

RESEARCH ARTICLE

Accessory mutations balance the marginal stability of the HIV-1 protease in drug resistance

Thomas R. Weikl¹  | Bahram Hemmateenejad^{1,2}

¹Max Planck Institute of Colloids and Interfaces, Department of Theory and Bio-Systems, Science Park Golm, Potsdam, Germany

²Department of Chemistry, Shiraz University, Shiraz, Iran

Correspondence

Thomas R. Weikl, Max Planck Institute of Colloids and Interfaces, Department of Theory and Bio-Systems, Science Park Golm, 14424 Potsdam, Germany.

Email: thomas.weikl@mpikg.mpg.de

Funding information

Shiraz University; Alexander von Humboldt Foundation

Peer Review

The peer review history for this article is available at <https://publons.com/publon/10.1002/prot.25826>.

Abstract

The HIV-1 protease is a major target of inhibitor drugs in AIDS therapies. The therapies are impaired by mutations of the HIV-1 protease that can lead to resistance to protease inhibitors. These mutations are classified into major mutations, which usually occur first and clearly reduce the susceptibility to protease inhibitors, and minor, accessory mutations that occur later and individually do not substantially affect the susceptibility to inhibitors. Major mutations are predominantly located in the active site of the HIV-1 protease and can directly interfere with inhibitor binding. Minor mutations, in contrast, are typically located distal to the active site. A central question is how these distal mutations contribute to resistance development. In this article, we present a systematic computational investigation of stability changes caused by major and minor mutations of the HIV-1 protease. As most small single-domain proteins, the HIV-1 protease is only marginally stable. Mutations that destabilize the folded, active state of the protease therefore can shift the conformational equilibrium towards the unfolded, inactive state. We find that the most frequent major mutations destabilize the HIV-1 protease, whereas roughly half of the frequent minor mutations are stabilizing. An analysis of protease sequences from patients in treatment indicates that the stabilizing minor mutations are frequently correlated with destabilizing major mutations, and that highly resistant HIV-1 proteases exhibit significant fractions of stabilizing mutations. Our results thus indicate a central role of minor mutations in balancing the marginal stability of the protease against the destabilization induced by the most frequent major mutations.

KEY WORDS

AIDS therapy, computation, correlation analysis, mutation-induced stability changes, protein mutations

1 | INTRODUCTION

The HIV-1 protease plays an essential role in HIV replication by cleaving newly synthesized polyprotein chains at several places into functional protein components of the virus. The cleavage of the polyprotein chain occurs at a tunnel-shaped active site enclosed by the two identical chains of the dimeric HIV-1 protease.^{1,2} In its unbound state,

the HIV-1 protease adopts a semiopen conformation that enables the entry of the polyprotein chain into the active-site tunnel. The binding of the polyprotein substrate induces a change to the closed conformation of the protease in which substrate cleavage occurs.^{3,4} Because of its essential role for HIV replication, the HIV-1 protease is a major target in AIDS therapy. Nine drugs approved for highly active antiretroviral therapies are inhibitors of the HIV-1 protease.^{1,5}

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Proteins: Structure, Function, and Bioinformatics* published by Wiley Periodicals, Inc.

These protease inhibitors bind in the active-site tunnel and, thus, prevent substrate binding.

During therapy, mutations of the HIV-1 protease can lead to resistance to protease inhibitors. The development of resistance typically involves several mutations.^{1,6} Mutations that usually occur first and individually reduce the susceptibility⁷ to one or several protease inhibitors are called "major" mutations.⁸ Mutations that occur later and individually do not substantially affect the susceptibility to inhibitors are called "minor" mutations. At present, 23 major mutations and 50 minor mutations involved in resistance development have been classified.⁸ Major mutations of the HIV-1 protease involved in drug resistance are predominantly located in the active-site tunnel or at the dimer interface and can directly interfere with drug binding by changing the shape of the active-site tunnel.^{1,5} Minor mutations, in contrast, are predominantly located distal to the active-site tunnel.

A central question is how minor mutations contribute to resistance development. Structural investigations of mutant proteases indicate that some minor mutations may indirectly affect the active-site tunnel via a propagation of structural changes from the distal site of mutation to the active site,^{5,9} via coordinated structural rearrangements of multiple mutations,¹⁰ or via coordinated changes in the structural dynamics that may affect the balance between substrate cleavage and drug binding.¹¹ Other minor mutations have been shown to increase the thermal stability of the HIV-1 protease and to compensate stability decreases caused by major mutations.¹² Highly drug-resistant mutants of the HIV-1 protease with a large number of mutations can be more stable than the wildtype (WT).^{13,14} The stability of the WT protease dimer depends on the monomer concentration and ranges from 4 to 10 kcal/mol for micro- to millimolar concentrations,^{15,16} which are typical stability values of marginally stable proteins.¹⁷⁻²¹

Here, we systematically investigate mutation-induced stability changes $\Delta\Delta G$ of the HIV-1 protease via calculations with the program Rosetta.^{22,23} The Rosetta prediction accuracies for $\Delta\Delta G$ exceed the accuracies of other programs,²³⁻²⁵ in particular for mutations of smaller amino acids into larger amino acids, which frequently occur in the HIV-1 protease. Our $\Delta\Delta G$ calculations indicate that the most frequent major mutations destabilize the HIV-1 protease, and that roughly half of the frequent minor mutations are stabilizing. We find that the stabilizing minor mutations are often correlated with destabilizing major mutations, and that highly resistant HIV-1 proteases exhibit significant fractions of stabilizing mutations. These results indicate a central role of minor mutations in maintaining the marginal stability of the HIV-1 protease.

