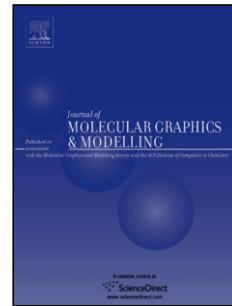


Accepted Manuscript

Title: Recovery of the wild type atomic flexibility in the HIV-1 protease double mutants

Author: Valderes De Conto Antônio S.K. Braz David Perahia Luis P.B. Scott



PII: S1093-3263(15)00072-8

DOI: <http://dx.doi.org/doi:10.1016/j.jmgm.2015.04.006>

Reference: JMG 6534

To appear in: *Journal of Molecular Graphics and Modelling*

Received date: 30-9-2014

Revised date: 9-4-2015

Accepted date: 17-4-2015

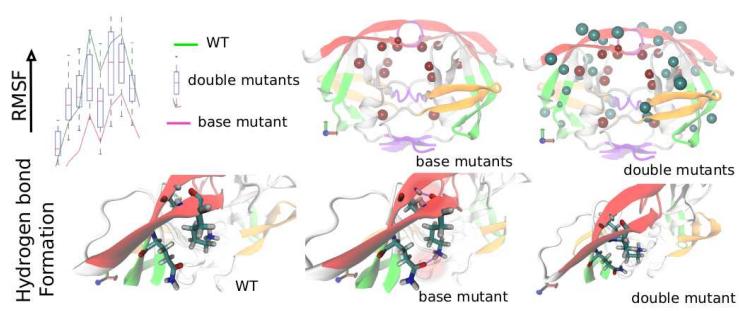
Please cite this article as: V. De Conto, A.S.K. Braz, D. Perahia, L.P.B. Scott, Recovery of the wild type atomic flexibility in the HIV-1 protease double mutants, *Journal of Molecular Graphics and Modelling* (2015), <http://dx.doi.org/10.1016/j.jmgm.2015.04.006>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

For Table of Contents Use Only

Title: Recovery of the wild type atomic flexibility in the HIV-1 protease double mutants

Valderes De Conto, Antônio S. K. Braz, David Perahia, Luis P. B. Scott



Recovery of the wild type atomic flexibility in the HIV-1 protease double mutants
Authors: Valderes De Conto, Antônio S. K. Braz, David Perahia, Luis P. B. Scott
Correspondence to: Luis P. B. Scott - e-mail: luisp37@gmail.com

HIGHLIGHTS

We compare the CA atoms flexibility between the WT, single and double mutants by NMA.

Flexibility profile of double mutants becomes more similar to that of WT.

Single mutants showed a significant alteration in h-bond formation compared to WT.

H-bond pattern of double mutants becomes more similar to that of WT.

1

2
3
Title: Recovery of the wild type atomic flexibility in the HIV-1 protease double mutants4
5
6
7 Valderes De Conto^a, Antônio S. K. Braz^a, David Perahia^b, Luis P. B. Scott^{a,b,*}
89
10 ^a Laboratório de Biologia Computacional e Bioinformática, Universidade Federal do ABC, Santo André,
11 Brasil12
13 ^b Laboratoire de Biologie et Pharmacologie Appliquée (LBPA), Ecole Normale Supérieure de Cachan,
14 Cachan, France

15 * Corresponding Author: Luis P. B. Scott

16 Address: Rua Catequese, 242, Bairro jardim, Santo André, São Paulo, CEP 09030-335, Brazil.

17 Phone: +551149960129

18
19 E-mail address: luisp37@gmail.com

2

20 **Abstract**

21

22 The emergence of drug resistant mutations due to the selective pressure exerted by antiretrovirals,
23 including protease inhibitors (PIs), remains a major problem in the treatment of AIDS. During PIs
24 therapy, the occurrence of primary mutations in the wild type HIV-1 protease reduces both the
25 affinity for the inhibitors and the viral replicative capacity compared to the wild type (WT) protein,
26 but additional mutations compensate for this reduced viral fitness. To investigate this phenomenon
27 from the structural point of view, we combined Molecular Dynamics and Normal Mode Analysis to
28 analyze and compare the variations of the flexibility of C-alpha atoms and the differences in
29 hydrogen bond (h-bond) network between the WT and double mutants. In most cases, the flexibility
30 profile of the double mutants was more often similar to that of the WT than to that of the related
31 single base mutants. All single mutants showed a significant alteration in h-bond formation
32 compared to WT. Most of the significant changes occur in the border between the flap and
33 cantilever regions. We found that all the considered double mutants have their h-bond pattern
34 significantly altered in comparison to the respective single base mutants affecting their flexibility
35 profile that becomes more similar to that of WT. This WT flexibility restoration in the double
36 mutants appears as an important factor for the HIV-1 fitness recovery observed in patients.

37

38 **Keywords:** Molecular modeling, Normal modes Analysis, HIV-1 protease, Double mutants, Protein
39 flexibility.

40

41

42

44 **Abbreviations:**

45 PIs, protease inhibitors; WT, wild type; h-bond, hydrogen bond; HIV-1, Human Immunodeficiency
46 virus type-1; RMSF, root mean square fluctuations

47 **1. Introduction**

48 According to the recent reports published by UNAIDS, there are about 35.3 million (32.2 -
49 38.8 million) people living with HIV- AIDS around the world [1]. Since the occurrence of first
50 cases of AIDS up until now, remarkable advances in treatment have been achieved. Although an
51 effective vaccine has not yet been developed, the highly active antiretroviral therapy (HAART) has
52 dramatically improved the prognosis of the patients. HIV protease is a valuable drug target, as
53 inhibition of PR activity results in immature noninfectious virions. Because of its essential role for
54 the replication of HIV [2], protease is one of the primary targets used to treat AIDS currently.

55 Human Immunodeficiency Virus type 1 (HIV-1) protease is a retropepsin and a member of
56 Retroviral aspartyl protease family (Pfam: PF00077). This enzyme is a symmetrical homodimer,
57 each subunit containing 99 amino acids residues (conventionally labeled as residues 1–99 for chain
58 A and 1'–99' for chain B). As shown in Fig. 1 HIV-1 protease contains a short α -helix and nine β
59 strands in each chain [3,4]. It cleaves the Gag and Gag-Pol polyproteins during viral maturation by
60 catalyzing the hydrolysis of specific peptide bonds at 9 different sites to produce each of the
61 following matrix, capsid, nucleocapsid, reverse transcriptase, protease and integrase proteins [5,6].
62 The active-site contains the conserved catalytic triad, Asp-Thr-Gly (amino acids 25-27), common to
63 aspartyl protease family [7]. The critical structural regions are the active site cavity, the dimer
64 interface and the flexible flaps. Substrates and PIs (protease inhibitors) bind in the active site cavity
65 between the catalytic residues and the flexible flaps. The protease active site cavity comprises
66 residues from both chains: 8, 23–30, 32, 45–50, 53, 56, 76, 80–82 and 84 [5,8]. The two sub-units
67 interact mostly by their amino- and carboxyl-terminal β -strands (residues 1–4 and 96–99,
68 respectively), their active site triplets tips, through the salt bridge formation between residues 29, 87
69 and 8', and by the tips of the flaps [5,9]. The flap region is composed by the flap elbows (residues
70 37-42), the flap antiparallel β sheet (residues 43-48 and 53-58) and flap tips (residues 49-52) [10].

72 PIs in combination with inhibitors targeting other viral components have led to a significant
73 decrease of HIV/AIDS related mortality and morbidity [11]. However, the long-term efficiency has
74 been threatened by the emergence of multiple drug-resistant and cross-resistant mutations within the
75 protease due to the selective pressure exerted by PIs [12,13,14]. These mutations can alter the
76 protease catalytic activity, inhibitor binding or the enzyme stability leading to drug resistance to the
77 inhibitors and respective shifts in viral fitness compared to the WT virus [15,16]. The high genetic
78 diversity observed for the HIV-1 protease in the infected host is a consequence of the inherent high
79 error rate of reverse transcriptase and the elevated replication levels of the virus *in vivo* [17]. This
80 great variability is at the origin of drug resistance evolution. In the HIV-1 protease, approximately
81 only 38 residues of the 99 present in each monomer are invariant (mutation frequencies lower than
82 0.5%) and nearly 36 residues are associated with resistance toward one or more inhibitors [18].

83 During PI therapy, primary substitutions (also referred to as major mutations) are selected in
84 an early stage and tend to alter the affinity towards inhibitors by introducing steric hindrance or by
85 removing essential molecular interactions [19], yet maintaining a sufficient but lower affinity
86 towards the natural substrate [20,21,22]. This can explain why the primary PI resistance mutations
87 rarely occur as natural polymorphisms in untreated people. Primary mutations include residues
88 located mostly in the binding cavity (30, 32, 46, 47, 48, 50, 54, 76, 82, 84, 88 and 90) [18].

