Selection of HIV-1 for Resistance to Fifth Generation Protease Inhibitors Reveals Two Independent Pathways to High-Level Resistance

Ean Spielvogel[1,2], Sook-Kyung Lee[2], Shuntai Zhou[2], Gordon J. Lockbaum[3+], Mina Henes[3#], Amy Sondegroth[1,2], Klajdi Kosovrasti[3], Ellen A. Nalivaika[3], Akbar Ali[3], Nese Kurt Yilmaz[3], Celia A. Schiffer[3*], Ronald Swanstrom[2,4,*]

[1]Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA

[2]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA

[3]Department of Biochemistry and Molecular Biotechnology, University of Massachusetts Chan Medical School, Worcester, MA 01605, USA

[4]Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA

ORCID:

Swanstrom: 0000-0001-7777-0773

Lockbaum: 0000-0003-2720-6984

Henes: 0000-0001-8667-541X

Kosovrasti: 0000-0001-7200-4068

Nalivaika: 0000-0002-2654-1513

Ali: 0000-0003-3491-791X

24 Kurt Yilmaz: 0000-0002-5036-676X

25 Schiffer: 0000-0003-2270-6613

26

27 *Corresponding Authors

28 Ronald Swanstrom: Phone: +1 919 966 5710; ron_swanstrom@med.unc.edu

29 Celia A. Schiffer: Phone: +1 508 856 8008; Celia.Schiffer@umassmed.edu

30

31 [+]GJL – present address: Accent Therapeutics, 65 Hayden Ave, Lexington, MA 02421

32 [#]MH – present address: Emory University School of Medicine,100 Woodruff Circle, Atlanta,

33 GA 30322 USA

Summary

Darunavir (DRV) is exceptional among potent HIV-1 protease inhibitors (PIs) in the high drug concentrations that are achived *in vivo*. Little is known about the *de novo* resistance pathway for DRV. We have selected for resistance to high drug concentrations against ten PIs and their structural precursor DRV. Mutations accumulated through two pathways (anchored by protease mutations I50V or I84V). Small chemical changes in the inhibitor P1' side chain led to preferential use of one pathway over the other. Changes in the inhbitor P2' side chain determined differences in potency that were retained in the resistant viruses. Viral variants from the two pathways showed differential selection of compensatory mutations in Gag cleavage sites. These results reveal the high level of selective pressure that is attainable with fourth generation PIs, and how features of the inhibitor affect both the resistance pathway and the residual potency in the face of resistance.

Keywords

HIV-1, protease, inhibitor, resistance, selection, evolution, infectivity, mutations, enzyme, compensatory

50 **Introduction**

51 Highly active antiretroviral therapy against HIV-1 with combinations of drugs effectively
52 block viral replication and preclude the evolution of drug resistance. Three major factors interplay
53 to define the emergence of resistance *in vivo*: i) the active drug concentration relative to its
54 inhibitory activity; ii) the level of resistance conferred by one or more mutations; and iii) the
55 fitness cost of the resistance mutations. For protease inhibitors (PIs), the number of mutations
56 selected increases with increasing drug concentration (Watkins et al., 2003). Thus the
57 concentration of this drug class *in vivo* is an important parameter in determining the genetic
58 barrier to resistance.

59 The early identification of the retroviral protease as a member of the aspartyl proteinase
60 family and the determination of a number of cleavage site sequences led to the development of
61 first generation PIs (PIs) that validated the HIV-1 protease as a drug target (Katoh et al., 1987;
62 Richards et al., 1989; Seelmeier et al., 1988). A second generation of PIs was quickly developed
63 for use in humans, becoming the third drug in a three-drug regimen that achieved sustained
64 suppression of viral load with no evolution of resistance (Gulick et al., 1997). The third generation
65 of PIs had improved properties with regard to side effects and efficacy. These properties have
66 been further enhanced with a fourth-generation PI, darunavir (DRV), which achieves a drug level
67 in plasma (>1 μM) that is 1,000 fold greater than its inhibitory activity in cell culture (Ali et al.,
68 2010; Nalam et al., 2013; Kurt Yilmaz et al., 2009). The potential efficacy of a fourth-generation
69 PI such as DRV has led to attempts to use this drug in monotherapy (Antinori et al., 2015; Arribas
70 et al., 2012; Paton et al., 2015; Valantin et al., 2012). Although the initial trials of monotherapy
71 showed a modest increase in the loss of viral suppression relative to standard ART, in the cases
72 of virologic failure there was no significant resistance to DRV detected in the rebound virus

73 (Antinori et al., 2015; Arribas et al., 2012; Paton et al., 2015; Valantin et al., 2012). Thus the
74 observed virologic failure is most easily attributed to issues of low drug exposure in specific
75 participants. While therapy has not yet moved to single agents, potent two drug therapies are
76 being explored. In this regard, several studies have shown that a PI can be part of a successful
77 two drug therapy regimen (Casado et al., 2020; Di Cristo et al., 2020; Hawkins et al., 2019; Huang
78 et al., 2019; Vizcarra et al., 2020).

79 Drug resistance to DRV is incompletely understood. The appearance of DRV-resistant
80 viruses during therapy failure has occurred in people who had previously experienced virologic
81 failure with other PIs (de Meyer et al., 2008; Descamps et al., 2009; Pellegrin et al., 2008;
82 Tremblay, 2008). In these cases the pathway to DRV resistance was likely influenced by the pre-
83 existing resistance to the previous PI. DRV resistance has been selected in cell culture (Aoki et
84 al., 2018; de Meyer et al., 2008; Koh et al., 2010). High level resistance was selected starting with
85 either a mixture of highly resistant variants that were allowed to recombine, or the selection was
86 started with a single PI-resistant strain; in both cases an I84V resistance pathway in the protease
87 was observed. However, there are no reports of selection of high level resistance starting with an
88 unmutated "wild type" virus. Thus there is little information on what a *do novo* pathway for
89 resistance to DRV would be.

90 An important concept in HIV-1 PI design is to avoid chemical moieties that extend outside
91 of the substrate envelope, the shared volume in the protease active site that is occupied by natural
92 protease substrates (King et al., 2004; Prabu-Jeyabalan et al., 2002). We have previously designed
93 a series of highly potent PIs based on DRV, UMASS-1 through -10, that still fit within the
94 substrate envelope (Nalam et al., 2013). The designed inhibitors have modified chemical moieties
95 that relative to DRV further fill the substrate envelope at the equivalent of the P1' and P2'

96 positions. All bind tightly to the wild type HIV-1 protease with a Ki of less than 5 pM. These
97 inhibitors retained robust binding to many multi-drug resistant protease variants and viral strains.
98 In an effort to define the *de novo* pathway to resistance for DRV and determine the effects
99 of these chemical changes in the inhibitors, we examined the evolutionary path used to attain high
100 level resistance to DRV and the UMASS-1 through -10 panel of PIs. Under continuous and
101 escalating selective pressure, representing between 50-95 passages, the virus evolved to
102 accumulate multiple mutations. In most cases, selection was carried out until the the final
103 concentration approximated that achieved by DRV *in vivo*. Relatively minor modifications in
104 inhibitor structure favored selection of one of two pathways to resistance, anchored by protease
105 mutation I50V or I84V. In addition, one side chain of the inhibitor increased the residual potency
106 against highly resistant strains. These results reveal the extremely high genetic barrier to
107 resistance to fourth-generation PIs at inhibitor concentrations that can be achieved *in vivo*, and
108 the complex evolutionary pathways required to achieve resistance. In addition, these results reveal
109 features of inhibitor design that both affect the selected pathway and determine residual potency
110 against resistant variants.
111

112 **Results**

113 **Panel of highly potent and analogous HIV-1 PIs**

114 HIV-1 PIs were designed by modifications to DRV to increase favorable interactions with
115 the protease within the substrate envelope thereby increasing potency while minimizing evolution
116 of resistance (Nalam et al., 2013). A panel of ten DRV analogues were chosen with enzymatic
117 inhibition constants (K_i) in the single or double-digit picomolar range to wild-type NL4-3 protease
118 and the I84V and I50V/A71V drug resistant variants, respectively (Table 1) (Lockbaum et al.,
119 2019; Mittal et al., 2013). These PIs contained modified P1' positions with (S)-2-methylbutyl or
120 2-ethyl-n-butyl moieties (R1-1 and R1-2, respectively) in combination with five diverse P2′
121 phenyl-sulfonamides (R2-1 to R2-5), with the inhibitors named UMASS- 1 through -10 (Table
122 1). These inhibitors and DRV were also tested in a cell culture-based viral inhibition assay. The
123 EC_{50} values (the effective concentration needed at the time of virus production to reduce
124 infectivity by 50%) for DRV and the UMASS analogues ranged from 2.4 to 9.1 nM, significantly
125 more potent than the second and third generation PIs tested (Fig. S1).

126 There are several patterns worth noting in the K_i and EC_{50} values. First, all inhibitors had
127 a K_i of less than 5 pM (limit of detection) against the wild type enzyme, precluding a comparison
128 of potency among them without using partially resistant mutant proteases (I84V and I50V/A71V).
129 Second, the single P1'/R1 position change to the larger 2-ethyl-n-butyl moiety (R1-2, UMASS-6)
130 increased potency against the I84V mutant protease relative to DRV while the intermediate sized
131 R1-1 moiety did not (UMASS-1). Third, four of the inhibitors with the R2 moieties R2-2 or R2-
132 5 were significantly more potent (UMASS-2, -5, -7, -10) than with the other R2 moieties. This
133 trend was most pronounced for activity against the I50V/A71V mutant enzyme in both the R1-1
134 and R1-2 backgrounds, and for the I84V mutant enzyme in the R1-1 background (Fig. S2). The

135 significant potency of the R2-2 and R2-5 containing inhibitors against the I50V/A71V mutant
136 enzyme while maintaining good potency against the I84V enzyme indicates that both the R1 and
137 R2 constituents contribute to potency, and to enhancing potency over DRV. This potency was
138 also seen in inhibition of viral infectivity as the EC_{50} values of the R2-2 and R2-5 containing
139 inhibitors were among the lowest measured (Table 1).
140

141 **Table 1. Structures, Ki Constants, and EC50 Values of the UMass Series of HIV-1 PIs**

Inhibitor	Structure	Ki (pM)			EC50 (nM)
		WT	I84V	I50V/A71V	WT
DRV		< 5.0ψ	25.6 ± 5.6[a]	74.5 ± 5.6[a]	7.7
UMASS-1	R2 R1	< 5.0ψ	26.1 ± 3.7[a]	110.3 ± 8.8[a]	5.9
UMASS-2		< 5.0	< 5.0	15.0 ± 2.7	2.4
UMASS-3		< 5.0	9.9 ± 2.7	79.9 ± 5.9	9.1
UMASS-4		< 5.0	10.5 ± 1.8	32.9 ± 3.0	3.2
UMASS-5		< 5.0	7.0 ± 1.7	7.8 ± 0.9	4.0
UMASS-6		< 5.0ψ	12.8 ± 3.1[a]	100.0 ± 9.9[a]	5.2
UMASS-7		< 5.0	12.1 ± 4.5	18.2 ± 3.0	3.1
UMASS-8		< 5.0	< 5.0	55.4 ± 4.0	4.2
UMASS-9		< 5.0	7.6 ± 1.6	42.3 ± 2.6	6.4
UMASS-10		< 5.0	14.3 ± 9.3	5.8 ± 1.1	4.1

142 [a]Previously reported in Lockbaum et al (ACS Infect Dis. 2019 Feb 8; 5(2): 316–325.)

143

144 **Selection for high level resistance during passage in cell culture follows two pathways**

145 To evaluate the potential of each inhibitor to select for mutations that would confer high-
146 level resistance and to compare these mutations across the inhibitor series, we grew HIV-1 under
147 conditions of escalating inhibitor concentration in cell culture through a lengthy period of
148 selection (50-95 passages). The selection experiments were performed under two separate starting
149 conditions: first when the starting virus was generated from the NL4-3 clone (this clone closely
150 approximates the clade B consensus for the protease amino acid sequence and will be referred to
151 as "wild type"), then again when the starting virus was a mixture of 26 isogenic viruses each with
152 a single mutation associated with drug resistance in the NL4-3 background. Notably, in the latter
153 case only about one-half of the mutations that appeared in the culture during selections were in
154 the mixture of starting mutations, indicating that even in the selections that were seeded with the
155 pool of single resistance mutations there was sufficient evolutionary capacity to explore mutations
156 at positions that started as the wild type sequence. The initial inhibitor concentration started at
157 low nanomolar concentrations and increased by a factor of 1.5 with each subsequent viral passage
158 (the inhibitor concentration was increased only after the virus spread efficiently through the
159 culture). All of the selections starting with wild type virus reached at least 5 μM of inhibitor
160 concentration (approximating the therapeutic concentration reached by DRV in the blood and in
161 the range of a 1,000 fold increase in the starting EC_{50}). For technical reasons, only 5 of the
162 selections starting with the mixture of mutants reached an inhibitor concentration as high as 400
163 nM and are reported here (Fig. S3).

