

University of Massachusetts Medical School

eScholarship@UMMS

University of Massachusetts Medical School Faculty Publications

2020-07-16

Structural Analysis of Potent Hybrid HIV-1 Protease Inhibitors Containing Bis-Tetrahydrofuran in a Pseudo-Symmetric Dipeptide Isostere

Linah Rusere

University of Massachusetts Medical School

Et al.

Let us know how access to this document benefits you.

Follow this and additional works at: https://escholarship.umassmed.edu/faculty_pubs



Part of the [Amino Acids, Peptides, and Proteins Commons](#), [Biochemistry Commons](#), [Enzymes and Coenzymes Commons](#), [Medicinal and Pharmaceutical Chemistry Commons](#), [Medicinal Chemistry and Pharmaceutics Commons](#), [Structural Biology Commons](#), and the [Viruses Commons](#)

Repository Citation

Rusere L, Lockbaum GJ, Henes M, Lee S, Spielvogel E, Rao DN, Kosovrasti K, Nalivaika EA, Swanstrom R, Yilmaz NK, Schiffer CA, Ali A. (2020). Structural Analysis of Potent Hybrid HIV-1 Protease Inhibitors Containing Bis-Tetrahydrofuran in a Pseudo-Symmetric Dipeptide Isostere. University of Massachusetts Medical School Faculty Publications. <https://doi.org/10.1021/acs.jmedchem.0c00529>. Retrieved from https://escholarship.umassmed.edu/faculty_pubs/1723

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in University of Massachusetts Medical School Faculty Publications by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.

Structural Analysis of Potent Hybrid HIV-1 Protease Inhibitors Containing Bis-Tetrahydrofuran in a Pseudo-Symmetric Dipeptide Isostere

Linah N Rusere, Gordon J. Lockbaum, Mina Henes, Sook-Kyung Lee, Ean Spielvogel, Desaboini Nageswara Rao, Klajdi Kosovrasti, Ellen A. Nalivaika, Ronald Swanstrom, Nese Kurt Yilmaz, Celia A. Schiffer, and Akbar Ali

J. Med. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jmedchem.0c00529 • Publication Date (Web): 16 Jul 2020

Downloaded from pubs.acs.org on July 27, 2020

Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Structural Analysis of Potent Hybrid HIV-1 Protease Inhibitors Containing Bis-Tetrahydrofuran in a Pseudo-Symmetric Dipeptide Isostere

*Linah N. Rusere,†,§ Gordon J. Lockbaum,†,§ Mina Henes,† Sook-
Kyung Lee,‡ Ean Spielvogel,‡ Desaboini Nageswara Rao,† Klajdi
Kosovrasti,† Ellen A. Nalivaika,† Ronald Swanstrom,‡ Nese Kurt
Yilmaz,† Celia A. Schiffer,†,* and Akbar Alit,**

†Department of Biochemistry and Molecular Pharmacology,
University of Massachusetts Medical School, Worcester,
Massachusetts 01605, United States

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

ABSTRACT

The design, synthesis, and X-ray structural analysis of hybrid HIV-1 protease inhibitors (PIs) containing *bis*-tetrahydrofuran (*bis*-THF) in a pseudo- C_2 -symmetric dipeptide isostere are described. A series of PIs were synthesized by incorporating *bis*-THF of darunavir on either side of the Phe-Phe isostere of lopinavir in combination with hydrophobic amino acids on the opposite P2/P2' position. Structure-activity relationship studies indicated that the *bis*-THF moiety can be attached at either the P2 or P2' position without significantly affecting potency. However, the group on the opposite P2/P2' position had a dramatic effect on potency depending on the size and shape of the side chain. Cocrystal structures of inhibitors with wild-type HIV-1 protease revealed that the *bis*-THF moiety retained similar interactions as observed in the darunavir-protease complex regardless of position on the Phe-Phe isostere. Analyses of cocrystal structures and molecular dynamics simulations provide insights for optimizing HIV-1 PIs containing *bis*-THF in non-sulfonamide dipeptide isosteres.

INTRODUCTION

Combinations of drugs targeting HIV-1 enzymes essential for viral replication have been highly effective in reducing viral load in infected individuals and significantly increasing their life expectancy.¹ HIV-1 protease inhibitors (PIs) have played an important role as a highly potent component of combination antiretroviral therapy (cART) for people living with HIV-1. There are nine HIV-1 PIs approved by the US FDA; however, clinical use of most of these has diminished over the years due to side effects, unfavorable pharmacokinetics, and drug resistance. Currently only three PIs, darunavir (DRV), lopinavir (LPV), and atazanavir (ATV) (**Figure 1**) - boosted with small doses of ritonavir (RTV) - are being used in cART as they have comparatively higher potency and better resistance profiles.^{2, 3}

Most HIV-1 PIs are substrate-based peptidomimetics containing a hydroxyl group as a transition state mimetic and diverse moieties that target the S1/S1' and S2/S2' subsites of the enzyme.^{4, 5} Among the PIs in clinical use, DRV is the most potent with the highest barrier to resistance and efficacy against multidrug-resistant HIV-1 strains.⁶ The excellent potency and resistance profiles of DRV are attributed to the *bis*-tetrahydrofuran (*bis*-THF) moiety, which forms strong hydrogen bonding interactions with the main-chain NH of Asp29 and Asp30 in the S2 subsite of the enzyme.⁷ Moreover, the extensive van der Waals (vdW) interactions of the

1
2
3 *bis*-THF moiety, along with the sulfonamide-based Phe-Leu dipeptide
4 isostere and the 4-aminobenzene group, also significantly
5 contribute to the potency of DRV.⁸ Due to these unique binding
6 characteristics, the P2 *bis*-THF moiety of DRV has emerged as a
7 privileged ligand for targeting the S2 subsite of HIV-1 protease
8 especially when incorporated into the (R)-
9 (hydroxyethylamino)sulfonamide isostere.⁹

10 Efforts to improve potency, pharmacokinetics, and resistance
11 profiles of HIV-1 PIs have led to the discovery of exceptionally
12 potent compounds exhibiting a wide range of properties.^{5, 10-14} In
13 recent years, these efforts have focused mainly on exploring DRV
14 analogues with similar sulfonamide-based dipeptide isosteres and
15 P2/P2' moieties. Many analogues of DRV with modifications at P1/P1'
16 and P2/P2' positions have been explored extensively,^{5, 15-18} including
17 novel *bis*-THF analogues showing improved hydrogen bonding and vdW
18 interactions with the protease.¹⁹⁻²² Moreover, detailed structural
19 analyses informing structure-based design of inhibitors with the
20 DRV scaffold have been reported numerous times including recent
21 structures determined by neutron crystallography.^{5, 7, 8, 23-25}

22 However, very few inhibitors containing the *bis*-THF moiety in
23 non-sulfonamide dipeptide isosteres have been explored.²⁶⁻²⁸
24 Previously, Chen and colleagues at Abbott Laboratories explored
25 PIs with *bis*-THF on either side of the pseudo-C₂-symmetric Phe-Phe
26 dipeptide isostere of LPV,^{26, 27} while Cannizzaro and coworkers at
27

1
2
3 Gilead Sciences designed similar analogues using the aza-dipeptide
4 isostere of ATV.²⁸ Although some of the Phe-Phe isostere-based
5
6
7
8 compounds, such as **4** and **5** (**Figure 1**), exhibited significant
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
antiviral activity against wild-type HIV-1, in the absence of
direct comparison, it is unclear if their potency is similar to
that of DRV.²⁶ Also, due to limited structure-activity relationship
studies, the S2/S2' subsite preference of the *bis*-THF moiety in the
context of Phe-Phe isostere remains unclear,^{26, 27} as is the identity
of an optimal moiety for targeting the opposite S2/S2' subsite.²⁶
Moreover, no cocrystal structures of such hybrid PIs have been
reported, limiting opportunities for structure-guided design and
optimization.

31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
In this study, we investigate HIV-1 PIs containing the *bis*-THF
moiety of DRV in the Phe-Phe dipeptide isostere of LPV with the
aim to identify compounds with potency similar to that of DRV but
with an alternate non-sulfonamide scaffold. Protease-inhibitor
molecular interactions are characterized using X-ray
crystallography and molecular dynamics (MD) simulations to aid
structure-guided optimization of these hybrid PIs. In addition,
interdependence of HIV-1 protease subsites is explored to provide
insight into subsite preferences and influence of different
inhibitor moieties on binding and potency.

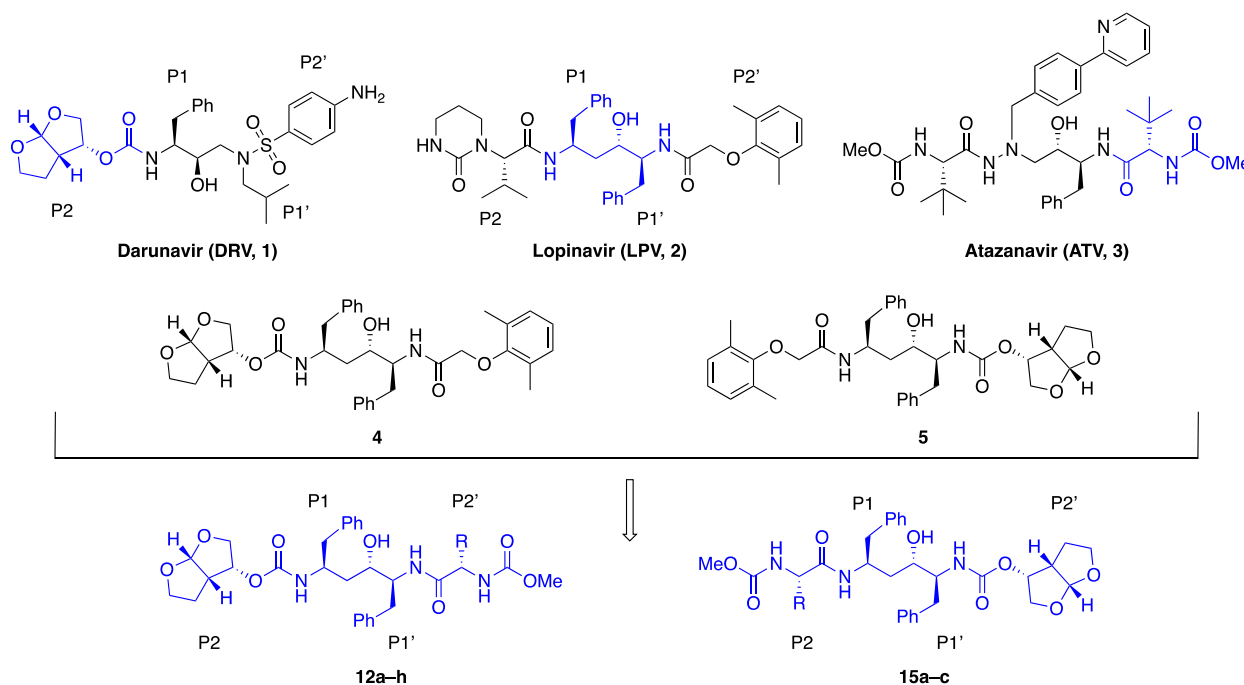


Figure 1. Structures of HIV-1 protease inhibitors darunavir (DRV) (1), lopinavir (LPV) (2), atazanavir (ATV) (3), previously reported hybrid compounds 4 and 5, and new hybrid compounds designed by incorporating the *bis*-THF moiety of DRV into the pseudo-symmetric Phe-Phe dipeptide isostere of LPV at the P2 (12a-h) or P2' position (15a-c), in combination with an ATV-like hydrophobic amino acid moiety at the P2' or P2 position, respectively. The canonical nomenclature is used to indicate the position of each inhibitor moiety.

The active site of HIV-1 protease exhibits subsite interdependence in substrate recognition and inhibitor binding,²⁹ but this feature of the protease has rarely been considered in inhibitor design. While identifying optimal moieties for targeting

1
2
3 each of the subsites in the protease active site is important,
4
5 understanding the subsite interdependence and selection of the
6
7 best combination of moieties are also crucial to achieving higher
8
9 potency especially against drug-resistant variants.¹⁹ The *bis*-THF
10
11 moiety potentially can be incorporated in diverse dipeptide
12
13 isosteres in combination with suitable ligands targeting the other
14
15 subsites. Thus, a series of hybrid HIV-1 PIs were synthesized by
16
17 incorporating the *bis*-THF moiety of DRV on either side of the
18
19 pseudo-*C*₂-symmetric Phe-Phe dipeptide isostere of LPV in
20
21 combination with ATV-like hydrophobic P2/P2' ligands of varying
22
23 size and shape (**Figure 1**). The hybrid PIs were tested for
24
25 inhibitory activity against wild-type and two drug-resistant
26
27 variants and for antiviral activity in cellular assays. We also
28
29 determined high-resolution X-ray crystal structures of all
30
31 designed hybrid compounds and previously reported PI **4** bound to
32
33 wild-type HIV-1 protease and characterized the binding
34
35 interactions of *bis*-THF in the context of Phe-Phe isostere. In
36
37 addition, molecular dynamics simulations were utilized to assess
38
39 the stability of protease-inhibitor interactions in solution. The
40
41 comprehensive structural and dynamic analyses of hybrid PIs have
42
43 provided insights into inhibitor binding and interdependence of
44
45 protease subsites.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

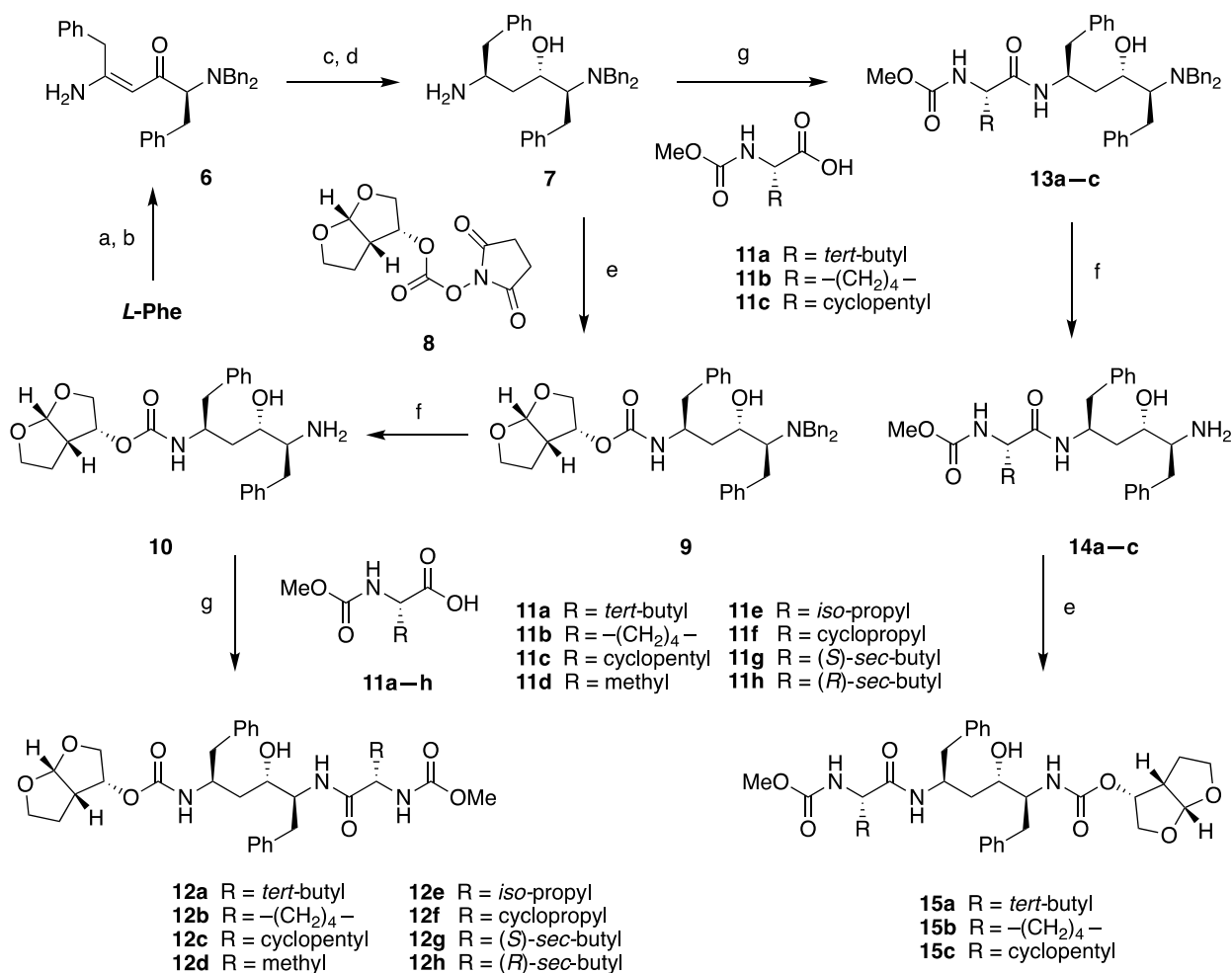
CHEMISTRY

The synthesis of hybrid PIs incorporating the *bis*-THF moiety at either the P2 or P2' position is outlined in **Scheme 1**. The required Phe-Phe dipeptide isostere core intermediate **7** was synthesized from phenylalanine following previously reported methods with minor modifications (see Supporting Information for details) (Haight, OPRD, 1999).³¹ Reaction of the dipeptide isostere **7** with *bis*-THF activated carbonate **8** in the presence of DIEA provided intermediate **9**. Debenzylation of **9** using ammonium formate and 10% palladium on activated carbon in methanol gave the corresponding deprotected amine **10**, which was subsequently coupled with the amino acid derivatives **11a-h** using HATU and DIEA to provide the target compounds **12a-h** with *bis*-THF at the P2 position. While *N*-(methoxycarbonyl)-protected amino acid derivatives **11a** and **11e** were commercially available, intermediates **11b-d** and **11f-h** were prepared from the corresponding amino acids by reaction with methyl chloroformate under basic conditions following a reported procedure.³²

The PIs **15a-c** with the *bis*-THF at the P2' position were synthesized using a similar reaction sequence. Coupling of the dipeptide isostere **7** with acids **11a-c** under HATU peptide coupling conditions gave the corresponding intermediates **13a-c**. Catalytic debenzylation of compounds **13a-c** followed by coupling with *bis*-

THF activated carbonate **8** in the presence of DIEA provided the target compounds **15a-c**. A pair of previously reported PIs **4** and **5** containing the dimethylphenoxy acetate and *bis*-THF moieties as the P2 or P2' ligands in the Phe-Phe isostere were also synthesized using the same reaction sequences.