89 Several protease single mutations decrease the viral replication severely. However, around
90 half of HIV-1 isolates containing single primary PI resistance mutations have developed additional
91 primary mutations during salvage PI therapy [23]. A continued drug pressure allows an
92 accumulation of a number of secondary mutations (also referred to as minor mutations) which
93 partially compensate for the reduced replicative capacity resulting from the incorporation of
94 primary mutations [24,25,18,26]. In many instances, the selection of secondary mutations restore
95 the catalytic efficiency of the drug-targeted resistant enzyme leading to a rebound in fitness [27,28].
96 Several proposed mechanisms have contributed to shed light on this issue. Chang & Torbett [29]
97 studied the issue of drug resistance in a thermodynamical view of protein stability advancing that

98 major mutations destabilize and accessory mutations re-stabilize drug resistant HIV-1 protease.
99 According to Piana et al. [30,31] secondary mutations can enhance the catalytic rate of protease
100 mutants by affecting the flexibility of the protein. Several mutants of HIV-1 protease (positively
101 selected, polymorphisms, drug resistant and compensatory) was studied by DEER (Double
102 Electron-Electron Resonance) and compared to MD results (Molecular Dynamic) [32,33,34,35,36].
103 Galiano et al. [34] found an alteration in protease flap conformation in the drug pressure selected
104 mutants. de Vera et al. [33] show that drug resistant mutations (primary and secondary) in HIV-1
105 protease alter the conformational distribution in the region of the flap region using DEER (Double
106 Electron-Electron Resonance) and MD. They found that drug resistance emerges when mutations
107 increase the population of “open-like” states, at the expense of the “closed-like” conformations,
108 while maintaining “semi-open” population close to WT. Other groups [35,36] found an alteration
109 in conformational sampling and dynamics in polymorphic HIV-1 protease subtypes.

110 Several studies have been devoted to the analysis of the structural effects of protease
111 mutations in the interaction with their inhibitors [37,38,39]. However, only few studies have been
112 devoted to a systematic investigation of the relationship between the flexibility profile variations
113 and the resistance mechanism/fitness loss using the Normal Mode Analysis (NMA) [40,41,42]. This
114 method is known to be particularly efficient for studying the flexibility and structural aspects of
115 proteins since it has a low computational cost compared to molecular dynamics (Batista et al. 2011;
116 Braz et al. 2012) [40,41].

117 In a previous article [41], we have presented an extensive study of the flexibility profiles of
118 70 positively selected HIV protease mutants [43] obtained by NMA. In the present article, we
119 present those of the highly occurring double mutants and compare them to those of the
120 corresponding primary mutants and that of the WT protease. We also correlate the atomic
121 fluctuations with h-bond networks that are formed during the normal mode displacements for each
122 mutant.

124 **2. Material and Methods**

125 In this work, we used a method that combines Molecular Dynamics (MD) and NMA to
126 generate energetically relaxed conformations of the HIV-1 subtype B protease, of its primary
127 mutants, and of the most frequent related double mutants along their individual lowest frequency
128 modes. The backbone flexibilities were calculated considering the sets of conformations for each
129 molecular model considered, and a Statistical Analysis was carried out to compare the flexibility
130 profiles between them. Furthermore, we analyzed the differences in the h-bond network among
131 these structures in order to have a better insight into the resulting flexibilities. We give in what
132 follows the main steps of this study.

133

134 **2.1. Selection of single and related double mutants**

135 All the single and double mutants of the HIV-1 M subtype B protease were chosen from the
136 *Stanford HIV Drug Resistance Database* by using the “Detailed Protease Mutation Query” available
137 online in <http://hivdb.stanford.edu/cgi-bin/PRMut.cgi> (accessed at the 30th March 2012). Seven
138 primary mutations appearing very frequently against commonly used PIs were first chosen and for
139 each one of them several additional mutations occurring with a high frequency were selected. We
140 constituted two groups of mutants called base and reference groups, the former containing at least
141 one primary mutation and one additional mutation, and the latter including only additional
142 mutations excluding the primary mutations. In the base group mutants appearing with a frequency
143 above or equal to 10% were first selected, and thereafter we proceeded to the choice of secondary
144 mutants appearing with a high statistically relevant frequency within this selection. We adopted the
145 following rules to chose the secondary mutations: 1) if the secondary mutation has a frequency of
146 occurrence less than 1% in the reference group, its frequency of occurrence in the base group must
147 be at least 10 times the reference frequency; 2) if the reference frequency is between 1 and 10%, the
148 frequency in the base group must be at least 5 times this value; 3) the corresponding values for
149 larger occurrences are 11-30 % in the reference group and 3 times or more in the base group, and 4)

150 above 30% at least 0.3 in the base group. To give an example, for the primary G48V mutation, the
151 additional L10I mutation had in the reference group a frequency of occurrence of 18.3 % (that does
152 not contain G48V mutation), but it had a frequency of 83.6 % within the base group, an increase of
153 ~ 4.6 times. It was therefore selected. It should be noted that although the HIV-1 protease mutations
154 are usually divided into primary and secondary, a scheme that has been proven to be useful, this
155 classification is not always clear-cut, for example some mutations have been placed into both
156 categories, depending on the specific PI [44].

157 Using the above-described rules, 104 couples of mutations were selected. The major
158 mutations considered were I47V, G48V, I50V, L76V, V82A, I84V and L90M. The additional
159 selected mutations were V32I-M46I-I54M-I54V-L10I-L24I-L33F-K43T-I50V-F53L-Q58E-A71V-
160 G73S-T74S-E34Q-K55R- K20R-I62V-L63P-L76V-V82A-I84V-L90M (see Table 1 for the
161 combinations).

162

163 2.2. Model construction and energy minimization

164 All mutants structures were constructed using PyMOL (PyMOL Molecular Graphics System,
165 Version 1.2r3pre, Schrödinger, LLC), based on the pdb structure 1BVE [45]. The first NMR model
166 was used as the input and the drug was deleted from the model before preparing each system. The
167 energy minimization protocols used for obtaining the structures appropriate for NMA calculation
168 are the same as described in a previous article [41]. All the PyMOL built structures were first
169 minimized with explicit solvent (spc model) by using GROMACS 4.5.3 [46] and the force field
170 GROMOS96.53a6 [47] in four stages as follows: i) steepest descent minimization with the whole
171 protein kept frozen; ii) steepest descent with only the backbone kept frozen; iii) steepest descent
172 with the whole protein free; iv) unconstrained conjugate gradient minimization for the whole
173 system.

174

175 2.3. Normal Mode Analysis

176 All the modeled structures were minimized again in vacuum using the CHARMM software
177 [48] with the force field charmm27 in order to prepare them for the all-atom normal mode
178 calculations. We computed the first 27 internal lowest frequency modes of the minimized structures
179 since they describe the most representative global conformational changes.

180

181 2.4. Backbone flexibility

182 The Root Mean Square Fluctuations (RMSF) of C-alpha atoms as well as the average values
183 of the backbone atoms per residue for all the constructed models were evaluated by considering the
184 corresponding set of normal modes using CHARMM. We refer to the two monomers as chain A
185 and chain B, with residues numbered from 1 to 99 in chain A and from 1' to 99' in chain B.

186 The flexibility profiles of double mutants were compared to that of WT and to their
187 respective base mutants. Box plot of the RMSFs data were performed by all double mutants groups.
188 The vertical line extends from the minimum to the maximum value. The central box represents the
189 value from lower to upper quartile (25-75 percentile) and the red horizontal middle line represents
190 the median. The outliers are displayed as separate points.

191 Statistical tests for flexibility profiles were performed using MATLAB ((MathWorks Inc.,
192 Natick, MA) Data was evaluated according normality by Lilliefors test. Doubles mutants were
193 compared to WT and to their respective base through the Kruskal–Wallis non-parametric tests by
194 the function 'kruskalwallis'. If the Kruskal-Wallis tests were deemed to be significant ($p\text{-value} <$
195 0.05), pairwise comparison using 'multcompare' was performed with $p\text{-value}$ of 0.05 to indicate
196 how significant of the difference is. Tukey's honestly significant difference criterion was used in the
197 pair-wise comparison.

198

199 2.5. Structures generation along the modes

10

200 Detailed structural analyses were performed for the most representative double mutants that
201 showed a restoration of the WT flexibility. Based on this criterion we selected the following
202 subgroup of mutants: G48V, G48V/K43T, G48V/L10I, G48V/T74S, L76V, L76V/I84V,
203 L76V/L63P, L76V/M46I, V82A, V82A/L24I, V82A/L10I, V82A/L90M, I84V, I84V/L63P,
204 I84V/L90M, I84V/L76V, L90M, L90M/M46I, L90M/V82A, L90M/I84V. The structures of these
205 mutants were displaced along each of the internal lowest frequency modes, from 7 to 17. These
206 modes were shown to provide the most representative motions related to the protease catalytic
207 activity [40]. For a given mutant a set of structures along each of its lowest frequency modes
208 considered were generated by displacing the reference energy minimized structure to values ranging
209 from -1.4 and +1.4 Å (a negative value referring to a reverse normal mode direction), by steps of
210 0.2 Å. For each displacement successive low temperature ($T = 30K$) Molecular Dynamics (MD)
211 simulations and energy minimizations were carried out in order to obtain energetically relaxed
212 structures. A generalized-Born implicit solvent force field was used for taking into consideration the
213 solvation energy [49]. The displacements were achieved through a normal mode restraining
214 potential added to the CHARMM standard potential using the VMOD command implemented in
215 the CHARMM package [50,51]. A similar procedure has been used to describe the opening of HIV-
216 1 Protease flaps [40].

217

218 2.6. Analysis of h-bonds in the displaced structures

219 H-bond networks in all the generated structures along the modes (see above) were
220 determined by default criterion of VMD software [52]. The possibility for two residues i and j in a
221 given mutant (or in WT) to form a h-bond was evaluated as the percentage of the number of times it
222 is formed over the whole set of displaced structures, and was called MHB(i,j). The corresponding
223 value in WT was called WTHB(i,j). A relative number of h-bond formation of a mutant with respect
224 to WT between two residues was defined as PHB(i,j) = MHB(i,j) – WTHB(i,j). They are presented
225 as color graphics maps using MATLAB in the Supplementary Pages. To confirm these H-bond

11

226 changes we have analyzed the structures using Baker & Hubbard (B&H) criteria by MDTraj 1.2
227 software (mdtraj.org/latest/) in this method we no consider h-bond for distance > 4.0 Å.