164 Resistance mutations selected in the protease coding domain during the escalating
165 selective pressure of increasing PI concentration were examined at various time-points using a

166 next generation sequencing (NGS) protocol that included Primer ID with the MiSeq platform
167 (Zhou et al., 2015). We first examined the sequence of the most abundant variant in each of the
168 two selection schemes present at the highest inhibitor concentration reached (Fig. 1A). For those
169 selections that reached 5 μM, 8 to 14 mutations were present in the most abundant variant.
170 Typically the next two most abundant variants in the population differed by a single amino acid,
171 and on average the three most abundant variants in the 5 μM cultures accounted for 88% of the
172 total viral population (Fig. S4). Evolution of resistance followed two pathways, one based on
173 I50V and one based on I84V. In three cultures, V2wt, V5wt, and V2mut (which reached either 4
174 or 5 μM final drug concentration; V2wt indicates the virus pool selected with UMASS-2 starting
175 with wild type virus), these two mutations, I84V and I50V, became linked on the same genome
176 (Fig. 1A) showing they are not mutually exclusive.

177 In addition to the sequence of the terminally selected virus, we validated that selection
178 occurred over the entire course of the selection protocol. First, as can be seen in Fig. 1B, the
179 number of mutations present at 100 nM drug concentration was approximately half the number
180 present after selection to greater than 1 μM. In comparing the number of mutations present in
181 cultures using the I84V versus the I50V pathway, we noted that on average there were more
182 mutations in the I50V viruses at the intermediate drug concentration compared to the I84V viruses
183 (100 nM, P=0.03), and this trend continued at the higher drug concentration. Next we compared
184 the EC_{50} of five virus cultures representing the I84V or the I50V pathways (highlighted in yellow
185 in Fig. 1), and including one that had both mutations linked, testing the level of resistance to all
186 of the UMASS-1 through -10 inhibitors after selection to 100 nM drug concentration or to 5 μM.
187 As can be seen in Fig. 1C, resistance (measured as an increase in the EC_{50}) was apparent after

188 selection to the 100 nM inhibitor concentration, and in each case the EC_{50} increased an average
189 of 100-fold after selection to the final inhibitor concentration of 5 μM.

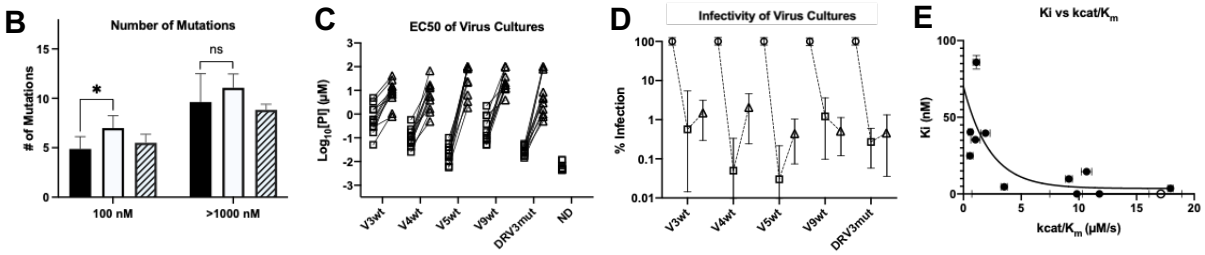
190 We were also interested in the fitness cost of these mutations. To measure this effect we
191 compared the infectivity of the virus in 10 culture supernatants (5 selections at each of two levels
192 of selection), normalizing infectivity on a reporter cell line for a given amount of the virion p24
193 CA protein in the culture supernatant. This was compared to the unmutated NL4-3 parent (given
194 a value of 100%). As can be seen in Fig. 1D, selection to 100 nM inhibitor concentration resulted
195 in a reduction in fitness/relative infectivity by between 100 and 1000-fold. Those cultures with
196 the biggest decrease in fitness rebounded to around 1% relative infectivity after selection to an
197 inhibitor level of 5 μM. This suggests that these high levels of resistance are linked to maintaining
198 a residual level of relative infectivity around 1% (as defined under these culture conditions).

199 We used another approach to explore the interplay between resistance and fitness; we
200 examined the effect of resistance on catalytic efficiency of the protease by measuring K_i and the
201 catalytic efficiency (kcat/Km) for a set of mutant proteases against the entire panel of UMASS
202 inhibitors and DRV. The average K_i value of each mutant protease (sequences shown in Table
203 S1) is plotted against the catalytic efficiency (kcat/K_m) for that enzyme in Fig. 1E. As can be seen,
204 there is a strong relationship between higher K_i values (i.e. resistance) and dramatic reductions in
205 catalytic efficiency, consistent with the loss of fitness associated with these viral populations.
206
207

A

Culture Name	Enzyme Name	Inhib. Conc. (nM)	R1	R2	PIs	Wild-Type Selection: Most Abundant Variants Detected at the Last Time Point	Abundance
V1wt	-	5000	1	1	UMASS-1	10I 16E 32I 33F 46I 54L 71V 76V 82I 84V	55.4%
V2wt	-	5000	1	2	UMASS-2	10F 33F 46I 47V 50V 71V 82I 84V	37.5%
V3wt	PR-3wt	5000	1	3	UMASS-3	10F 33F 46I 47V 50V 53L 63P 71V 76S 82I 85V 89I	85.7%
V4wt	PR-4wt	5000	1	4	UMASS-4	10F 11I 13V 32I 33F 43T 46L 54L 71V 82I 84V 89M 91S 92R	87.3%
V5wt	PR-5wt	5000	1	5	UMASS-5	10F 15V 46I 47V 50V 53L 71V 82I 84V 89T	67.9%
V6wt	-	5000	2	1	UMASS-6	10F 13V 33F 46I 47A 50V 53L 71V	45.1%
V7wt	-	5000	2	2	UMASS-7	10F 13V 33F 46I 47V 50V 53L 54L 71V 82I	94.5%
V8wt	-	5000	2	3	UMASS-8	10F 12K 33F 46I 47V 50V 53L 54L 63P 71V 82I 85V	55.4%
V9wt	PR-9wt	5000	2	4	UMASS-9	10F 13V 33F 45R 46I 47V 50V 53L 54L 66F 71V 74A 76S 82I	92.3%
V10wt	-	5000	2	5	UMASS-10	10F 13V 33F 43T 46I 47V 50V 54L 71V 82L 88S	91.2%
DRVwt	-	1350	-	1	DRV	32I 41I 46I 47V 82I 84V 85V	62.8%
NDwt	PR-WT	0			ND		WT 82.9%

Culture Name		Inhib. Conc. (nM)	R1	R2	PIs	Mutant Selection: Most Abundant Variants Detected at the Last Time Point	Abundance
V2mut	PR-2mut	4000	1	2	UMASS-2	10I 32I 33F 45I 46I 50V 71V 82I 84V	34.8%
V4mut	-	1275	1	4	UMASS-4	10F 28S 46I 84V	9.4%
V6mut	-	5000	2	1	UMASS-6	13V 16E 32I 33F 45I 46I 54L 71V 76V 82F 84V	62.8%
V7mut	-	400	2	2	UMASS-7	10F 46I 50V 63P 85V	21.0%
V8mut	PR-8mut	1300	2	3	UMASS-8	10F 46I 47V 50V 53L 63P 72V 73S 82I 85V	70.0%
DRV3mut	PR-Dmut	5000	-	1	DRV-3	10F 32I 33F 37D 46I 63P 74S 76V 84V 89M	51.1%
NDmut	-	0			ND	93L	36.7%



208

209 **Fig. 1. Features of Viruses Selected To High Level Resistance Against DRV and the UMASS-**
210 **1 Through -10 Inhibitor Series.** Infected cultures were put under drug selection starting with
211 either wild type virus or a pool of mutant viruses. **A)** Most abundant variants present at final
212 timepoint. Resistance mutations in the viruses selected starting with the wild type virus are in
213 blue, and the selection starting with the pool of mutant viruses in in red. The culture name
214 indicates a virus pool (V), the UMASS PI number (V1) and whether the selection started with
215 wild type or mutants (V1wt). PR sequences that were used to generate purified recombinant
216 protease are indicated (e.g. PR-3wt). The final PI concentration reached and the R1 and R2
217 moieties are indicated along with the PI. **B)** Number of mutations present in the 3 most abundant
218 variants of each selection at two different PI concentrations. Cultures containing I84V resistance
219 mutation are in black (n=4), I50V-containing cultures are in white (n=7), and cultures with
220 I50V+I84V linked are shown with hatched box (n=3). Unpaired t-test was used to compare the

221 number of mutations in the 84V vs the 50V cultures. **C)** EC_{50} values for highlighted selections in
222 panel A reaching 100 nM (square) and 5000 nM (triangle). **D)** Relative infectivity values for the
223 same selections shown relative to the wild type virus (circle). Relative infectivity was measured
224 with normalized input amounts of HIV-1 p24 CA protein. **E)** Enzyme inhibition constants (K_i) of
225 end-point PR variants versus catalytic efficiency ($kcat/K_m$). Open circle represents wildtype NL4-
226 3. Trendline is for visualization purposes only. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
227
228
229 **Features of the selection process assessed by next generation sequencing (NGS) of**
230 **longitudinal samples**
231 Each of the viral cultures started with wild type virus showed an accumulation of protease
232 mutations with increasing selective pressure. NGS analysis revealed very few fixed variants in
233 the cultures started with wild type virus until the inhibitor concentration reached approximately
234 3-5 nM (approaching the EC_{50} value for the wild type virus, Table 1); in contrast the cultures that
235 were started with the mutant library selected for the outgrowth of a subset of those mutants by 1
236 nM inhibitor concentration (Fig. 2A). In all cultures, the transition through the EC_{50} concentration
237 provided significant selective pressure in fixing mutations. Multiple unfixed mutations were
238 observed in each culture after the drug concentration exceeded the EC_{50} values above 3 nM,
239 highlighting the high genetic diversity in the culture. Additional mutations became linked on each
240 viral genome at higher drug concentrations.
241 Deep sequencing revealed that mutations accumulated in complex patterns. We assessed
242 the sequence complexity of each culture by calculating the Shannon Entropy to allow comparison
243 of changes in the diversity in each culture over time and as a function of increasing selective

244 pressure (Fig. 2A). In the cultures that showed the early appearance of the I84V mutation, this
245 was associated with a peak in entropy, reflecting high genetic diversity, followed by a decrease
246 in entropy when the I84V mutation became fixed (Fig. S5A). The introduction of the I50V
247 mutation was generally not associated with a drop in entropy, rather these populations maintained
248 high genetic diversity even at higher drug concentrations (Fig. S5B). We interpret these patterns
249 as indicative of I84V conferring some level of resistance without a dramatic loss in fitness,
250 allowing a more homogeneous culture (i.e. less entropy). In contrast, I50V may confer a higher
251 level of resistance but at a greater fitness cost thus supporting greater diversity in the culture either
252 as compensatory mutations or as other combinations of mutations with lesser resistance but higher
253 fitness. In this regard, we previously showed I50V significantly reduces the fitness of the virus
254 relative to the fitness loss of a virus with I84V as single mutations (Henderson et al., 2012).
255 When we examined the sequence diversity as assessed by Shannon Entropy for all
256 selections that reached at least 1 μM in inhibitor concentration, we found that cultures starting
257 with the mixture of resistant viruses averaged a nearly two-fold higher entropy value compared
258 to the cultures where the selection started with just the virus generated from the NL4-3/wild type
259 clone (3.0 vs 1.6, $P<0.0001$ Mann-Whitney test; Fig. 2B, C, D). This was unexpected, as both
260 sets of selections passed through many genetic bottlenecks. This result is most easily explained if
261 the rates of recombination were fairly high throughout the culture period so that most sequence
262 variants were maintained at least at a low level.
263