Scheme 1. Synthesis of Protease Inhibitors Incorporating *Bis*-Tetrahydrofurn in a Pseudo-Symmetric Phe-Phe Dipeptide Isostere



1
2
3
4 Reagents and conditions: (a) K_2CO_3 , KOH, BnCl, H_2O , 100 °C, 5 h;
5 (b) $NaNH_2$, CH_3CN , MTBE, 0 °C, 2 h, then BnMgCl (2 M in THF), RT,
6 14 h; (c) $NaBH_4$, MsOH, *i*-PrOH, EGDME, 0 °C, 12 h; (d) $NaBH_4$, TEA,
7 DMA, 0–15 °C, 2 h; (e) DIEA, CH_3CN , RT, 24 h; (f) Pd/C, HCO_2NH_4 ,
8 MeOH, 50 °C, 12 h; (g) *N*-(methoxycarbonyl)-capped amino acid, HATU,
9 DIEA, DMF, RT, 4 h.
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

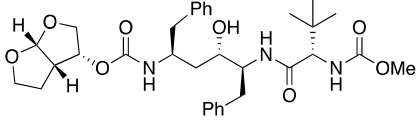
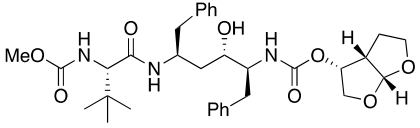
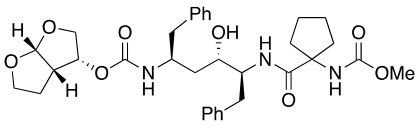
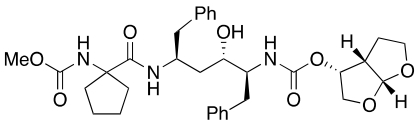
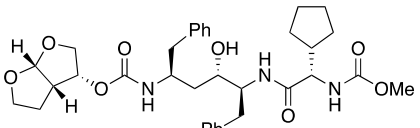
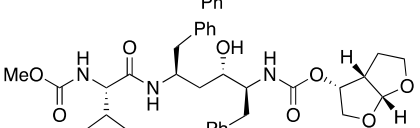
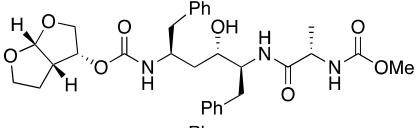
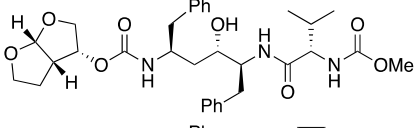
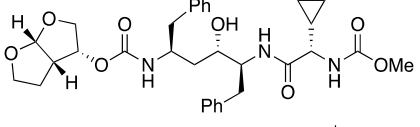
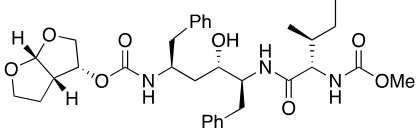
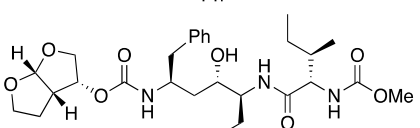
RESULTS AND DISCUSSION

We explored hybrid HIV-1 PIs containing the *bis*-THF moiety of DRV in the Phe-Phe isostere of LPV, with the goal to characterize the binding interactions of this P2 ligand in the context of a pseudo- C_2 -symmetric dipeptide isostere. The *bis*-THF moiety was used in combination with diverse *N*-(methoxycarbonyl)-capped amino acid derivatives similar to the P2/P2' moiety of ATV to optimize hydrogen bonding and vdW interactions in both the S2 and S2' subsites of HIV-1 protease. To explore subsite preference for the *bis*-THF moiety, pairs of compounds were synthesized with the *bis*-THF moiety attached at either the P2 (**12a-c**) or P2' (**15a-c**) position with respect to the central hydroxyl group of the Phe-Phe isostere. Considering potential subsite interdependence, incorporating a relatively small and conformationally flexible moiety at the other P2/P2' position was expected to allow the *bis*-THF moiety to optimally interact with the backbone atoms in the S2/S2' subsite, likely improving overall inhibitor binding with the protease. Structure-activity relationships were explored by incorporating hydrophobic amino acids of varying size and shape to identify an optimal ligand for targeting the other S2/S2' subsite.

Enzyme Inhibition and Antiviral Assays. The potency of PIs was assessed using biochemical and antiviral assays. The enzyme inhibition constants (K_i) were determined against wild-type

1
2
3 protease and two drug-resistant variants (I84V and I50V/A71V)
4
5 using a highly sensitive fluorogenic assay (**Table 1**).³³ Both I50V
6
7 and I84V are major protease inhibitor resistance mutations that
8
9 reduce susceptibility to LPV and DRV. For a subset of compounds,
10
11 antiviral potencies (EC₅₀) were determined against wild-type HIV-
12
13 1 (NL4-3 strain) using a cell-based antiviral assay. LPV and DRV
14
15 were used as controls in all assays.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Enzyme Inhibitory Activity and Antiviral Potency of Hybrid HIV-1 Protease Inhibitors

Inhibitor	Structure	K_i (nM) ^a			EC_{50} (nM) ^b
		WT	I84V	150V/A71V	WT
12a		0.018 ± 0.003	0.208 ± 0.014	0.273 ± 0.043	11
15a		0.013 ± 0.001	0.372 ± 0.021	0.072 ± 0.005	39
12b		46 ± 3	236 ± 39	484 ± 70	NT
15b		7.3 ± 0.5	191 ± 30	40 ± 3	NT
12c		0.235 ± 0.016	0.71 ± 0.10	2.35 ± 0.39	NT
15c		0.161 ± 0.011	1.18 ± 0.26	1.39 ± 0.12	NT
12d		0.019 ± 0.004	0.069 ± 0.013	0.073 ± 0.003	66
12e		0.074 ± 0.006	0.302 ± 0.022	0.295 ± 0.082	55
12f		0.571 ± 0.076	1.82 ± 1.50	1.66 ± 0.48	NT
12g		0.055 ± 0.004	0.198 ± 0.014	0.560 ± 0.094	70
12h		0.027 ± 0.004	0.091 ± 0.013	0.029 ± 0.003	240

4		0.017 ± 0.002	0.069 ± 0.005	0.072 ± 0.002	38
5		0.037 ± 0.005	0.259 ± 0.055	NT	16
LPV		< 0.005	0.051 ± 0.004	0.061 ± 0.011	12
DRV		< 0.005	0.025 ± 0.006	0.075 ± 0.006	5

^aEnzyme inhibition constants against wild-type HIV-1 protease and drug resistant variants; ^bAntiviral potency against wild-type HIV-1; NT = Not Tested

Compounds **12a** and **15a**, with the *bis*-THF moiety at the P2 and P2' position, respectively, and the *N*-(methoxycarbonyl)-*tert*-leucine at the opposite P2/P2' position, showed excellent inhibitory potencies against wild-type protease and the drug resistant variants. Compared to DRV and LPV, both compounds were less active, particularly against the I84V variant. Interestingly, compounds **12a** and **15a** exhibited similar K_i values, though the latter was 4-fold more active against the I50V/A71V variant, indicating no clear S2/S2' subsite preference for the *bis*-THF moiety. However, in antiviral assays compound **12a** was 3-fold more potent than **15a** against wild-type HIV-1. The compound **12a**, with the *bis*-THF moiety at the P2 position, exhibited excellent antiviral potency (EC_{50} = 11 nM) comparable to that of LPV (EC_{50} = 12 nM) and DRV (EC_{50} = 5 nM). A similar potency profile was observed for previously reported

1
2
3 compounds **4** and **5** containing the *bis*-THF and the dimethylphenoxy
4
5 acetate moieties as the P2 and P2' ligands. Compound **5** exhibited
6
7 2-fold better antiviral potency than **4**, indicating a reverse
8
9 subsite preference for the *bis*-THF moiety relative to the **12a/15a**
10
11 pair. The small variation in the antiviral potencies of compounds
12
13 **12a/15a** and **4/5** indicate that the position of the *bis*-THF moiety
14
15 with respect to the central hydroxyl group of the Phe-Phe isostere,
16
17 though not critical for binding, may affect potency in cellular
18
19 assays.
20
21
22

23
24 Although the potency of the hybrid compounds was not particularly
25
26 sensitive to the position of the *bis*-THF moiety, compounds with
27
28 different hydrophobic amino acid derivatives at the opposite P2/P2'
29
30 position showed markedly different inhibitory potencies. The
31
32 compounds with the cycloleucine moiety, **12b** and **15b**, were
33
34 significantly less active (>2500-fold and 560-fold, respectively)
35
36 than the corresponding *tert*-leucine analogues **12a** and **15a**. The
37
38 dramatic loss of potency is likely due to the altered shape and
39
40 flexibility of the hydrophobic amino acid moiety. This was further
41
42 supported by the fact that the corresponding cyclopentylglycine
43
44 analogues **12c** and **15c**, though also an order of magnitude less
45
46 active than **12a** and **15a**, still exhibited sub-nanomolar to low nM
47
48 potency against wild-type protease and drug-resistant variants.
49
50 Together these data confirm that the hydrophobic amino acid moiety
51
52
53
54
55
56
57
58
59
60

1
2
3 at the opposite P2/P2' position significantly influence potency in
4
5 both biochemical and antiviral assays.
6
7

8 Since potency of the compounds with *bis*-THF moiety attached to
9
10 the Phe-Phe isostere largely depended on the hydrophobic amino
11
12 acid moiety at the opposite side of the core scaffold, we
13
14 investigated size and shape requirements for optimally targeting
15
16 the S2' subsite. Analogues of compound **12a** were synthesized with
17
18 diverse hydrophobic amino acid derivatives as P2' ligands while
19
20 keeping the *bis*-THF moiety at the P2 position. The alanine analogue
21
22 **12d** exhibited excellent inhibitory potency with K_i values against
23
24 resistant variants comparable to that of LPV and DRV. The
25
26 corresponding valine analogue **12e** was 4-fold less active against
27
28 wild-type protease but maintained similar inhibitory potency as
29
30 **12a** against the resistant variants. However, compared to **12a**
31
32 compounds **12d** and **12e** were 6- and 5-fold less active in antiviral
33
34 assays, respectively, indicating that factors other than binding
35
36 affinity are responsible for lower antiviral potency. The
37
38 cyclopropyl analogue **12f** was considerably less active than **12a**,
39
40 with K_i values similar to the cyclopentyl analogue **12c**, strongly
41
42 suggesting that cyclic moieties at the P2/P2' position are
43
44 unfavorable. Replacement of *tert*-leucine with an isoleucine
45
46 (compound **12g**) or alloisoleucine (compound **12h**) moiety at the P2'
47
48 position also resulted in slightly lower activity against wild-
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 type protease but similar or better potency against resistant
4
5 variants. This further confirmed earlier findings that the P2/P2'
6
7 group opposite to the *bis*-THF moiety on the Phe-Phe core has a
8
9 significant impact on potency. These results are in contrast to
10
11 the sulfonamide-based dipeptide isostere of DRV, where diverse P2'
12
13 modifications result in relatively minor changes in potency.⁶ The
14
15 identity of the P2/P2' moiety opposite to the *bis*-THF on the Phe-
16
17 Phe core considerably impacts potency in both biochemical and
18
19 antiviral assays.
20
21
22
23

24 **Analysis of Protease-Inhibitor Complexes.** To characterize
25
26 protease-inhibitor molecular interactions, twelve high-resolution
27
28 (1.8–2.0 Å) crystal structures of hybrid PIs and compound **4** bound
29
30 to wild-type HIV-1 protease of the NL4-3 strain were determined.
31
32 A cocrystal structure of LPV was also determined bound to the same
33
34 protease enzyme for direct comparison. The crystallographic data
35
36 collection and refinement statistics are summarized in **Table S1**.
37
38 Following convention, protease chains were assigned chain A (non-
39
40 prime) or chain B (prime) depending on the interactions between
41
42 the central hydroxyl group of the Phe-Phe isostere and the
43
44 catalytic aspartates Asp25/Asp25'.
45
46
47
48

49 LPV and DRV achieve low picomolar affinity to wild-type HIV-1
50
51 protease by making a number of hydrogen bonds and vdW interactions
52
53 with the active site residues (**Figure S1**). The interactions of the
54
55
56
57
58
59
60

1
2
3 transition state mimetic secondary hydroxyl group with the
4 catalytic aspartates, together with the vdW interactions of the
5 adjacent P1/P1' moieties, are critical for overall inhibitor
6 binding. Together these interactions determine the positions of
7 the P2/P2' moieties in the active site, which in turn determine
8 the patterns of hydrogen bonding and vdW interactions of these
9 moieties with the protease. DRV contains a sulfonamide-based Phe-
10 Leu dipeptide isostere core with the (*R*)-configuration of the
11 secondary hydroxyl group, whereas LPV has a pseudo- C_2 -symmetric
12 Phe-Phe dipeptide isostere with the (*S*)-configuration of the
13 hydroxyl group. Though chemically quite distinct, the core
14 scaffolds of LPV and DRV make largely similar hydrogen bonding and
15 vdW interactions with the protease. These inhibitors, however,
16 differ in their interactions with the protease in the S2 and S2'
17 subsites.

18
19 The *bis*-THF moiety of DRV forms direct hydrogen bonds with the
20 main-chain NH of Asp29 and Asp30 in the S2 subsite, while the 4-
21 aminobenzene makes a direct hydrogen bond with the main-chain
22 carbonyl of Asp30' as well as a water-mediated interactions with
23 the side-chain carboxylate of the same residue in the S2' subsite
24 (**Figure S1**). The cyclic urea moiety of LPV makes direct hydrogen
25 bonds with the main-chain NH and side-chain carboxylate of Asp29.
26 The same moiety also makes water-mediated interactions with the
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 main-chain carbonyl of Gly27 and the side-chain carboxylate of
4
5 Asp29 in the S2 subsite. However, in contrast to the P2' moiety of
6
7 DRV, the dimethylphenoxy acetate moiety of LPV only makes
8
9 hydrophobic interactions in the S2' subsite (**Figure S1**). The hybrid
10
11 compounds combine the pseudo-C₂-symmetric Phe-Phe isostere of LPV
12
13 with the *bis*-THF moiety of DRV and an ATV-like amino acid moiety
14
15 to enhance direct hydrogen bonding and vdW interactions in the S2
16
17 and S2' subsites.
18
19
20

21
22 The cocrystal structures of hybrid compounds revealed similar
23
24 binding conformations of the Phe-Phe isostere and the *bis*-THF
25
26 moiety as observed in the protease complexes with LPV and DRV,
27
28 respectively (**Figure 2**). Importantly, the *bis*-THF moiety is
29
30 positioned to interact with the main-chain NH of residues Asp29
31
32 and Asp30, mimicking the binding interactions of DRV (**Figure 3**).
33
34 This binding conformation is maintained in all hybrid compounds
35
36 including **4** regardless of the position of the *bis*-THF moiety. As
37
38 a result, within each compound series (with the *bis*-THF moiety at
39
40 the P2 and P2' position respectively), the structures aligned well
41
42 with variations occurring only at the hydrophobic amino acid moiety
43
44
45
46
47 (**Figure 4**).
48
49
50
51
52
53
54
55
56
57
58
59
60

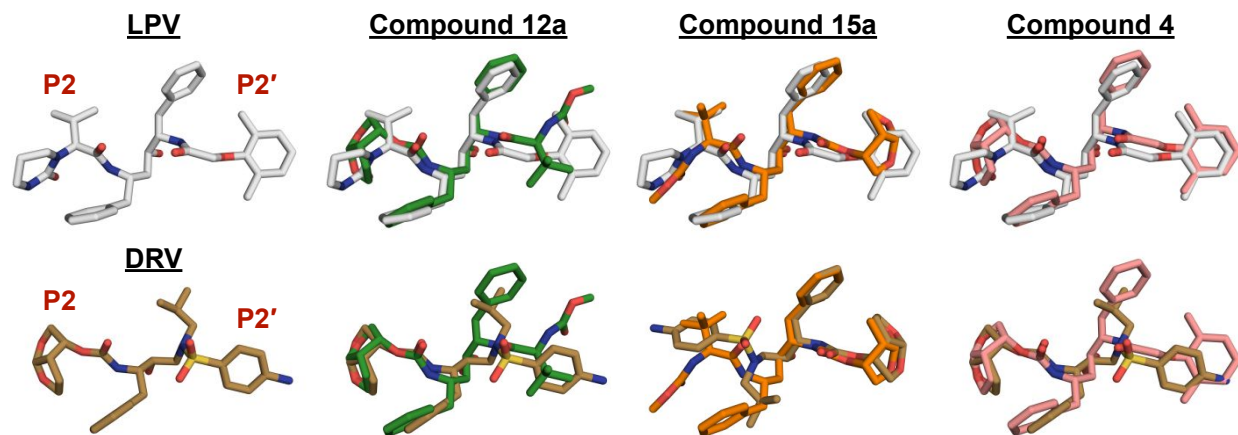


Figure 2. Overlays of protease cocrystal structures bound to LPV (PDB 6PJB) (grey, top) and DRV (6DGX) (brown, bottom) with compounds **12a** (*bis*-THF at P2; PDB 6PJD) (green), **15a** (*bis*-THF at P2'; PDB 6PJM) (orange), and **4** (PDB 6PJC) (salmon) revealed similar binding conformations of the Phe-Phe isostere and the *bis*-THF moiety as observed in the protease complexes with LPV and DRV, respectively. Note that DRV is flipped 180° to compare with *bis*-THF of **15a** at the P2' position.

