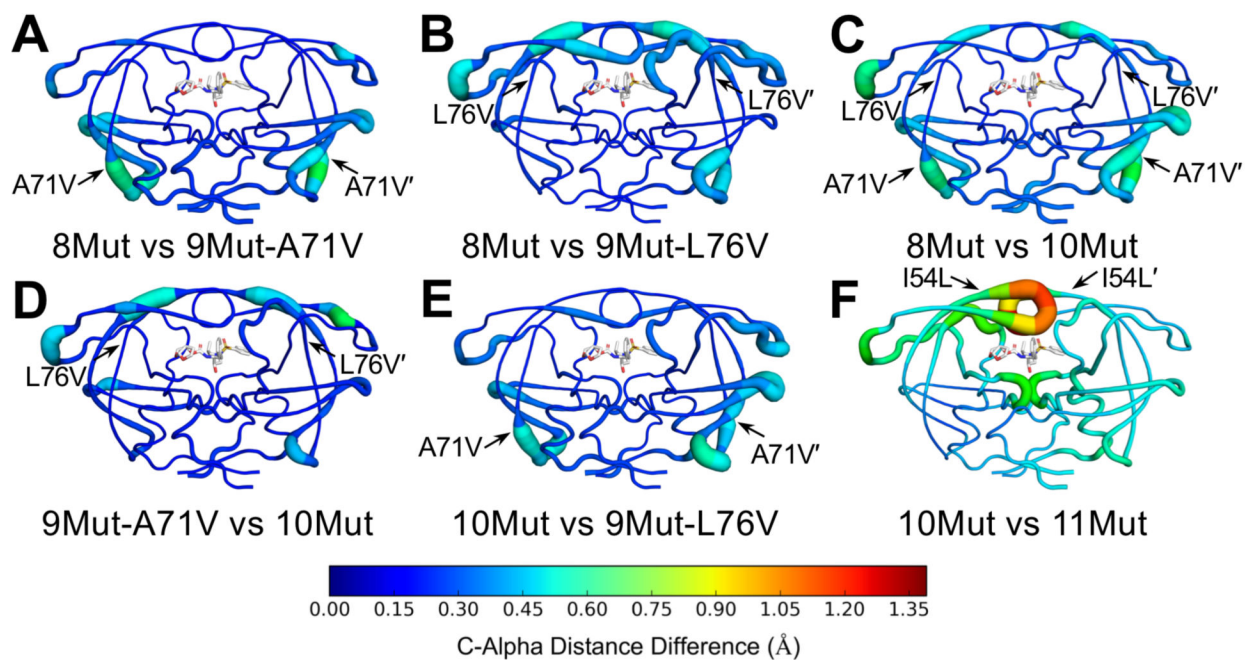
Figure 2.
Mean internal C_α distance differences between (A) 8Mut and 9Mut-A71V, (B) 8Mut and 9Mut-L76V, (C) 8Mut and 10Mut, (D) 9Mut-A71V and 10Mut, (E) 10Mut and 9Mut-L76V, and (F) 10Mut and 11Mut protease variants from MD simulations plotted onto protease structure. In all panels, DRV is shown as sticks in the active site. Putty thickness and color indicate distance difference where hot colors indicate large changes.