2 | METHODS

We have chosen the high-resolution and high-quality structure 2PCO²⁶ for the stability calculations in the unbound state of the HIV-1 protease. The resolution of this structure is 1.4 Å, which is significantly higher than the resolution of other unbound structures. For the stability calculations in the bound state, we have chosen the three high-resolution and high-quality structures 4EJD,²⁷ 4EJK,²⁷ and 4E43²⁸

with the same sequence as 2PCO. We have performed the stability calculations with the recommended Rosetta protocol 16, which includes limited backbone flexibility.²³ Backbone flexibility appears to be of particular importance for reliable predictions of stability changes for mutations of smaller amino acids into larger amino acids,^{23,29} which frequently occur in the HIV-1 protease. Rosetta results are multiplied with a factor 0.57 kcal/mol to obtain units of kcal/mol.²³ A comparison of calculated and experimentally measured $\Delta\Delta G$ values for a set of 1210 mutations indicates that Rosetta results have an average prediction error between 0.5 and 1 kcal/mol (see Figure 2 in Reference 23). A comparison of experimentally determined melting-temperature changes ΔT_m and $\Delta\Delta G$ values calculated for the high-quality structure 2PCO points towards an average prediction error of our calculations of about 0.5 kcal/mol, because the average deviations in $\Delta\Delta G$ from the regression line of Figure 1 are 0.4 kcal/mol. We have performed seven Rosetta runs for each considered mutant, and have discarded the two smallest and two largest $\Delta\Delta G$ values of these runs to exclude outliers. The reported results are averages of the remaining three $\Delta\Delta G$ values. For all possible single-residue mutations of the unbound HIV-1 protease, the mean of the SDs of these remaining three $\Delta\Delta G$ values is 0.1 kcal/mol and, thus, significantly smaller than the average prediction error of Rosetta. The mean SD for the $\Delta\Delta G$ values of all seven Rosetta runs per mutation is 0.5 kcal/mol due to outliers. For some HIV-1 mutations, in particular for some mutations of smaller amino acids into larger amino acids such as the mutation A71V (see Section 3), the required reconfiguration of the backbone or neighboring sidechain appears to be beyond the scope of the recommended Rosetta protocol, which leads to steric overlaps. To avoid unreliable results from steric overlaps, we exclude $\Delta\Delta G$ results for destabilizing single-residue mutations with average Lennard-Jones repulsive energies between atoms in different residues that exceed 1.5 kcal/mol. The average Lennard-Jones repulsive energies are calculated for the Rosetta runs with the selected three $\Delta\Delta G$ values.

3 | RESULTS AND DISCUSSION

3.1 | Calculated stability changes are highly correlated with experimentally determined changes of the melting temperature

We first validate our $\Delta\Delta G$ calculations by a comparison to experimentally measured changes in the melting temperature. Chang and Torbett³⁰ have measured the changes ΔT_m of the melting temperature induced by selected single and double mutants of the HIV-1 protease in the absence of substrate or drug molecules. The common major mutations V82A, I84V, and L90M, which are associated with resistance against several approved inhibitors,⁷ lead to a decrease of the melting temperature T_m relative to the WT. This decrease in the melting temperature reflects a thermal destabilization and is significantly more pronounced for the double mutant I84V-L90M, compared to the single mutants I84V and L90M. In contrast, double mutants that contain the major mutation I84V and one of the minor mutations L10I, L63P, and A71V have melting temperatures ΔT_m close to the

WT value, which indicates that these minor mutations compensate the thermal destabilization induced by I84V. The double mutant I84V-V77I exhibits a melting temperature close to the single mutant I84V.

In Figure 1, the melting-temperature changes ΔT_m measured by Chang and Torbett are plotted against stability changes $\Delta\Delta G$ calculated with Rosetta²³ for the open conformation of the HIV-1 protease, which is the ground-state conformation in the unbound state. The double-mutant A71V-I84V is excluded from the figure because of steric overlaps in Rosetta, which occur also for the single mutant A71V. These steric overlaps indicate that the reconfiguration of the backbone and/or neighboring sidechains induced by the "small-to-large" mutation A71V is beyond the scope of the recommended Rosetta protocol for calculating stability changes.²³ The calculated $\Delta\Delta G$ values of all other single and double mutants considered by Chang and Torbett are highly correlated with the experimentally measured melting-temperature changes ΔT_m . For the major mutation L90M, the relative large change $\Delta T_m \approx -5^\circ\text{C}$ of the melting temperature is consistent with the relative large stability change of $\Delta\Delta G \approx 3.2 \text{ kcal/mol}$ calculated with Rosetta. For the major mutations V82A and I84V, the smaller decreases in the melting temperature observed by Chang and Torbett are in agreement with smaller $\Delta\Delta G$ values, compared to L90M. For the double mutants L10I-I84V and L63P-I84V, both the measured melting temperature and the calculated stability are closer to the WT values than for the single mutant I84V, which reflects a compensatory, stabilizing effect of the minor mutations L10I and L63P.

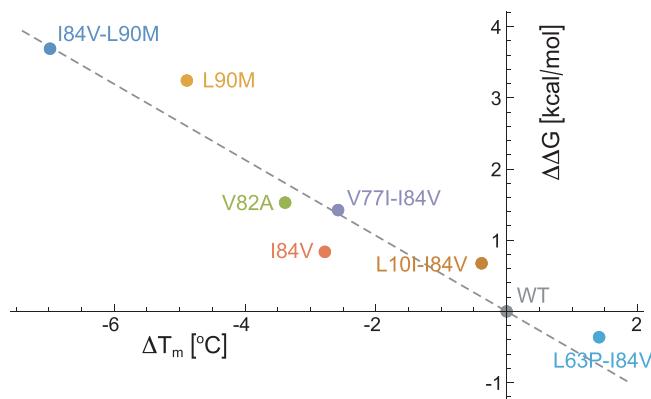


FIGURE 1 Calculated stability changes $\Delta\Delta G$ vs experimentally determined melting-temperature changes ΔT_m induced by single- and double-residue mutations in the unbound state of the HIV-1 protease. Values of ΔT_m are from Chang and Torbett.³⁰ The slope of the dashed regression line is $-0.53 \text{ (kcal/mol)}/^\circ\text{C}$, the Pearson correlation coefficient is -0.95 , and the P -value is $9.4 \cdot 10^{-4}$ without the wildtype data point. The deviations in $\Delta\Delta G$ from the regression line range from 0 to 0.6 kcal/mol, with an average of 0.4 kcal/mol. Rosetta results are multiplied with a factor 0.57 kcal/mol to obtain units of kcal/mol.²³ A comparison of calculated and experimentally measured $\Delta\Delta G$ values for a set of 1210 mutations indicates that Rosetta calculations have an average prediction error between 0.5 and 1 kcal/mol (see Figure 2 in Reference 23). Errors of the experimentally determined melting-temperature changes ΔT_m have not been reported [Color figure can be viewed at wileyonlinelibrary.com]

From the slope $-0.53 \text{ (kcal/mol)}/^\circ\text{C}$ of the regression curve in Figure 1 and from the measured values $\Delta T_m = -0.6^\circ\text{C}$ for A71V-I84V and $\Delta T_m = -2.8^\circ\text{C}$ for I84V, we obtain the estimate $\Delta\Delta G = -1.2 \text{ kcal/mol}$.