In all complexes, the central hydroxyl group of the Phe-Phe core maintains the key hydrogen bonds with the side-chain carboxylates of the catalytic residues Asp25/Asp25'. Nevertheless, subtle differences in hydrogen bonding and vdW interactions were observed between the Phe-Phe cores in LPV and the hybrid compounds. Notably, one of the four water-mediated interactions with the flap residue Ile50 is consistently longer (2.9–3.4 Å) than what is observed in the LPV structure (2.7 Å) (**Table S2**). The two amide NH groups on

1
2
3 each side of the Phe-Phe core could potentially interact with the
4 backbone carbonyl of Gly27 and Gly27', but in the LPV structure
5
6 only the Gly27 is close enough to form a hydrogen bond (2.9 Å
7
8 versus 3.9 Å) (**Figure 3**). This phenomenon is also observed in the
9
10 protease complex with compound **4** (2.9 Å versus 3.7 Å) indicating
11
12 that the large hydrophobic P2' moiety prevents the additional
13
14 hydrogen bond from occurring with Gly27'. In contrast, in all hybrid
15
16 compound structures the core NH groups are within 3.5 Å from the
17
18 backbone carbonyls of Gly27 and Gly27' (**Table S2**), although
19
20 simultaneous interactions with both Gly27/Gly27' may not be stable.
21
22 Nevertheless, the combination of *bis*-THF and a flexible,
23
24 hydrophobic amino acid moiety allow both core NH groups to interact
25
26 with the main-chain carbonyls of residues Gly27/Gly27'.
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

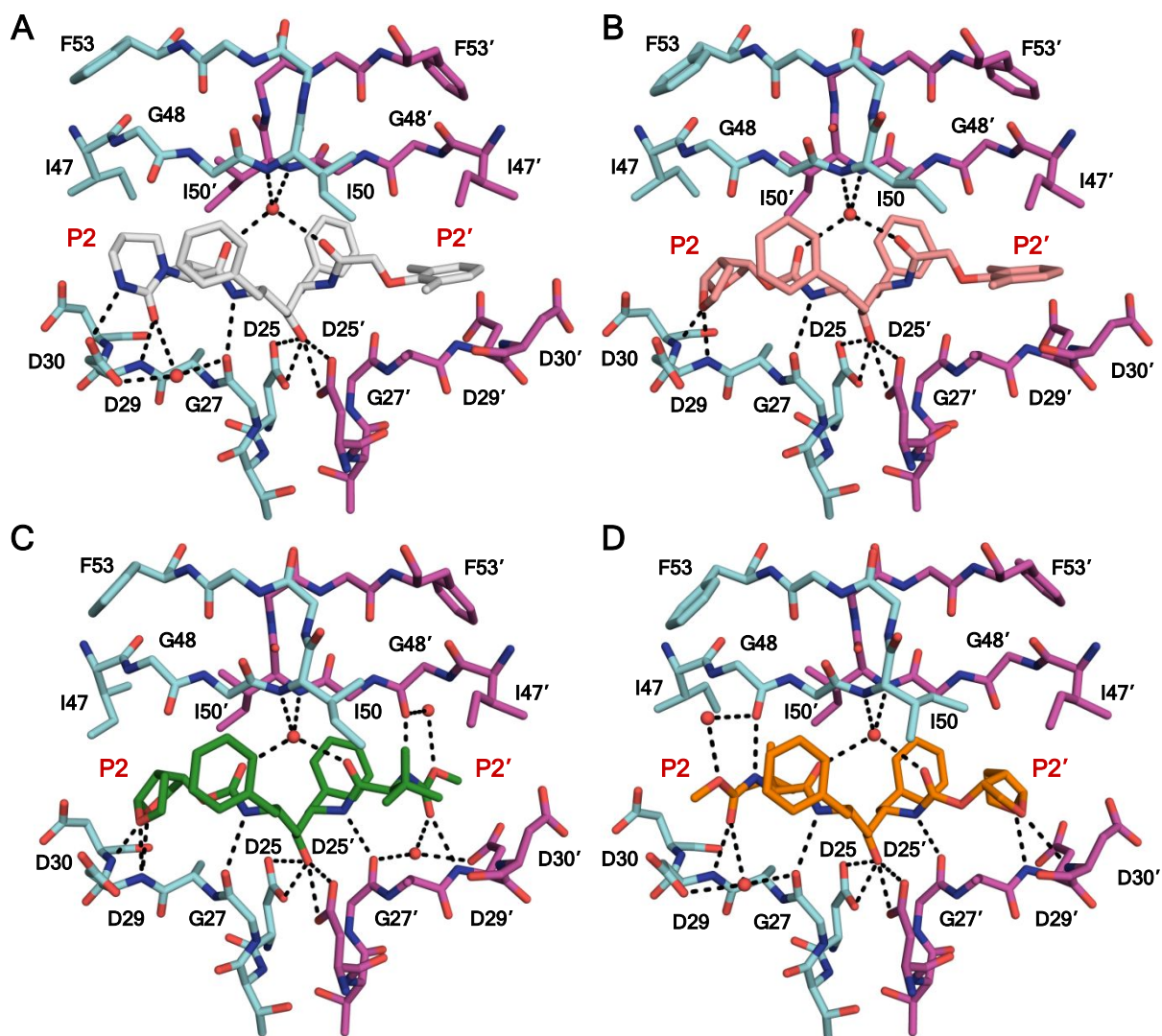


Figure 3. Crystal structures of wild-type HIV-1 protease in complex with LPV (PDB 6PJB) (A) and hybrid compounds **4** (PDB 6PJC) (B), **12a** (PDB 6PJD) (C), and **15a** (PDB 6PJM) (D). The *bis*-THF moiety maintains the hydrogen bonding interactions with the main-chain NH of Asp29/Asp29' and Asp30/Asp30' in the S2/S2' subsites regardless of its position on the Phe-Phe core. Hydrogen bonds and water-mediated interactions as determined by PyMol are shown as black dashed lines with waters shown as red spheres.

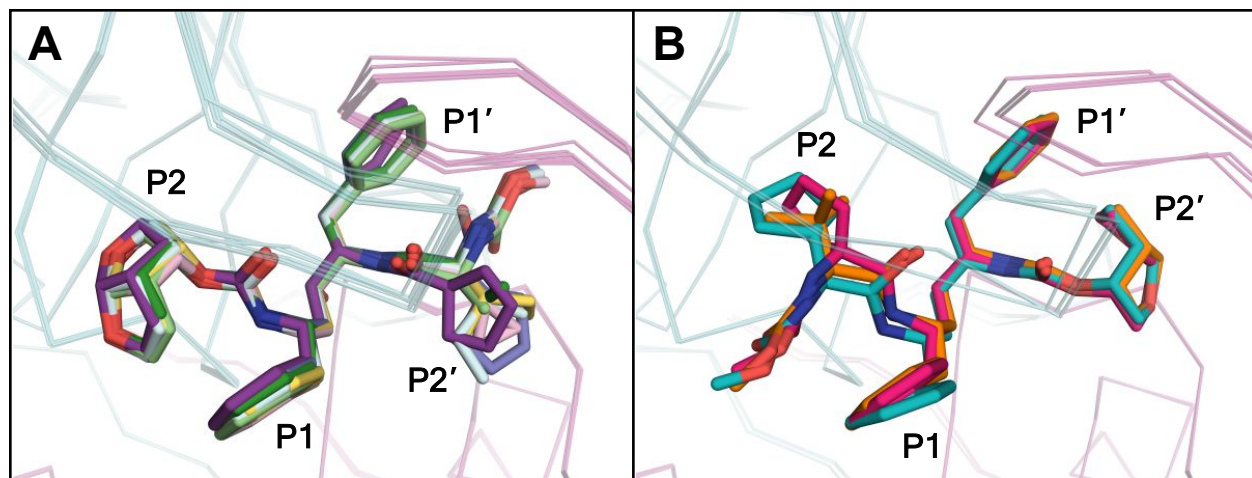


Figure 4. Superposition of HIV-1 protease complexes with compounds **12a-12h** (A) and compounds **15a-15c** (B), showing similar binding conformations of the *bis*-THF moiety regardless of the position and minor differences at the Phe-Phe isostere, but major differences at the P2/P2' amino acid moiety. The protease active site is shown in a ribbon representation, with bound inhibitors depicted as sticks.

The Phe-Phe isostere in hybrid compounds, in addition to maintaining key hydrogen bonding interactions with the protease, makes a number of favorable vdW interactions with the hydrophobic residues that make up the S1/S1' subsites. Despite the same phenylalanine side chains at the P1 and P1' positions, there are a few notable differences in vdW contacts with the protease (**Figure S2**). Compared to LPV, the P1/P1' phenyl group adjacent to the amino acid moiety is shifted towards Val82 resulting in improved vdW contacts with that residue, particularly for compounds with larger

1
2
3 P2/P2' amino acid moieties. These localized gains in vdW contacts
4
5 are slightly lower for compounds **15a-c** with *bis*-THF at the P2'
6
7 position than their corresponding analogues **12a-c**. The shift of
8
9 the P1/P1' phenyl group towards Val82, creating more vdW contacts,
10
11 resulted in a nearly equal and opposite loss of contacts with Ile50
12
13 in the same subsite. The packing of the phenyl group adjacent to
14
15 the *bis*-THF moiety resulted in a small increase in vdW contacts
16
17 with the 80's loop (residues 81, 82 and 84) but decrease in
18
19 contacts with flap residue Gly49. For compounds **15a-c**, gains in
20
21 vdW contacts are smaller and the losses are greater. Overall,
22
23 despite localized losses and gains in vdW interactions for both
24
25 series of compounds, the packing of the P1 and P1' groups remain
26
27 generally similar to LPV.
28
29
30
31
32

33 The cocrystal structures of hybrid compounds (and **4**) also
34
35 revealed similar binding interactions of the *bis*-THF moiety as in
36
37 DRV-protease complex regardless of the position, with the two ring
38
39 oxygen atoms positioned to form hydrogen bonds with the main-chain
40
41 NH of Asp29 and Asp30. However, there are subtle differences in
42
43 the position and puckering of the *bis*-THF rings, causing the
44
45 carbamate-linked THF ring oxygen atom to slightly shift away from
46
47 the backbone NH of Asp30 compared to DRV (3.1-3.4 Å versus 2.9 Å)
48
49 (**Figure 5**). Whereas the oxygen atom of the other THF ring maintains
50
51 a similar distance to the backbone NH of Asp29 (2.7-2.9 Å versus
52
53
54
55
56
57
58
59
60

1
2
3 2.8 Å) (**Table S2**). For compounds **12a-c**, these subtle changes in
4 the *bis*-THF moiety arise from the differences in the binding of
5 core isosteres, where the P1 benzyl group and the carbamate are
6 shifted towards the B-chain compared to their position in DRV.
7 This is likely due to the additional carbon in the core or to
8 accommodate the larger P1' group. In compounds **15a-c**, the core
9 scaffold from the central hydroxyl to the *bis*-THF moiety is
10 identical. However, the opposite stereochemistry of the central
11 hydroxyl causes a subtle shift preventing the *bis*-THF moiety from
12 optimally interacting with Asp30' (**Figure 5**). Despite these minor
13 differences in the position of the *bis*-THF moiety, the hybrid
14 compounds make similar hydrogen bonding interactions with the
15 protease in the S2/S2' subsite compared to DRV.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

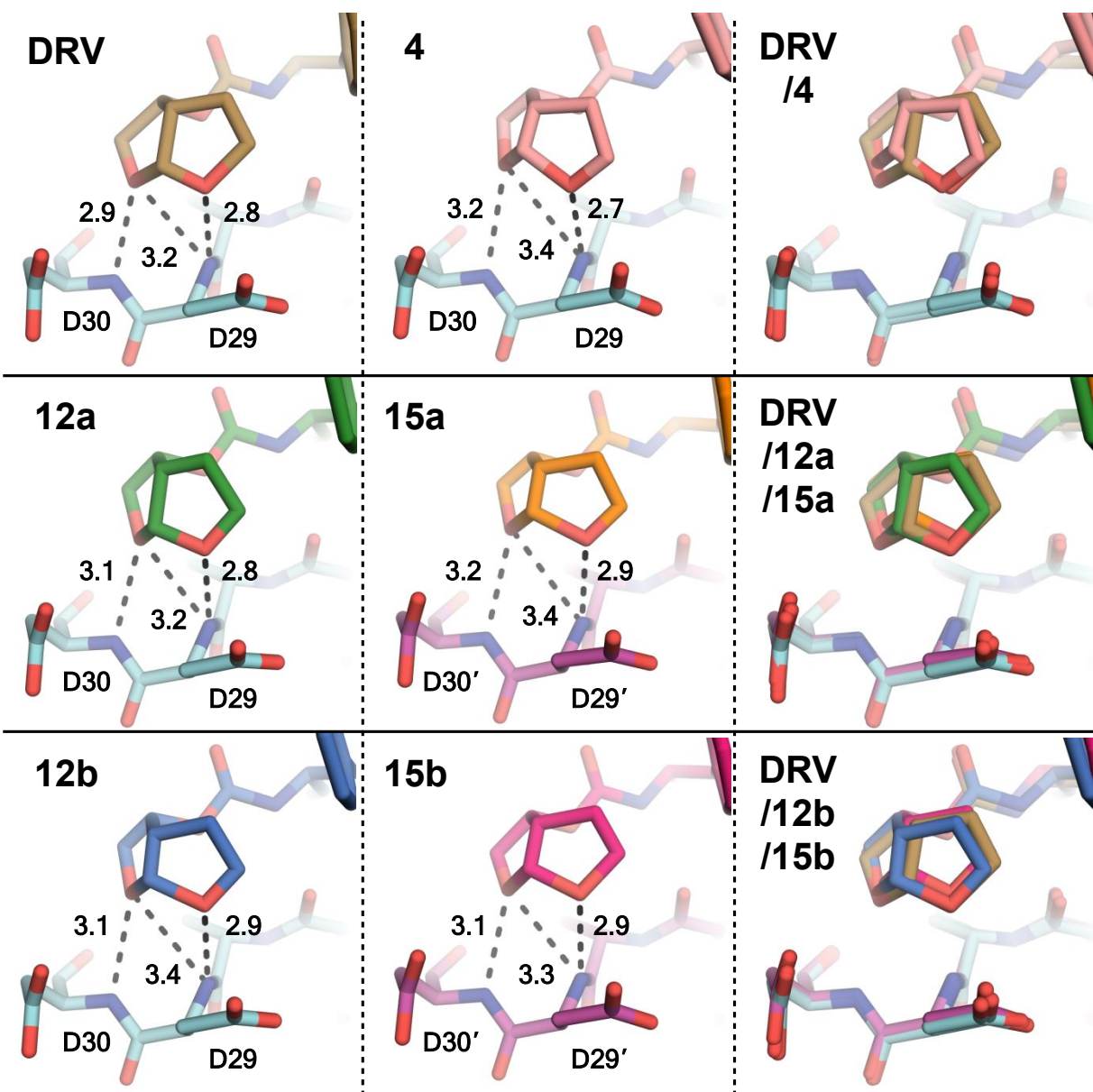


Figure 5. Comparison of the binding conformations of the *bis*-THF moiety and protease-inhibitor intramolecular distances with the main-chain NH of Asp29 and Asp30 in the S2/S2' subsite of HIV-1 protease. As observed in the DRV-protease complex (PDB 6DGX), the *bis*-THF moiety retained similar interactions with the protease in complexes with compound **4** (PDB 6PJC) (top), the most potent hybrid

1
2
3 compounds **12a** (PDB 6PJD) and **15a** (PDB 6PJM) (middle), and the least
4
5 potent hybrid compounds **12b** (PDB 6PJE) and **15b** (PDB 6PJN) (bottom),
6
7 irrespective of position on the Phe-Phe core.
8
9