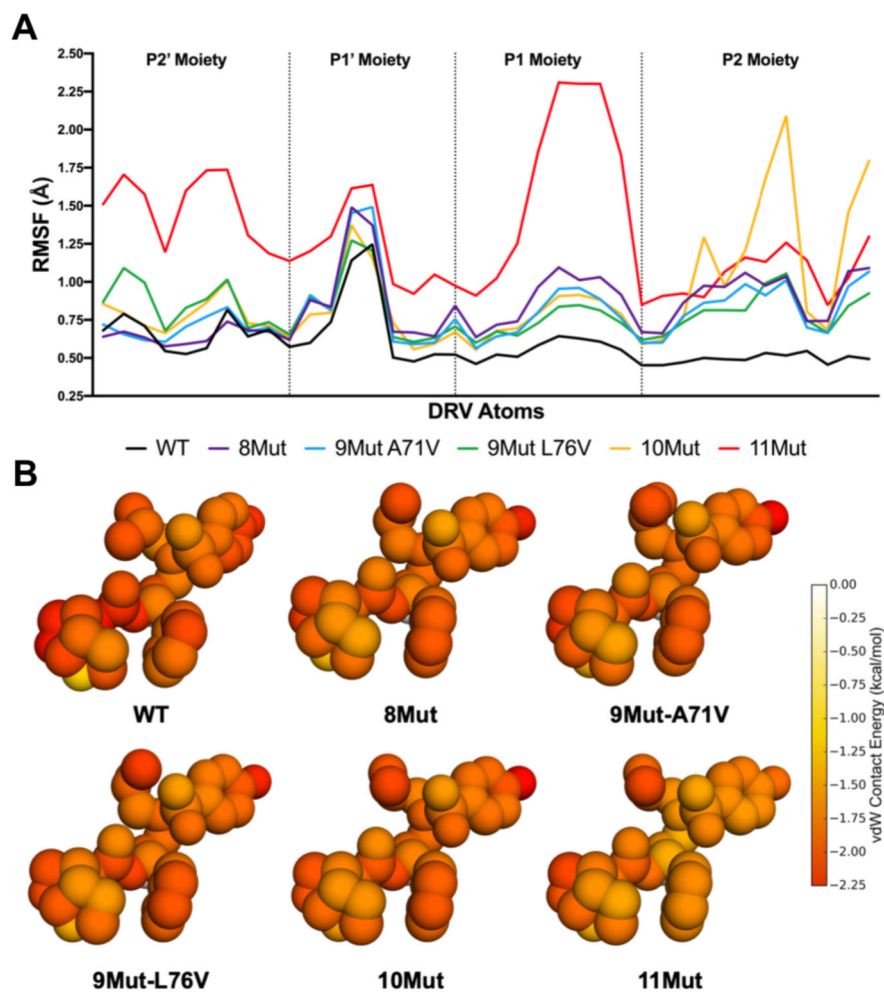
Figure 3.
A) Root-mean-square fluctuation (RMSF) of DRV atoms grouped by moiety monitored during MD simulations bound to WT and resistant HIV-1 protease variants (color scheme as in Figure 1). **B)** Packing around DRV in complex with WT protease and resistant variants. Total per atom protease–DRV vdW contact energies mapped onto the respective DRV crystal structure, with red indicating more contacts.
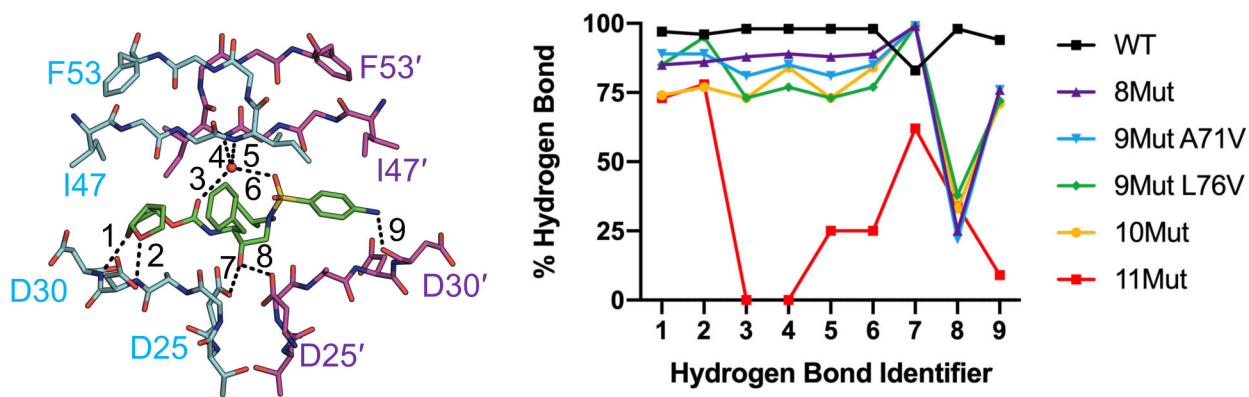
Figure 4.