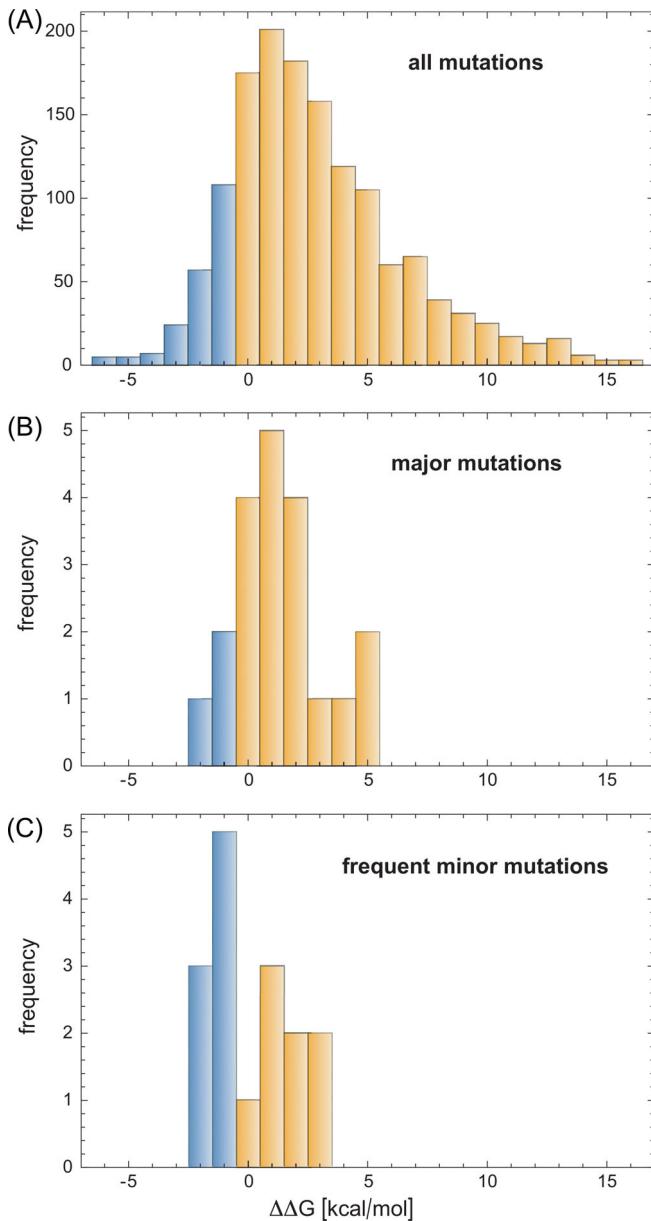


FIGURE 2 Distributions of calculated stability changes in the unbound, open state of the HIV-1 protease (A) for 1427 single-residue mutations without steric clashes in Rosetta out of $99 \cdot 19 = 1881$ possible mutations, (B) for the major mutations of Table 1, and (C) for the frequent minor mutations of Table 2. Blue bars indicate stabilizing mutations with $\Delta\Delta G \leq -0.5 \text{ kcal/mol}$. The mutation-induced stability changes $\Delta\Delta G$ in the unbound state have been calculated for the high-resolution and high-quality X-ray structure with protein data bank code 2PC0.²⁶ The 1427 mutations and stability changes in (A) are listed in the Supporting Information. The stability changes for the major and frequent minor mutations in (B) and (C) are given in Tables 1 and 2. The distribution in (C) includes the $\Delta\Delta G_{\text{open}}$ values estimated from experimental data for the mutations L10I and A71V (see Section 3.1) [Color figure can be viewed at wileyonlinelibrary.com]

mol for the stability change of the minor mutation A71V, assuming additivity of the $\Delta\Delta G$ values for the single mutations of the residues A71 and I84, which are not in direct contact. For the mutation L10I, we obtain the estimate $\Delta\Delta G = -1.3$ kcal/mol from the experimentally determined ΔT_m values -0.4°C and -2.8°C for L10I-I84V and I84V, which indicates that the Rosetta result $\Delta\Delta G = -0.2$ kcal/mol underestimates the magnitude of the stabilizing effect of the mutation L10I.

TABLE 1 Calculated stability changes for major mutations

Mutation	f_{naive}	f_{PI}	$\Delta\Delta G_{\text{open}}$	$\Delta\Delta G_{\text{closed}}$
D30N	0.1	4.8	-2.5	-1.3 (0.3)
V32I	0	4.0	2.3	0.6 (1.2)
M46I	0.3	17.2	1.1	-1.1 (0.4)
M46L	0.3	7.6	-0.5	-0.3 (0.2)
I47A	0	0.6	4.6	4.0 (0.4)
I47V	0	3.4	1.5	1.2 (0.4)
G48V	0	2.5	-	1.7 (3.1)
I50L	0	1.8	-0.2	-0.1 (0.5)
I50V	0	1.5	-0.6	1.3 (0.3)
I54L	0	2.1	1.3	3.5 (0.7)
I54M	0	1.6	-0.4	3.8 (0.9)
Q58E	0.4	5.3	3.8	3.7 (0.3)
T74P	0	1.4	-1.1	-0.1 (0.3)
L76V	0	3.4	-	-
V82A	0.2	18.7	1.5	1.2 (0.4)
V82F	0	1.5	0.6	-0.4 (1.8)
V82L	0	0.2	-0.3	0.3 (0.4)
V82S	0	0.9	2.1	1.8 (0.2)
V82T	0	1.8	0.2	0.7 (0.5)
N83D	0	0.8	5.0	4.0 (0.4)
I84V	0.1	9.2	0.8	0.9 (0.1)
N88S	0	1.2	1.0	0.7 (0.1)
L90M	0.4	22.8	3.2	3.2 (0.3)