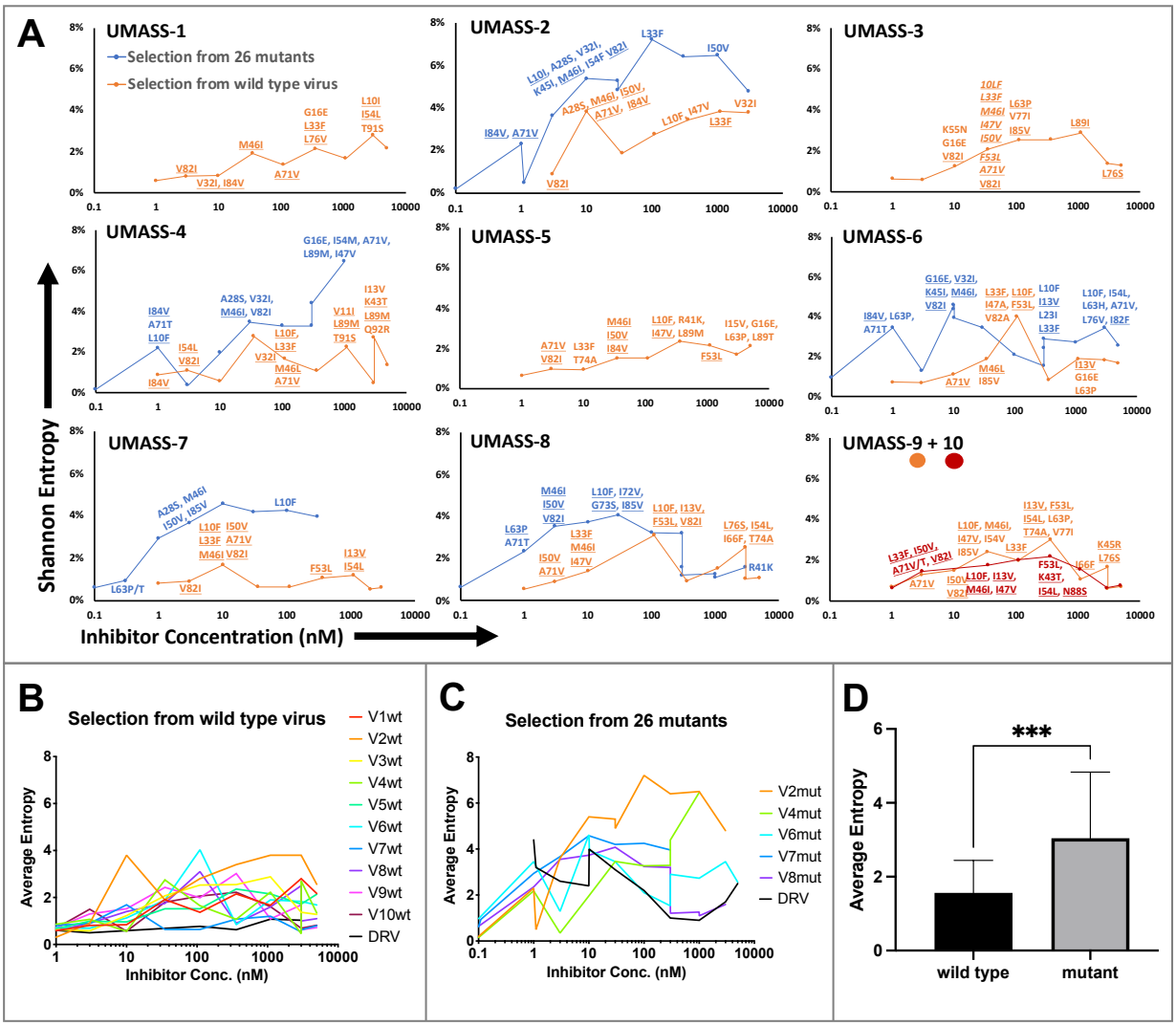


Fig. 2. Evaluation of Sequence Diversity In the Viral Population During Increasing Selective Pressure With PI Inhibitors. A) Shannon entropy was calculated using sequence diversity in the protease region. Abundant mutations are shown and fixed mutations are underlined. Selections from wild type are in orange (n=6) and from the mutant pool (n=11) in blue. Two selections from wild type virus are shown in orange (UMASS-9) and red (UMASS-10). **B)** Compiled Shannon entropy values from selections starting with wild type virus that reached concentrations of >1000 nM. **C)** Compiled Shannon entropy values from selections starting with the pool of mutant viruses that reached concentrations of at least 400 nM. **D)** Selections starting

273 with the pool of 26 mutants showed higher entropy when averaged over all analyzed drug
274 concentrations ($p < 0.001$ using the unpaired t-test).
275
276 **The chemical nature of R1, but not R2, determines the resistance pathway**
277 We next examined if the inhibitor structure influenced the resistance pathway chosen. We
278 found that the R1 group, i.e. the (S)-2-methylbutyl (R1-1) or the 2-ethyl-n-butyl (R1-2), largely
279 defined the resistance pathway observed; for these inhibitors the R1 group takes the position of
280 P1' in the protease substrate analog, occupying the S1' subsite. With the UMASS-1 through -5
281 series (the smaller R1-1 group), the I84V mutation appeared first in 6 of 7 cultures. In contrast,
282 the UMASS-6 through -10 series with the larger R1-2 group, the I50V mutation appeared first in
283 8 of 9 cultures ($P=0.009$, Fisher's Exact Test; in this analysis we included two cultures that did
284 not reach at least 400 nM inhibitor concentration but did fix an initial set of mutations to increase
285 our sample size, and we did not include two cultures where both I84V and I50V were initially
286 fixed together [data not shown]). We considered the possibility that the starting mixture of mutant
287 viruses in the first selection might skew the pathway selected, especially since the mutant pool
288 included I84V but not I50V. However, in only 1 of the 8 cultures with sufficient data from both
289 selections was there a switch from the I84V pathway to the I50V pathway between the first and
290 second selections (cultures of UMASS-6 with an R1-2 group). Thus, we conclude that the P1'
291 side chain of the inhibitor is a strong determinant of the resistance pathway selected. It is notable
292 that one inhibitor could select for different pathways in two separate selections (also seen with
293 DRV, see below) even when both major mutations are maintained in the viral population (Fig.
294 S6). This suggests that either pathway can provide some level of resistance to most if not all of

295 these inhibitors, and that the chance addition of the initial compensatory mutations may determine
296 which pathway becomes the major resistant population.

297 To examine the potential for linked mutations and to infer the order in which mutations
298 accumulated in the protease gene to confer high level resistance, the abundance data from multiple
299 selections that ended in one or the other pathway were pooled and compared. In this analysis,
300 shown in Fig. 3, summary data for the selections resulting in the I84V pathway point up, with
301 I84V reaching 100% penetrance by definition. Similarly, the summary data for those selections
302 that fixed I50V are shown pointing down, with I50V reaching 100% penetrance. Several
303 mutations are uniquely linked or at least strongly favored in each pathway, with I84V being linked
304 to V32I, and I50V being linked to I47V, F53L, and I13V. A number of mutations appear in both
305 pathways, although not with equal frequency: L10F, L33F, M46I, I54L, A71V, and V82I. Finally,
306 other mutations appear less frequently making it difficult to assign them to one of these categories.
307 These results show that while some mutations are largely linked to one pathway, other mutations
308 are often shared between the two pathways. Also, the variation in frequency of appearance of
309 shared mutations in the two pathways suggests different levels of impact on resistance and/or
310 fitness in the I84V vs the I50V background for these mutations (e.g. L10F, M46I and A71V). In
311 particular, the larger impact of I50V on fitness compared to I84V mutations (Henderson et al.,
312 2012) is consistent with the earlier appearance and greater levels of fixation of the shared
313 compensatory mutations L10F, L33F, M46I and A71V in the cultures that followed the I50V
314 pathway. We next considered the possibility that the R2 constituent would provide additional
315 selective pressure in the form of additional resistance mutations; however, similar analyses did
316 not reveal a significant pattern of sequence selection based on the potency associated with the R2
317 group (data not shown).

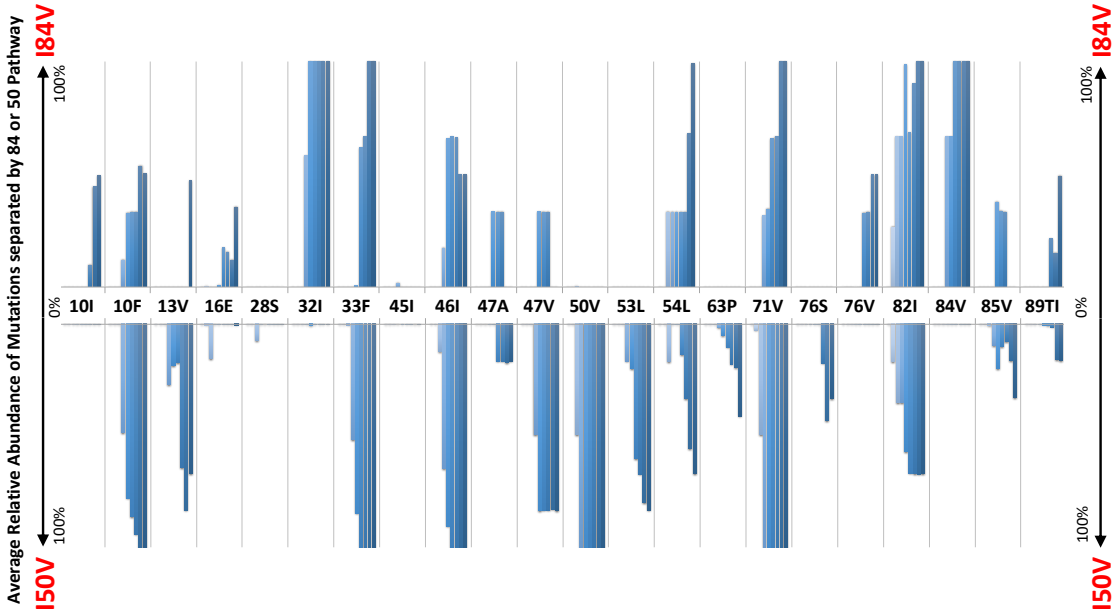


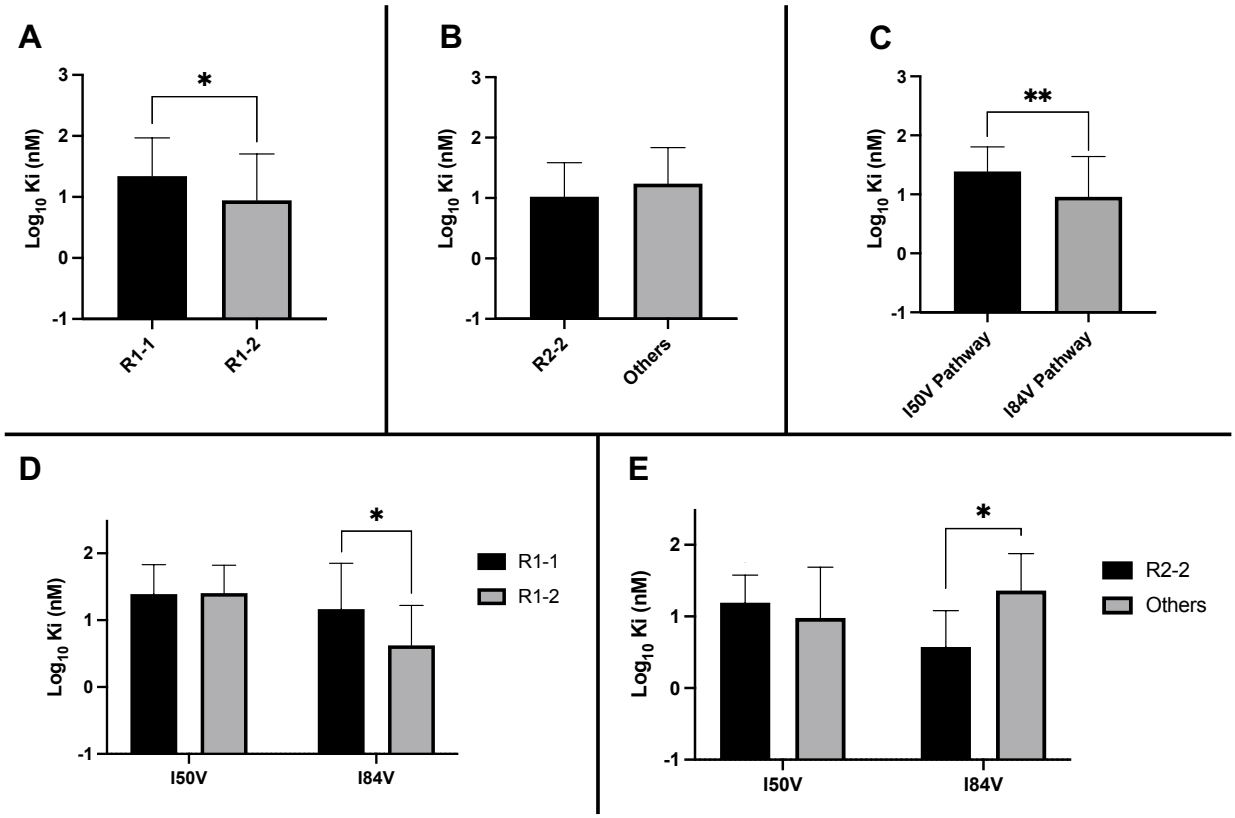
Fig. 3. Patterns of Accumulation of Amino Acid Substitutions Associated with Resistance During Selection. Abundance data from selections for each indicated amino acid substitution were pooled and examined sequentially at different levels of drug concentration. Selections that reached the maximum concentration of at least 3000 nM were assessed longitudinally to allow for all timepoints to be averaged equally. Mutations from selections resulting in the I84V pathway (V1wt, V4wt, V6mut, and DRV3mut) point up, with I84V reaching 100% penetrance by definition. Similarly, those selections that fixed I50V only (V3wt, V6-10wt) are shown pointing downward, with I50V reaching 100% penetrance.

The chemical nature of R1 and R2 determine residual potency among the resistant variants

We considered several variables in examining the nature of the interaction between inhibitors and the resistant proteases: the extent to which the inhibitor structure (either R1 or R2) affected residual potency; and the extent to which the pathway (I84V or I50V) conferred the

333 greatest resistance. As shown in Fig. 4A, the inhibitors with the smaller R1-1 group showed a
334 trend to be less potent than the inhibitors with the larger R1-2 group when tested against all of the
335 resistant proteases. Inhibitors with the R2-2 group showed a trend toward greater residual potency
336 against the mutant enzymes compared to the other inhibitors (Fig. 4B). In looking at the individual
337 pathways, the highly resistant enzymes with the I84V mutation remained more sensitive to the
338 entire group of inhibitors than the enzymes with the I50V mutation (Fig. 4C).