10 The *bis*-THF moiety of DRV is critical to its potency, not only
11
12 for the hydrogen bonding interactions with Asp29 and Asp30, but
13
14 also vdW packing with residues in the S2 subsite, particularly
15
16 Val32 and Ile47. In hybrid compounds, the minor change in the
17
18 position and puckering of the *bis*-THF moiety causes reduced vdW
19
20 interactions with Ala28 and Asp30, but increased interactions with
21
22 Ile47 and Gly48 (**Figure S3**). The main-chain carbonyl of Gly48 is
23
24 oriented towards the hydrophobic region of *bis*-THF, so enhanced
25
26 contacts with this residue may be unfavorable. In contrast,
27
28 compounds **4**, **15a** and **15b** do not make these unfavorable contacts
29
30 with Gly48. With the exception of enhanced interactions with Ile47,
31
32 the *bis*-THF moiety in most hybrid compounds make less optimal vdW
33
34 interactions with the protease compared to those of DRV.
35
36
37
38
39

40 The major difference between the hybrid compounds and the
41
42 previously reported Phe-Phe core-based *bis*-THF containing PIs **4**
43
44 and **5** is the introduction of an ATV-like amino acid moiety to
45
46 target the opposite S2/S2' subsite. The amino acid side chain
47
48 interacts with hydrophobic protease residues, while the *N*-
49
50 (methoxycarbonyl)-capping group allows for additional hydrogen
51
52 bonding interactions with the protease. As already noted, LPV's
53
54
55
56
57
58
59
60

1
2
3 P2' moiety does not form any hydrogen bonds, and DRV's P2' moiety
4
5 makes one direct and one water-mediated interactions in the S2'
6
7 subsite (**Figure S1**). In most hybrid compounds, the carbonyl oxygen
8
9 of the carbamate makes one direct hydrogen bond with the backbone
10
11 NH of Asp29' (Asp29), and the carbamate NH makes another direct
12
13 hydrogen bond with the backbone carbonyl of Gly48' (Gly48) in the
14
15 flaps (**Figure 6**). There are also a number of coordinated water
16
17 molecules that allow for additional water-mediated interactions.
18
19 The carbonyl oxygen of the carbamate coordinates the same water-
20
21 mediated interactions as the cyclic urea moiety of LPV, with the
22
23 backbone carbonyl of Gly27' (Gly27) and the side-chain carboxylate
24
25 of Asp29' (Asp29) (**Figure 6**). This network of polar interactions
26
27 is maintained by all amino acid moieties in both the S2' and S2
28
29 subsites except the cycloleucine in compound **12b**.
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

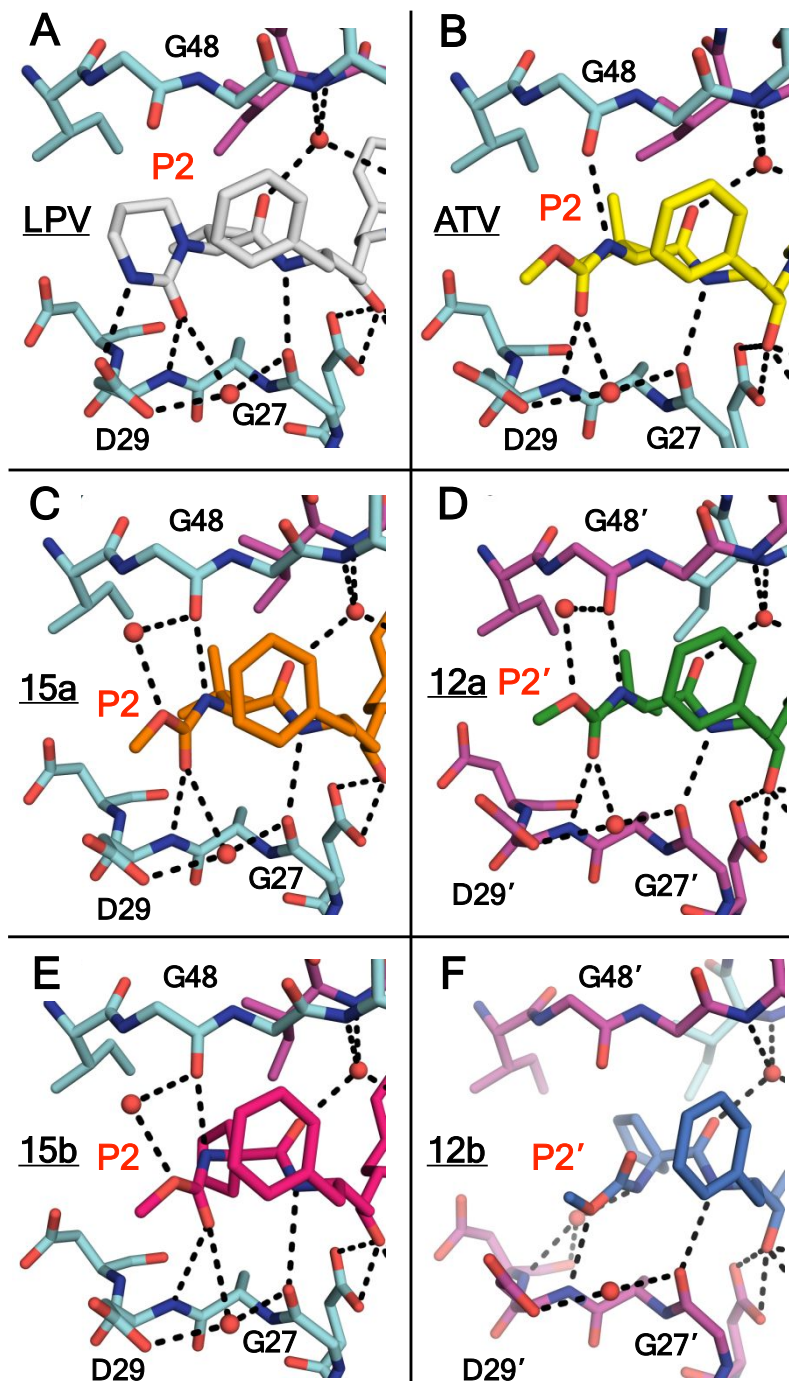


Figure 6. Comparison of protease-inhibitor direct hydrogen bonds and water-mediated interactions in the S2/S2' subsites for the cyclic urea moiety of LPV (PDB 6PJB) (A), *N*-(methoxycarbonyl)-*tert*-leucine of ATV (PDB 3EKY) (B), and the *N*-(methoxycarbonyl)-

1
2
3 capped amino acid moieties of hybrid compound **15a** (PDB 6PJM) (C),
4
5 **12a** (PDB 6PJD) (D), **15b** (PDB 6PJN) (E), and **12b** (PDB 6PJE) (F).
6
7 Polar interactions are shown as black dashed lines (as determined
8
9 by PyMol) with waters shown as red spheres.

10
11
12
13 The carbamate of compound **12b** is flipped relative to all other
14
15 inhibitors losing the direct hydrogen bond to flap residue Gly48'
16
17 (**Figure 6F**). This flip positions the carbonyl oxygen of the
18
19 carbamate close to the carbonyl oxygen of Gly48' (3.3 Å), creating
20
21 a very unfavorable interaction and preventing compound **12b** from
22
23 utilizing the same water-mediated interactions as LPV and most
24
25 hybrid compounds. Instead **12b** forms a hydrogen bond with a
26
27 "backside" water in the S2' subsite, creating a water-mediated
28
29 interaction with the main-chain carbonyl and NH of Asp30' (**Figure**
30
31 **6F**). The marked changes in the hydrogen bonding interactions in
32
33 the S2/S2' subsite likely underlie the greatly reduced potency of
34
35 **12b** compared to all other hybrid compounds. Unlike **12b**, the P2
36
37 carbamate group in compound **15b** maintained a similar conformation
38
39 as in the *tert*-leucine analogue **15a** but with a slight change in
40
41 the position of the carbonyl oxygen away from Asp29. This shift
42
43 resulted in increased hydrogen bond distance between the P2
44
45 carbamate carbonyl and the main-chain NH of Asp29 compared to **15a**
46
47 (3.3 Å versus 2.9 Å). However, the reduced potency of **15b** could
48
49 not result from minor differences in hydrogen bonding interactions
50
51
52
53
54
55
56
57
58
59
60

1
2
3 alone, highlighting the significance of non-polar interactions for
4
5 potency.
6

7
8 The protease-inhibitor vdW interactions in the S2/S2' subsite
9
10 varied significantly depending on the size and shape of the
11
12 hydrophobic amino acid moiety (**Figure S4**). Compounds **12a** and **15a**
13
14 with the *tert*-leucine make highly distributed vdW contacts with a
15
16 number of residues. In contrast, compared to *tert*-leucine, the
17
18 cycloleucine side chain in **12b** predominantly interacts with Ile50
19
20 and I84V, while losing significant contacts with residues 28-30
21
22 (**Figure 7**). The cycloleucine side chain of **15b** packs in the same
23
24 hydrophobic area but makes less vdW contacts than **12b**. Compound
25
26 **12c** with cyclopentyl glycine also has enhanced vdW contacts with
27
28 Ile50 but without experiencing losses at other residues as observed
29
30 for **12b**. In general, the cyclic hydrophobic amino acid derivatives
31
32 make more localized vdW interactions in the S2/S2' subsite while
33
34 the acyclic moieties make more distributed contacts. Thus, in
35
36 addition to altered polar interactions, the differences in vdW
37
38 contacts due to the size and shape of the hydrophobic amino acid
39
40 moiety are likely responsible for the varied inhibitor potencies.
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

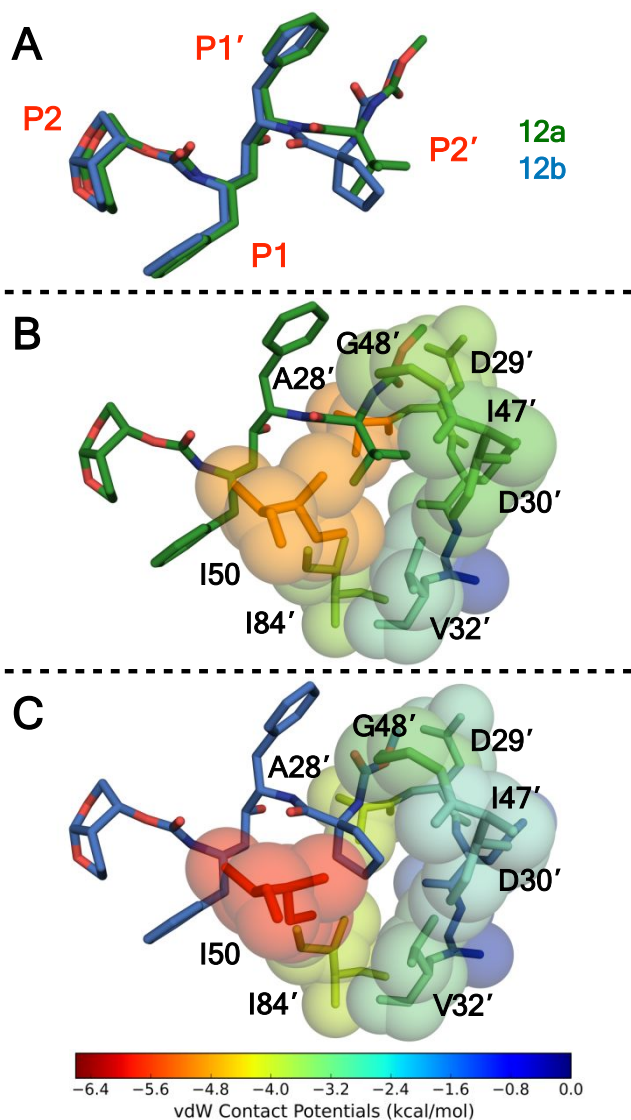


Figure 7. Comparison of binding conformation (A) and van der Waals packing with residues in the S2' subsite for the most and least potent inhibitors **12a** (PDB 6PJD) (B, green sticks) and **12b** (PDB 6PJE) (C, blue sticks). Compound **12a** makes more extensive contacts with the protease as estimated by relative vdW contact energy compared to **12b** (-91 versus -86 kcal/mol), with major difference occurring at the S2' subsite (-27 versus -23 kcal/mol). Protease

1
2
3 residues are colored red to blue for highest to lowest vdW
4
5 interactions with the inhibitor.
6
7

8 **Molecular Dynamics (MD) Simulations.** Starting from the cocrystal
9
10 structures, molecular dynamics (MD) simulations were utilized to
11
12 interrogate the stability of protease-inhibitor interactions in
13
14 the dynamic ensemble of the complexes. Three replicates of fully
15
16 hydrated 100 ns MD simulations were performed on each protease-
17
18 inhibitor complex. All simulations reached convergence (**Figure**
19
20 **S5**).
21
22
23

24 The protease active site and flaps dynamics were analyzed by
25
26 measuring C_{α} distances between specific residues. The analysis
27
28 focused on the most and least potent pairs of compounds containing
29
30 *bis*-THF at either the P2 or P2' position (**12a/15a** and **12b/15b**)
31
32 compared to LPV. The cocrystal structures suggested the flaps and
33
34 the 80's loops surrounding the inhibitor may be perturbed when
35
36 bound to different compounds. To monitor dynamics of the flaps,
37
38 distance distributions between C_{α} atoms of Ile50-Ile84', Ile50'-
39
40 Ile84 and Ile50-Ile50' were calculated during the MD trajectories.
41
42 Distances for Ile84-Ile84' and Pro81-Pro81' were used to probe the
43
44 expansion or narrowing of the "lower" and "upper" side wall of the
45
46 active site, respectively (**Figure S6**). In protease complexes with
47
48 hybrid compounds, the active site remained largely unchanged in
49
50 the "lower" part of the active site as indicated by a single,
51
52
53
54
55
56
57
58
59
60

1
2
3 narrow distribution of Ile84-Ile84' distance centered around 15.2
4
5 Å. However, compared to LPV the shift in the Pro81-Pro81' distance
6
7 distribution with compounds **12a**, **15a**, **12b**, and **15b** indicated the
8
9 active site was narrower in the "upper" portion close to the flaps.
10
11 In addition, the active site was asymmetrically longer (Ile50'-
12
13 Ile84 or Ile50-Ile84'), sampling bimodal distance distributions,
14
15 suggesting the sampling of a semi-open conformation of the flaps. This was supported
16
17 by the Ile50-Ile50' distance distribution, with one major peak around 5.3
18
19 Å (closed flaps) and another ~10 Å corresponding to semi-open flaps
20
21 with compounds **15a** and **15b** (**Figure S6**). Increased distance between
22
23 residues Ile50 and Ile50' in the relatively short time
24
25 (nanoseconds) during MD simulations has been previously suggested
26
27 to indicate flap opening.^{34, 35} Thus, perturbed flaps dynamics and
28
29 narrowing of the active site may indicate unfavorable binding of
30
31 these compounds compared to LPV, which is supported by the *K_i* data.
32
33
34
35
36
37

38 Next, the stability of inhibitor binding was analyzed by
39
40 calculating root-mean-square fluctuations (RMSF) of each atom
41
42 (**Figure 8**). In all complexes, the *bis*-THF moiety often displayed
43
44 the lowest RMSF amongst all moieties, regardless of its position
45
46 on the Phe-Phe core. Notably, even in the least active compounds
47
48 with nanomolar potency (**12b** and **15b**), the *bis*-THF moiety showed
49
50 fluctuations similar to that observed in **12a** and **15a**. Moreover,
51
52 the phenylalanine side chain adjacent to the *bis*-THF moiety
53
54
55
56
57
58
59
60

1
2
3 displayed lower RMSF compared to the distal phenylalanine side
4 chain, despite the pseudo-symmetric nature of the Phe-Phe core.
5
6 The P2/P2' amino acid moieties showed varying RMSF profiles,
7
8 independent of the size or shape of the hydrophobic side chain.
9
10 The phenylalanine adjacent to the amino acid moiety consistently
11 displayed greater RMSF, with the exception of **12e**. The varying
12 fluctuations of the P1/P1' moiety depending on the size and shape
13 of the P2/P2' moiety indicate interdependence between the S1/S1' and
14 S2/S2' subsites. The P2/P2' amino acid moiety also affected the
15 fluctuations of the *bis*-THF moiety but to a much lesser extent. This data
16 suggests that in hybrid compounds the *bis*-THF moiety and the
17 adjacent P1/P1' moiety remain relatively stable in the protease
18 active site regardless of the position on the Phe-Phe core. In
19 contrast, the fluctuations of the amino acid and the adjacent
20 P1/P1' moieties appear to destabilize overall inhibitor binding.
21 Thus, due to interdependence between the S1/S1' and S2/S2' subsites,
22 the optimization of the *bis*-THF containing hybrid PIs, would
23 require modification of the amino acid moiety together with the
24 adjacent P1/P1' group.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