Note: The mutations in the table are classified as major by Wensing et al.⁸ for at least one protease inhibitor. The mutation frequencies f_{naive} before treatment and f_{PI} during treatment with protease inhibitors are the percentages of 105 599 sequences of untreated patients and 26 838 sequences of treated patients, respectively, that contain the mutation (one sequence per patient, last in treatment). These sequences are provided by the Stanford University HIV Drug Resistance Database.^{38,39} The calculated stability changes $\Delta\Delta G_{\text{open}}$ and $\Delta\Delta G_{\text{closed}}$ in the open and closed conformation of the HIV-1 protease are given in units of kcal/mol. The stability change $\Delta\Delta G_{\text{open}}$ is calculated for the high-resolution and high-quality X-ray structure with protein data bank code 2PCO.²⁶ The stability change $\Delta\Delta G_{\text{closed}}$ is the mean of three values calculated for the three high-resolution and high-quality structures 4EJD,²⁷ 4EJK,²⁷ and 4E43²⁸ with the same sequence as 2PCO. The numbers in brackets are the SDs of these three values. To avoid unreliable results from steric overlaps, we exclude Rosetta results for destabilizing mutations with overall Lennard-Jones repulsive energies between atoms in different residues that exceed 1.5 kcal/mol.

3.2 | Roughly half of the frequent minor mutations are stabilizing

Tables 1 and 2 summarize the calculated stability changes for all major and frequent minor mutations of the HIV-1 protease.⁸ We have performed Rosetta calculations for both the open conformation and the closed conformation of the HIV-1 protease. The stability changes $\Delta\Delta G_{\text{closed}}$ of the closed conformation are calculated for the “empty” state of the protease without substrate or drug molecules. The overall mutation-induced free energy change in the bound state is the sum of the stability change $\Delta\Delta G_{\text{closed}}$ of the protein and the binding-free energy change $\Delta G_{\text{binding}}$ of a drug or substrate molecule. In addition to the stability changes, Tables 1 and 2 contain the frequencies f_{naive} and f_{PI} of major and minor mutations before and during treatment with protease inhibitors.

We focus first on the calculated stability changes $\Delta\Delta G_{\text{open}}$ for the open conformation because the protease is in general more vulnerable to stability loss in this conformation. The closed protease conformation of the bound state is additionally stabilized by the binding free energy $\Delta G_{\text{binding}}$ of the substrate or drug molecule. According to the calculated $\Delta\Delta G_{\text{open}}$ values of Table 1, the four most frequent major

TABLE 2 Calculated stability changes for frequent minor mutations

Mutation	f_{naive}	f_{PI}	$\Delta\Delta G_{\text{open}}$	$\Delta\Delta G_{\text{closed}}$
L10F	0.1	6.9	-2.5	-0.7 (0.1)
L10I	6.8	22.8	-1.3 ^a	-0.9 (0.5)
L10V	4.2	9.0	0.2	-0.3 (0.4)
G16E	9.6	7.1	-	-
K20I	9.7	7.2	-1.0	-2.5 (0.5)
K20R	9.7	16.6	-1.9	0.4 (0.3)
L33F	0.5	9.8	-0.7	-0.9 (0.3)
M36I	56.1	51.7	-2.2	0.1 (0.7)
I54V	0.1	19.9	0.5	1.8 (0.4)
D60E	11.1	11.3	1.4	1.4 (0.4)
I62V	14.1	26.8	1.5	0.9 (0.2)
L63P	35.7	54.1	-1.1	-1.1 (0.9)
I64V	8.0	14.0	2.9	1.9 (0.7)
A71T	4.1	6.7	-	-
A71V	3.2	20.2	-1.2 ^a	-
G73S	0	5.4	-	-
V77I	17.9	20.2	0.8	-0.3 (0.2)
L89M	45.8	19.7	2.0	3.0 (0.8)
I93L	45.8	39.9	2.5	2.6 (0.4)

Note: The mutations in the table are classified as minor by Wensing et al.⁸ The mutation frequencies f_{naive} before treatment and f_{PI} during treatment with protease inhibitors and the stability changes $\Delta\Delta G_{\text{open}}$ and $\Delta\Delta G_{\text{closed}}$ in the open and closed conformation of the HIV-1 protease have been calculated as in Table 1. The stability changes are given in units of kcal/mol.

^aEstimated from experimental data (see Section 3.1).

mutations M46I, V82A, I84V, and L90M are destabilizing. Among the remaining, less frequent major mutations, the four mutations D30N, M46 L, I50V, and T74P are stabilizing with stability changes $\Delta\Delta G_{\text{open}} \leq -0.5$ kcal/mol. The remaining major mutations are either destabilizing with values $\Delta\Delta G_{\text{open}} \geq 0.5$ kcal/mol, or rather neutral with stability changes in the range $-0.5 < \Delta\Delta G_{\text{open}} < 0.5$ kcal/mol. The calculated stability changes thus indicate that major mutations are predominantly destabilizing, in particular the four most common major mutations M46I, V82A, I84V, and L90M. We use threshold values of -0.5 and 0.5 kcal/mol for stabilizing and destabilizing mutations, respectively, because we estimate the average prediction error of our $\Delta\Delta G$ calculations to be about ± 0.5 kcal/mol (see Section 2).

Among the minor mutations that occur with frequencies $f_{\text{PI}} \geq 5\%$, in contrast, the eight mutations L10F, L10I, K20I, K20R, L33F, M36I, L63P, and A71V are stabilizing with values $\Delta\Delta G_{\text{open}} \leq -0.5$ kcal/mol, the seven mutations I54V, D60E, I62V, I64V, V77I, L89M, and I93L are destabilizing with values $\Delta\Delta G_{\text{open}} \geq 0.5$ kcal/mol, and the mutation L10 V is rather neutral with $\Delta\Delta G_{\text{open}} = 0.2$ kcal/mol. The stability changes calculated with Rosetta thus indicate that about half of the more frequent minor mutations are stabilizing rather than destabilizing. Besides the three stabilizing mutations L10I, L63P, A71V identified by Chang and Torbett from experimentally measured changes of the melting temperature (see Figure 1), our calculated $\Delta\Delta G_{\text{open}}$ values point towards the five rather frequent additional stabilizing minor mutations L10F, K20I, K20R, L33F, and M36I.