Resistant HIV-1 protease variants lose hydrogen bonding with inhibitor DRV. **A**) Hydrogen bonds with DRV (green sticks) indicated as black dashed lines on the WT protease crystal structure. The "flap water" molecule is depicted as a red sphere. Numbers correspond to hydrogen bonds shown in panel B. **B**) Hydrogen bonding frequencies from MD simulations. The water-mediated hydrogen-bonding network between the protease flaps and DRV is severely disrupted in 11Mut.
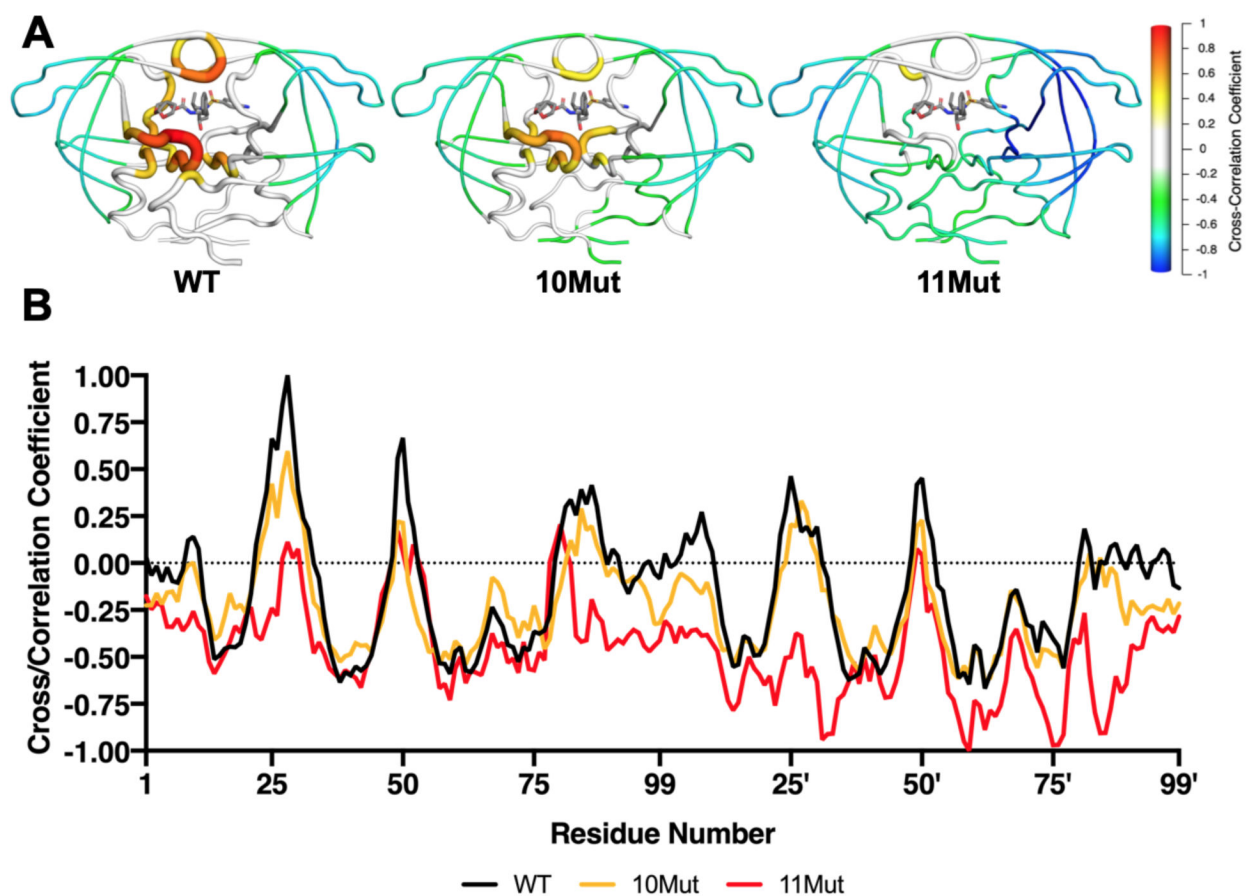
Figure 5.