228

229 **3. Results and Discussion**

230 We previously reported the effects of single mutations on the flexibility in HIV-1 protease
231 and its relation with the fitness towards the inhibitors [41]. Now we present and discuss the results
232 obtained on double mutants of the HIV-1 protease using the protocol described in Methods section.
233 We combined here molecular dynamics and NMA to analyze through a statistical analysis the
234 variations of the backbone flexibility and h-bond formation for 7 protease single mutations (base
235 mutants) and 104 associated double mutants.

236 The single and double mutants have a myriad of little effects in intensity and direction of
237 normal mode vectors. An example is given for mode 7 in Supplementary Page S1.

238 The C-alpha RMSF values of each group of double mutants associated to a base mutation
239 are presented in Fig. 2. The base mutants considered (I47V, G48V, I50V, L76V, V82A, I84V and
240 L90M) and the associated doubles mutants are listed in Table 1. Doubles mutants were compared to
241 WT and to their respective base mutant through the Kruskal-Wallis non-parametric tests.
242 Interestingly, we observed that the double mutants flexibility profile is frequently more similar to
243 that of the WT than to that of the related single base mutant. The significantly differences between
244 the regions elbow, cantilever, interface and flap were highlighted with colored blur vertical
245 rectangles in the Fig. 2. The binding significantly differences are show as blue arrows. Pairwise
246 comparisons are provided in the Supplementary Pages S2 and S3.

247 The Table 2 shows the RMSF values and the standard deviations calculated for different
248 regions (flaps, cantilever, fulcrum, active site and interface) of the WT, base and double mutants.
249 The single mutant presents a significant variation of the flexibility when compared with the WT
250 while the double mutants show clearly a recovery of the WT flexibility mainly in the flaps region.

12

251 The flaps movement is essential for substrate binding and products release. All individual
252 base mutations analyzed in this study showed some kind of alteration in the flexibility of the flaps
253 in comparison with WT HIV-1 group M subtype B protease. Base mutants I47V, G48V, I84V,
254 L90M and 76V show a decrease of the flap flexibility, while I50V and V82A base mutants showed
255 an increase of the flap flexibility. However, the majority of the analyzed double mutants show a flap
256 flexibility behavior more similar to that of the WT (Fig. 2, Supplementary Pages S11-S15 and Table
257 2).

258 Our results about the mutation effects on drug resistance corroborate those presented by
259 other authors. Using different methodologies like DEER, NMR and MD, some groups found that
260 flap flexibility is altered in HIV-1 protease mutants [53,34,36]. They showed that a limited
261 conformational opening of the flaps alters the ability of the inhibitor and /or substrate to enter into
262 the active site and, on the other hand, the longer average distance between the flaps in semi-open
263 conformations might increase the free energy cost for the flap closing in the presence of inhibitor
264 and /or substrate. The interface region have flexibility alteration in almost of single mutants and, in
265 similar way as flaps, your flexibility is restore to WT behavior in double mutants (Fig. 2,
266 Supplementary Pages S10-S15 and Table 2).

267 For residues 16-20 (C-terminal of fulcrum region - chain A) we can observe an increase of
268 the flexibility in the great majority of double mutants in comparison with the WT and the single
269 related base mutants. This may be indicative of a common compensatory mechanism in these
270 mutants in response to a common alteration induced by base mutations (Fig. 2, Supplementary
271 Pages S11-S15 and Table 2).

272 In a previous work [41] we observed that residues 8 and 81 (both present in the binding site)
273 display an increase of flexibility in the majority of positively selected mutants. In this analysis, the
274 majority of double mutants do not show any significant alteration of flexibility for these residues in
275 comparison to base mutants (Fig2, Supplementary Pages S11-S15). This may point out a deficiency
276 with regard to drug resistance. The residues 8 and 81 in the base mutant V82A display a very large

13

277 increase of flexibility and in the associated double mutants a largely low value but still significantly
278 higher than in the WT (data not shown). In the case of I47V, this base mutant does not have any
279 flexibility alteration for the residue 81 in comparison to WT but the majority of the related double
280 mutants display a significantly increased flexibility for this residue (Fig. 2, Supplementary Page
281 S10).

282 Two complementary views can be taken into account for describing HIV-1 protease drug
283 resistance and fitness recovery based on thermodynamic considerations (entropy and enthalpy) or
284 on structural dynamics (changes in flexibility). In a thermodynamic approach an alteration of the
285 population of thermodynamic states have to be taken into account in the drug resistance [33]; in the
286 structural dynamics view a variation of exchange rates among conformational states have to be
287 considered [54]. Protein motions are fundamental for enzyme activity and are connected to catalytic
288 chemical step in lowering the free energy barrier [55,56,57]. Mutations located far from the active
289 site, which do not present significant structural variations in enzyme-inhibitor complexes, give an
290 indirect evidence that protein motions might be altered affecting the reaction rate and transition
291 state stabilization [58,55,59,60,61,62].

292 To unravel the probable causes of this flexibility behavior present in base and double
293 mutants we investigated the h-bond formation or loss in the structures generated along the modes
294 for a selected group of single and double mutants and compared the results to those of the WT. We
295 present in Table 3 the frequencies of h-bond formation calculated by VMD and B&H by MDTraj
296 for few pairs of residues in WT, base and double mutants. These methods present similar results
297 major part h-bonds except for 56'-45'. The VMD criteria found this H-bond only in single mutants
298 in intermediary frequency values, but by B&H we found two h-bonds (both residues are donor and
299 acceptor) in high frequency values for WT and all mutants (Table 3). The criteria for identifying h-
300 bonds are somewhat arbitrary [63] and B & H criteria accept other range of angles in h-bonds than
301 VMD default criteria.

302 We observe that some h-bonds, such as those between residues 56'-45', 58'-45' and 30'-74'

14

303 that are not present in the WT, appear in the base mutants, but disappear in the double mutants.
304 Conversely, there are cases in which h-bond present in WT such as those between residues 35'-55'
305 and 60'-74', disappear in base mutants and appear in some double mutants at various extents. A
306 very significant result to be noted is that h-bond between between residues 58' and 45'(side chain)
307 that are formed in all the base mutants (not present in WT). We found a similar results by VMD
308 default criteria between residues 56' and 45' (backbone) but B&H criteria not confirm this h-bond
309 changes.

310 The percentage of h-bonds formation between pairs of residues in the structures generated
311 along the normal mode 7 for the WT (WTHB map), and the differences between all the mutants and
312 the WT (PHB maps) are provided in the Supplementary Pages S4 to S9. In all base mutants
313 significant alteration in h-bond formation or loss may be observed in comparison to WT. Most of
314 the changes are concentrated in the flap and cantilever regions. All the double mutants considered
315 have their h-bond network significantly altered in comparison to their respective base mutant with a
316 probable effect that their flexibility profile becomes more similar to that of the WT (see Table 3 and
317 Supplementary Pages S4 to S9).

318 The h-bonds that are formed are close to the flap cantilever region (Fig. 3), and could induce
319 important structural alterations that might be responsible for the flexibility changes and fitness loss,
320 both restored in double mutants. We observe in particular that a h-bond is formed between the
321 residues 30' and 74' in base mutants L76V, I84V and L90M, that is lost in all the derived double
322 mutants (see Table 3). This constitutes another possible case of flexibility and fitness restoration.

323 In other interesting cases the h-bonds between the residues 56-45, 59-35 and 58-45 are lost
324 in double mutants compared to the related base mutants (see Supplementary Pages S4 to S9)
325 resulting in an increase of the flexibility in the fulcrum region of chain A (Fig. 2 and Fig. 3).

326 The hydrogen bonds between residues 61'-74' (flap - cantilever) present in WT was lost in
327 all the base and double mutants (with exception of G48V and G48V/K43T) (not shown). A similar
328 case of loss of h-bond can be observed between residues 35' and 55' in the majority of mutants (with

15

329 the exception of the double mutant I84V/L76V). This feature of loss of h-bonds in mutants can be
330 closely related to drug resistance mechanism. We note also that V82A shows large changes in h-
331 bond formation or loss, the h-bond between residues 35' and 55' and between 60' and 74' disappear
332 in this mutant while new h-bonds appear between 56' and 45'(backbone) and between 58' and
333 45'(side chain), in a similar behavior to the others single mutants (see Table 3). However, h-bonds
334 between residues 54' and 47' and between 59' and 61' are exclusive to V82A base mutant (see
335 Supplementary Page S7) as far as its associated double mutants are concerned almost all the formed
336 new h-bonds in the base mutant disappear. This fact relates closely to large flexibility changes
337 observed in the base mutant.

338 Concerning G48V mutant we can observe a significant number of alterations in h-bond
339 formation and loss. Interestingly the associated double mutants display a hydrogen bonding pattern
340 very similar to that of WT. They lose the h-bonds between the residues 11 and 22 (fulcrum),
341 between residues 35 and 59 (loop-cantilever), and between 58 and 60 (cantilever) when compared
342 to the base mutant. The losses of such h-bonds are important for restoring flexibility and fitness.
343 The same situation exists for the base mutant I84V and its double mutants. For the mutants L76V
344 and L90M we observe a small change in h-bond formation.