Figure 2 compares the distribution of stability change $\Delta\Delta G_{\text{open}}$ for the major mutations of Table 1 and the frequent minor mutations of Table 2 to the distribution of stability changes for 1427 single-residue mutations without steric clashes in Rosetta out of all $99 \cdot 19 = 1881$ possible single-residue mutations. The distribution of stability changes for all possible mutations without steric clashes has a characteristic single-peaked and slightly asymmetric shape that has also been found for other proteins.³¹ Out of the 1427 single-residue mutations of Figure 2A, 206 mutations are stabilizing with $\Delta\Delta G_{\text{open}} \leq -0.5$ kcal/mol, which amounts to a fraction of 14%. The distribution of stability changes for major mutations in Figure 2B has a shape that is rather similar to the shape of the distribution in Figure 2A, with the caveat that the distribution of Figure 2B only includes the relatively small number of 21 mutations with reliable $\Delta\Delta G_{\text{open}}$ values from Table 1. Out of these 21 mutations, four are stabilizing with $\Delta\Delta G_{\text{open}} \leq -0.5$ kcal/mol, which corresponds to a fraction of 19% that is comparable to the fraction of stabilizing mutations in Figure 2A. The distribution of stability changes for frequent minor mutations in Figure 2C includes eight stabilizing mutations with $\Delta\Delta G_{\text{open}} \leq -0.5$ kcal/mol out of 16 mutations with reliable $\Delta\Delta G_{\text{open}}$ values in Table 2, which is a fraction of 50%. This fraction of stabilizing frequent minor mutations is significantly larger than the fraction of stabilizing mutations in Figure 2A,B. A lower threshold of -1 kcal/mol for stabilizing mutations leads to the same conclusion, because seven out of 16 frequent minor mutations with reliable $\Delta\Delta G_{\text{open}}$ values, that is, a fraction of 44%, have values of $\Delta\Delta G_{\text{open}} \leq -1$ kcal/mol. In contrast, the fractions of mutations with $\Delta\Delta G_{\text{open}} \leq -1$ kcal/mol among all possible mutations and among major mutations are both 10%.

3.3 | Stabilizing minor mutations are frequently correlated with destabilizing major mutations

Pair correlations of mutations associated with resistance to protease inhibitors can be determined from the protease sequences of patients in treatment.^{32–34} We calculate the correlation coefficient ϕ of two mutations from the currently available 26 838 sequences of treated patients (one sequence per patient). The ϕ coefficient is the appropriate correlation coefficient for binary variables (absence or presence of a mutation), ranges from -1 to 1 , and is equivalent in interpretation to the Pearson correlation coefficient. For two mutations X and Y, the ϕ coefficient is $\phi = (N_{XY}N_{00} - N_{X0}N_{0Y})/([N_{X0} + N_{XY}] + [N_{00} + N_{0Y}] + [N_{00} + N_{X0}] + [N_{0Y} + N_{XY}])$, where N_{XY} is the number of sequences with both mutation X and Y, N_{00} is the number of sequences with neither X nor Y, and N_{X0} and N_{0Y} are the numbers of sequences with only mutation X and Y, respectively. Table 3 lists all correlated pairs of major and frequent minor mutations X and Y with $\phi \geq 0.2$. Besides the ϕ coefficients, Table 3 includes the stability changes $\Delta\Delta G_X$ and $\Delta\Delta G_Y$ for the single mutations X and Y in the open conformation of the HIV-1 protease. Missing values indicate Lennard-Jones repulsive energies larger than 1.5 kcal/mol, which point towards steric overlaps in the Rosetta calculations that lead to unreliable results.

Among the 26 pairs of correlated major and frequent minor mutations in the table, 23 pairs exhibit reliable $\Delta\Delta G$ values for both the major mutation X and the minor mutation Y. In a majority of 16 of these 23 pairs, a destabilizing major mutation X with $\Delta\Delta G_X \geq 0.5$ kcal/mol is correlated with a stabilizing minor mutation Y with $\Delta\Delta G_Y \leq -0.5$ kcal/mol. For 15 of the 16 pairs with a destabilizing major and a stabilizing minor mutation, the major mutation is one of the four most common major mutations M46I, V82A, I84V, and L90M. For the mutation M46I, three of the four minor mutations correlated with $\phi \geq 0.2$ are stabilizing. For V82A, four out of the five correlated minor mutations are stabilizing. For I84V, all four correlated frequent minor mutations Y are stabilizing, and for L90M, four of the six correlated minor mutations Y with reliable $\Delta\Delta G_Y$ values are stabilizing. In the 16 pairs with a destabilizing major mutation and stabilizing minor mutation, the minor mutations L10I, L33F, and A71V occur four times, L10F occurs twice, and K20R and L63P occur once. For the mutations of Table 3, the sum $\Delta\Delta G_X + \Delta\Delta G_Y$ of the stability changes for the single-residue mutants does not deviate from the stability change $\Delta\Delta G_{XY}$ of the double mutants by more than ± 0.5 kcal/mol, which is within the numerical accuracy of our Rosetta calculations. The stability changes $\Delta\Delta G_X$ and $\Delta\Delta G_Y$ for the single mutants X and Y thus appear to be additive within numerical accuracy.

The correlations of major and frequent minor mutations of Table 3 are highly significant with P-values smaller than 10^{-100} , but typically do not exceed values of $\phi \simeq 0.4$, except for the correlated pair V82A and I54V with $\phi = 0.54$. This range of correlation coefficients clearly below the maximum value of 1 is plausible for correlations that result from stabilization because a destabilizing major mutation can be stabilized by different minor mutations and, thus, does not require a particular minor mutation. The relatively large correlation coefficient $\phi = 0.54$ for the pair V82A and I54V may result from a correlated structural change.