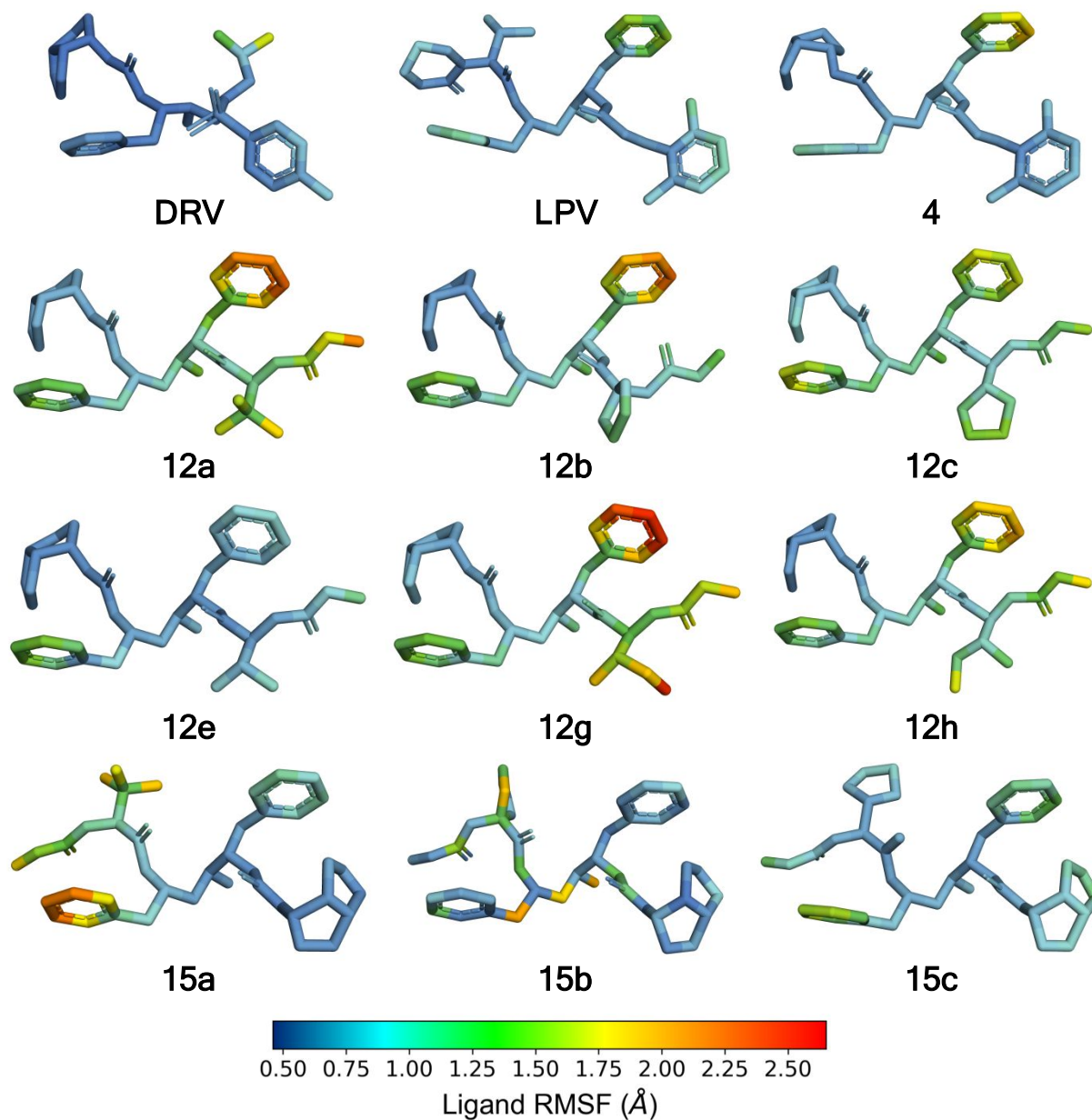


Figure 8. Inhibitor root-mean-square fluctuation (RMSF) of each non-hydrogen atom of DRV, LPV, and hybrid compounds mapped onto their crystal structure. Warmer colors indicate larger fluctuations.

The relationship between inhibitor and protease conformational dynamics was examined via simultaneous monitoring of inhibitor

1
2
3 bond rotations and protease residue distances. The dihedral angles
4
5 of all inhibitor rotatable bonds were calculated throughout the MD
6
7 simulations. Given the bimodal sampling of the flap distances
8
9 (Ile50-Ile50', Ile50-Ile84' and Ile50'-Ile84), we examined whether
10
11 the semi-open conformation of flaps was associated with certain
12
13 inhibitor bond rotations. In protease complexes with compounds **12a**
14
15 and **15a**, the two most potent inhibitors, the separation of the
16
17 flap tips (Ile50-Ile50') was associated with the conformational
18
19 sampling of the dihedral controlling the amino acid moiety (**Figure**
20
21 **9**). In this semi-open flap conformation, the Ile50/Ile50' residues at
22
23 the top of the flaps also lost intra-protease vdW contacts with residues 32, 47-49 and 84 (**Figure**
24
25 **S7**). Loss of intra-protease vdW interactions could result in a less stable protease-inhibitor
26
27 complex, leading to an overall decrease in inhibitor binding affinity.
28
29
30
31
32

33 In addition, protein dynamics were compared by calculating the
34
35 root-mean-square fluctuations (RMSF) of C α atoms. Consistent with
36
37 our observations from the distance distributions, there were
38
39 significant increases in fluctuations at the flaps and 80's loop
40
41 (**Figure S8**). In complex with LPV, the flaps and active site show
42
43 little fluctuations, comparable to DRV. However, compounds
44
45 containing *bis*-THF at the P2' position showed increased
46
47 fluctuations at both flaps, with fluctuations asymmetrically
48
49 greater at the B chain flap, contacting the amino acid moiety. The
50
51 increased protein fluctuation could impact the stability of
52
53 protease-inhibitor hydrogen bonding interactions.
54
55
56
57
58
59
60

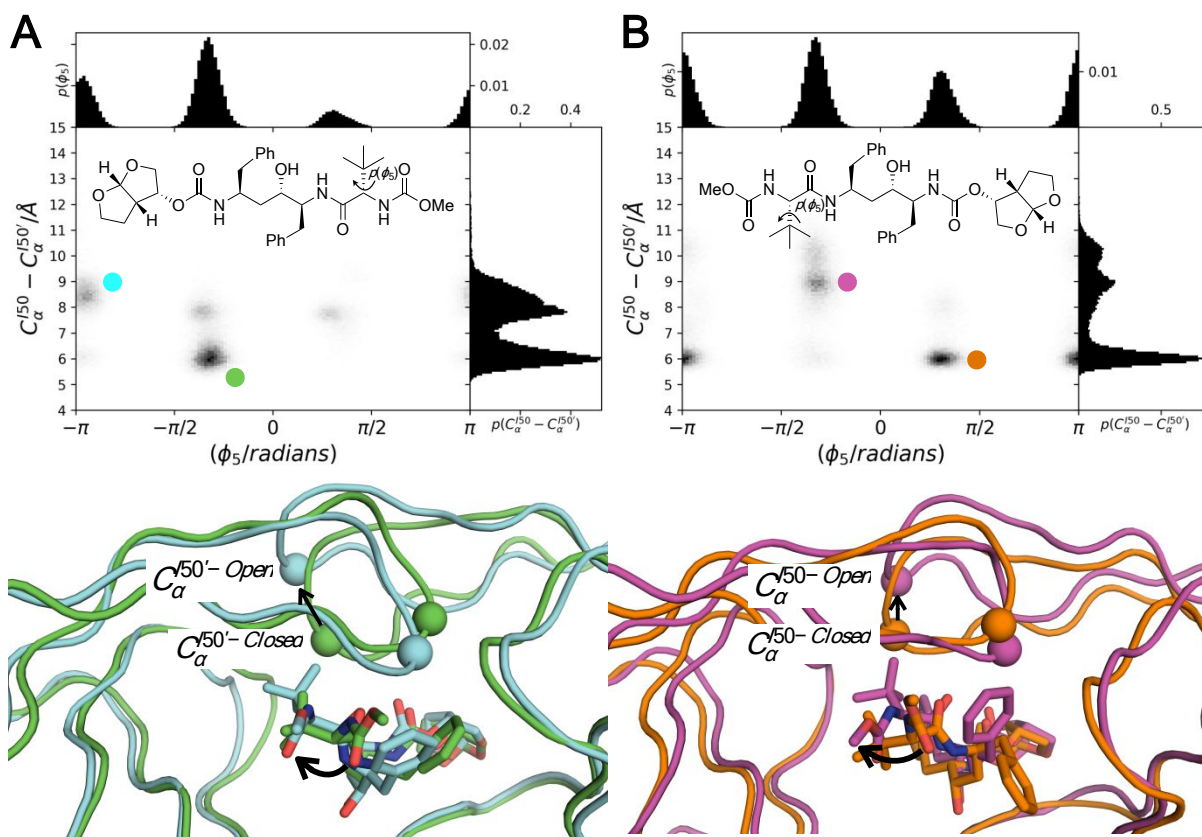


Figure 9. The dihedral angle (ϕ_5) responsible for conformational sampling of the *tert*-leucine moiety on **12a** (A) and **15a** (B) is associated with increased Ile50-Ile50' distance sampling.

In the cocrystal structures, the hybrid compounds make a number of direct and water-mediated interactions with the protease active site, including the central hydroxyl with catalytic residues, coordinated water with Ile50/50', the Phe-Phe core nitrogen atoms with Gly27, *bis*-THF with Asp29 and Asp30, carbamate with Asp29 and Gly48, and the coordinated S3' water with Gly27' and Asp29'. To assess if the increased protein and ligand fluctuations impact hydrogen bonding stability, the patterns and frequencies of these hydrogen bonds were monitored throughout the MD simulations (**Figure 10**, **Figure**

1
2
3 **S9**). In MD simulations, LPV and the hybrid compounds did not
4 maintain all hydrogen bonds observed in the cocrystal structures.
5 LPV and all compounds with *bis*-THF at the P2 position (**4** and **12a-h**)
6 maintained a hydrogen bond with Asp25' (66-99% frequency) but did
7 not form a hydrogen bond with Asp25. Whereas compounds **15a** and **15b**
8 containing *bis*-THF at the P2 position formed the hydrogen bonds
9 with Asp25 and Asp25' at roughly 100% and 50% frequency,
10 respectively. In contrast, the hydrogen bond between compound **15c**
11 and Asp25 was unstable, while the hydrogen bond with Asp25' was highly
12 stable (92% frequency). Hybrid compounds maintained the network of
13 water-mediated interactions with the main-chain NH of Ile50 and
14 Ile50' at varying stability. The interactions with Ile50 were
15 observed to be less stable compared to the ones with Ile50' (21-
16 65% versus 49-78%). For most of the hybrid compounds, the hydrogen
17 bonds between the Phe-Phe core NH groups and the main-chain
18 carbonyls of Gly27/Gly27' were either not observed or were very unstable.

19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40 The *bis*-THF moiety in all hybrid compounds maintained two
41 hydrogen bonds with the main-chain NH of Asp29/Asp29' and
42 Asp30/Asp30' at high frequency, regardless of the position on the
43 Phe-Phe isostere (**Figure 10, Figure S9**). In contrast, the P2/P2' amino
44 acid moiety had varying effects on hydrogen bonding interactions
45 in the S2/S2' subsite. The carbamate nitrogen of hybrid compounds
46 formed a hydrogen bond with the backbone carbonyl of Gly48' at
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 moderate to high frequency, but this interaction was not observed
4
5 for the least potent compound **12b** and was less stable for **15b**. The
6
7 carbonyl oxygen of the P2' amino acid moiety in compounds **12a** and
8
9 **12c** formed relatively stable water-mediated interactions with the
10
11 sidechain of Asp30'. In contrast, the carbonyl oxygen of the P2
12
13 moiety in compounds **15a** and **15c** formed three low-frequency (14-
14
15 39%) water-mediated interactions, bridging the main-chain carbonyl
16
17 of Gly27 and the side-chain carboxylate of Asp29. These data suggest that
18
19 for most hybrid compounds the key hydrogen bonding interactions of the *bis*-THF moiety and the
20
21 Phe-Phe isostere with protease remain relatively stable. In addition, the size and shape of the P2/P2'
22
23 amino acid moiety not only influences the interactions in the S2/S2' subsites but can also propagate
24
25 changes in the water-mediate interactions between the inhibitors and Ile50/Ile50'.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

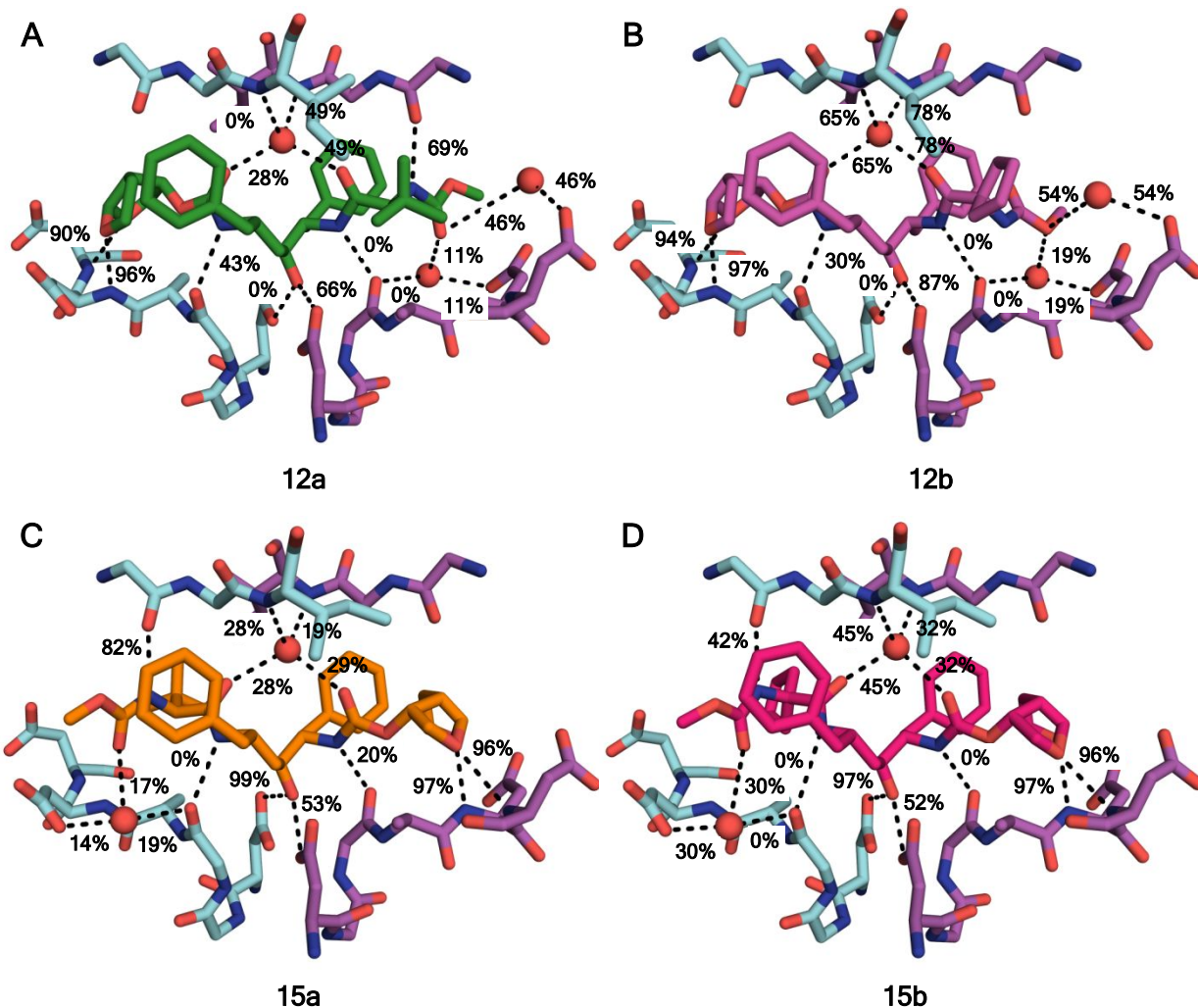


Figure 10. Protease-inhibitor hydrogen bond frequencies from MD simulations mapped onto the cocrystal structures of compounds **12a** (A), **12b** (B), **15a** (C), and **15b** (D). The protease side chains and inhibitors are shown as sticks; chains A and B are colored cyan and magenta, respectively.