345 Six control mutations were carried out for each base mutant in order to verify the
346 relationship between h-bond and flexibility. These mutations were made to prevent the formation of
347 the h-bond between side-chains of residues 58 'and 45'. The K45 'and Q58' residues were replaced
348 by residues with large side-chains, hydrophobic and non-aromatic (I, L and M), unable to form h-
349 bonds with their side chains, making it possible to analyze the effect of the lack of the h-bond on the
350 flexibility of the protein. Normal modes were calculated for each base mutant and compared with
351 the previous RMSF results. In all cases analyzed, the loss of the h-bond in the mutant base lead to
352 RMSD values more similar to wild type and double mutant. This indicates that the absence of h-
353 bonding results in recovery of mobility. The results are shown in Supplementary Pages S11 to S15.

354

355 **4. Conclusions**

356 In the literature there are several examples of viral fitness restoration associated with a
357 combination of primary and secondary mutations [64,25,65,66,26]. Our results point out to the
358 important observation that the flexibility behavior of the double mutants is often more similar to
359 that of the WT protease than to its base single mutant. This is intimately correlated to the formation
360 or loss of h-bonds that take place more often in the flap and cantilever regions. Such structural
361 alterations are bound to have an effect on the stability and catalytic activity of the protease, and the
362 fitness recovery observed in patients during PI therapy.

363 Another important point worth to mention is that h-bonds formed between side-chain atoms
364 of 58' and 45' in all primary mutants studied are lost in almost all the associated double and control
365 mutants, the h-bond patterns becoming similar to that of the WT. These h-bonds that are located in
366 the flap and cantilever regions near the active site are certainly associated to a decrease of the
367 flexibility and fitness loss (see Supplementary Page S7). Tiefenbrunn et al. [67] used a library of
368 brominated fragments to make a screening in HIV-1 protease crystals for finding new drug binding
369 sites in this surface. They found 3 new sites: TL-3 site, Flap-site, and Exosite. The Flap-site is
370 composed of the residues 42, 44, 45, 45, 55 and 56 (in both chains). This site could be very
371 interesting for drug design because it overlaps with our described 56'-45' and 58-45' hydrogen
372 bonds. Designing new drugs interacting at the same time with the active site and with these regions
373 could contribute effectively to a reduction of the fitness of drug resistant mutants that could arise.

374 **Acknowledgements**

375 The authors gratefully acknowledge the technical support of Patricia C. Tuffanetto. This
376 work was supported by grants from the Fapesp - Fundação de Amparo à Pesquisa do Estado de São
377 Paulo (Brazil), CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico (Brazil)
378 and the CNRS -Centre National de la Recherche Scientifique (France).

17

379 **References**

- 380 [1] UNAIDS 2013 The Joint United Nations Program on HIV/AIDS (UNAIDS). Global report:
381 UNAIDS report on the global AIDS epidemic 2013. UNAIDS / JC2502/1/E; UNAIDS, Geneva,
382 Switzerland, 2013.
- 383
- 384 [2] N.E. Kohl, E.A. Emini, W.A. Schleif, L.J. Davis, J.C. Heimbach, R.A. Dixon, E.M. Scolnick,
385 I.S. Sigal, Active human immunodeficiency virus protease is required for viral infectivity, Proc.
386 Natl. Acad. Sci. U.S.A. 85 (1988) 4686-4690.
- 387
- 388 [3] M.A. Navia, P.M. Fitzgerald, B.M. McKeever, C.T. Leu, J.C. Heimbach, W.K. Herber, I.S.
389 Sigal, P.L. Darke, J.P. Springer, Three-dimensional structure of aspartyl protease from human
390 immunodeficiency virus HIV-1, Nature 337 (1989) 615-620.
- 391
- 392 [4] A. Wlodawer, M. Miller, M. Jaskolski, B.K. Sathyaranayana, E. Baldwin, I.T. Weber, L.M. Selk,
393 L. Clawson, J. Schneider, S.B. Kent, Conserved folding in retroviral proteases: crystal structure of a
394 synthetic HIV-1 protease, Science 245 (1989) 616-621.
- 395
- 396 [5] J.M. Louis, I.T. Weber, J. Tozser, G.M. Clore, A.M. Gronenborn, HIV-1 protease: maturation,
397 enzyme specificity, and drug resistance, Adv. Pharmacol. 49 (2000) 111-146.
- 398
- 399 [6] S.C. Pettit, S.F. Michael, R. Swanstrom, The specificity of the HIV-1 protease, Perspect. Drug.
400 Discov. Des. 1 (1993) 69-83.
- 401
- 402 [7] H. Toh, M. Ono, K. Saigo, T. Miyata, Retroviral protease-like sequence in the yeast transposon
403 Ty 1, Nature 315 (1985) 691.

18

404

405 [8] I.T. Weber, J. Agniswamy, HIV-1 Protease: Structural Perspectives on Drug Resistance, *Viruses*
406 1 (2009) 1110-1136. doi:10.3390/v1031110

407

408 [9] I.T. Weber, Comparison of the crystal structures and intersubunit interactions of human
409 immunodeficiency and Rous sarcoma virus proteases, *J. Biol. Chem.* 265 (1990) 10492-10496.

410

411 [10] V. Hornak, A. Okur, R.C. Rizzo, C. Simmerling, HIV-1 protease flaps spontaneously open and
412 reclose in molecular dynamics simulations, *Proc. Natl. Acad. Sci. USA* 103 (2006) 915-920.
413 doi:10.1073/pnas.0508452103

414

415 [11] F.J. Palella, Jr., K.M. Delaney, A.C. Moorman, M.O. Loveless, J. Fuhrer, G.A. Satten, D.J.
416 Aschman, S.D. Holmberg, Declining morbidity and mortality among patients with advanced human
417 immunodeficiency virus infection. HIV Outpatient Study Investigators, *N. Engl. J. Med.* 338 (1998)
418 853-860.

419

420 [12] A. Ali, R.M. Bandaranayake, Y. Cai, N.M. King, M. Kolli, S. Mittal, J.F. Murzycki, M.N.
421 Nalam, E.A. Nalivaika, A. Ozen, M.M. Prabu-Jeyabalan, K. Thayer, C.A. Schiffer, Molecular Basis
422 for Drug Resistance in HIV-1 Protease, *Viruses* 2 (2010) 2509-2535. doi:10.3390/v2112509

423

424 [13] M.W. Tang, R.W. Shafer, HIV-1 antiretroviral resistance: scientific principles and clinical
425 applications, *Drugs* (2012) 72:e1-25. doi: 10.2165/11633630-00000000-00000

426

427 [14] I.T. Weber, Y. Zhang, J Tözsér, HIV-1 Protease and AIDS Therapy, In: *Proteases in Biology*
428 and Diseases, U. Lendeckel and M.N. Hooper, Eds., Springer, Berlin, 2009, Vol. 8, pp. 25-45.

19

429

430 [15] D. Boden, M. Markowitz, Resistance to human immunodeficiency virus type 1 protease
431 inhibitors, *Antimicrob. Agents Chemother.* 42 (1998) 2775-2783.

432

433 [16] F. Clavel, A.J. Hance, HIV drug resistance, *N. Engl. J. Med.* 350 (2004) 1023–1035.

434

435 [17] J.M. Coffin, HIV population dynamics in vivo: implications for genetic variation, pathogenesis,
436 and therapy, *Science* 267 (1995) 483-489.

437

438 [18] V.A. Johnson, V. Calvez, H.F. Gunthard, R. Paredes, D. Pillay, R. Shafer, A.M. Wensing,
439 D.D., Richman, 2011 update of the drug resistance mutations in HIV-1, *Top. Antivir. Med.* 19
440 (2011) 156-164.

441

442 [19] A.C. Anderson, Winning the arms race by improving drug discovery against mutating targets,
443 *ACS Chem. Biol.* 7 (2012) 278-288. doi: 10.1021/cb200394t

444

445 [20] J.H. Condra, W.A. Schleif, O.M. Blahy, L.J. Gabryelski, D.J. Graham, J.C. Quintero, A.
446 Rhodes, H.L. Robbins, E. Roth, M. Shivaprakash, D. Titus, T. Yang, H. Teplert, K.E. Squires, P.J.
447 Deutsc, E.A. Emini, In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors,
448 *Nature* 374 (1995) 569-571.

449

450 [21] J. Martinez-Picado, A.V. Savara, L. Sutton, R.T. D'Aquila, Replicative fitness of protease
451 inhibitor-resistant mutants of human immunodeficiency virus type 1, *J. Virol.* 73 (1999) 3744-3752.

452

20

453 [22] N.A. Roberts, Drug-resistance patterns of saquinavir and other HIV proteinase inhibitors, Aids
454 9 Suppl. 2 (1995) 27-S32.

455

456 [23] R. Kantor, W.J. Fessel, A.R. Zolopa, D. Israelski, N. Shulman, J.G. Montoya, M. Harbour, J.M.
457 Schapiro, R.W. Shafer, Evolution of primary protease inhibitor resistance mutations during protease
458 inhibitor salvage therapy, *Antimicrob. Agents Chemother.* 46 (2002) 1086-1092.