339 We next linked the two pathways to the specific structural features of the inhibitors. The
340 resistant proteases with I50V had similar K_i values to both the R1-1 and R1-2 inhibitors, while
341 the proteases with I84V were more sensitive (lower K_i) to the larger R1-2 inhibitors (Fig. 4D).
342 Similarly, the increased potency of the R2-2 inhibitors over the rest of the inhibitors was seen
343 against the enzymes carrying I84V but not those with I50V (Fig. 4E). These results are consistent
344 with the smaller R1-1 inhibitors selecting for the I84V pathway and the larger R1-2 inhibitors
345 selecting for the I50V pathway, and with the R2-2 inhibitors adding additional potency that is
346 retained even after selection for resistance, most apparent with the I84V pathway.
347



348

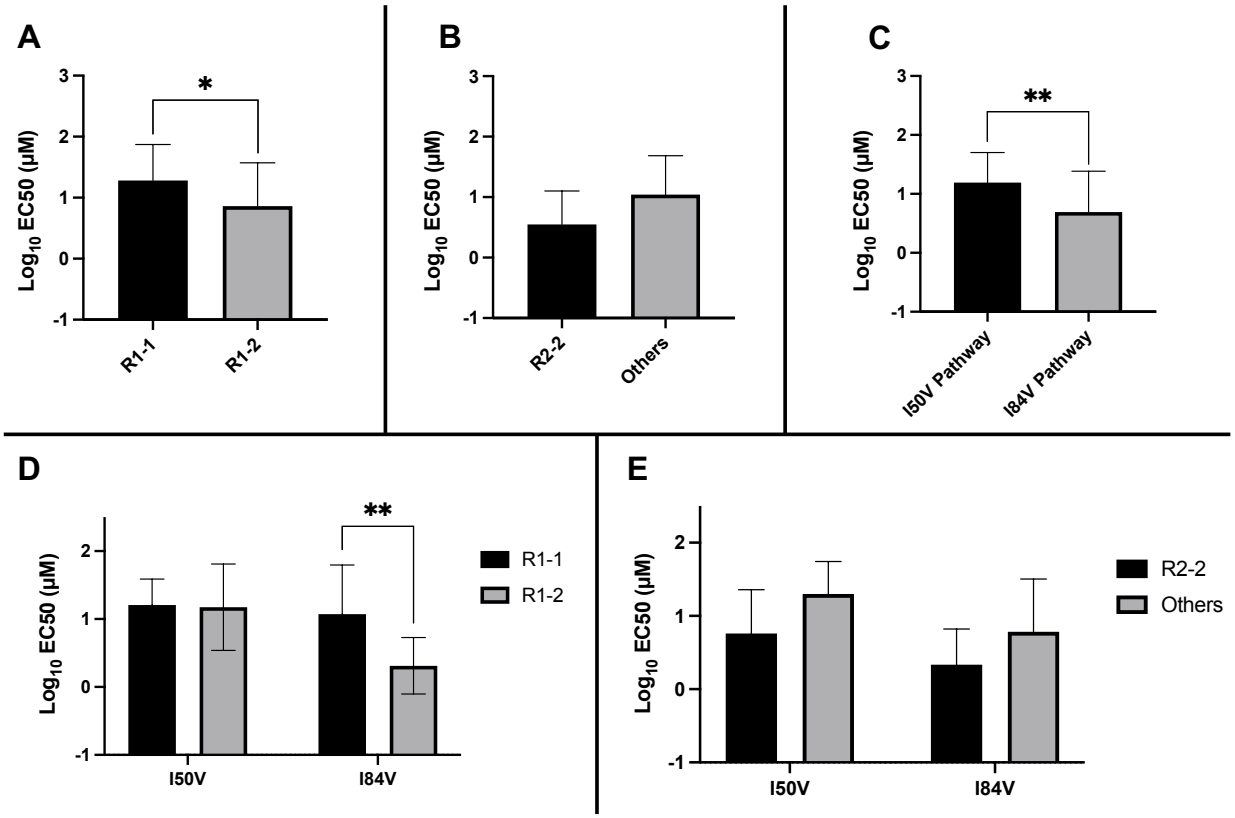
349 **Fig. 4. Analysis of Ki Values For Mutant Enzymes.** Ki values were determined against all
350 UMASS inhibitors (Table 1). Brackets above the bars represent significant p-values between the
351 two groups using the unpaired t-test. Data were pooled in different ways for the following
352 analyses: **A)** The Ki values against the inhibitors with the larger R1-2 moiety (n=34) were more
353 potent against resistant proteases compared to the inhibitors with the smaller R1-1 moiety (n=35)
354 **B)** Ki values for the inhibitors with the R2-2 group (n=10) showed a trend toward being more
355 potent against the highly mutated proteases compared to the other R2 groups (n=40). **C)** Ki values
356 for enzymes with the I50V mutation (n=33) showed greater resistance to the inhibitors compared
357 to enzymes with the I84V mutation (n=22). **D)** The R1-2 moiety provided increased potency to
358 enzymes with the I84V mutation (n=20) but not the I50V mutation (n=20). **E)** The R2-2 moiety
359 was more potent against the enzymes with I84V mutation (n=10) compared to the enzymes with
360 the I50V mutation (n=10). The unpaired t-test was used to assess differences in K_i values.

361

362

363 Residual potency dependence of inhibitor structure could also be seen in the EC_{50} values
364 of the resistant virus pools. When the EC_{50} values against the virus pools using the I84V or the
365 I50V pathway were compared for all of the inhibitors, the inhibitors with the larger R1-2 group
366 were more potent, retaining on average lower EC_{50} values compared to the EC_{50} values for the
367 inhibitors with the smaller R1-1 group (Fig. 5A). Inhibitors with the R2-2 group showed a trend
368 toward greater potency compared to the other R2 groups (Fig. 5B), consistent with what was
369 observed in enzymatic assays. When we considered the effects based on the selection pathway
370 we observed that viruses in the I50V pathway had a higher level of resistance to these inhibitors
371 than the viruses in the I84V pathway (Fig. 5C). The higher level of resistance for the viruses using
372 the I50V pathway was due to the fact that these viruses were similarly resistant to the inhibitors
373 with either the R1-1 or the R1-2 groups; in contrast, the viruses using the I84V pathway conferred
374 a greater level of resistance to the inhibitors with the smaller R1-1 group while the inhibitors with
375 the larger R1-2 group retained a higher level of potency (Fig. 5D). Finally, there was a trend for
376 viruses in either pathway to be less resistant to the inhibitors with the R2-2 group (Fig. 5E).

377



378

379 **Fig. 5. Analysis of EC_{50} Values For Mutant Virus Cultures.** EC_{50} values were determined for
380 a subset of the selected virus cultures against a panel of inhibitors (n=50). For **(A-E),** the EC_{50}
381 data were pooled using the same methods as in Fig. 4. The highest level of resistance recorded
382 was 100 μM. The Mann-Whitney rank-sum test was used to assess differences in EC_{50} values.

383

384 **DRV favors the I84V pathway**

385 DRV has a butyl group at the R1 position, smaller than the R1-1 methylbutyl in the
386 UMASS-1 through -5 series. We carried out five selections with DRV where the final inhibitor
387 concentration reached greater than 1 μM. One DRV selection was carried out in parallel with each
388 of the two different selections with the UMASS inhibitors (Fig. 1A); both of these selections
389 resulted in the appearance of the I84V mutation as defining the resistance pathway. Three
390 additional selections were done in parallel with DRV and starting with the mixture of the 26

391 isogenic mutants; all three selections reached the level of 5 μM as the highest drug concentration.
392 Two of these selections used the I84V pathway to resistance, while one of the selections used the
393 I50V pathway to resistance (data not shown). Thus four of five independent selection with DRV
394 favored the resistance pathway I84V, associated with a smaller R1 constituent. These results add
395 further evidence for the size of the R1 moiety strongly influencing the resistance pathway.
396

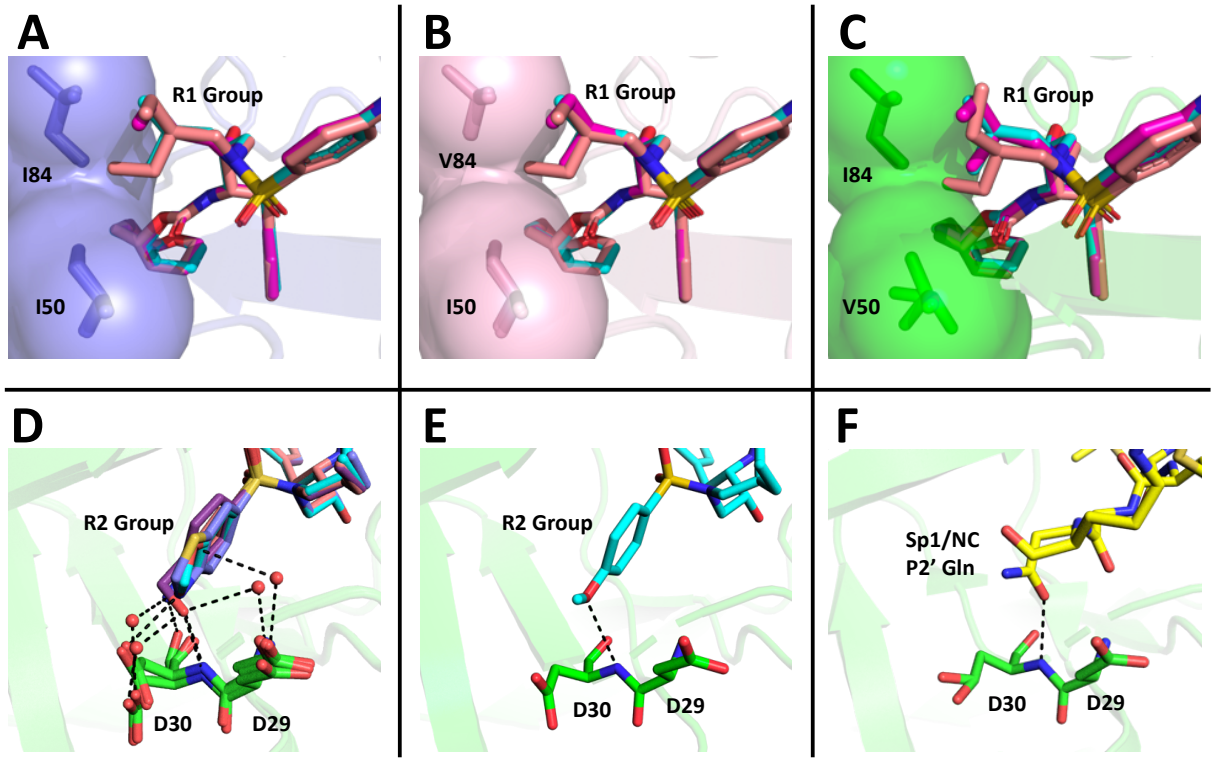
397 **Discussion/Structural Interpretation**

398 DRV is a notable PI both for its tight binding to the HIV-1 protease and for its ability,
399 with boosting doses of ritonavir, to reach micromolar levels of drug concentration in the blood.
400 These features are important as resistance to PIs typically require multiple mutations affecting
401 both resistance and the rescue of enzyme activity/fitness lost with primary resistance mutations
402 (Arribas et al., 2012; Clemente et al., 2004; Henes et al., 2019; Mahalingam et al., 2002;
403 Muzammil et al., 2003; Ragland et al., 2014; Wensing et al., 2010). While we were successful in
404 selecting for resistance to DRV in cell culture, the spectrum of mutations observed largely overlap
405 those seen in previous studies based on testing individual mutations for their effect on DRV
406 sensitivity (de Meyer et al., 2008; King et al., 2004; Rhee et al., 2003). In this work, we have
407 explored chemical modifications to the DRV scaffold. We found that modification of the P1'/R1
408 chemical moiety to be a larger aliphatic group favored the use of the I50V resistance pathway
409 compared to the smaller DRV moiety or an intermediate-sized moiety, forcing the virus to use
410 the more deleterious I50V mutation compared to the I84V mutation used with the smaller P1'
411 moieties. Conversely, we identified a subset of the P2'/R2 groups that contributed to residual
412 potency even in the resistant proteases but did not select for additional resistance mutations. Thus
413 chemical modifications at both of these inhibitor positions yielded improvements over DRV.

414 **P1'/R1 Modification.** DRV has a relatively small butyl group at this position that most
415 often selected for an I84V mutation. When we extended this moiety by a single methyl group to
416 (S)-2-methylbutyl we obtained a similar resistance selection profile. However, by extending it
417 with an additional methyl group to (S)-2-methylbutyl or 2-ethyl-n-butyl created an inhibitor that
418 switched the preferred selection to the I50V pathway. In Fig. 6A-C are representations of
419 structures of a subset of the inhibitors with the different P1'/R1 groups with the wild type protease,
420 the I84V mutant protease, and the I50V mutant protease. As we have previously discussed
421 (Lockbaum et al., 2019), in these structures it is apparent that the P1'/R1-1 group is directed at
422 the protease I84 sidechain, consistent with the shortening of this sidechain as a major resistance
423 mechanism. Conversely, the longer P1'/R1-2 group is oriented into the space between I84 and
424 I50, thus drawing in I50V as the pathway to resistance. As noted in the longitudinal analysis of
425 the selection pathway, the I50V resistance pathway presents additional challenges to the virus in
426 the more rapid accumulation of compensatory mutations (Fig. 3). Under the circumstances of
427 rapidly declining viral load during the initiation of therapy this need for additional mutations
428 would represent an enhanced genetic barrier.