CONCLUSIONS

We have explored hybrid HIV-1 PIs containing the *bis*-THF moiety of DRV on either side of the Phe-Phe dipeptide isostere of LPV to ascertain whether this moiety could maintain key interactions with the protease and improve potency against wild-type and primary drug-resistant variants of the enzyme. Most of the hybrid compounds retained picomolar biochemical potency irrespective of the position of *bis*-THF relative to the central hydroxyl group, likely due to the pseudo- C_2 -symmetric nature of the Phe-Phe isostere. But in both compound series the identity of the other P2/P2' moiety significantly affected potency, favoring relatively flexible, hydrophobic moieties. The cocrystal structures of hybrid compounds revealed that the *bis*-THF moiety maintains the key hydrogen bonding interactions with the protease in the S2/S2' subsite as observed in the DRV-protease complex regardless of its position on the Phe-Phe isostere. In contrast, the binding of the hydrophobic amino acid moiety in the other S2/S2' subsite was greatly influenced by the size, shape, and flexibility of the hydrophobic group, which affected hydrogen bonding and vdW interactions with the protease. Moreover, this moiety appears to influence binding of the adjacent P1/P1' group as well as the water-mediated interactions between the inhibitor and flap residues, indicating dynamic interdependence between the protease subsites. This is further supported by the relatively stable binding

1
2
3 interactions of the *bis*-THF moiety and the adjacent P1/P1' group.
4
5 The dynamic interdependence between subsites in the protease
6
7 active site **can be exploited** to optimize inhibitor potency against drug-
8
9 resistant protease variants. The detailed structural
10
11 characterization of hybrid HIV-1 PIs containing *bis*-THF in non-
12
13 sulfonamide dipeptide isosteres offers opportunities for
14
15 structure-guided optimization of these promising inhibitors.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

EXPERIMENTAL SECTION

General. All reactions were performed in oven-dried round-bottom flasks fitted with rubber septa under argon atmosphere unless otherwise noted. All reagents and solvents, including anhydrous solvents, were purchased from commercial sources and used as received. Flash column chromatography was performed on an automated Teledyne ISCO CombiFlash Rf+ system equipped with a UV-vis detector using disposable Redisep Gold high performance silica gel columns or was performed manually using silica gel (230-400 mesh, EMD Millipore). Thin-layer chromatography (TLC) was performed using silica gel (60 F₂₅₄) coated aluminum plates (EMD Millipore), and spots were visualized by exposure to ultraviolet light (UV), exposure to iodine adsorbed on silica gel, and/or staining with alcohol solutions of phosphomolybdic acid (PMA) and ninhydrin followed by brief heating. ¹H NMR and ¹³C NMR spectra were acquired on Varian Mercury 400 MHz and Bruker Avance III HD 500 MHz NMR instruments. Chemical shifts are reported in ppm (δ scale) with the residual solvent signal used as a reference and coupling constant (J) values are reported in hertz (Hz). Data are presented as follows: chemical shift, multiplicity (s = singlet, d = doublet, dd = doublet of doublet, dt = doublet of triplet, t = triplet, m = multiplet, br s = broad singlet), coupling constant in Hz, and integration. High-resolution mass spectra (HRMS) were recorded on a Thermo Scientific Orbitrap Velos Pro mass

1
2
3 spectrometer coupled with a Thermo Scientific Accela 1250 UPLC and
4 an autosampler using electrospray ionization (ESI) in the positive
5 mode. The purity of final compounds was determined by analytical
6 HPLC and was found to be $\geq 95\%$ pure. HPLC was performed on a Agilent
7 1200 system equipped with a multiple wavelength detector and a
8 manual injector under the following conditions: column, Phenomenex
9 Hypersil-BDS-5u-C18 (5 μm , 4.6 mm \times 250 mm, 130 \AA); solvent A, H_2O
10 containing 0.1% trifluoroacetic acid (TFA); solvent B, CH_3CN
11 containing 0.1% TFA; gradient, 20% B to 100% B over 15 min followed
12 by 100% B over 5 min; injection volume, 20 μL ; flow rate, 1 mL/min.
13 The wavelengths of detection were 254 nm and 280 nm. Retention
14 times and purity data for each target compound are provided in the
15 Experimental Section.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31

32 **((3R,3aS,6aR)-Hexahydrofuro[2,3-b]furan-3-yl ((2S,4S,5S)-5-**
33 **(dibenzylamino)-4-hydroxy-1,6-diphenylhexan-2-yl) carbamate (9)**. A
34 solution of the Phe-Phe core intermediate **7** (3.00 g, 6.46 mmol) in
35 anhydrous CH_3CN (50 mL) was cooled to 0 $^\circ\text{C}$ and treated with
36 diisopropylethylamine (DIEA) (3.37 mL, 19.35 mmol) followed by
37 *bis*-THF activated carbonate **8** (1.93 g, 7.10 mmol). The reaction
38 mixture was warmed to room temperature and stirred for 24 h,
39 concentrated under reduced pressure and dried under high vacuum.
40 The residue was purified by flash column chromatography (RediSep
41 Gold, 80 g, gradient elution with 0-100% ethyl acetate/hexanes) to
42 provide the intermediate **9** (4.01 g, 100%) as a white solid. ^1H NMR
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

(500 MHz, CDCl₃) δ 7.34–7.17 (m, 16 H), 7.10 (d, J = 7.5 Hz, 2 H), 7.01 (d, J = 7.0 Hz, 2 H), 5.69 (d, J = 5.0 Hz, 1 H), 5.29 (d, J = 6.0 Hz, 1 H), 5.12 (q, J = 6.5 Hz, 1 H), 4.52 (br s, 1 H), 4.02 (dd, J = 9.0, 7.0 Hz, 1 H), 3.95–3.73 (m, 6 H), 3.65–3.58 (m, 1 H), 3.38 (d, J = 13.0 Hz, 2 H), 3.10 (dd, J = 14.0, 5.5 Hz, 1 H), 3.04–2.97 (m, 1 H), 2.85–2.76 (m, 2 H), 2.66 (dd, J = 13.0, 6.5 Hz, 1 H), 2.54 (dd, J = 14.5, 7.0 Hz, 1 H), 1.87–1.75 (m, 2 H), 1.52 (dd, J = 14.5, 3.5 Hz, 1 H), 1.11 (dt, J = 14.5, 8.5 Hz, 1 H) ppm; ¹³C NMR (125 MHz, CDCl₃) δ 155.41, 140.15, 138.74, 138.18, 129.61, 129.15, 128.86, 128.69, 128.40, 127.53, 126.47, 126.40, 109.48, 73.00, 71.03, 69.77, 69.74, 64.31, 53.95, 52.28, 45.49, 41.35, 38.15, 32.15, 26.09 ppm; HRMS (ESI) m/z : [M + H]⁺ calcd for C₃₉H₄₅N₂O₅, 621.3323; found 621.3306.

(3R,3aS,6aR)-Hexahydrofuro[2,3-*b*]furan-3-yl ((2S,4S,5S)-4-hydroxy-5-((S)-2-((methoxycarbonyl)amino)-3,3-dimethylbutanamido)-1,6-diphenylhexan-2-yl) carbamate (12a). A solution of compound **9** (0.50 g, 0.81 mmol) in anhydrous MeOH (10 mL) was treated with ammonium formate (0.30 g, 4.83 mmol) and 10% Pd/C (0.1 g) and the reaction mixture was stirred at 50 °C for 14 h. The reaction mixture was filtered through a pad of Celite and the filtrate was concentrated under reduced pressure. The resulting deprotected amine **10** was dissolved in anhydrous CH₂Cl₂ (10 mL) and the solution was cooled to 0 °C. *N*-(methoxycarbonyl)-*L*-*tert*-leucine **11a** (0.18 g, 0.97 mmol), DIEA (0.280 mL, 1.61 mmol)

1
2
3 and HATU (0.48 g, 1.26 mmol) were added. The reaction mixture was
4
5 stirred at room temperature for 16 h, concentrated under reduced
6
7 pressure, and dried under high vacuum. The residue was purified by
8
9 flash column chromatography (RediSep Gold, 24 g, gradient elution
10
11 with 0-15% methanol/dichloromethane) to provide the target
12
13 compound **12a** (0.27 g, 55%) as a white solid. ¹H NMR (500 MHz,
14
15 CDCl₃) δ 7.25-7.13 (m, 8 H), 7.07 (d, *J* = 7.0 Hz, 2 H), 6.57 (d, *J*
16
17 = 9.0 Hz, 1 H), 5.65 (d, *J* = 5.0 Hz, 1 H), 5.46 (d, *J* = 9.0 Hz, 1
18
19 H), 5.12 (d, *J* = 8.0 Hz, 1 H), 5.04 (q, *J* = 6.5 Hz, 1 H), 4.15 (q,
20
21 *J* = 7.0 Hz, 1 H), 4.06-3.96 (m, 1 H), 3.93 (dd, *J* = 9.5, 6.5 Hz,
22
23 1 H), 3.87-3.63 (m, 4 H), 3.66 (s, 3 H, overlapping), 2.96-2.81
24
25 (m, 3 H), 2.80 (s, 2 H), 2.66 (dd, *J* = 14.0, 7.5 Hz, 1 H), 1.74-
26
27 1.55 (m, 4 H), 0.93 (s, 9 H) ppm; ¹³C NMR (125 MHz, CDCl₃) δ 170.93,
28
29 157.18, 155.58, 137.95, 137.39, 129.39, 128.67, 128.60, 126.82,
30
31 126.60, 109.44, 73.47, 70.97, 69.74, 69.67, 63.44, 54.83, 52.59,
32
33 50.95, 45.51, 41.72, 40.19, 38.35, 34.21, 26.71, 26.02 ppm; HRMS
34
35 (ESI) *m/z*: [M + H]⁺ calcd for C₃₃H₄₆N₃O₈, 612.3280; found 612.3276;
36
37
38
39
40
41
42
43 Anal. HPLC: *t*_R 10.86 min, purity 99%.

44 **(3*R*, 3*aS*, 6*aR*)-Hexahydrofuro[2,3-*b*]furan-3-yl** **((2*S*, 4*S*, 5*S*)-4-**
45
46 **hydroxy-5-(1-((methoxycarbonyl)amino)cyclopentane-1-**
47
48 **carboxamido)-1,6-diphenylhexan-2-yl) carbamate** (**12b**). The same
49
50 procedure was used as described above for compound **12a**. Compound
51
52 **9** (1.00 g, 1.61 mmol) was treated with ammonium formate (0.61 g,
53
54 9.66 mmol) and 10% Pd/C (0.25 g) to give the corresponding
55
56
57
58
59
60

1
2
3 deprotected amine **10**, which was coupled with *N*-(methoxycarbonyl)-
4 cycloleucine **11b** (0.36 g, 1.93 mmol) using DIEA (1.12 ml, 6.44
5 mmol) and HATU (0.96 g, 2.51 mmol) to provide the target compound
6
7
8
9
10 **12b** (0.66 g, 67%) as a white solid. ¹H NMR (500 MHz, CDCl₃) δ 7.28-
11 7.22 (m, 4 H), 7.21-7.16 (m, 4 H), 7.14 (d, *J* = 7.0 Hz, 2 H), 6.61
12 (d, *J* = 8.5 Hz, 1 H), 5.66 (d, *J* = 5.0 Hz, 1 H), 5.07 (app q, *J* =
13 6.5 Hz, 2 H), 4.96 (br s, 1 H), 4.07 (q, *J* = 7.5 Hz, 1 H), 4.02-
14 3.93 (m, 2 H), 3.90-3.84 (m, 1 H), 3.76-3.65 (m, 3 H), 3.63 (s, 3
15 H), 3.35 (br s, 1 H), 2.99-2.93 (m, 1 H), 2.90-2.84 (m, 2 H), 2.83
16 (dd, *J* = 14.5, 7.0 Hz, 1 H, overlapping), 2.75 (dd, *J* = 12.5, 7.0
17 Hz, 1 H), 2.27-2.18 (m, 1 H), 2.00-1.91 (m, 1 H), 1.86-1.59 (m, 10
18 H) ppm; ¹³C NMR (125 MHz, CDCl₃) δ 174.05, 156.34, 155.58, 138.29,
19 137.85, 129.53, 129.34, 128.57, 126.64, 126.55, 109.45, 73.31,
20 70.91, 69.74, 67.49, 55.07, 52.53, 51.16, 45.50, 41.84, 39.19,
21 37.80, 36.60, 26.00, 24.11 ppm; HRMS (ESI) *m/z*: [M + H]⁺ calcd for
22 C₃₃H₄₄N₃O₈, 610.3123; found 610.3117; Anal. HPLC: *t*_R 9.74 min, purity
23 96%.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

41
42 **(3*R*, 3*aS*, 6*aR*)-Hexahydrofuro[2,3-*b*]furan-3-yl ((2*S*, 4*S*, 5*S*)-5-((*S*)-**
43 **2-cyclopentyl-2-((methoxycarbonyl)amino)acetamido)-4-hydroxy-**
44 **1,6-diphenylhexan-2-yl) carbamate (12c)**. The same procedure was
45 used as described above for compound **12a**. Compound **9** (0.50 g, 0.81
46 mmol) was treated with ammonium formate (0.30 g, 4.83 mmol) and
47 10% Pd/C (0.13 g) to give the corresponding deprotected amine **10**,
48 which was coupled with *N*-(methoxycarbonyl)-*L*-cyclopentylglycine
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **11c** (0.19 g, 0.97 mmol) using DIEA (0.566 ml, 3.22 mmol) and HATU
4 (0.48 g, 1.26 mmol) to provide the target compound **12c** (0.32 g,
5
6 64%) as a white solid. ¹H NMR (500 MHz, CDCl₃) δ 7.29–7.22 (m, 4
7
8 H), 7.21–7.16 (m, 4 H), 7.11 (d, *J* = 7.5 Hz, 2 H), 6.35 (d, *J* =
9
10 6.5 Hz, 1 H), 5.67 (d, *J* = 5.0 Hz, 1 H), 5.12 (d, *J* = 7.5 Hz, 1
11
12 H), 5.07 (q, *J* = 6.5 Hz, 1 H), 4.94 (d, *J* = 8.5 Hz, 1 H), 4.11
13
14 (app q, *J* = 8.5 Hz, 1 H), 4.03–3.93 (m, 2 H), 3.92–3.83 (m, 2 H),
15
16 3.77–3.66 (m, 3 H), 3.67 (s, 3 H, overlapping), 3.30 (br s, 1 H),
17
18 2.99–2.93 (m, 1 H), 2.91–2.85 (m, 2 H), 2.82 (dd, *J* = 14.5, 6.5
19
20 Hz, 1 H), 2.70 (dd, *J* = 13.5, 7.5 Hz, 1 H), 2.22–2.11 (m, 1 H),
21
22 1.76–1.46 (m, 10 H), 1.27–1.14 (m, 2 H) ppm; ¹³C NMR (125 MHz,
23
24 CDCl₃) δ 172.09, 157.16, 155.62, 138.05, 137.55, 129.43, 129.36,
25
26 128.66, 128.62, 126.80, 126.62, 109.46, 73.47, 70.93, 70.15,
27
28 69.73, 59.47, 54.91, 52.63, 51.02, 45.53, 41.94, 41.81, 39.98,
29
30 38.17, 29.59, 28.73, 26.00, 25.51, 25.28 ppm; HRMS (ESI) *m/z*: [M
31
32 + H]⁺ calcd for C₃₄H₄₆N₃O₈, 624.3280; found 624.3273; Anal. HPLC:
33
34 *t*_R 10.98 min, purity 96%.

35
36
37 **(3*R*, 3*aS*, 6*aR*)-Hexahydrofuro[2,3-*b*]furan-3-yl ((2*S*, 4*S*, 5*S*)-4-**
38
39 **hydroxy-5-((*S*)-2-((methoxycarbonyl)amino)propanamido)-1,6-**
40
41 **diphenylhexan-2-yl) carbamate (12d)**. The same procedure was used as
42
43 described above for compound **12a**. Compound **9** (0.45 g, 0.73 mmol)
44
45 was treated with ammonium formate (0.28 g, 4.36 mmol) and 10% Pd/C
46
47 (0.11 g) to give the corresponding deprotected amine **10**, which was
48
49 coupled with *N*-(methoxycarbonyl)-*L*-alanine **11d** (0.14 g, 0.95 mmol)
50
51
52
53
54
55
56
57
58
59
60

1
2
3 using DIEA (0.377 ml, 2.18 mmol) and HATU (0.47 g, 1.23 mmol) to
4 provide the target compound **12d** (0.20 g, 47%) as a white solid. ¹H
5 NMR (500 MHz, CDCl₃) δ 7.28–7.22 (m, 4 H), 7.21–7.16 (m, 4 H), 7.11
6 (d, *J* = 7.5 Hz, 2 H), 6.39 (d, *J* = 9.0 Hz, 1 H), 5.66 (d, *J* = 5.0
7 Hz, 1 H), 5.17–5.09 (m, 1 H), 5.06 (q, *J* = 6.0 Hz, 1 H), 4.99 (d,
8 *J* = 8.0 Hz, 1 H), 4.16–4.03 (m, 2 H), 3.99–3.91 (m, 2 H), 3.85
9 (td, *J* = 8.0, 2.0 Hz, 1 H), 3.77–3.65 (m, 3 H), 3.67 (s, 3 H,
10 overlapping), 2.97–2.91 (m, 1 H), 2.88 (d, *J* = 7.5 Hz, 2 H), 2.81
11 (dd, *J* = 13.5, 5.5 Hz, 1 H), 2.71 (dd, *J* = 14.0, 8.0 Hz, 1 H),
12 1.75–1.58 (m, 4 H), 1.25 (d, *J* = 7.0 Hz, 3 H) ppm; ¹³C NMR (125
13 MHz, CDCl₃) δ 172.73, 156.74, 155.66, 138.05, 137.61, 129.44,
14 129.36, 128.65, 126.79, 126.66, 109.45, 73.48, 70.96, 70.03,
15 69.73, 54.92, 52.61, 51.01, 45.54, 41.82, 39.87, 37.99, 26.00,
16 18.76 ppm; HRMS (ESI) *m/z*: [M + H]⁺ calcd for C₃₀H₄₀N₃O₈, 570.2810;
17 found 570.2811.; Anal. HPLC: *t*_R 9.14 min, purity 99%.