459

460 [24] C. Alteri, A. Artese, G. Beheydt, M.M. Santoro, G. Costa, L. Parrotta, A. Bertoli, C. Gori, N.
461 Orchi, E. Girardi, A. Antinori, S. Alcaro, A. d'Arminio Monforte, K. Theys, A.M. Vandamme, F.
462 Ceccherini-Silberstein, V. Svicher, C.F. Perno, Structural modifications induced by specific HIV-1
463 protease-compensatory mutations have an impact on the virological response to a first-line
464 lopinavir/ritonavir-containing regimen, *J. Antimicrob. Chemother.* 68 (2013) 2205-2209. doi:
465 10.1093/jac/dkt173

466

467 [25] G. Croteau, L. Doyon, D. Thibeault, G. McKercher, L. Pilote, D. Lamarre, Impaired fitness of
468 human immunodeficiency virus type 1 variants with high-level resistance to protease inhibitors, *J.*
469 *Virol.* 71 (1997) 1089-1096.

470

471 [26] M. Nijhuis, R. Schuurman, D. de Jong, J. Erickson, E. Gustchina, J. Albert, P. Schipper, S.
472 Gulnik, C.A. Boucher, Increased fitness of drug resistant HIV-1 protease as a result of acquisition
473 of compensatory mutations during suboptimal therapy, *AIDS* 13 (1999) 2349-2359.

474

475 [27] Z. Chen, Y. Li, H.B. Schock, D. Hall, E. Chen, L.C. Kuo, Three-dimensional structure of a
476 mutant HIV-1 protease displaying cross-resistance to all protease inhibitors in clinical trials, *J. Biol.*
477 *Chem.* 270 (1995) 21433-21436.

478

21

479 [28] M.E. Quiñones-Mateu, E.J. Arts, HIV-1 Fitness: Implications for Drug Resistance, Disease
480 Progression, and Global Epidemic Evolution, In: HIV Sequence Compendium 2001, C. Kuiken, B.
481 Foley, B. Hahn, P. Marx, F. McCutchan, et al., Eds., Theoretical Biology and Biophysics Group,
482 Los Alamos National Laboratory, New Mexico, 2001, pp 134-170.

483

484 [29] M.W. Chang, B.E. Torbett, Accessory mutations maintain stability in drug-resistant HIV-1
485 protease, *J. Mol. Biol.* 410 (2011) 756-760. doi: 10.1016/j.jmb.2011.03.038

486

487 [30] S. Piana, P. Carloni, U., Rothlisberger, Drug resistance in HIV-1 protease: Flexibility-assisted
488 mechanism of compensatory mutations, *Protein Sci.* 11 (2002) 2393-2402.

489

490 [31] S. Piana, P. Carloni, M. Parrinello, Role of conformational fluctuations in the enzymatic
491 reaction of HIV-1 protease, *J. Mol. Biol.* 319 (2002) 567-583.

492

493 [32] I.M. de Vera, M.E. Blackburn, G.E. Fanucci, Correlating conformational shift induction with
494 altered inhibitor potency in a multidrug resistant HIV-1 protease variant, *Biochemistry* 51 (2012)
495 7813-7815. doi: 10.1021/bi301010z

496

497 [33] I.M. de Vera, A.N. Smith, M.C. Dancel, X. Huang, B.M. Dunn, G.E. Fanucci, Elucidating a
498 relationship between conformational sampling and drug resistance in HIV-1 protease, *Biochemistry*
499 52 (2013) 3278–3288. doi: 10.1021/bi400109d

500

501 [34] L. Galiano, F. Ding, A.M. Veloro, M.E. Blackburn, C. Simmerling, G.E. Fanucci, Drug
502 pressure selected mutations in HIV-1 protease alter flap conformations, *J. Am. Chem. Soc.* 131
503 (2009) 430-431. doi:10.1021/ja807531v

22

504

505 [35] X. Huang, M.D. Britto, J.L. Kear, B.D. Christopher, J.R. Rocca, C. Simmerling, R. McKenna,
506 M. Bieri, P.R. Gooley, B.M. Dunn, G.E. Fanucci, The role of select subtype polymorphisms on
507 HIV-1 protease conformational sampling and dynamics, *J. Biol. Chem.* 289 (2014) 17203-17214.
508 doi: 10.1074/jbc.M114.571836.

509

510 [36] J.L. Kear, M.E. Blackburn, A.M. Veloro, B.M. Dunn, G.E. Fanucci, Subtype
511 polymorphisms among HIV-1 protease variants confer altered flap conformations and flexibility, *J.*
512 *Am. Chem. Soc.* 131 (2009) 14650-14651. doi: 10.1021/ja907088a

513

514 [37] M. Kozísek, J. Bray, P. Rezáčová, K. Saskovaá, J. Brynda, J. Pokorna, F. Mammano, L.
515 Rulšek, J. Konvalinka, Molecular analysis of the HIV-1 resistance development: enzymatic
516 activities, crystal structures, and thermodynamics of nelfinavir-resistant HIV protease mutants, *J.*
517 *Mol. Biol.* 374 (2007) 1005-1016.

518

519 [38] J. Martinez-Picado, A.V. Savara, L. Shi, L. Sutton, R.T. D'Aquila, Fitness of human
520 immunodeficiency virus type 1 protease inhibitor-selected single mutants, *Virology* 275 (2000)
521 318-322.

522

523 [39] S.K. Sadiq, S. Wan, P.V. Coveney, Insights into a mutation-assisted lateral drug escape
524 mechanism from the HIV-1 protease active site, *Biochemistry* 46 (2007) 14865-14877.

525

526 [40] P.R. Batista, P. Gaurav, P.G. Pascutti, P.M. Bisch, D. Perahia, C.H. Robert, Free energy
527 profiles along consensus normal modes provide insight into HIV-1 protease flap opening, *J. Chem.*
528 *Theory Comput.* 7 (2011) 2348–2352. doi: 10.1021/ct200237u

529

23

530 [41] A.S. Braz, P. Tufanetto, D. Perahia, L.P. Scott, Relation between flexibility and positively
531 selected HIV-1 protease mutants against inhibitors, *Proteins* 80 (2012) 2680-2691. doi:
532 10.1002/prot.24151

533

534 [42] S.E. Dobbins, V.I. Lesk, M.J. Sternberg, Insights into protein flexibility: The relationship
535 between normal modes and conformational change upon protein-protein docking, *Proc. Natl. Acad.*
536 *Sci. U.S.A.* 105 (2008) 10390-10395. doi: 10.1073/pnas.0802496105

537

538 [43] L. Chen, A. Perlina, C.J. Lee, Positive selection detection in 40,000 human immunodeficiency
539 virus (HIV) type 1 sequences automatically identifies drug resistance and positive fitness mutations
540 in HIV protease and reverse transcriptase, *J. Virol.* 78 (2004) 3722-3732.

541

542 [44] C. Dykes, L.M. Demeter, Clinical significance of human immunodeficiency virus type 1
543 replication fitness, *Clin. Microbiol. Rev.* 20 (2007) 550-578. doi: 10.1128/CMR.00017-07

544

545 [45] T. Yamazaki, A.P. Hinck, Y.X. Wang, L.K. Nicholson, D.A. Torchia, P. Wingfield, S.J. Stahl,
546 J.D. Kaufman, C.H. Chang, P.J. Domaille, P.Y. Lam, Three-dimensional solution structure of the
547 HIV-1 protease complexed with DMP323, a novel cyclic urea-type inhibitor, determined by nuclear
548 magnetic resonance spectroscopy, *Protein Sci.* 5 (1996) 495-506.

549

550 [46] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: a message passing parallel
551 molecular dynamics implementation, *Comput. Phys. Commun.* 91 (1995) 43–56.

552

553 [47] C. Oostenbrink, A. Villa, A.E. Mark, W.F. van Gunsteren, A biomolecular force field based on
554 the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and
555 53A6, *J. Comput. Chem.* 25 (2004) 1656-1676.

556

557 [48] B.R. Brooks, R.E. Brucolieri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus,

24

558 CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, J.
559 Comp. Chem. 4 (1983) 187–217.

560

561 [49] W. Im, M.S. Lee, C.L. Brooks, 3rd, Generalized born model with a simple smoothing function,
562 J. Comput. Chem. 24 (2003) 1691-1702.

563

564 [50] N. Floquet, J.D. Marechal, M.A. Badet-Denisot, C.H. Robert, M. Dauchez, D. Perahia, Normal
565 mode analysis as a prerequisite for drug design: application to matrix metalloproteinases inhibitors,
566 FEBS Lett. 580 (2006) 5130-5136.

567

568 [51] N. Floquet, P. Durand, B. Maigret, B. Badet, M.A. Badet-Denisot, D. Perahia, Collective
569 motions in glucosamine-6-phosphate synthase: influence of ligand binding and role in ammonia
570 channelling and opening of the fructose-6-phosphate binding site, J. Mol. Biol. 385 (2009) 653-
571 664. doi: 10.1016/j.jmb.2008.10.032

572

573 [52] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14
574 (1996) 33-38.

575

576 [53] Y. Cai, N.K. Yilmaz, W. Myint, R. Ishima, C.A. Schiffer, Differential Flap Dynamics in Wild-
577 type and a Drug Resistant Variant of HIV-1 Protease Revealed by Molecular Dynamics and NMR
578 Relaxation, J. Chem. Theory Comput. 10 (2012) 3452-3462. doi: [10.1021/ct300076y](https://doi.org/10.1021/ct300076y)

579

580 [54] A.L. Perryman, J.H. Lin, J.A. McCammon, HIV-1 protease molecular dynamics of a wild-type
581 and of the V82F/I84V mutant: possible contributions to drug resistance and a potential new target
582 site for drugs, Protein 13 (2004) 1108-1123.