429 **PR2'/R2 Modification.** The inhibitors with two of the R2 modifications (R2-2 and R2-5)
430 showed greater potency against the mildly resistant proteases (Fig. S2) and against the wild type
431 virus (Table 1). This effect was also evident as a trend for inhibitors with the R2-2 group against
432 the highly resistant proteases (Fig. 4E) and the highly resistant viruses (Fig. 5E). When we
433 examined the interactions between protease and inhibitor there is a clear structural explanation
434 (Fig. 6D-F). The less potent inhibitors interact with the protease side chains and backbone at D29
435 and D30 through a network of water molecules. In contrast, with the more potent R2-2 group the
436 water molecules are replaced with direct interactions between inhibitor and protease in the S2'

437 subsite (Fig. 6E). This direct interaction more closely mimics the interaction made by the
438 glutamine P2' side chain in the optimized protease cleavage site at the SP1/NC boundary (Fig.
439 6F). Thus the ability to replace the water-mediated interactions with a more rigid framework that
440 is interacting with the protease backbone deep in the S2' pocket appears to be a unique feature of
441 the inhibitors with R2-2. Improved potency in another series of PIs was also reported for this R2-
442 2 P2' moiety (Zhu et al., 2020).
443



444

445 **Fig. 6. Structural Interpretation Of Protease Inhibitor Resistance and Of Residual Inhibitor**
446 **Potency. (A-C)** Hydrophobic packing in the S1' subsite in complex with R1 structural groups;
447 darunavir is shown in cyan, UMASS-1 in magenta, and UMASS-6 in salmon.**A)** The two forms
448 of R1 and DRV are shown packing against wild type protease at I84 and I50 (DRV, UMASS-1,
449 UMASS-6 in wild type protease [PDB: 6DGX (cyan), 6DGY (magenta), 6DGZ (salmon)]. **B)**

450 Those same inhibitors packing against the I84V mutant [PDB: 6DH0 (cyan), 6DH1 (magenta),
451 6DH2 (salmon)]. **C)** Those same inhibitors packing against the I50V protease variant [PDB: ,
452 6DH6 (salmon), 6DH7 (magneta), 6DH8 (salmon)]. **(D-F)** Binding interactions with Asp29 and
453 Asp30 of the protease S2' subsite in complex with R2 structural groups. Inhibitor/enzyme
454 interactions are shown with black dashed lines representing hydrogen bonds (<3.0 Å). **D)**
455 UMASS-1, UMASS-3, UMASS-4, and UMASS-5 in WT protease [PDB: 3O99 (salmon), 3O9B
456 (purple), 309C (orange), 309D (slate blue)]. Water molecules are shown as red dots. **E**) UMASS-
457 2 in WT protease (PDB: 3O9A). **F)** SP1/NC S2' subsite representation of peptide substrate
458 complexed with an inactive form of wild type protease (PDB: 1KJ7) showing an interaction
459 between the P2' glutamine and the protein backbone at Asp29 and Asp30.
460
461 **Protease Mutation Networks In the Resistant Pathways.** In most of the selections
462 either the I84V or the I50V mutation was largely fixed by the time the inhibitor concentration
463 reached 10 nM in the culture (Fig. 2, Fig. S6). The viral cultures went through cyclical changes
464 in population diversity as subsequent mutations were added (Fig. 2A). Because of the large
465 number of selections it was possible to pool data and see trends in the way mutations were added
466 in these two pathways. Resistance mutations that were detected in multiple selections summed to
467 at least 16 positions (Fig. 1A). While these mutations are well known, their relationships to each
468 other, and as a function of selective pressure, are less well understood. The data in Fig. 3 show
469 strong linkage between I84V and V32I, while I50V shows linkage with I47V. Other mutations
470 appeared often in both pathways but could be found earlier and/or more fequently in one pathway:
471 in the I84V pathway I54L and V82I were favored; in the I50V pathway L10F, I13V, L33F, M46I,
472 L63P, and A71V were favored. While the role of some of these changes may be in restoring

473 fitness (Wensing et al., 2010), with a bigger fitness loss due to I50V (Henderson et al., 2012), we
474 have demonstrate in previous studies that the mutations outside the active site have play a key
475 role in conferring high levels of resistance (Henes et al., 2019; Ragland et al., 2017; Matthew et
476 al., 2021; Leidner et al., 2021).

477 In summary, using a number of lengthy selections for resistance to a series of structurally
478 related PIs we have been able to identify the size of the hydrophobic side chain at the P1' position
479 of the inhibitor as the major determinant for selection of the I84V pathway with a smaller P1' side
480 chain, including for DRV, versus a larger side chain driving selection of the I50V pathway. Also,
481 one of the P2' groups tested gave a higher level of potency even in the face of high level resistance
482 by creating a direct interaction with the protease backbone. These changes increase the genetic
483 barrier for the evolution of resistance and emphasize the potential utility of what a fifth generation
484 HIV-1 PI could add to regimens with reduced drug complexity.
485

486 **Materials and Methods**

487 **Cell lines and viruses**

488 CEMx174 cells were maintained in RPMI 1640 medium with 10% fetal calf serum and penicillin-
489 streptomycin. TZM-bl and 293T cells were maintained in Dulbecco's modified Eagle-H medium
490 supplemented with 10% fetal calf serum and penicillin-streptomycin. A wild-type virus stock
491 NL4-3 was prepared by transfection of the pNL4-3 plasmid (purified using the Qiagen Plasmid
492 Maxikit) into HeLa cells. For the mixture of isogenic mutant viruses the following NL4-3 variants
493 were created, each with a single mutation in the protease with this mixture forming the virus pool
494 for the initiation of selection with mutant viruses: L10I, K20R, K20I, L24I, D30N, V32I, M36I,
495 M46I, M46L, I47V, G48V, F53L, I54V, I62V, L63P, A71T, A71V, G73S, V77I, V82A, V82T,
496 I84V, N88D, N88S, L90M, I93L (Henderson et al., 2012).

497

498 **Selections**

499 An aliquot of 3 x 10^6 CEMx174 cells was incubated at 37°C for 2 to 3 h with 250 µl of a virus
500 stock generated from the HIV-1 infectious molecular clone pNL4-3. The culture volume was then
501 brought to 10 ml with RPMI medium. Each flask received one of the following inhibitors at
502 escalating concentrations: UMass1, UMass2, UMass3, UMass4, UMass5, UMass6, UMass7,
503 UMass8, UMass9, UMass10, DRV and no drug (ND). After 48 h and every 48 h after, the cells
504 were pelleted by centrifugation and 10ml of fresh medium and inhibitors were added. When the
505 culture had undergone extensive cytopathic effect (CPE) indicative of viral replication, the
506 supernatant medium and the cells were harvested separately and stored at -80°C. The virus-
507 containing supernatant was used to start the next round of infection, and after several rounds at
508 the initial concentration, the inhibitor concentration was increased 1.5-fold at each subsequent

509 round of virus passage. The level of resistance (50% inhibitory concentration [EC50]) of the
510 single inhibitor-selected virus pools was determined by a TZM infection assay in which the PI is
511 added to productively infected cells and the titers of supernatant virus made in the presence of the
512 inhibitor are determined.
513
514 **TZM Infection Assay**
515 PI dilutions were prepared by taking 10 µM stocked and performing a 5-fold serial dilution using
516 RPMI media (final drug concentration is 100 µM). One dilution of drug was added to each well
517 of a 24 well plate and repeated so each virus would have a full set of dilutions. Viruses for the
518 assay were made by seeding 3 x 10^6 CEM cells in a 24 well plate and incubating with 250 µl of
519 virus at 37°C for 2 to 3 h before bringing the culture to 10ml with RPMI media. After 48 h the
520 medium was changed and repeated every 48 h after until the culture had undergone CPE. Infected
521 CEM cells were collected and diluted so that 1ml of cells could be plated in each well containing
522 a unique drug dilution. Then 24 hours later the virus supernatant was collected from each well
523 followed by filtering through a 0.45 µM filter then placed in -80°C. Viruses were thawed and
524 added to 96 well plates in triplicate. TZM-bl cells were collected and diluted to a concentration
525 of $2x10^5$ cells/ml, 100 µl were added on top of the pre-plated viruses. Plates were kept in 37°C,
526 5% CO_2 in an incubator for 48 hours. After 48 hours, the cells in the plates were lysed by removing
527 the medium, washing two times with 100 µl PBS and then lysed with 1x lysis buffer (made from
528 5x Promega Firefly Lysis Buffer). Plates were frozen for at least 24 hours and then thawed for 2
529 hours before analyzing with Promega Firefly Luciferase Kit on a luminometer.
530
531 **DNA preparation and amplification of the protease-coding region**

532 Total cellular DNA was isolated from infected cell pellets by using the QIAamp blood kit
533 (Qiagen). The protease-coding domain of viral DNA was amplified by nested PCR. The PCR
534 conditions are available on request. PCR products were purified by using QIAquick PCR
535 purification kit (Qiagen) and directly sequenced or cloned into the pT7Blue vector (Novagen) and
536 sequenced.
537
538 **Primer-ID Deep Sequencing of viral RNA**
539 We used the PID protocol to prepare MiSeq PID libraries with multiplexed primers. Viral RNA
540 was extracted from plasma samples using the QIAamp viral RNA mini kit (Qiagen, Hilden,
541 Germany). Complementary DNA (cDNA) was synthesized using a cDNA primer mixture
542 targeting protease (PR) with a block of random nucleotides in each cDNA primer serving as the
543 PID, and SuperScript III RT (ThermoFisher). After 2 rounds of bead purification of the cDNA,
544 we amplified the cDNA using a mixture of a forward primer that targeted the upstream coding
545 region, followed by a second round of PCR to incorporate the Illumina adaptor sequences. Gel-
546 purified libraries were pooled and sequenced using the MiSeq 300 base paired-end sequencing
547 protocol (Illumina). The sequencing covered the HIV-1 PR region (HXB2 2648–2914, 3001–
548 3257).
549 We used the Illumina bcl2fastq pipeline for the initial processing and constructed template
550 consensus sequences (TCSs) with TCS pipeline version 1.33
551 (https://github.com/SwanstromLab/PID). We then aligned TCSs to an HXB2 reference to remove
552 sequences not at the targeted region or that had large deletions.
553
554 **Protease expression and purification**

555 The highly mutated, resistant, protease variant genes were purchased on a pET11a plasmid with
556 codon optimization for protein expression in *E. coli* (Genewiz). A Q7K mutation was included to
557 prevent autoproteolysis (Rose et al., 1993). The expression, isolation, and purification of WT and
558 mutant HIV-1 proteases used for enzymatic assays were carried out as previously described
559 (Henes et al., 2019; King et al., 2002; Ozen et al., 2014). Briefly, the gene encoding the desired
560 HIV-1 protease was subcloned into the heat-inducible pXC35 expression vector (ATCC) and
561 transformed into *E. coli* TAP-106 cells. Cells grown in 6 L of Terrific Broth were lysed with a
562 cell disruptor twice, and the protein was purified from inclusion bodies (Hui et al., 1993).
563 Inclusion bodies, isolated as a pellet after centrifugation, were dissolved in 50% acetic acid
564 followed by another round of centrifugation at 19,000 rpm for 30 minutes to remove insoluble
565 impurities. Size exclusion chromatography was carried out on a 2.1-L Sephadex G-75 Superfine
566 (Sigma Chemical) column equilibrated with 50% acetic acid to separate high molecular weight
567 proteins from the desired protease. Pure fractions of HIV-1 protease were refolded using a 10-
568 fold dilution of refolding buffer (0.05 M sodium acetate at pH 5.5, 5% ethylene glycol, 10%
569 glycerol, and 5 μM DTT). Folded protein was concentrated to 0.5–3 mg/ml and stored. The stored
570 protease was used in K_M and K_i assays.
571
572 **Enzymatic assays**
573 ***K_m Assay.*** K_m values were determined as previously described (Henes et al., 2019; Lockbaum et
574 al., 2019; Matayoshi et al., 1990; Windsor and Raines, 2015). Briefly, a 10-amino acid substrate
575 containing the natural MA/CA cleavage site with an EDANS/DABCYL FRET pair was dissolved
576 in 8% DMSO at 40 nM and 6% DMSO at 30 nM. The 30 nM substrate was 4/5 serially diluted
577 from 30 nM to 6 nM. HIV-1 protease was diluted to 120 nM and, and 5 μl were added to the 96-

578 well plate to obtain a final concentration of 10 nM. Fluorescence was observed using a
579 PerkinElmer Envision plate reader with an excitation at 340 nm and emission at 492 nm, and
580 monitored for 200 counts. A FRET inner filter effect correction was applied as previously
581 described (Liu et al., 1999). Data corrected for the inner filter effect was analyzed with Prism7.
582
583 ***K_i Assay.*** Enzyme inhibition constants (K_i values) were determined as previously described
584 (Henes et al., 2019; Lockbaum et al., 2019; Matayoshi et al., 1990; Windsor and Raines, 2015).
585 Briefly, in a 96-well plate, inhibitors were serially diluted down from 2000-10,000 nM depending
586 on protease resistance. All samples were incubated with 5 nM protein for 1 hour. A 10-amino
587 acid substrate containing an optimized protease cleavage site (Windsor and Raines, 2015),
588 purchased from Bachem, with an EDANS/DABCYL FRET pair was dissolved in 4% DMSO at
589 120 μM. Using a PerkinElmer Envision plate reader, 5 μl of the 120 μM substrate were added to
590 the 96-well plate to a final concentration of 10 μM. Fluorescence was observed with an excitation
591 at 340 nm and emission at 492 nm and monitored for 200 counts. Data was analyzed with Prism7.
592
593 **Data will be submitted to the NIH short read archive.**
594
595
596
597
598
599
600

601 **Acknowledgements**

602 This work was supported by grants from NIGMS (P01-GM109767, R01-GM135919) and NIAID
603 (R01-AI140970) of the NIH. This research received infrastructure support from the UNC CFAR
604 (P30-AI050410), and the UNC Lineberger Comprehensive Cancer Center (P30-CA016086). We
605 also acknowledge receipt of reagents from the NIH HIV Reagent Program: TZM-bl cells (J.
606 Kappes and X. Wu contributors) and CEMx174 cells (P. Cresswell contributor). The support of
607 the UNC High Throughput Sequencing Facility is also acknowledged.