Major mutation X	Minor mutation Y	ϕ	$\Delta\Delta G_X$	$\Delta\Delta G_Y$	$\Delta\Delta G_{XY}$
M46I	L10F	0.29	1.1	-2.5	-1.0
M46I	L10I	0.20	1.1	-1.3 ^a	
M46I	I54V	0.31	1.1	0.5	2.1
M46I	A71V	0.26	1.1	-1.2 ^a	-
M46L	L33F	0.20	-0.5	-0.7	-1.1
M46L	I54V	0.28	-0.5	0.5	0.0
M46L	A71V	0.20	-0.5	-1.2 ^a	-
G48V	L10I	0.20	-	-1.3 ^a	
I54L	L33F	0.23	1.3	-0.7	0.5
L76V	I54V	0.21	-	0.5	-
V82A	L10I	0.32	1.5	-1.3 ^a	
V82A	K20R	0.27	1.5	-1.9	0.2
V82A	L33F	0.32	1.5	-0.7	0.8
V82A	I54V	0.54	1.5	0.5	2.1
V82A	A71V	0.34	1.5	-1.2 ^a	-
I84V	L10F	0.27	0.8	-2.5	-2.2
I84V	L10I	0.20	0.8	-1.3 ^a	
I84V	L33F	0.21	0.8	-0.7	0.3
I84V	A71V	0.26	0.8	-1.2 ^a	-
L90M	L10I	0.28	3.2	-1.3 ^a	
L90M	L33F	0.23	3.2	-0.7	2.3
L90M	I54V	0.33	3.2	0.5	3.7
L90M	I62V	0.28	3.2	1.5	4.9
L90M	L63P	0.31	3.2	-1.1	2.3
L90M	A71V	0.41	3.2	-1.2 ^a	-
L90M	G73S	0.37	3.2	-	-

TABLE 3 Correlated major and frequent minor mutations

Note: The calculated stability changes $\Delta\Delta G_X$ and $\Delta\Delta G_Y$ for the single-residue mutations X and Y in the open, unbound conformation of the HIV-1 protease are from Tables 1 and 2, respectively. The stability changes $\Delta\Delta G_{XY}$ for the double mutants are calculated with Rosetta for the open, unbound conformation in the same way as $\Delta\Delta G_X$ and $\Delta\Delta G_Y$. The sum $\Delta\Delta G_X + \Delta\Delta G_Y$ of the stability changes for the single residue mutants agrees with $\Delta\Delta G_{XY}$ within the numerical accuracy of the calculations, that is, within ± 0.5 kcal/mol, which indicates that the stability changes are additive for the mutations in this table. As in Tables 1 and 2, hyphens indicate that $\Delta\Delta G$ values have been excluded due to steric overlaps in Rosetta calculations. The $\Delta\Delta G$ values for double mutants that include the mutation L10I are not listed because the value $\Delta\Delta G_Y = -1.3$ kcal/mol for the single-residue mutation L10I has been estimated based on experimental data (see Section 3.1 and Table 2). The calculation of the correlation coefficient ϕ is described in the text.

^aEstimated from experimental data (see Section 3.1 and Table 2).

Previous correlation analyses have been based on smaller number of sequences.^{32–34} Based on 3475 sequences of treated patients, Rhee et al³³ have identified 92 pairs of positively correlated protease mutations with statistical significance. This set of 92 pairs includes 17 of the 26 correlated pairs of major and frequent minor mutations with correlation coefficient $\phi \geq 0.2$ listed in Table 3, which we have identified based on 26 838 sequences of treated patients. Varghese et al³⁴ have analyzed 11 351 sequences from treated and untreated patients that include one or more resistance-associated mutation. Based on their analysis, Varghese et al have identified 75 pairs of correlated mutations with correlation coefficient larger than 0.1, which include nine of the 26 correlated pairs of major and frequent

minor mutations of Table 3. In contrast to previous analyses, we have focussed here on pairs of major and frequent minor mutations, and have combined the correlation analysis with an analysis of mutation-induced stability changes.

3.4 | Highly resistant HIV-1 proteases exhibit significant fractions of individually stabilizing mutations

In standard essays,^{35,36} the resistance of protease mutants to inhibitors is inferred from the half-maximal inhibitory concentration IC_{50} at which virus replication is inhibited by 50%. The fold resistance is the ratio of

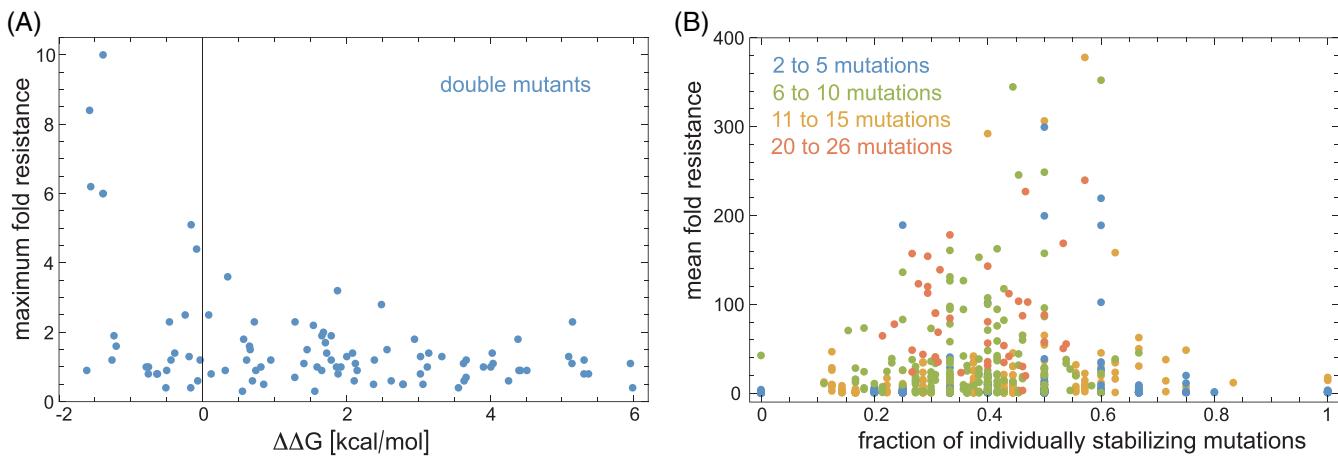


FIGURE 3 (A) Maximum fold resistance of eight clinically approved protease inhibitors vs calculated stability changes $\Delta\Delta G$ in the open, unbound protease conformation for double mutants in the high-quality filtered dataset of the Stanford HIV Drug Resistance Database with overall 1808 protease isolates from patients. Mutants with unreliable $\Delta\Delta G$ values are excluded (see Section 2). For double mutants containing the mutation A71V, the stability change $\Delta\Delta G$ is calculated by superposition of stability changes for single mutants. (B) Average fold resistance vs fraction of individually stabilizing mutations for all protease isolates in the high-quality filtered dataset with more than one mutation and reliable $\Delta\Delta G$ values for all individual mutations of the isolates