583

584 [55] Y. Fan, A. Cembran, S. Ma, J. Gao, Connecting protein conformational dynamics with
585 catalytic function as illustrated in dihydrofolate reductase, Biochemistry 52 (2013) 2036-2049. doi:

25

586 10.1021/bi301559q

587

588 [56] D.R. Glowacki, J.N. Harvey, A.J. Mulholland, Taking Ockham's razor to enzyme dynamics
589 and catalysis, *Nat. Chem.* 4 (2012) 169-176. doi: 10.1038/nchem.1244

590

591 [57] Z.D. Nagel, J.P. Klinman, A 21st century revisionist's view at a turning point in enzymology,
592 *Nat. Chem. Biol.* 5 (2009) 543-550. doi: 10.1038/nchembio.204

593

594 [58] D.D. Boehr, D. McElheny, H.J. Dyson, P.E. Wright, The dynamic energy landscape of
595 dihydrofolate reductase catalysis, *Science* 313 (2006) 1638-1642.

596

597 [59] M. Ghanem, L. Li, C. Wing, V.L. Schramm, Altered thermodynamics from remote mutations
598 altering human toward bovine purine nucleoside phosphorylase, *Biochemistry* 47 (2008) 2559-2564.
599 doi: 10.1021/bi702132e.

600

601 [60] K.A. Henzler-Wildman, V. Thai, M. Lei, M. Ott, M. Wolf-Watz, T. Fenn, E. Pozharski, M.A.
602 Wilson, G.A. Petsko, M. Karplus, C.G. Hübner, D. Kern, Intrinsic motions along an enzymatic
603 reaction trajectory, *Nature* 450 (2007) 838-844.

604

605 [61] V.J. Hilser, Biochemistry. An ensemble view of allostery, *Science* 327 (2010) 653-654. doi:
606 10.1038/nature13001

607

608 [62] J.P. Loria, R.B. Berlow, E.D. Watt, Characterization of enzyme motions by solution NMR
609 relaxation dispersion, *Acc. Chem. Res.* 41 (2008) 214-221. doi: 10.1021/ar700132n.

610

611 [63] P. Gasparotto, and M. Ceriotti, Recognizing molecular patterns by machine learning: an
612 agnostic structural definition of the hydrogen bond, *J. Chem. Phys.* 141 (2014) 174110. doi:

26

613 10.1063/1.4900655

614

615

616 [64] J.D. Barbour, T. Wrin, R.M. Grant, J.N. Martin, M.R. Segal, C.J. Petropoulos, S.G. Deeks,
617 Evolution of phenotypic drug susceptibility and viral replication capacity during long-term
618 virologic failure of protease inhibitor therapy in human immunodeficiency virus-infected adults, J.
619 Virol. 76 (2002) 11104-11112.

620

621 [65] G.J. Henderson, N. Parkin, R. Swanstrom, A systematic analysis of the fitness effects of
622 mutations associated with resistance to protease inhibitors, Antivir. Ther. 10 (2005) Suppl 1:S165.

623

624 [66] J. Martinez-Picado, A.V. Savara, L. Sutton, R.T. D'Aquila, Replicative fitness of protease
625 inhibitor-resistant mutants of human immunodeficiency virus type 1, J. Virol. 73 (1999) 3744-3752.

626

627 [67] T. Tiefenbrunn, S. Forli, M. Happer, A. Gonzalez, Y. Tsai, M. Soltis, J.H. Elder, A.J. Olson,
628 C.D. Stout, Crystallographic fragment-based drug discovery: use of a brominated fragment library
629 targeting HIV protease, Chem. Biol. Drug. Des. 83 (2014) 141-148. doi: 10.1111/cbdd.12227

630

631

27

632

633 **Table 1**
 634 HIV-1 Protease Double Mutations Combination list
 635

Base mutation	Additional mutation
I47V	L10I K20R V32I L33F E34Q K43T M46I F53L I54M K55R Q58E I62V L63P A71V G73S V82A I84V L90M
G48V	L10I K20R L33F E34Q K43T M46I I50V I54V Q58E I62V L63P A71V T74S V82A I84V L90M
I50V	L10I K20R L33F E34Q M46I G48V F53L I54V K55R Q58E L63P A71V T74S V82A L90M
L76V	L10I K20R L24I L33F E34Q M46I I54V Q58E I62V L63P A71V I84V L90M
V82A	L10I K20R L24I V32I L33F K43T M46I I47V G48V I50V I54V K55R Q58E I62V L63P A71V G73S T74S I84V L90M
I84V	L10I K20R L24I V32I L33F E34Q M46I I47V G48V I54V Q58E I62V L63P A71V G73S L76V V82A I90M
L90M	L10I K20R V32I L33F K43T M46I I47V G48V I50V I54V K55R Q58E I62V L63P G73S L76V V82A I84V

636

637

638

639

640

641

642

643

644

645

646

28

647 **Table 2**

648 RMSF values (mean and standard deviation-std) for different regions (flaps, cantilever, fulcrum,
649 active site and interface) of the wild type (WT), selected base mutants (BM) and selected double
650 mutants.

651

652

653 **Table 3**

654 Frequency of hydrogen bond formation calculated by criteria of VMD and H&B (Baker & Hubbard)
655 involving few residues in wild type (WT), base mutants (BM) and double mutants.

656

657

658 Tables 2 and 3 were attached in xls format.

29

659 **Figure captions**

660

661 **Fig. 1. HIV-1 Protease Structure (1BVE).** Color indicates distinct regions: Flap region in red
662 (residues 43–58), cantilever in green (residues 59–75), fulcrum in orange (residues 10–23), dimer
663 interface in violet (1–4, 50–52 and 96–99). A) yellow spheres indicate location of the binding site, B)
664 Red spheres indicate location of the base mutations: I47V, G48V, I50V, L76V, V82A, I84V,
665 L90M. C) Cyan spheres indicate location of the following additional mutations: V32I, M46I,
666 I54M, I54V, L10I, L24I, L33F, K43T, F53L, Q58E, A71V, G73S, T74S, E34Q, K55R,
667 K20R, I62V, L63P

668

669 **Fig. 2. double mutations of:** A)I47V, B)G48V, C)I50V, D)L76V and E)I84V. Quartile data
670 analysis of grouped double mutants RMSF values compared with I47V single mutant and wild type
671 (WT) RMSF values for Chain A and Chain-B: The blue rectangles, indicate IQR (Inter-Quartile
672 Range, with 50% of data), the red horizontal lines inside rectangle represent median values; the
673 vertical lines above rectangle indicates other 25% of data, and vertical lines below indicates the last
674 25% of data; the red plus sign represent outliers values. The green line represents RMSF values of
675 WT and the pink line those of the base single mutant. In the bottom: the pink arrows highlight
676 active site cavity; interface, flap, cantilever and fulcrum regions are marked, respectively, by pink,
677 red, green and orange rectangles. The list of double mutants for each base mutant is given in Table
678 1. The colored blurred background indicate regions (interface, flap, cantilever and fulcrum) with
679 significant flexibility differences, the blue arrows indicate significant flexibility alteration in binding
680 site.

681

682 **Fig. 3. Normal mode 7:** Hydrogen bonds possibilities (gain and loss) in HIV-1 protease mutants in
683 the structures generated along the normal mode 7 involving the residues 45', 56' and 56' and 58'
684 represented in licorice. Circles point to the hydrogen bonds formed by these residues. These
685 residues are located in the border of flap (red), cantilever (green) and fulcrum (orange) regions near
686 the active-site. Hydrogen bonds are present only in single mutants (G48V, V82A, L90M)

Table I Double mutation HIV-1 Protease Combination list

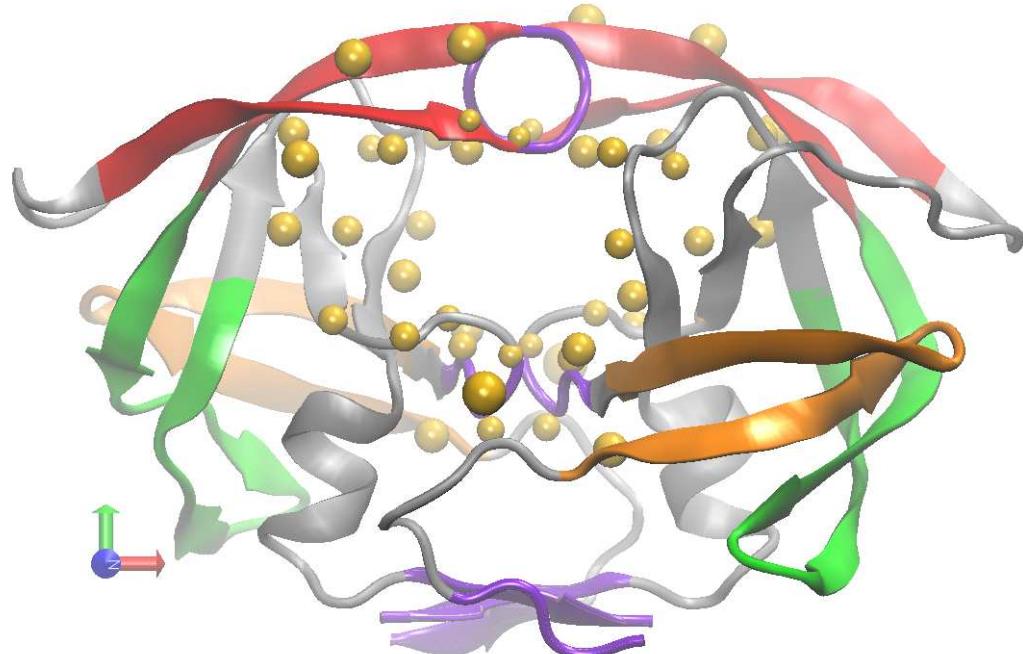
Base mutation	Additional mutation
I47V	L10I K20R V32I L33F E34Q K43T M46I F53L I54M K55R Q58E I62V L63P A71V G73S V82A I84V L90M
G48V	L10I K20R L33F E34Q K43T M46I I50V I54V Q58E I62V L63P A71V T74S V82A I84V L90M
I50V	L10I K20R L33F E34Q M46I G48V F53L I54V K55R Q58E L63P A71V T74S V82A L90M
L76V	L10I K20R L24I L33F E34Q M46I I54V Q58E I62V L63P A71V I84V L90M
V82A	L10I K20R L24I V32I L33F K43T M46I I47V G48V I50V I54V K55R Q58E I62V L63P A71V G73S T74S I84V L90M
I84V	L10I K20R L24I V32I L33F E34Q M46I I47V G48V I54V Q58E I62V L63P A71V G73S L76V V82A L90M
L90M	L10I K20R V32I L33F K43T M46I I47V G48V I50V I54V K55R Q58E I62V L63P G73S L76V V82A I84V