Cross-correlation of fluctuations between HIV-1 protease and DRV from MD simulations.
A) Mean cross-correlation coefficient for each residue mapped onto protease structure,
shown as cartoon putty for WT, 10Mut, and 11Mut variants. **B**) Cross-correlations for WT,
10Mut, and 11Mut plotted. As resistance mutations accumulate, positive correlations
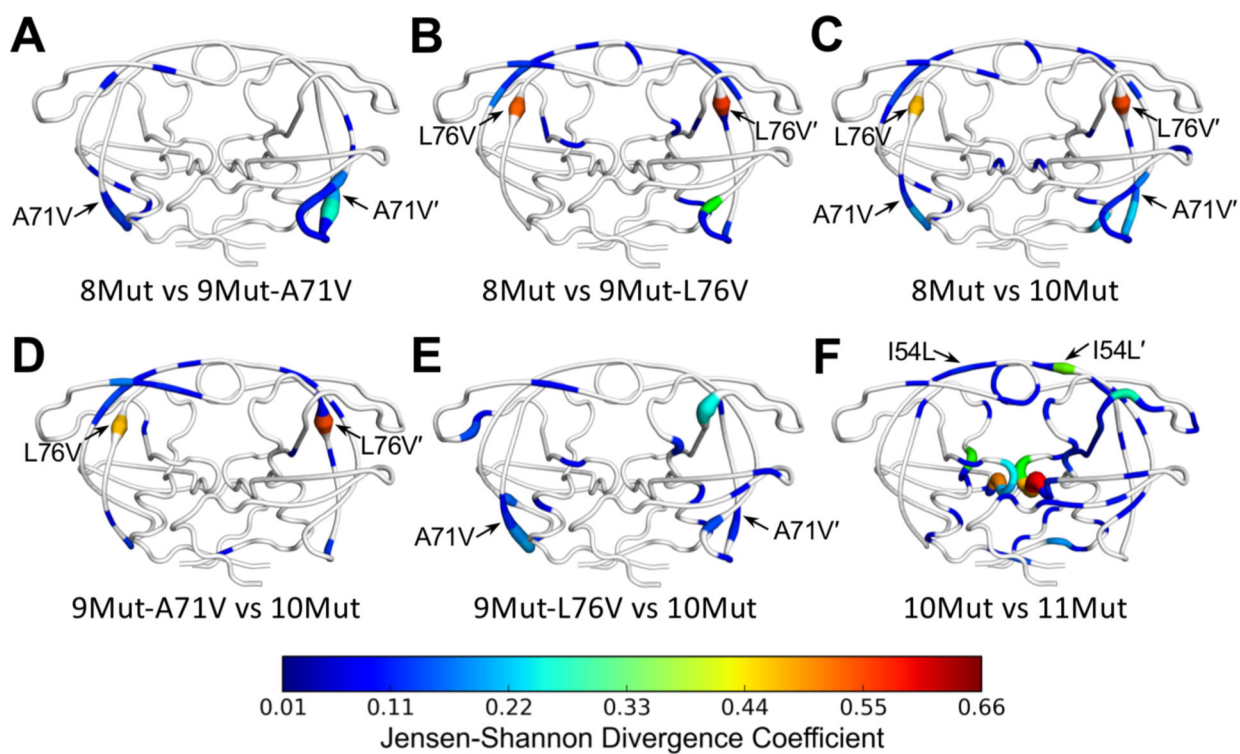decrease while negative correlations increase.

Figure 6.
Comparison of protease dynamics via Jensen-Shannon divergence of dihedral angles between (A) 8Mut and 9Mut-A71V, (B) 8Mut and 9Mut-L76V, (C) 8Mut and 10Mut, (D) 9Mut-A71V and 10Mut, (E) 9Mut-L76V and 10Mut, and (F) 10Mut and 11Mut from MD simulations, mapped onto protease structure. Tube thickness and warmer colors indicate larger perturbation of the dynamic ensemble. Residues in white had no significant difference between the two variants. Arrows denote location of mutation(s) different between the variants compared.