the IC_{50} values for the mutant and WT. A fold resistance of 1 indicates unchanged resistance, while fold resistances larger than 1 indicate increased resistance. In a high-quality filtered dataset, the Stanford HIV Drug Resistance Database provides fold resistances of eight clinically approved inhibitors for 1808 protease isolates from patients.

In Figure 3A, the maximum fold resistance of the double mutants among the 1808 isolates is plotted against the calculated stability change $\Delta\Delta G$ for the open, unbound protease conformation. The large majority of the mutants have fold resistances around the WT value 1. The stability changes $\Delta\Delta G$ for these mutants range from -2 to 6 kcal/mol. Double mutants with significantly larger fold resistances up to 10 tend to be stabilizing with $\Delta\Delta G$ values between -2 and 0.

For protease isolates with more than two mutations, the calculation of stability changes is less reliable due to addition of errors and possible concerted structural rearrangements. For those mutants, we simply determine the fraction $n_{st}/(n_{st} + n_{de})$ of individually stabilizing mutations where n_{st} is the number of mutations with calculated stability changes $\Delta\Delta G \leq -0.5$ kcal/mol for the single-residue mutant, and n_{de} is the number of individually destabilizing mutations with $\Delta\Delta G \geq 0.5$ kcal/mol. The fraction of individually stabilizing mutations ranges from 0 for mutants without any such mutations to 1 for mutants that exhibit only individually stabilizing mutations or individually neutral mutations with stability changes $-0.5 < \Delta\Delta G < 0.5$ kcal/mol in single-residue mutants. For highly resistant mutants with mean fold resistances larger than 100 for the eight inhibitors, the fraction of individually stabilizing mutations ranges from about 0.25 to 0.65, with an average of 0.43 (see Figure 2B). Figure 3 thus indicates that maintaining or increasing the stability of the HIV-1 protease plays an important role for drug resistance. Mutants with high mean fold resistances contain

significant fractions of individually stabilizing mutations that balance the destabilizing effects of major, resistance-inducing mutations.

3.5 | Mutations at position I54 destabilize the closed conformation relative to the open conformation

The calculated stability changes $\Delta\Delta G_{open}$ and $\Delta\Delta G_{closed}$ for the open and closed state of HIV-1 can be different because of the conformational variations between the states. A striking example is the major and minor mutations of residue I54 located in the “flaps” of the HIV-1 protease, which close during binding. The calculated stability changes of the two major mutations I54L and I54M and the frequent minor mutation I54V indicate that all three mutations clearly destabilize the closed conformation relative to the open conformation (see Tables 1 and 2). A theoretical analysis indicates that the catalytic rate in the presence of inhibitors can be increased by a factor $\exp((\Delta\Delta G_{closed} - \Delta\Delta G_{open})/RT)$ if product unbinding is rate-limiting, which is plausible for the HIV-1 protease.³⁷ The mutations of the distal residue I54 thus may contribute to resistance by destabilizing the closed conformation relative to the open conformation. This contribution to resistance due to mutation-induced shifts of the conformational equilibrium is drug-independent, in agreement with an analysis of fold resistances, which indicates that all mutations of I54 are associated with decreased susceptibility to six to seven of the eight considered protease inhibitors.⁷

4 | CONCLUSIONS

In this article, we have investigated the stability changes $\Delta\Delta G$ of the HIV-1 protease induced by major and minor mutations with the program Rosetta.^{22,23} Our results indicate that roughly half of

the frequent minor mutations of the HIV-1 protease are stabilizing (see Table 1). These stabilizing minor mutations are often correlated with the most frequent, destabilizing major mutations (see Table 3) and, thus, appear to play a central role in maintaining the marginal stability of the HIV-1 protease. Highly resistant HIV-1 proteases exhibit significant fractions of stabilizing mutations (see Figure 3B).

ACKNOWLEDGMENTS

Support for B.H. from the Alexander von Humboldt Foundation via a Georg Forster Research Fellowship and from Shiraz University is gratefully acknowledged.