WT	BM G48V			Double mutants K43T L10I T74S			BM L76V			BM I84V			Double mutants L24I L10I L90M			BM I84V			Double mutants L63P L90M L76V			BM L90M			Double mutants M46I V82A I84V		
RMSF																											
all																											
mean	0.32	0.25	0.33	0.33	0.34	0.28	0.35	0.35	0.36	0.48	0.33	0.33	0.33	0.28	0.34	0.33	0.35	0.32	0.32	0.33	0.33						
std	0.10	0.08	0.11	0.12	0.11	0.09	0.10	0.11	0.14	0.31	0.11	0.11	0.11	0.08	0.12	0.11	0.10	0.14	0.10	0.11	0.11						
flaps																											
mean	0.36	0.28	0.37	0.36	0.38	0.33	0.44	0.42	0.42	0.55	0.42	0.39	0.43	0.34	0.41	0.42	0.44	0.31	0.39	0.43	0.42						
std	0.07	0.04	0.09	0.06	0.05	0.06	0.08	0.10	0.15	0.20	0.09	0.09	0.15	0.06	0.09	0.09	0.08	0.04	0.07	0.15	0.09						
cantilever																											
mean	0.31	0.27	0.32	0.34	0.34	0.28	0.34	0.39	0.34	0.40	0.31	0.33	0.31	0.26	0.33	0.31	0.34	0.33	0.36	0.31	0.31						
std	0.09	0.08	0.10	0.19	0.07	0.09	0.09	0.11	0.10	0.15	0.09	0.10	0.08	0.05	0.10	0.08	0.09	0.20	0.11	0.08	0.08						
fulcrum																											
mean	0.34	0.27	0.37	0.32	0.38	0.29	0.37	0.35	0.39	0.48	0.35	0.36	0.33	0.27	0.39	0.36	0.37	0.32	0.34	0.33	0.36						
std	0.08	0.10	0.10	0.07	0.17	0.09	0.12	0.08	0.13	0.25	0.13	0.12	0.09	0.05	0.16	0.13	0.12	0.12	0.11	0.09	0.13						
interface																											
mean	0.42	0.29	0.44	0.47	0.35	0.31	0.43	0.48	0.55	0.54	0.40	0.42	0.44	0.30	0.44	0.44	0.43	0.35	0.40	0.44	0.44						
std	0.12	0.05	0.15	0.11	0.07	0.07	0.10	0.12	0.23	0.19	0.11	0.10	0.17	0.07	0.12	0.13	0.10	0.08	0.09	0.17	0.13						
bind site																											
mean	0.28	0.22	0.28	0.28	0.29	0.25	0.32	0.31	0.31	0.44	0.30	0.30	0.33	0.27	0.31	0.31	0.32	0.27	0.28	0.33	0.31						
std	0.08	0.05	0.09	0.07	0.08	0.08	0.11	0.10	0.13	0.20	0.10	0.11	0.14	0.07	0.10	0.11	0.11	0.07	0.10	0.14	0.11						

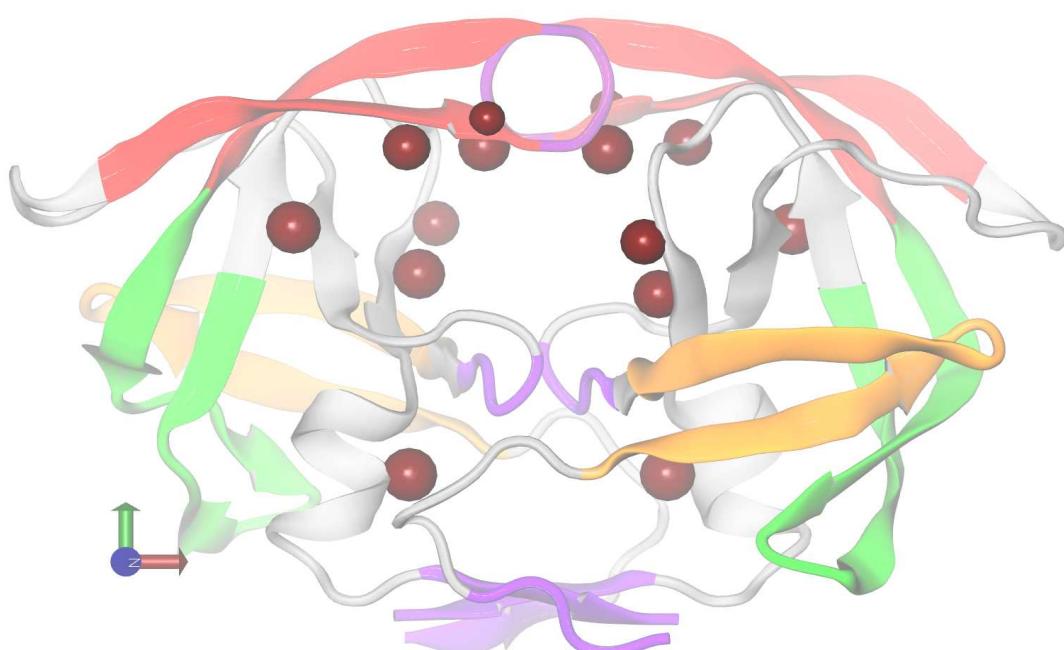
met	hydrogen bonds	WT	BM		BM			BM			BM			BM			BM					
		G48V	Double mutants	K43T L10I T74S	L76V	Double mutants	I84V L63P M46I	V82A	Double mutants	L24I L10I L90M	I84V	Double mutants	L63P L90M L76V	L90M	Double mutants	M46I V82A I84V						
56'-45'																						
VMD	freq	0	32	1	0	1	83	8	0	0	66	0	0	0	57	0	0	8	75	0	0	0
B&H	freq	96	99	86	100	93	100	100	100	99	98	89	96	100	98	100	100	100	100	100	100	100
58'-45'																						
VMD	freq	0	96	89	0	5	94	0	0	0	82	0	0	0	93	0	0	0	86	0	0	0
B&H	freq	0	97	99	0	0	100	0	0	0	97	0	0	0	98	0	0	0	99	0	0	0
30'-74'																						
VMD	freq	0	5	0	2	81	94	0	1	0	8	5	1	0	88	0	1	0	95	0	0	1
B&H	freq	0	0	0	0	92	97	0	0	0	0	0	0	0	90	0	0	0	98	0	0	0
35'-55'																						
VMD	freq	100	11	0	0	0	0	100	100	0	2	0	0	0	2	0	0	100	3	0	0	0
B&H	freq	100	11	0	0	0	0	100	100	0	0	0	0	0	0	0	0	100	0	0	0	0
60'-74'																						
VMD	freq	99	77	74	3	4	0	10	1	2	9	2	4	0	1	0	0	10	0	1	0	0
B&H	freq	100	89	81	0	0	0	34	0	0	0	0	0	0	0	0	0	34	0	0	0	0