608

609 **Author Contributions**

610 ES, S-KL, SZ, GJL, MH, EAN, AS, and KK carried out the experiments. ES, GJL, AA, NKY,
611 CAS, and RS analyzed the data. ES, NKY, CAS, and RS wrote the manuscript.

612

613 **Declaration of Interests**

614 UNC Chapel Hill is pursuing IP protection of Primer ID technology, for which RS in included as
615 a co-inventor and has received nominal royalties. The remaining authors declare no competing
616 interests.

617

618 **References**

619 Ali, A., Bandaranayake, R.M., Cai, Y., King, N.M., Kolli, M., Mittal, S., Murzycki, J.F., Nalam,
620 M.N., Nalivaika, E.A., Ozen, A., *et al.* (2010). Molecular Basis for Drug Resistance in HIV-1
621 Protease. Viruses *2*, 2509-2535.

622 Antinori, A., Clarke, A., Svedhem-Johansson, V., Arribas, J.R., Arenas-Pinto, A., Fehr, J.,
623 Gerstoft, J., Horban, A., Clotet, B., Ripamonti, D., *et al.* (2015). Week 48 efficacy and central
624 nervous system analysis of darunavir/ritonavir monotherapy versus darunavir/ritonavir with two
625 nucleoside analogues. AIDS *29*, 1811-1820.

626 Aoki, M., Das, D., Hayashi, H., Aoki-Ogata, H., Takamatsu, Y., Ghosh, A.K., and Mitsuya, H.
627 (2018). Mechanism of Darunavir (DRV)'s High Genetic Barrier to HIV-1 Resistance: A Key V32I
628 Substitution in Protease Rarely Occurs, but Once It Occurs, It Predisposes HIV-1 To Develop
629 DRV Resistance. mBio *9 e02425-17*.

630 Arribas, J.R., Clumeck, N., Nelson, M., Hill, A., van Delft, Y., and Moecklinghoff, C. (2012).
631 The MONET trial: week 144 analysis of the efficacy of darunavir/ritonavir (DRV/r) monotherapy
632 versus DRV/r plus two nucleoside reverse transcriptase inhibitors, for patients with viral load <
633 50 HIV-1 RNA copies/mL at baseline. HIV Med *13*, 398-405.

634 Casado, J.L., Vizcarra, P., Blanco, J.L., Montejano, R., Negredo, E., Espinosa, N., Montero, M.,
635 Mena, A., Palacios, R., Lopez, J.C., *et al.* (2020). Maintenance of virologic suppression and
636 improvement in comorbidities after simplification to raltegravir plus boosted darunavir among
637 treatment-experienced HIV-infected patients. Int J STD AIDS *31*, 467-473.

638 Clemente, J.C., Moose, R.E., Hemrajani, R., Whitford, L.R., Govindasamy, L., Reutzel, R.,
639 McKenna, R., Agbandje-McKenna, M., Goodenow, M.M., and Dunn, B.M. (2004). Comparing

640 the accumulation of active- and nonactive-site mutations in the HIV-1 protease. Biochemistry *43*,
641 12141-12151.

642 de Meyer, S., Vangeneugden, T., van Baelen, B., de Paepe, E., van Marck, H., Picchio, G.,
643 Lefebvre, E., and de Bethune, M.P. (2008). Resistance profile of darunavir: combined 24-week
644 results from the POWER trials. AIDS Res Hum Retroviruses *24*, 379-388.

645 Descamps, D., Lambert-Niclot, S., Marcelin, A.G., Peytavin, G., Roquebert, B., Katlama, C.,
646 Yeni, P., Felices, M., Calvez, V., and Brun-Vezinet, F. (2009). Mutations associated with
647 virological response to darunavir/ritonavir in HIV-1-infected PI-experienced patients. J
648 Antimicrob Chemother *63*, 585-592.

649 Di Cristo, V., Adorni, F., Maserati, R., Annovazzi Lodi, M., Bruno, G., Maggi, P., Volpe, A.,
650 Vitiello, P., Abeli, C., Bonora, S.*, et al.* (2020). 96-week results of a dual therapy with
651 darunavir/ritonavir plus rilpivirine once a day vs triple therapy in patients with suppressed
652 viraemia: virological success and non-HIV related morbidity evaluation. HIV Res Clin Pract *21*,
653 34-43.

654 Gulick, R.M., Mellors, J.W., Havlir, D., Eron, J.J., Gonzalez, C., McMahon, D., Richman, D.D.,
655 Valentine, F.T., Jonas, L., Meibohm, A.*, et al.* (1997). Treatment with indinavir, zidovudine, and
656 lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral
657 therapy. N Engl J Med *337*, 734-739.

658 Hawkins, K.L., Montague, B.T., Rowan, S.E., Beum, R., McLees, M.P., Johnson, S., and
659 Gardner, E.M. (2019). Boosted darunavir and dolutegravir dual therapy among a cohort of highly
660 treatment-experienced individuals. Antivir Ther *24*, 513-519.

661 Henderson, G.J., Lee, S.K., Irlbeck, D.M., Harris, J., Kline, M., Pollom, E., Parkin, N., and
662 Swanstrom, R. (2012). Interplay between single resistance-associated mutations in the HIV-1
663 protease and viral infectivity, protease activity, and inhibitor sensitivity. Antimicrob Agents
664 Chemother *56*, 623-633.

665 Henes, M., Lockbaum, G.J., Kosovrasti, K., Leidner, F., Nachum, G.S., Nalivaika, E.A., Lee,
666 S.K., Spielvogel, E., Zhou, S., Swanstrom, R., *et al.* (2019). Picomolar to Micromolar:
667 Elucidating the Role of Distal Mutations in HIV-1 Protease in Conferring Drug Resistance. ACS
668 Chem Biol *14*, 2441-2452.

669 Huang, Y., Huang, X., Chen, H., Wu, H., and Chen, Y. (2019). Efficacy and Safety of Raltegravir-
670 Based Dual Therapy in AIDS Patients: A Meta-Analysis of Randomized Controlled Trials. Front
671 Pharmacol *10*, 1225.

672 Hui, J.O., Tomasselli, A.G., Reardon, I.M., Lull, J.M., Brunner, D.P., Tomich, C.S., and
673 Heinrikson, R.L. (1993). Large scale purification and refolding of HIV-1 protease from
674 Escherichia coli inclusion bodies. J Protein Chem *12*, 323-327.

675 Katoh, I., Yasunaga, T., Ikawa, Y., and Yoshinaka, Y. (1987). Inhibition of retroviral protease
676 activity by an aspartyl proteinase inhibitor. Nature *329*, 654-656.

677 King, N.M., Melnick, L., Prabu-Jeyabalan, M., Nalivaika, E.A., Yang, S.S., Gao, Y., Nie, X.,
678 Zepp, C., Heefner, D.L., and Schiffer, C.A. (2002). Lack of synergy for inhibitors targeting a
679 multi-drug-resistant HIV-1 protease. Protein Sci *11*, 418-429.

680 King, N.M., Prabu-Jeyabalan, M., Nalivaika, E.A., and Schiffer, C.A. (2004). Combating
681 susceptibility to drug resistance: lessons from HIV-1 protease. Chem Biol *11*, 1333–1338.

682 Koh, Y., Amano, M., Towata, T., Danish, M., Leshchenko-Yashchuk, S., Das, D., Nakayama,
683 M., Tojo, Y., Ghosh, A.K., and Mitsuya, H. (2010). In vitro selection of highly darunavir-resistant
684 and replication-competent HIV-1 variants by using a mixture of clinical HIV-1 isolates resistant
685 to multiple conventional PIs. J Virol *84*, 11961-11969.

686 Kurt Yilmaz, N., Izadkhashti, A., Price, R.W., Mallon, P.W., De Meulder, M., Timmerman, P.,
687 and Gisslen, M. (2009). Darunavir concentrations in cerebrospinal fluid and blood in HIV-1-
688 infected individuals. AIDS Res Hum Retroviruses *25*, 457-461.

689 Leidner, F., Kurt Yilmaz, N., & Schiffer, C. A. (2021). Deciphering Complex Mechanisms of
690 Resistance and Loss of Potency through Coupled Molecular Dynamics and Machine Learning. J
691 Chem Theory Comput *17*, 2054–2064.

692 Liu, Y., Kati, W., Chen, C.M., Tripathi, R., Molla, A., and Kohlbrenner, W. (1999). Use of a
693 fluorescence plate reader for measuring kinetic parameters with inner filter effect correction. Anal
694 Biochem *267*, 331-335.

695 Lockbaum, G.J., Leidner, F., Rusere, L.N., Henes, M., Kosovrasti, K., Nachum, G.S., Nalivaika,
696 E.A., Ali, A., Kurt Yilmaz, N., and Schiffer, C.A. (2019). Structural Adaptation of Darunavir
697 Analogues against Primary Mutations in HIV-1 Protease. ACS Infect Dis *5*, 316-325.

698 Mahalingam, B., Boross, P., Wang, Y.F., Louis, J.M., Fischer, C.C., Tozser, J., Harrison, R.W.,
699 and Weber, I.T. (2002). Combining mutations in HIV-1 protease to understand mechanisms of
700 resistance. Proteins *48*, 107-116.

701 Matayoshi, E.D., Wang, G.T., Krafft, G.A., and Erickson, J. (1990). Novel fluorogenic substrates
702 for assaying retroviral proteases by resonance energy transfer. Science *247*, 954-958.

703 Matthew, A. N., Leidner, F., Lockbaum, G. J., Henes, M., Zephyr, J., Hou, S., Rao, D. N., Timm,
704 J., Rusere, L. N., Ragland, D. A., Paulsen, J. L., Prachanronarong, K., Soumana, D. I., Nalivaika,
705 E. A., Kurt Yilmaz, N., Ali, A., & Schiffer, C. A. (2021). Drug Design Strategies to Avoid
706 Resistance in Direct-Acting Antivirals and Beyond. Chemical reviews *121*, 3238–3270.

707 Mittal, S., Bandaranayake, R.M., King, N.M., Prabu-Jeyabalan, M., Nalam, M.N., Nalivaika,
708 E.A., Kurt Yilmaz, N., and Schiffer, C.A. (2013). Structural and thermodynamic basis of
709 amprenavir/darunavir and atazanavir resistance in HIV-1 protease with mutations at residue 50. J
710 Virol *87*, 4176-4184.

711 Muzammil, S., Ross, P., and Freire, E. (2003). A major role for a set of non-active site mutations
712 in the development of HIV-1 protease drug resistance. Biochemistry *42*, 631-638.

713 Nalam, M.N., Ali, A., Reddy, G.S., Cao, H., Anjum, S.G., Altman, M.D., Kurt Yilmaz, N., Tidor,
714 B., Rana, T.M., and Schiffer, C.A. (2013). Substrate envelope-designed potent HIV-1 PIs to avoid
715 drug resistance. Chem Biol *20*, 1116–1124.

716 Ozen, A., Lin, K.H., Kurt Yilmaz, N., and Schiffer, C.A. (2014). Structural basis and distal effects
717 of Gag substrate coevolution in drug resistance to HIV-1 protease. Proc Natl Acad Sci U S A *111*,
718 15993-15998.

719 Paton, N.I., Stohr, W., Arenas-Pinto, A., Fisher, M., Williams, I., Johnson, M., Orkin, C., Chen,
720 F., Lee, V., Winston, A.*, et al.* (2015). PI monotherapy for long-term management of HIV
721 infection: a randomised, controlled, open-label, non-inferiority trial. Lancet HIV *2*, e417-426.