ORCID

Thomas R. Weikl  <https://orcid.org/0000-0002-0911-5328>

REFERENCES

1. Wensing AMJ, van Maarseveen NM, Nijhuis M. Fifteen years of HIV protease inhibitors: raising the barrier to resistance. *Antiviral Res.* 2010;85(1):59-74.
2. Prabu-Jeyabalan M, Nalivaika E, Schiffer CA. Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. *Structure.* 2002; 10(3):369-381.
3. Furfine ES, D'Souza E, Ingold KJ, Leban JJ, Spector T, Porter DJ. Two-step binding mechanism for HIV protease inhibitors. *Biochemistry.* 1992;31:7886-7891.
4. Gustchina A, Weber IT. Comparison of inhibitor binding in HIV-1 protease and in non-viral aspartic proteases: the role of the flap. *FEBS Lett.* 1990;269:269-272.
5. Weber IT, Agniswamy J. HIV-1 protease: structural perspectives on drug resistance. *Viruses.* 2009;1(3):1110-1136.
6. Weber IT, Kneller DW, Wong-Sam A. Highly resistant HIV-1 proteases and strategies for their inhibition. *Future Med Chem.* 2015;7(8):1023-1038.
7. Rhee SY, Taylor J, Fessel WJ, et al. HIV-1 protease mutations and protease inhibitor cross-resistance. *Antimicrob Agents Chemother.* 2010;54(10):4253-4261.
8. Wensing AM, Calvez V, Günthard HF, et al. 2015 update of the drug resistance mutations in HIV-1. *Top Antivir Med.* 2015;23(4):132-141.
9. Ragland DA, Nalivaika EA, Nalam MNL, et al. 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(34):11956-11963.
10. 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(13):2819-2828.
11. Cai Y, Yilmaz NK, Myint W, Ishima R, Schiffer CA. Differential flap dynamics in wild-type and a drug resistant variant of HIV-1 protease revealed by molecular dynamics and NMR relaxation. *J Chem Theory Comput.* 2012;8(10):3452-3462.
12. Louis JM, Zhang Y, Sayer JM, Wang YF, Harrison RW, Weber IT. The L76V drug resistance mutation decreases the dimer stability and rate of autoprocessing of HIV-1 protease by reducing internal hydrophobic contacts. *Biochemistry.* 2011;50(21):4786-4795.
13. Louis JM, Tözsér J, Roche J, Matúz K, Aniana A, Sayer JM. Enhanced stability of monomer fold correlates with extreme drug resistance of HIV-1 protease. *Biochemistry.* 2013;52(43):7678-7688.
14. 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.
15. Xie D, Gulnik S, Erickson JW. Dissection of binding energy with native and ligand-bound protein stabilities: determining the affinity of ultratight-binding inhibitors of HIV-1 protease and its drug-resistance mutants. *J Am Chem Soc.* 2000;122(46):11533-11534.
16. Grant SK, Deckman IC, Culp JS, et al. Use of protein unfolding studies to determine the conformational and dimeric stabilities of HIV-1 and SIV proteases. *Biochemistry.* 1992;31(39):9491-9501.
17. Tokuriki N, Tawfik DS. Stability effects of mutations and protein evolvability. *Curr Opin Struct Biol.* 2009;19(5):596-604.
18. Taverna DM, Goldstein RA. Why are proteins marginally stable? *Proteins.* 2002;46(1):105-109.
19. DePristo MA, Weinreich DM, Hartl DL. Missense meanderings in sequence space: a biophysical view of protein evolution. *Nat Rev Genet.* 2005;6(9):678-687.
20. Bloom JD, Labthavikul ST, Otey CR, Arnold FH. Protein stability promotes evolvability. *Proc Natl Acad Sci U S A.* 2006;103(15): 5869-5874.
21. Zeldovich KB, Chen P, Shakhnovich EI. Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proc Natl Acad Sci U S A.* 2007;104(41):16152-16157.
22. Leaver-Fay A, Tyka M, Lewis SM, et al. ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol.* 2011;487:545-574.
23. Kellogg EH, Leaver-Fay A, Baker D. Role of conformational sampling in computing mutation-induced changes in protein structure and stability. *Proteins.* 2011;79:830-838.
24. Potapov V, Cohen M, Schreiber G. Assessing computational methods for predicting protein stability upon mutation: good on average but not in the details. *Protein Eng Des Sel.* 2009;22:553-560.
25. Guerois R, Nielsen JE, Serrano L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J Mol Biol.* 2002;320:369-387.
26. Heaslet H, Rosenfeld R, Giffin M, et al. Conformational flexibility in the flap domains of ligand-free HIV protease. *Acta Crystallogr D Biol Crystallogr.* 2007;63(Pt 8):866-875.
27. Tiefenbrunn T, Forli S, Baksh MM, et al. Small molecule regulation of protein conformation by binding in the flap of HIV protease. *ACS Chem Biol.* 2013;8(6):1223-1231.
28. Perryman AL, Zhang Q, Souter HH, et al. Fragment-based screen against HIV protease. *Chem Biol Drug Des.* 2010;75(3):257-268.
29. Benedict A, Becker CM, de Groot BL, Caflisch A, Böckmann RA. Predicting free energy changes using structural ensembles. *Nat Methods.* 2009;6:3-4.
30. Chang MW, Torbett BE. Accessory mutations maintain stability in drug-resistant HIV-1 protease. *J Mol Biol.* 2011;410(4):756-760.
31. Tokuriki N, Stricher F, Schymkowitz J, Serrano L, Tawfik DS. The stability effects of protein mutations appear to be universally distributed. *J Mol Biol.* 2007;369:1318-1332.
32. Svicher V, Ceccherini-Silberstein F, Erba F, et al. Novel human immunodeficiency virus type 1 protease mutations potentially involved in resistance to protease inhibitors. *Antimicrob Agents Chemother.* 2005; 49(5):2015-2025.
33. Rhee SY, Liu TF, Holmes SP, Shafer RW. HIV-1 subtype B protease and reverse transcriptase amino acid covariation. *PLoS Comput Biol.* 2007;3(5):836-843.
34. Varghese V, Mitsuya Y, Fessel WJ, et al. Prototypical recombinant multi-protease-inhibitor-resistant infectious molecular clones of human immunodeficiency virus type 1. *Antimicrob Agents Chemother.* 2013; 57(9):4290-4299.
35. Petropoulos CJ, Parkin NT, Limoli KL, et al. A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1. *Antimicrob Agents Chemother.* 2000;44(4):920-928.

36. Zhang J, Rhee SY, Taylor J, Shafer RW. Comparison of the precision and sensitivity of the Antivirogram and PhenoSense HIV drug susceptibility assays. *J Acquir Immune Defic Syndr.* 2005;38(4):439-444.
37. Weikl TR, Hemmateenejad B. How conformational changes can affect catalysis, inhibition and drug resistance of enzymes with induced-fit binding mechanism such as the HIV-1 protease. *Biochim Biophys Acta.* 2013;1834(5):867-873.
38. 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(1):298-303.
39. Rhee SY, Kantor R, Katzenstein DA, et al. HIV-1 pol mutation frequency by subtype and treatment experience: extension of the HIVseq program to seven non-B subtypes. *Aids.* 2006;20(5):643-651.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Weikl TR, Hemmateenejad B. Accessory mutations balance the marginal stability of the HIV-1 protease in drug resistance. *Proteins.* 2019;1-9. <https://doi.org/10.1002/prot.25826>