18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37 **(3*R*,3*aS*,6*aR*)-Hexahydrofuro[2,3-*b*]furan-3-yl** **((2*S*,4*S*,5*S*)-4-**
38
39 **hydroxy-5-((*S*)-2-((methoxycarbonyl)amino)-3-methylbutanamido)-**
40 **1,6-diphenylhexan-2-yl) carbamate (12e)**. The same procedure was
41 used as described above for compound **12a**. Compound **9** (0.44 g, 0.70
42 mmol) was treated with ammonium formate (0.27 g, 4.22 mmol) and
43 10% Pd/C (0.11 g) to give the corresponding deprotected amine **10**,
44 which was coupled with *N*-(methoxycarbonyl)-*L*-valine **11e** (0.15 g,
45 0.84 mmol) using DIEA (0.364 ml, 2.11 mmol) and HATU (0.42 g, 1.10
46 mmol) to provide the target compound **12e** (0.29 g, 70%) as a white
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 solid. ^1H NMR (500 MHz, CDCl_3) δ 7.29–7.23 (m, 4 H), 7.22–7.17 (m,
4 4 H), 7.10 (d, $J = 7.5$ Hz, 2 H), 6.30 (d, $J = 9.0$ Hz, 1 H), 5.67
5 (d, $J = 5.0$ Hz, 1 H), 5.12–5.04 (m, 2 H), 4.93 (d, $J = 8.5$ Hz, 1
6 H), 4.12 (q, $J = 7.5$ Hz, 1 H), 4.01–3.93 (m, 2 H), 3.92–3.83 (m,
7 2 H), 3.75–3.67 (m, 3 H), 3.68 (s, 3 H, overlapping), 3.26 (br s,
8 1 H), 2.99–2.93 (m, 1 H), 2.91–2.84 (m, 2 H), 2.81 (dd, $J = 14.5$,
9 6.5 Hz, 1 H), 2.70 (dd, $J = 13.5$, 7.5 Hz, 1 H), 2.15–2.05 (m, 1
10 H), 1.77–1.57 (m, 4 H), 0.90 (d, $J = 7.0$ Hz, 3 H), 0.78 (d, $J =$
11 6.5 Hz, 3 H) ppm; ^{13}C NMR (125 MHz, CDCl_3) δ 171.58, 157.28, 155.67,
12 138.00, 137.49, 129.42, 129.35, 128.67, 126.82, 126.66, 109.45,
13 73.50, 70.93, 70.12, 69.72, 61.01, 54.88, 52.67, 51.09, 45.52,
14 41.86, 40.04, 38.26, 30.50, 26.01, 19.54, 17.54 ppm; HRMS (ESI)
15 m/z : $[\text{M} + \text{H}]^+$ calcd for $\text{C}_{32}\text{H}_{44}\text{N}_3\text{O}_8$, 598.3123; found 598.3122; Anal.
16 HPLC: t_{R} 10.23 min, purity 99%.

17
18
19 **(3*R*,3*aS*,6*aR*)-Hexahydrofuro[2,3-*b*]furan-3-yl ((2*S*,4*S*,5*S*)-5-((*S*)-**
20 **2-cyclopropyl-2-((methoxycarbonyl)amino)acetamido)-4-hydroxy-**
21 **1,6-diphenylhexan-2-yl) carbamate (12*f*)**. The same procedure was
22 used as described above for compound **12a**. Compound **9** (0.46 g, 0.75
23 mmol) was treated with ammonium formate (0.28 g, 4.49 mmol) and
24 10% Pd/C (0.12 g) to give the corresponding deprotected amine **10**,
25 which was coupled with *N*-(methoxycarbonyl)-*L*-cyclopropylglycine
26 **11f** (0.18 g, 1.05 mmol) using DIEA (0.390 ml, 2.25 mmol) and HATU
27 (0.42 g, 1.05 mmol) to provide the target compound **12f** (0.27 g,
28 61%) as a white solid. ^1H NMR (500 MHz, CDCl_3) δ 7.29–7.22 (m, 4
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 H), 7.21–7.15 (m, 4 H, overlapping), 7.10 (d, $J = 7.5$ Hz, 2 H),
4
5 6.46 (d, $J = 7.5$ Hz, 1 H), 5.66 (d, $J = 5.0$ Hz, 1 H), 5.49 (d, J
6
7 = 6.5 Hz, 1 H), 5.10–5.02 (m, 2 H), 4.10 (app, q, $J = 8.0$ Hz, 1
8
9 H), 4.02–3.91 (m, 2 H), 3.89–3.82 (m, 1 H), 3.78–3.64 (m, 3 H),
10
11 3.65 (s, 3 H, overlapping), 3.54–3.47 (br s, 1 H), 3.46–3.39 (m,
12
13 1 H), 2.98–2.91 (m, 1 H), 2.89 (d, $J = 7.5$ Hz, 2 H), 2.81 (dd, $J =$
14
15 14.0, 6.0 Hz, 1 H), 2.70 (dd, $J = 13.5, 7.5$ Hz, 1 H), 1.75–1.57
16
17 (m, 4 H), 0.96 (br s, 1 H), 0.55–0.44 (m, 2 H), 0.43–0.33 (m, 2 H)
18
19 ppm; ^{13}C NMR (125 MHz, CDCl_3) δ 171.65, 156.96, 155.58, 138.08,
20
21 137.65, 129.44, 129.36, 128.62, 128.59, 126.72, 126.62, 109.43,
22
23 73.41, 70.94, 69.70, 59.32, 54.96, 52.56, 50.82, 45.54, 41.63,
24
25 39.78, 37.95, 25.98, 14.59, 3.68, 3.40 ppm; HRMS (ESI) m/z : $[\text{M} +$
26
27 $\text{H}]^+$ calcd for $\text{C}_{32}\text{H}_{42}\text{N}_3\text{O}_8$, 596.2967; found 596.2957; Anal. HPLC: t_R
28
29 10.54 min, purity 99%.

30
31
32
33
34
35 **(3*R*, 3*aS*, 6*aR*)-Hexahydrofuro[2,3-*b*]furan-3-yl ((2*S*, 4*S*, 5*S*)-4-**
36
37 **hydroxy-5-((2*S*, 3*S*)-2-((methoxycarbonyl)amino)-3-**
38
39 **methylpentanamido)-1,6-diphenylhexan-2-yl) carbamate (12*g*).** The
40
41 same procedure was used as described above for compound **12a**.
42
43 Compound **9** (0.44 g, 0.70 mmol) was treated with ammonium formate
44
45 (0.27 g, 4.22 mmol) and 10% Pd/C (0.11 g) to give the corresponding
46
47 deprotected amine **10**, which was coupled with *N*-(methoxycarbonyl)-
48
49 *L*-isoleucine **11g** (0.15 g, 0.84 mmol) using DIEA (0.364 ml, 2.11
50
51 mmol) and HATU (0.42 g, 1.10 mmol) to provide the target compound
52
53 **12g** (0.29 g, 40%) as a white solid. ^1H NMR (500 MHz, CDCl_3) δ 7.29–
54
55
56
57
58
59
60

1
2
3 7.23 (m, 4 H), 7.22–7.17 (m, 4 H), 7.11 (d, $J = 7.5$ Hz, 2 H), 6.26
4
5 (d, $J = 9.0$ Hz, 1 H), 5.67 (d, $J = 5.5$ Hz, 1 H), 5.07 (q, $J = 6.5$
6
7 Hz, 1 H), 5.02 (d, $J = 8.0$ Hz, 1 H), 4.89 (d, $J = 8.0$ Hz, 1 H),
8
9 4.13 (q, $J = 7.0$ Hz, 1 H), 4.00–3.90 (m, 3 H), 3.89–3.84 (m, 1 H),
10
11 3.75–3.67 (m, 3 H), 3.68 (s, 3 H, overlapping), 3.16 (br s, 1 H),
12
13 2.99–2.93 (m, 1 H), 2.91–2.84 (m, 2 H), 2.81 (dd, $J = 15.5, 7.5$
14
15 Hz, 1 H), 2.71 (dd, $J = 14.0, 8.0$ Hz, 1 H), 1.90–1.80 (m, 1 H),
16
17 1.77–1.56 (m, 4 H), 1.33–1.22 (m, 1 H), 1.03–0.93 (m, 1 H), 0.87
18
19 (d, $J = 7.0$ Hz, 3 H), 0.84 (t, $J = 7.5$ Hz, 3 H) ppm; ^{13}C NMR (125
20
21 MHz, CDCl_3) δ 171.58, 157.21, 155.66, 137.99, 137.51, 129.42,
22
23 129.36, 128.68, 128.64, 126.82, 126.64, 109.45, 73.49, 70.91,
24
25 70.24, 69.73, 60.40, 56.02, 54.80, 52.67, 51.10, 45.52, 43.96,
26
27 41.90, 40.06, 38.25, 36.82, 26.00, 24.48, 18.78, 17.40, 15.85,
28
29 12.65, 11.57 ppm; HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ calcd for $\text{C}_{33}\text{H}_{46}\text{N}_3\text{O}_8$,
30
31 612.3280; found 612.3275; Anal. HPLC: t_{R} 10.82 min, purity 99%.

32
33
34
35
36
37 **((3R,3aS,6aR)-Hexahydrofuro[2,3-b]furan-3-yl ((2S,4S,5S)-4-**
38
39 **hydroxy-5-((2S,3R)-2-((methoxycarbonyl)amino)-3-**
40 **methylpentanamido)-1,6-diphenylhexan-2-yl) carbamate (12h).** The
41
42 same procedure was used as described above for compound **12a**.
43
44 Compound **9** (0.53 g, 0.86 mmol) was treated with ammonium formate
45
46 (0.33 g, 5.17 mmol) and 10% Pd/C (0.13 g) to give the corresponding
47
48 deprotected amine **10**, which was coupled with *N*-(methoxycarbonyl)-
49
50 *L*-alloisoleucine **11h** (0.15 g, 0.84 mmol) using DIEA (0.450 ml,
51
52 2.59 mmol) and HATU (0.59 g, 1.55 mmol) to provide the target
53
54
55
56
57
58
59
60

1
2
3 compound **12h** (0.30 g, 58%) as a white solid. ^1H NMR (500 MHz,
4 CDCl_3) δ 7.29–7.23 (m, 4 H), 7.22–7.17 (m, 4 H), 7.11 (d, J = 7.0
5 Hz, 2 H), 6.28 (d, J = 9.0 Hz, 1 H), 5.67 (d, J = 5.5 Hz, 1 H),
6 5.07 (q, J = 6.5 Hz, 1 H), 4.96 (d, J = 7.5 Hz, 1 H), 4.93 (d, J
7 = 8.0 Hz, 1 H), 4.16–4.09 (m, 1 H), 4.07 (dd, J = 7.0, 4.5 Hz, 1
8 H), 3.97 (dd, J = 9.5, 6.5 Hz, 2 H), 3.90–3.84 (m, 1 H), 3.75–3.67
9 (m, 3 H), 3.69 (s, 3 H, overlapping), 2.99–2.93 (m, 1 H), 2.90–
10 2.84 (m, 2 H), 2.81 (dd, J = 15.5, 7.5 Hz, 1 H), 2.72 (dd, J =
11 14.0, 8.0 Hz, 1 H), 1.96–1.88 (m, 1 H), 1.76–1.57 (m, 4 H), 1.40–
12 1.30 (m, 1 H), 1.21–1.11 (m, 1 H), 0.90 (t, J = 7.0 Hz, 3 H), 0.67
13 (d, J = 6.5 Hz, 3 H) ppm; ^{13}C NMR (125 MHz, CDCl_3) δ 171.84, 157.29,
14 155.66, 138.02, 137.54, 129.42, 129.34, 128.67, 126.81, 126.67,
15 109.45, 73.48, 70.89, 70.36, 69.73, 59.12, 54.88, 52.71, 51.16,
16 45.52, 41.96, 39.95, 38.22, 36.82, 26.52, 26.00, 14.17, 11.80 ppm;
17 HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ calcd for $\text{C}_{33}\text{H}_{46}\text{N}_3\text{O}_8$, 612.3280; found
18 612.3277; Anal. HPLC: t_{R} 10.87 min, purity 95%.

19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39 **Methyl ((S)-1-(((2S,4S,5S)-5-(dibenzylamino)-4-hydroxy-1,6-**
40 **diphenylhexan-2-yl) amino)-3,3-dimethyl-1-oxobutan-2-yl) carbamate**
41 **(13a)**. A solution of the Phe-Phe core intermediate **7** (1.00 g, 2.16
42 mmol) in DMF (15 mL) was cooled to 0 °C and treated with *N*-
43 (methoxycarbonyl)-*L*-*tert*-leucine **11a** (0.48 g, 2.58 mmol), DIEA
44 (1.1 mL, 6.30 mmol) and HATU (1.28 g, 3.36 mmol). The resulting
45 reaction mixture was stirred at room temperature for 4 h,
46 concentrated under reduced pressure, and dried under high vacuum.
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The residue was purified by flash column chromatography (RediSep Gold, 24 g, gradient elution with 0-10% methanol/dichloromethane) to give the intermediate **13a** (1.20 g, 87%) as a white solid. ¹H NMR (500 MHz, CDCl₃) δ 7.34-7.29 (m, 4 H), 7.28-7.23 (m, 4 H), 7.22-7.14 (m, 8 H), 7.06 (d, *J* = 7.5 Hz, 2 H), 7.00 (d, *J* = 6.5 Hz, 2 H), 6.55 (d, *J* = 6.0 Hz, 1 H), 5.42 (d, *J* = 9.0 Hz, 1 H), 4.65 (br s, 1 H), 4.06-3.97 (m, 1 H), 3.91-3.79 (m, 3 H), 3.66 (s, 3 H), 3.53 (t, *J* = 8.5 Hz, 1 H), 3.36 (d, *J* = 13.5 Hz, 2 H), 3.06 (dd, *J* = 14.5, 5.5 Hz, 1 H), 2.92 (dd, *J* = 13.5, 5.0 Hz, 1 H), 2.76 (q, *J* = 7.0 Hz, 1 H), 2.59-2.46 (m, 2 H), 1.50 (dd, *J* = 14.5, 5.0 Hz, 1 H), 1.17-1.08 (m, 1 H), 0.99 (s, 9 H) ppm; ¹³C NMR (125 MHz, CDCl₃) δ 170.57, 157.02, 140.03, 138.77, 138.10, 129.54, 129.15, 129.06, 128.80, 128.64, 128.36, 127.49, 126.40, 126.29, 69.83, 64.22, 63.31, 53.90, 52.34, 51.54, 41.12, 37.42, 34.78, 32.01, 26.60 ppm; HRMS (ESI) *m/z*: [M + H]⁺ calcd for C₄₀H₅₀N₃O₄, 636.3796; found 636.3793.

Methyl (1-(((2*S*,4*S*,5*S*)-5-(dibenzylamino)-4-hydroxy-1,6-diphenylhexan-2-yl) carbamoyl) cyclopentyl) carbamate (13b). The same procedure was used as described above for **13a**. The Phe-Phe core intermediate **7** (1.00 g, 2.16 mmol) was coupled with *N*-(methoxycarbonyl)-cycloleucine **11b** (0.47 g, 2.50 mmol) using DIEA (1.1 ml, 6.30 mmol) and HATU (1.21 g, 3.18 mmol) to provide **13b** (1.30 g, 95%) as a white solid. ¹H NMR (500 MHz, CDCl₃) δ 7.33-7.27 (m, 4 H), 7.26-7.15 (m, 12 H), 7.11 (d, *J* = 7.5 Hz, 2 H),

1
2
3 7.04 (d, $J = 7.0$ Hz, 2 H), 6.80 (br s, 1 H), 5.00 (br s, 1 H),
4
5 4.32 (br s, 1 H), 4.04 (m, 1 H), 3.93 (d, $J = 13.0$ Hz, 2 H), 3.55
6
7 (s, 4 H), 3.39 (d, $J = 13.0$ Hz, 2 H), 3.05 (dd, $J = 13.5, 5.0$ Hz,
8
9 1 H), 2.81 (dd, $J = 13.5, 5.0$ Hz 1 H), 2.74 (app q, $J = 7.5$ Hz, 1
10
11 H), 2.70–2.59 (m, 2 H), 2.28–2.17 (m, 1 H), 2.12–2.04 (m, 1 H),
12
13 1.93–1.78 (m, 1 H), 1.77–1.65 (m, 4 H), 1.51 (dd, $J = 14.5, 3.5$
14
15 Hz, 1 H), 1.35–1.24 (m, 1 H) ppm; ^{13}C NMR (125 MHz, CDCl_3) δ 173.64,
16
17 156.09, 140.40, 139.32, 138.38, 129.69, 129.27, 129.10, 128.74,
18
19 128.57, 128.30, 127.33, 126.28, 70.03, 67.31, 64.31, 54.27, 52.20,
20
21 50.93, 41.33, 38.17, 37.23, 36.65, 31.65, 24.30, 24.26 ppm; HRMS
22
23 (ESI) m/z : $[\text{M} + \text{H}]^+$ calcd for $\text{C}_{40}\text{H}_{48}\text{N}_3\text{O}_4$, 634.3640; found 634.3615.
24
25
26
27

28 **Methyl ((S)-1-cyclopentyl-2-((2S,4S,5S)-5-(dibenzylamino)-4-**
29
30 **hydroxy-1,6-diphenylhexan-2-yl)amino)-2-oxoethyl)carbamate (13c).**
31

32 The same procedure was used as described above for **13a**. The Phe-
33
34 Phe core intermediate **7** (1.04 g, 2.23 mmol) was coupled with *N*-
35
36 (methoxycarbonyl)-cyclopentylglycine **11c** (0.45 g, 2.23 mmol) using
37
38 DIEA (1.17 ml, 6.69 mmol) and HATU (1.02 g, 2.68 mmol) to give **13c**
39
40 (1.30 g, 90%) as a white solid. ^1H NMR (500 MHz, CDCl_3) δ 7.34–
41
42 7.28 (m, 4 H), 7.27–7.22 (m, 4 H), 7.21–7.14 (m, 8 H), 7.06 (d, J
43
44 = 7.0 Hz, 2 H), 7.00 (d, $J = 6.5$ Hz, 2 H), 6.57 (d, $J = 5.5$ Hz, 1
45
46 H), 5.22 (d, $J = 8.0$ Hz, 1 H), 4.66 (br s, 1 H), 4.06–3.98 (m, 1
47
48 H), 3.97–3.90 (m, 1 H), 3.86 (d, $J = 13.0$ Hz, 2 H), 3.64 (s, 3 H),
49
50 3.53 (t, $J = 8.5$ Hz, 1 H), 3.36 (d, $J = 13.0$ Hz, 2 H), 3.05 (dd,
51
52 $J = 14.5, 5.5$ Hz, 1 H), 2.88 (dd, $J = 13.5, 5.0$ Hz, 1 H), 2.76 (q,
53
54
55
56
57
58
59
60