Figure1

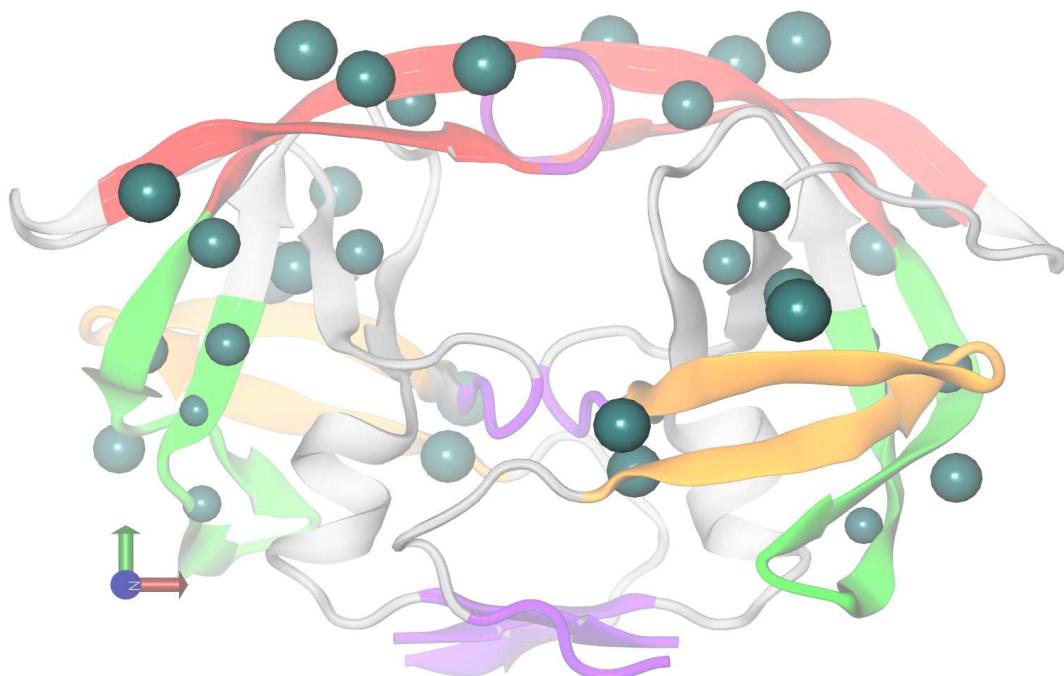
a



b



c



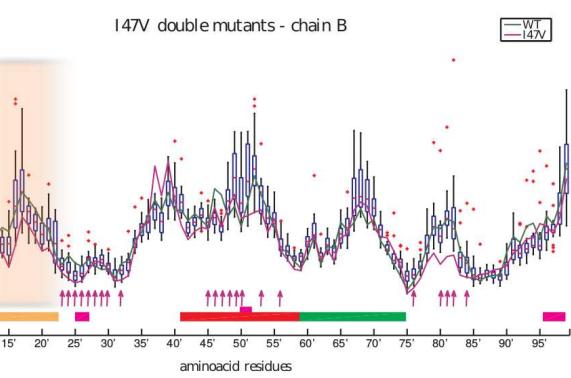
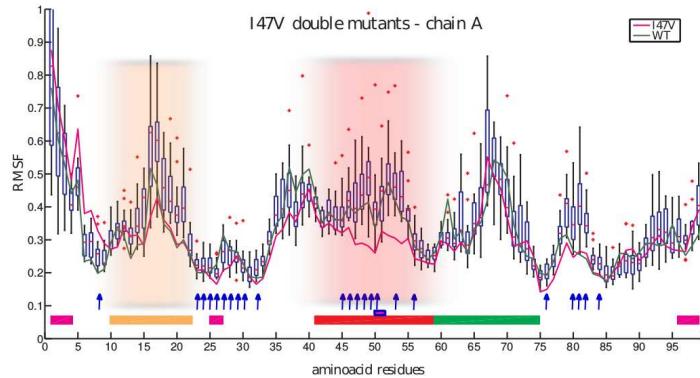
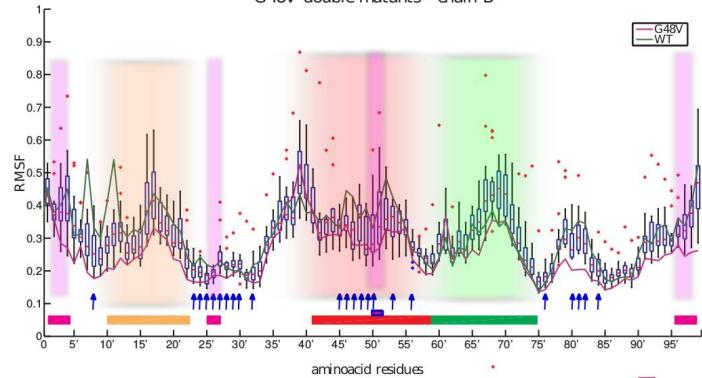
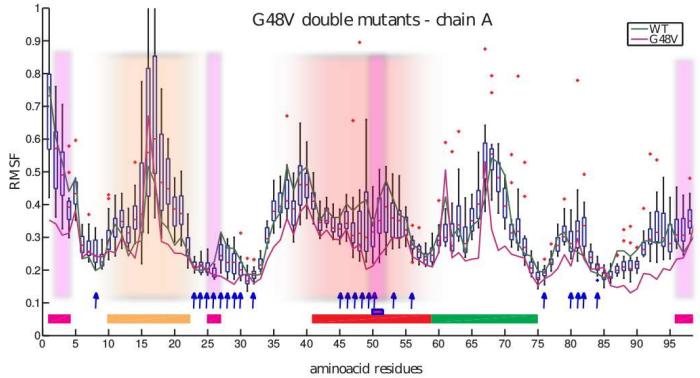
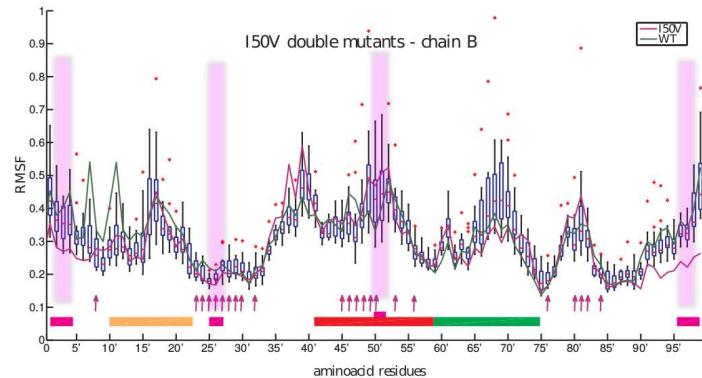
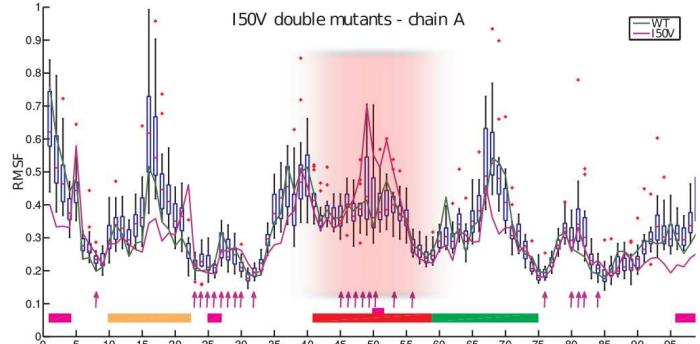
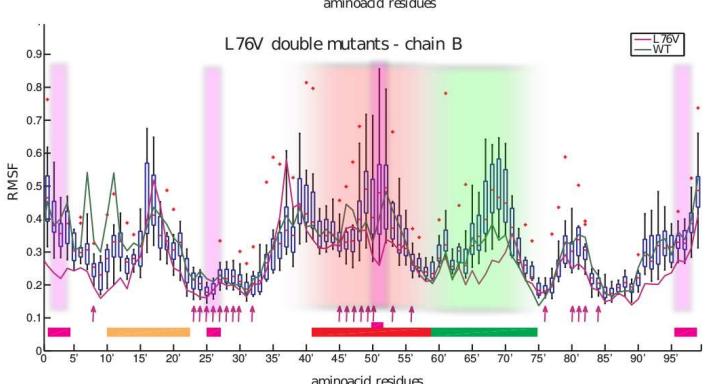
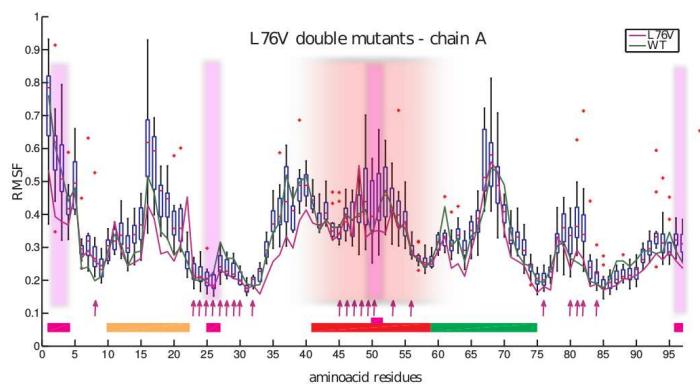
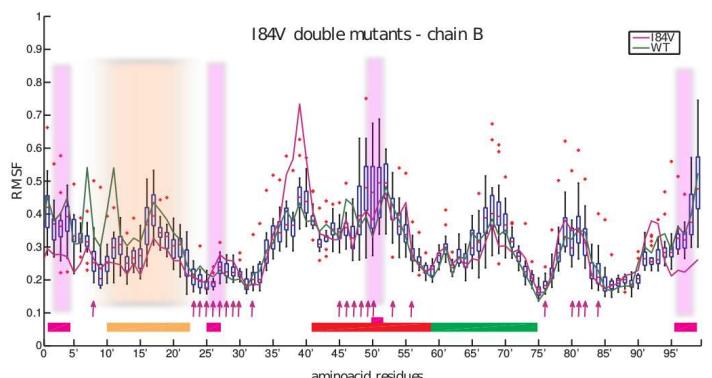
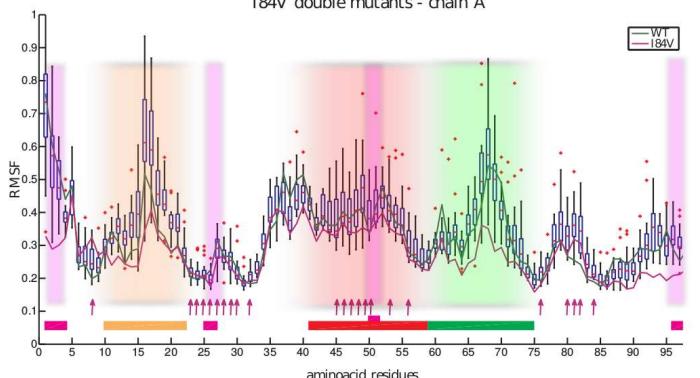
A**B****C****D****E**

Figure3