722 Pellegrin, I., Wittkop, L., Joubert, L.M., Neau, D., Bollens, D., Bonarek, M., Girard, P.M., Fleury,
723 H., Winters, B., Saux, M.C.*, et al.* (2008). Virological response to darunavir/ritonavir-based
724 regimens in antiretroviral-experienced patients (PREDIZISTA study). Antivir Ther *13*, 271-279.

725 Pettit, S.C., Henderson, G.J., Schiffer, C.A., and Swanstrom, R. (2002). Replacement of the P1
726 amino acid of human immunodeficiency virus type 1 Gag processing sites can inhibit or enhance
727 the rate of cleavage by the viral protease. J Virol *76*, 10226-10233.

728 Prabu-Jeyabalan, M., Nalivaika, E.A., and Schiffer, C.A. (2002). Substrate shape determines
729 specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate
730 complexes. Structure *10*, 369-381.

731 Ragland, D.A., Nalivaika, E.A., Nalam, M.N., Prachanronarong, K.L., Cao, H., Bandaranayake,
732 R.M., Cai, Y., Kurt Yilmaz, N., and Schiffer, C.A. (2014). Drug resistance conferred by mutations
733 outside the active site through alterations in the dynamic and structural ensemble of HIV-1
734 protease. J Am Chem Soc *136*, 11956-11963.

735 Ragland, D. A., Whitfield, T. W., Lee, S. K., Swanstrom, R., Zeldovich, K. B., Kurt Yilmaz, N.,
736 & Schiffer, C. A. (2017). Elucidating the Interdependence of Drug Resistance from Combinations
737 of Mutations. J Chem Theory Comput., *13*, 5671–5682.

738 Rhee, S.Y., Gonzales, M.J., Kantor, R., Betts, B.J., Ravela, J., and Shafer, R.W. (2003). Human
739 immunodeficiency virus reverse transcriptase and protease sequence database. Nucleic Acids Res
740 *31*, 298-303.

741 Richards, A.D., Roberts, R., Dunn, B.M., Graves, M.C., and Kay, J. (1989). Effective blocking
742 of HIV-1 proteinase activity by characteristic inhibitors of aspartic proteinases. FEBS Lett *247*,
743 113-117.

744 Rose, J.R., Salto, R., and Craik, C.S. (1993). Regulation of autoproteolysis of the HIV-1 and HIV-
745 2 proteases with engineered amino acid substitutions. J Biol Chem *268*, 11939-11945.

746 Seelmeier, S., Schmidt, H., Turk, V., and von der Helm, K. (1988). Human immunodeficiency
747 virus has an aspartic-type protease that can be inhibited by pepstatin A. Proc Natl Acad Sci U S
748 A *85*, 6612-6616.

749 Tremblay, C.L. (2008). Combating HIV resistance - focus on darunavir. Ther Clin Risk Manag
750 *4*, 759-766.

751 Valantin, M.A., Lambert-Niclot, S., Flandre, P., Morand-Joubert, L., Cabie, A., Meynard, J.L.,
752 Ponscarme, D., Ajana, F., Slama, L., Curjol, A.*, et al.* (2012). Long-term efficacy of
753 darunavir/ritonavir monotherapy in patients with HIV-1 viral suppression: week 96 results from
754 the MONOI ANRS 136 study. J Antimicrob Chemother *67*, 691-695.

755 Vizcarra, P., Blanco, J.L., Montejano, R., Negredo, E., Espinosa, N., Casado, J.L., and group,
756 B.I.s. (2020). Lack of impact of PI resistance-associated mutations on the outcome of HIV-1-
757 infected patients switching to darunavir-based dual therapy. Infect Dis (Lond) *52*, 202-206.

758 Watkins, T., Resch, W., Irlbeck, D., and Swanstrom, R. (2003). Selection of high-level resistance
759 to human immunodeficiency virus type 1 PIs. Antimicrob Agents Chemother *47*, 759-769.

760 Wensing, A.M., van Maarseveen, N.M., and Nijhuis, M. (2010). Fifteen years of HIV PIs: raising
761 the barrier to resistance. Antiviral Res *85*, 59-74.

762 Windsor, I.W., and Raines, R.T. (2015). Fluorogenic Assay for Inhibitors of HIV-1 Protease with
763 Sub-picomolar Affinity. Sci Rep *5*, 11286.

764 Zhou, S., Jones, C., Mieczkowski, P., and Swanstrom, R. (2015). Primer ID Validates Template
765 Sampling Depth and Greatly Reduces the Error Rate of Next-Generation Sequencing of HIV-1
766 Genomic RNA Populations. J Virol *89*, 8540-8555.

767 Zhu, M., Dou, Y., Ma, L., Dong, B., Zhang, F., Zhang, G., Wang, J., Zhou, J., Cen, S., and Wang,
768 Y. (2020). Novel HIV-1 PIs with Morpholine as the P2 Ligand to Enhance Activity against DRV-
769 Resistant Variants. ACS Med Chem Lett *11*, 1196-1204.

770

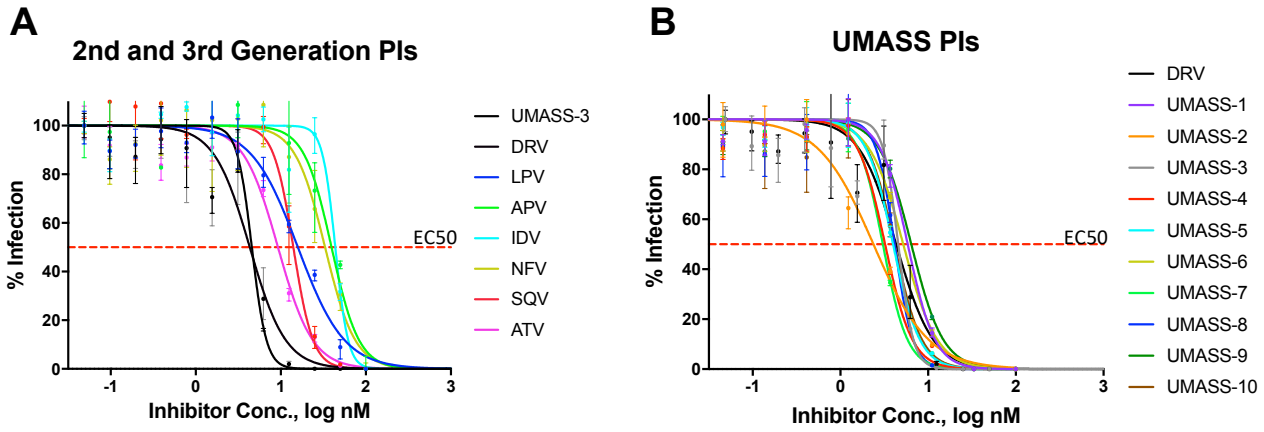


Fig. S1. EC50 Inhibition curves for **A)** old and new protease inhibitors and **B)** fourth generation analogs of DRV (UMass 1 – UMass 10).

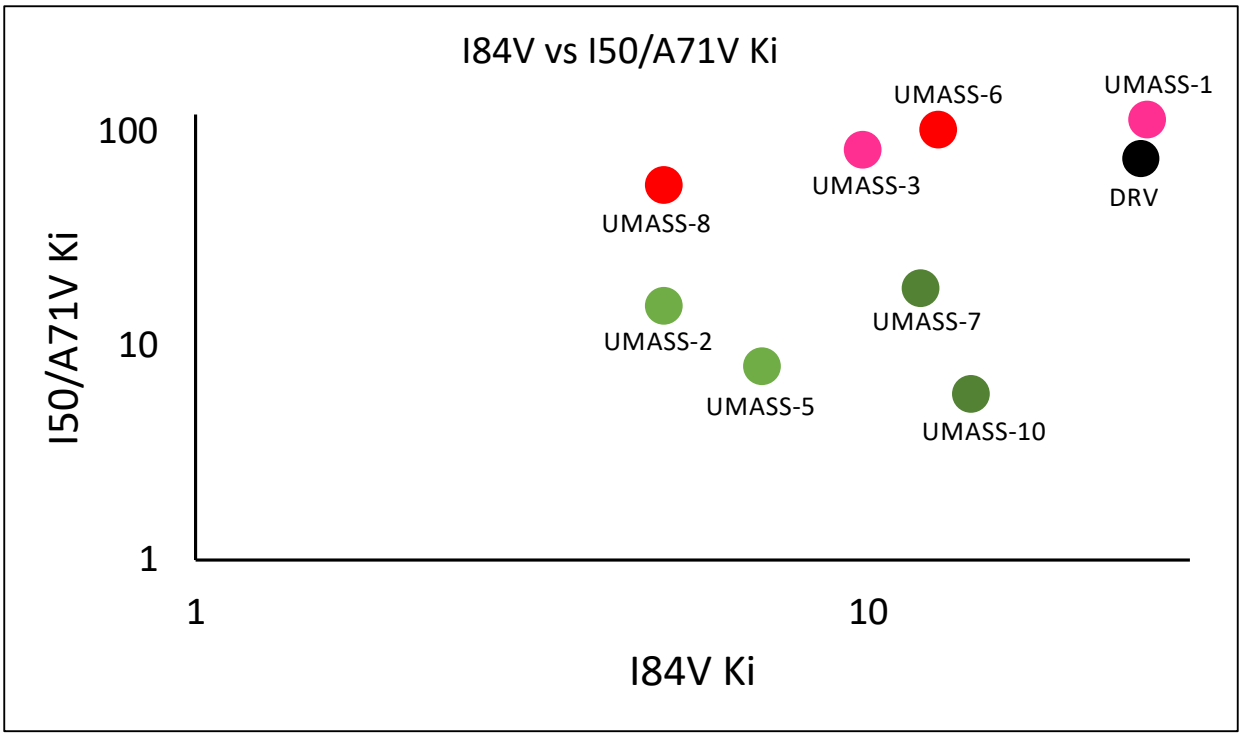
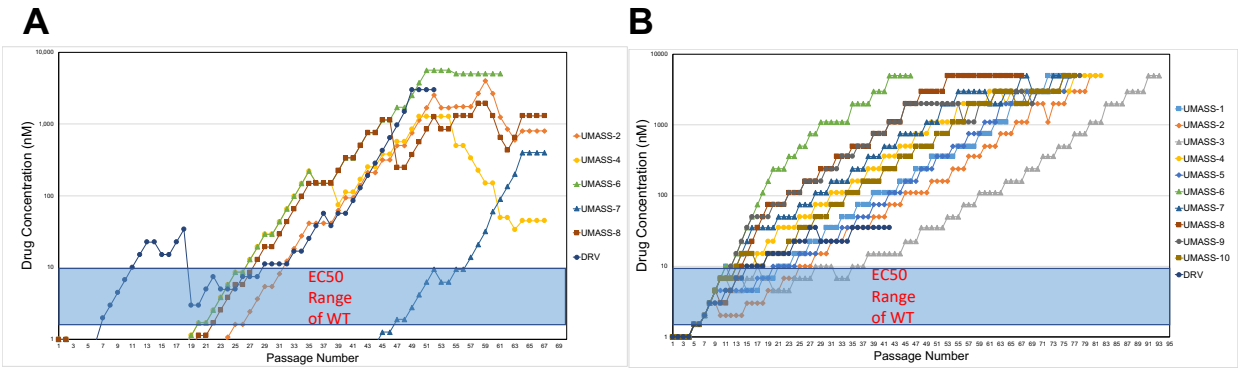


Fig S2. Enzymatic inhibition constant (Ki) values for UMASS-1-10 in the presence of pro with I84V or I50V/A71V mutations. UMASS-1, -3, -6, -8 show less potency towards these variants and is shaded in a red tint matching R2 groups. UMASS-2, -5, -7, -10 show increased potency towards these variants and is shaded in a green tint matching R2 groups



Selection with 26 Mutants Selection with Wild-type

Fig. S3. **(A-B)** Viral selections with DRV and UMASS-1-10 have increasing inhibitor concentrations during passaging. Passages are increased when extensive CPE is observed during PI selection. Virus generated from the **A)** 26 single mutants or **B)** NL4-3 molecular clone was passaged in the presence of increasing inhibitor concentration as described in the text. The period of time (in days) until the virus-infected culture displayed maximal CPE (massive syncytia) was 4-7 days on average.

Passage #	Drug Conc. (nM)	R1	R2	PIs	Wild-type Selection: Most Abundant Variants Detected at the Last Time Point	Abundance
75	5000	1	1	UMASS-1	10I 16E 32I 33F 46I 54L 71V 76V 82I 84V	55.4%
					10I 32I 33F 46I 54L 71V 76V 82I 84V	21.9%
					10I 16E 32I 33F 46I 54L 71V 76V 82I 84V 91S	6.2%
82	5000	1	2	UMASS-2	10F 33F 46I 47V 50V 71V 82I 84V	37.5%
					28S 32I 33F 46I 47V 50V 71V 82I 84V	33.5%
					33F 46I 47V 50V 71V 82I 84V	4.2%
93	5000	1	3	UMASS-3	10F 33F 46I 47V 50V 53L 63P 71V 76S 82I 85V 89I*	85.7%
					10F 33F 46I 47V 50V 53L 63P 71V 76S 82I 85V 89I* 91S	1.1%
					10F 33F 46I 47V 50V 53L 63P 71V 76S 82I	0.9%
81	5000	1	4	UMASS-4	10F 11I 13V 32I 33F 43T 46L 54L 71V 82I 84V 89M 91S 92R	87.3%
					10F 11I 32I 33F 43T 46L 54L 71V 82I 84V 89M 91S	1.3%
					10F 11I 32I 33F 46L 54L 71V 82I 84V 89M 91S	1.1%
77	5000	1	5	UMASS-5	10F 15V 46I 47V 50V 53L 71V 82I 84V 89T	67.9%
					10F 46I 47V 50V 53L 71V 82I 84V 89T	10.9%
					10F 46I 47V 50V 53L 63P 71V 82I 84V 89M	8.8%
46	5000	2	1	UMASS-6	10F 13V 33F 46I 47A 50V 53L 71V	45.1%
					10F 13V 33F 46I 47A 50V 53L 63P 71V	42.3%
					10F 13V 16E 33F 46I 47A 50V 53L 71V	2.2%
76	5000	2	2	UMASS-7	10F 13V 33F 46I 47V 50V 53L 54L 71V 82I	94.5%
					10F 13V 33F 46I 47V 50V 53L 54L 71V 82I 87K	0.2%
					10F 13V 33F 46I 47V 50V 53L 54L 67Y 71V 82I	0.2%
67	5000	2	3	UMASS-8	10F 12K 33F 46I 47V 50V 53L 54L 63P 71V 82I 85V	55.4%
					10F 33F 46I 47V 50V 53L 54L 63P 71V 82I 85V	41.9%
					10F 12K 33F 46I 47V 50V 53L 54L 63P 71V 82I 85V 87K	0.1%
78	5000	2	4	UMASS-9	10F 13V 33F 45R 46I 47V 50V 53L 54L 66F 71V 74A 76S 82I	92.3%
					10F 13V 33F 41K 45R 46I 47V 50V 53L 54L 66F 71V 74A 76S 82I	0.7%
					10F 13V 33F 45R 46I 47V 50V 53L 54L 66F 71V 74A 76S 82I 87K	0.4%
77	5000	2	5	UMASS-10	10F 13V 33F 43T 46I 47V 50V 54L 71V 82L 88S	91.2%
					10F 13V 33F 43T 46I 47V 50V 54L 71V 77I 82L 88S	1.6%
					10F 13V 33F 43T 46I 47V 50V 54L 71V 82L 87K 88S	0.2%
43	1350	-	1	DRV	32I 41I 46I 47V 82I 84V 85V	62.8%
					10I 32I 41I 46I 47V 82I 84V 85V	10.3%
					32I 41I 46I 47V 82L 84V 85V	10.1%
75	0			ND	WT	82.9%
					36I	1.2%
					41K	1.0%

Passage #	Drug Conc. (nM)	R1	R2	PIs	Mutant Selection: Most Abundant Variants Detected at the Last Time Point	Abundance
53	4000	1	2	UMASS-2	10I 32I 33F 45I 46I 50V 71V 82I 84V	34.8%
					10I 28S 32I 33F 46I 71V 82I 84V	29.2%
					10I 28S 32I 33F 45I 46I 50V 71V 82I 84V	2.7%
45	1275	1	4	UMASS-4	10F 28S 46I 84V	9.4%
					10F 32I 46I 54M 71V 82I 84V 89M	8.2%
					10F 32I 46I 54M 84V	7.0%
53	5000	2	2	UMASS-6	13V 16E 32I 33F 45I 46I 54L 71V 76V 82F 84V	62.8%
					13V 16E 32I 33F 45I 46I 54L 82F 84V	16.9%
					13V 16E 32I 33F 45I 46I 71V 76V 82F 84V	6.1%
58	400	2	2	UMASS-7	10F 46I 50V 63P 85V	21.0%
					28S 46I 63P 85V	18.9%
					10F 28S 46I 63P 85V	16.3%
61	1300	2	2	UMASS-8	10F 46I 47V 50V 53L 63P 72V 73S 82I 85V	70.0%
					10F 41K 46I 47V 50V 53L 63P 72V 73S 82I 85V	19.3%
					10F 46I 47V 50V 63P 72V 73S 82I 85V	1.4%
39	3000	-	1	DRV-3	10F 32I 33F 37D 46I 63P 74S 76V 84V 89M	51.1%
					10F 32I 33F 46I 63P 74S 76V 84V 89M	22.9%
					10F 32I 33F 46I 63P 76V 84V 89M	8.6%
61				ND	93L	36.7%
					36I 93L	12.6%
					WT	12.3%

Fig. S4. (A-B) Selection cultures for fourth generation analogs of DRV (UMASS-1 – UMASS-10), DRV and no inhibitor controls. Top 3 most abundant viral variants present at the final timepoint of each selection. **A)** Mutations arising from selections beginning with wild-type virus are shown in blue while **B)** mutations occurring in selections starting from 26-mutants are shown in red. Abundance of these variants is labeled as a percentage on the right.

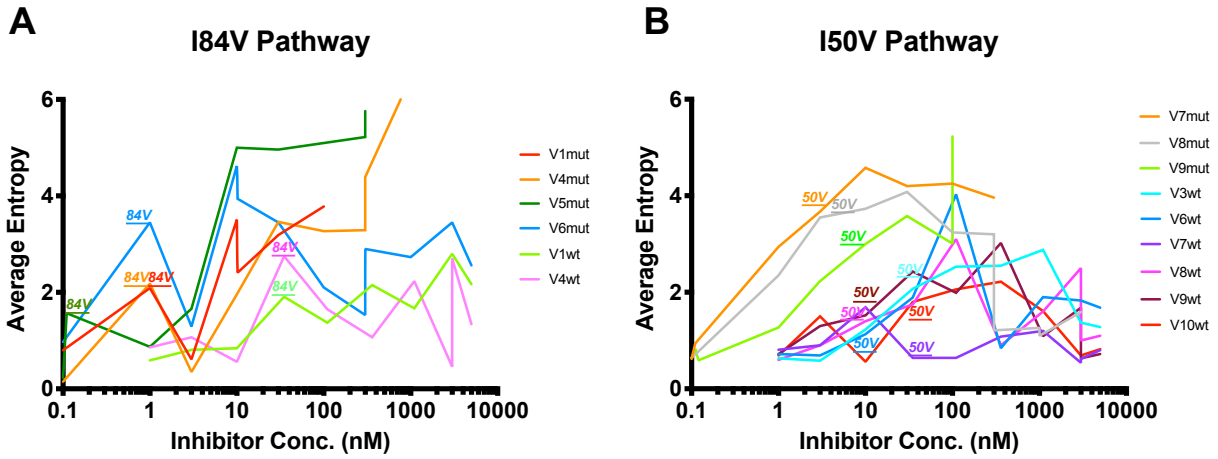


Fig. S5. (A-B) Shannon's entropy of viral cultures that utilize either **A)** I84V or **B)** I50V as their resistance pathway. When I84V or I50V arise in the selection cultures above the poisson distribution, that incident is labeled above the graph. I84V mutations are associated with increases/decreases in entropy while I50V mutations are not associated with major entropy differences.

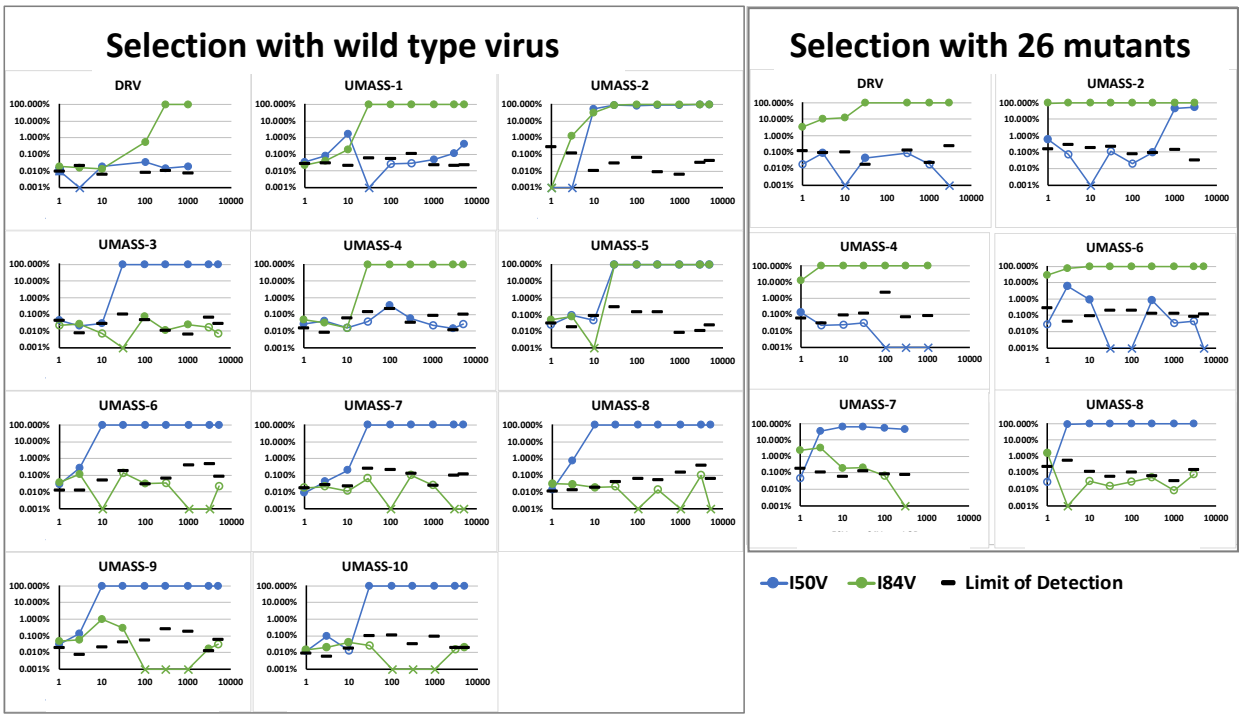


Fig. S6. Abundance of I50V and I84V mutations at each drug concentration of all UMASS inhibitors that reached at least 1000 nM concentrations in culture. Abundance limit of detection cutoffs are shown as black lines on each timepoint to resemble our TCS poisson distribution calculation as shown in Zhou et al, 2015.

A

	Inhib. Conc. (nM)	Pathway	NC P2	SP2 -	p6 P1'	p6 P5'
			A		L	P
V1*wt*	5000	I84V	V			L
V4*mut*	1275	I84V	V		F	
V4*wt*	5000	I84V	I			I
V6*mut*	5000	I84V	V			
V3*wt*	5000	I50V			F	L
V6*wt*	5000	I50V				L
V7*wt*	5000	I50V			F	
V8*mut*	1300	I50V				L
V8*wt*	5000	I50V			F	
V9*wt*	5000	I50V			F	
V10*wt*	5000	I50V			F	
V2*mut*	4000	I50V+I84V				L
V2*wt*	5000	I50V+I84V			F	L
V5*wt*	5000	I50V+I84V			F	L

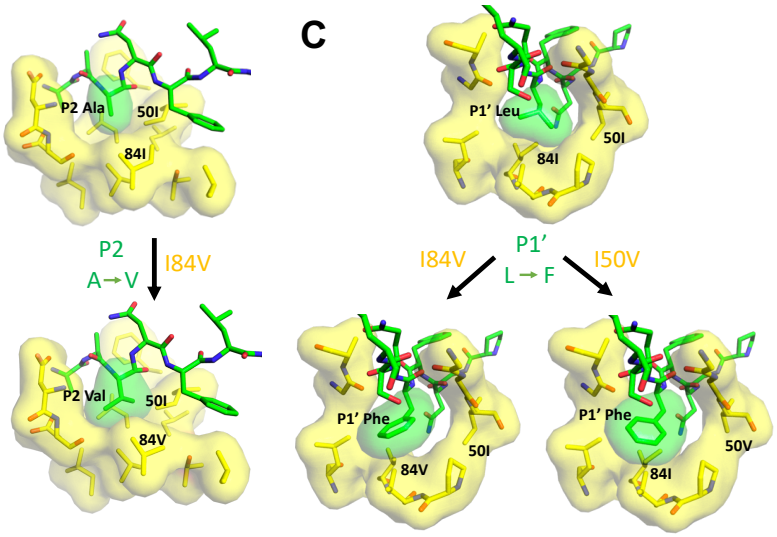


Fig. S7. (A-B) Protease Cleavage Site Mutations Observed After Selection For High Level Resistance. **A)** Culture names, final inhibitor concentration, and resistance pathway are shown. Changes in amino acid sequence are shown for the NC/SP2 and SP2/p6 cleavage sites. **B)** P2 substitution in the NC/SP2 cleavage site with the I84V mutation modeled. Subsite structure was modeled using the original structure (PDB: 1KJH) and mutating the P2 position in the subsite pocket from Alanine to Valine. **C)** P1' change in the SP2/p6 cleavage site with either the I84V mutation or the I50V mutations modeled. Subsite structure was modeled using the original structure (PDB: 1KJF) and mutating the P1' position from Leucine to Phenylalanine.