1
2
3 $J = 6.5$ Hz, 1 H), 2.59–2.47 (m, 2 H), 2.23 (app q, $J = 8.0$ Hz, 1
4
5 H), 1.74–1.54 (m, 4 H), 1.53–1.43 (m, 3 H), 1.40–1.22 (m, 2 H),
6
7 1.15–1.06 (m, 1 H) ppm; ^{13}C NMR (125 MHz, CDCl_3) δ 171.64, 157.03,
8
9 140.10, 138.79, 138.14, 129.56, 129.16, 129.11, 128.83, 128.66,
10
11 128.38, 127.51, 126.42, 126.33, 70.00, 64.20, 58.96, 53.88, 52.38,
12
13 51.53, 42.75, 41.40, 37.59, 32.05, 29.29, 28.55, 25.46, 25.20 ppm;
14
15 HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ calcd for $\text{C}_{41}\text{H}_{50}\text{N}_3\text{O}_4$, 648.3796; found
16
17 648.3795.
18
19
20

21 **(3R,3aS,6aR)-Hexahydrofuro[2,3-b]furan-3-yl** **((2S,3S,5S)-3-**
22
23 **hydroxy-5-((S)-2-((methoxycarbonyl)amino)-3,3-**
24
25 **dimethylbutanamido)-1,6-diphenylhexan-2-yl) carbamate (15a).** A
26
27 solution of intermediate **13a** (1.20 g, 1.89 mmol) in anhydrous MeOH
28
29 (15 mL) was treated with ammonium formate (0.71 g, 11.32 mmol) and
30
31 10% Pd/C (0.3 g). The resulting reaction mixture was stirred at 50
32
33 °C for 15 h, cooled to room temperature, filtered through a pad of
34
35 Celite, and the filtrate was concentrated under reduced pressure.
36
37 The resulting deprotected amine was dissolved in acetonitrile (10
38
39 mL) and the solution was cooled to 0 °C. DIEA (0.98 ml, 5.66 mmol)
40
41 was added followed by *bis*-THF activated carbonate **8** (0.56 g, 2.08
42
43 mmol). The reaction mixture was warmed up to room temperature and
44
45 stirred for 24 h. The solvents were removed under reduced pressure,
46
47 and the residue was purified by automated flash chromatography
48
49 using a silica gel column (RediSep Gold, 40 g, gradient elution
50
51 with 0–15% methanol/dichloromethane) to provide the target
52
53
54
55
56
57
58
59
60

1
2
3 compound **15a** (0.36 g, 31%) as a white solid. ¹H NMR (500 MHz,
4 CDCl₃) δ 7.29–7.22 (m, 4 H), 7.21–7.16 (m, 4 H), 7.10 (d, *J* = 7.0
5 Hz, 2 H), 6.07 (d, *J* = 6.0 Hz, 1 H), 5.67 (d, *J* = 5.0 Hz, 1 H),
6 5.36–5.22 (m, 1 H), 5.10 (d, *J* = 9.5 Hz, 1 H), 5.05 (q, *J* = 6.5
7 Hz, 1 H), 4.22–4.13 (m, 1 H), 3.98 (dd, *J* = 9.5, 6.5 Hz, 1 H),
8 3.87 (td, *J* = 8.5, 3.0 Hz, 1 H), 3.82–3.58 (m, 6 H), 3.67 (s, 3 H,
9 overlapping), 2.98–2.90 (m, 1 H), 2.88–2.73 (m, 4 H), 1.73–1.57
10 (m, 4 H), 0.91 (s, 9 H) ppm; ¹³C NMR (125 MHz, CDCl₃) δ 171.09,
11 157.19, 155.83, 138.13, 137.12, 129.41, 129.32, 128.73, 128.64,
12 126.91, 126.61, 109.47, 73.39, 70.98, 70.35, 69.78, 63.35, 56.60,
13 52.65, 49.38, 45.49, 41.43, 40.37, 38.87, 34.35, 26.65, 25.98 ppm;
14 HRMS (ESI) *m/z*: [M + H]⁺ calcd for C₃₃H₄₆N₃O₈, 612.3280; found
15 612.3278; Anal. HPLC: *t*_R 10.70 min, purity 97%.

16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33 **((2*S*,3*S*,5*S*)-3-**
34
35 **hydroxy-5-(1-((methoxycarbonyl)amino)cyclopentane-1-**
36
37 **carboxamido)-1,6-diphenylhexan-2-yl) carbamate (15b)**. The same
38 procedure was used as described above for **15a**. Compound **13b** (0.70
39 g, 1.10 mmol) was treated with ammonium formate (0.42 g, 6.63 mmol)
40 and 10% Pd/C (0.17 g) to give the corresponding deprotected amine,
41 which was subsequently treated with DIEA (0.380 ml, 2.20 mmol) and
42 *bis*-THF activated carbonate **8** (0.33 g, 1.21 mmol) to provide the
43 target compound **15b** (0.23 g, 35%) as a white solid. ¹H NMR (500
44 MHz, CDCl₃) δ 7.29–7.22 (m, 4 H), 7.21–7.16 (m, 4 H), 7.11 (d, *J* =
45 7.0 Hz, 2 H), 6.49 (d, *J* = 7.0 Hz, 1 H), 5.67 (d, *J* = 5.5 Hz, 1
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 H), 5.14 (d, $J = 9.5$ Hz, 1 H), 5.07 (q, $J = 6.5$ Hz, 1 H), 4.97 (s,
4
5 1 H), 4.23 (br s, 1 H), 3.99 (dd, $J = 9.5, 6.5$ Hz, 1 H), 3.94 (br
6
7 s, 1 H), 3.88 (td, $J = 8.0, 2.0$ Hz, 1 H), 3.85–3.69 (m, 4 H),
8
9 3.65 (s, 3 H), 2.98–2.91 (m, 1 H), 2.89 (dd, $J = 14.0, 7.5$ Hz, 1
10
11 H, overlapping), 2.83 (dd, $J = 15.5, 8.0$ Hz, 2 H), 2.71 (dd, J
12
13 = 13.5, 7.5 Hz, 1 H), 2.19–2.11 (m, 1 H), 1.94–1.59 (m, 11 H)
14
15 ppm; ^{13}C NMR (125 MHz, CDCl_3) δ 173.71, 156.67, 155.77, 138.34,
16
17 137.62, 129.52, 129.40, 128.56, 126.73, 126.49, 109.51, 73.29,
18
19 71.11, 70.40, 69.81, 67.39, 57.24, 52.62, 49.96, 45.50, 41.98,
20
21 40.64, 39.01, 37.14, 36.65, 26.03, 24.01, 23.81 ppm; HRMS (ESI)
22
23 m/z : $[\text{M} + \text{H}]^+$ calcd for $\text{C}_{33}\text{H}_{44}\text{N}_3\text{O}_8$, 610.3123; found 610.3118; Anal.
24
25 HPLC: t_R 7.73 min, purity 95%.

26
27
28
29
30 **(3R,3aS,6aR)-Hexahydrofuro[2,3-b]furan-3-yl ((2S,3S,5S)-5-((S)-**
31
32 **2-cyclopentyl-2-((methoxycarbonyl)amino)acetamido)-3-hydroxy-**
33
34 **1,6-diphenylhexan-2-yl)carbamate (15c)**. The same procedure was
35
36 used as described above for **15a**. Compound **13c** (0.70 g, 1.08 mmol)
37
38 was treated with ammonium formate (0.41 g, 6.48 mmol) and 10% Pd/C
39
40 (0.17 g) to give the corresponding deprotected amine, which was
41
42 subsequently treated with DIEA (0.570 ml, 3.24 mmol) and *bis*-THF
43
44 activated carbonate **8** (0.32 g, 1.19 mmol) to provide the target
45
46 compound **15c** (0.26 g, 39%) as a white solid. ^1H NMR (500 MHz,
47
48 CDCl_3) δ 7.29–7.22 (m, 4 H), 7.21–7.16 (m, 4 H), 7.11 (d, $J = 7.5$
49
50 Hz, 2 H), 6.29 (br s, 1 H), 5.67 (d, $J = 5.0$ Hz, 1 H), 5.19–5.07
51
52 (m, 2 H), 5.05 (q, $J = 6.0$ Hz, 1 H, overlapping), 4.24–4.15 (m, 1
53
54
55
56
57
58
59
60

1
2
3 H), 3.98 (dd, $J = 9.5, 6.0$ Hz, 1 H), 3.87 (td, $J = 8.0, 2.0$ Hz, 1
4
5 H), 3.85–3.78 (m, 2 H, overlapping), 3.77–3.63 (m, 3 H), 3.67 (s,
6
7 3 H, overlapping), 2.99–2.91 (m, 1 H), 2.90–2.71 (m, 4 H), 2.22–
8
9 2.09 (m, 1 H), 1.83–1.45 (m, 10 H), 1.22–1.07 (m, 2 H) ppm; ^{13}C NMR
10
11 (125 MHz, CDCl_3) δ 172.11, 157.29, 155.83, 138.24, 137.34, 129.45,
12
13 129.34, 128.63, 128.60, 126.80, 126.56, 109.48, 73.36, 71.00,
14
15 70.33, 69.79, 59.46, 56.68, 52.71, 49.33, 45.54, 41.70, 41.46,
16
17 40.51, 38.92, 29.48, 28.71, 25.97, 25.47, 25.22 ppm; HRMS (ESI)
18
19 m/z : $[\text{M} + \text{H}]^+$ calcd for $\text{C}_{34}\text{H}_{46}\text{N}_3\text{O}_8$, 624.3279; found 624.3282; Anal.
20
21 HPLC: t_R 10.83 min, purity 99%.

22 23 24 25 26 **Protease Gene Construction.**

27
28 Protease gene construction was carried out as previously
29
30 described.^{36, 37} The NL4-3 strain has four naturally occurring
31
32 polymorphisms in the protease relative to the SF2 strain. In
33
34 short, the protease variant genes (I84V, I50V/A71V) were
35
36 constructed using QuikChange site-directed mutagenesis (Genewiz)
37
38 onto NL4-3 wild-type protease on a pET11a plasmid with codon
39
40 optimization for protein expression in *Escherichia coli*. A Q7K
41
42 mutation was included to prevent autoproteolysis.³⁸

43
44
45
46 **Protein Expression and Purification.** The expression, isolation,
47
48 and purification of WT and mutant HIV-1 proteases used for the
49
50 kinetic assays and crystallization were carried out as previously
51
52 described.^{36, 37} Briefly, the gene encoding the HIV protease was
53
54 subcloned into the heat-inducible pXC35 expression vector (ATCC)
55
56
57
58
59
60

1
2
3 and transformed into E. coli TAP-106 cells. Cells grown in 6 L of
4 Terrific Broth were lysed with a cell disruptor and the protein
5 was purified from inclusion bodies.³⁹ The inclusion body
6 centrifugation pellet was dissolved in 50% acetic acid followed by
7 another round of centrifugation to remove impurities. Size
8 exclusion chromatography was used to separate high molecular
9 weight proteins from the desired protease. This was carried out on
10 a 2.1 L Sephadex G-75 superfine column (Millipore Sigma)
11 equilibrated with 50% acetic acid. The cleanest fractions of HIV
12 protease were refolded into a 10-fold dilution of 0.05 M sodium
13 acetate at pH 5.5, 5% ethylene glycol, 10% glycerol, and 5 mM DTT.
14 Folded protein was concentrated down to 1-2 mg/mL and stored. This
15 stored protease was used in K_i assays. For crystallography, a final
16 purification was performed with a Pharmacia Superdex 75 FPLC column
17 equilibrated with 0.05 M sodium acetate at pH 5.5, 5% ethylene
18 glycol, 10% glycerol, and 5 mM DTT. Protease fractions purified
19 from the size exclusion column were concentrated to 1-2 mg/mL using
20 an Amicon Ultra-15 10-kDa device (Millipore) for crystallization.

21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44 **Enzyme Inhibition Assays.** The enzyme inhibition assays were
45 carried out as previously described.^{33, 40} To determine the enzyme
46 inhibition constant (K_i), in a 96-well plate, each inhibitor was
47 serially diluted, including a no drug control, and incubated with
48 0.35 nM protein for 1 hour. A 10-amino acid substrate containing
49 an optimized protease cleavage site with an EDANS/DABCYL FRET pair
50
51
52
53
54
55
56
57
58
59
60

1
2
3 was dissolved in 4% DMSO at 120 μ M. Using the Envision plate
4 reader, 5 μ L of the 120 μ M substrate was added to the 96-well plate
5
6 to a final concentration of 10 μ M. The fluorescence was observed
7
8 with an excitation at 340 nm and emission at 492 nm and monitored
9
10 for 200 counts, for approximately 60 min. Data was analyzed with
11
12 Prism7. DRV was used as a control in all assays.
13
14
15

16
17 **Antiviral Assays.** 293T and TZM-BL⁴¹ cells (NIH AIDS Research and
18 Reference Reagent Program) were maintained in Dulbecco's modified
19 Eagle's medium supplemented with 10% fetal calf serum in the
20 presence of penicillin and streptomycin at 37 °C with 5% CO₂. To
21 determine the concentration of drugs achieving 50% inhibition of
22 infection compared with the drug-free control, 4.5 \times 10⁶ 293T cells
23 were seeded onto a 10-cm plate 24 h before transfection. Cells
24 were transfected with 8 μ g of either the wild-type plasmid,
25 infectious molecular clone pNL-CH derived from the pNL4-3 clone of
26 HIV-1 using FuGENE 6 transfection reagent (Roche). The culture
27 supernatant of 293T cells transfected with wild-type or PI-
28 resistant HIV-1 variant was removed 18 h after transfection and
29 the cells were washed with 1 \times PBS. The 293T cells were collected
30 and transferred to wells of a 24-well plate. Briefly, each drug
31 was serially diluted in the culture medium and the dilutions were
32 added to the wells of a 24-well plate. The 293T cells (0.5 \times 10⁶
33 per well) collected from the transfection were added to wells
34 containing various concentrations of drug. The culture supernatant
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 containing virus particles was harvested 18 h after the 293T cells
4 were reseeded in the presence of drug. This supernatant was
5 filtered through a 0.45- μm -pore-size membrane (Millipore) to
6 remove cell debris then used to infect 2×10^4 TZM-BL cells in a
7 96-well plate following a procedure previously described.⁴² The
8 culture supernatant was removed from each well 48 h post-infection,
9 and the cells were washed with $1 \times$ PBS. For the luciferase assay,
10 infected TZM-BL cells were lysed in $1 \times$ reporter lysis buffer
11 (Promega) and the cells were kept at -80°C . After one freeze-thaw
12 cycle, the cell lysates were transferred into a 96-well assay plate
13 (Costar), and luciferase activity was measured using a luminometer
14 (Promega). The culture supernatant harvested from 293T cells
15 reseeded in the absence of drugs was used as a drug-free control.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

EC₅₀ was determined based on a dose-response curve generated using GraphPad Prism (version 7.0).

Protein Crystallization. The condition reliably producing cocrystals of NL4-3 WT protease bound to PIs was discovered and optimized as previously described.^{40, 43} Briefly, all cocrystals were grown at room temperature by hanging drop vapor diffusion method in a 24-well VDX hanging-drop trays (Hampton Research) with a protease concentration of 1.4-1.7 mg/mL with 3-fold molar excess of inhibitors and mixed with the precipitant solution at a 1:2 ratio. The reservoir solution was 23-27% (w/v) ammonium sulfate with 0.1 M *bis*-Tris-methane buffer at pH 5.5, and the

1
2
3 crystallization drops were set with 2 μ L of well solution and 1 μ L
4
5 of protein-inhibitor solution and micro-seeded with a cat whisker.
6
7 Diffraction quality crystals were obtained within 1 week. As data
8
9 were collected at 100 K, cryogenic conditions contained the
10
11 precipitant solution supplemented with 25% glycerol.
12
13

14 **X-Ray Data Collection and Structure Solution.** X-ray diffraction
15
16 data were collected and solved as previously described.^{36, 40, 43}
17
18 Diffraction quality crystals were flash frozen under a cryostream
19
20 when mounting the crystals either at our in-house Rigaku_Saturn944
21
22 X-ray system or the Chicago APS Synchrotron Beamline 23-1D-D. The
23
24 cocrystal diffraction intensities from the Rigaku system were
25
26 indexed, integrated, and scaled using HKL3000.⁴⁴ Structures were
27
28 solved using molecular replacement with PHASER.⁴⁵ Model building
29
30 and refinement were performed using Coot⁴⁶ and Phenix.⁴⁷ Ligands
31
32 were designed in Maestro and the output sdf files were used in the
33
34 Phenix program eLBOW⁴⁸ to generate cif files containing atomic
35
36 positions and constraints necessary for ligand refinement.
37
38 Iterative rounds of crystallographic refinement were carried out
39
40 until convergence was achieved. To limit bias throughout the
41
42 refinement process, five percent of the data were reserved for the
43
44 free R-value calculation.⁴⁹ MolProbity⁵⁰ was applied to evaluate
45
46 the final structures before deposition in the PDB. Structure
47
48 analysis, superposition and figure generation was done using
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 PyMOL.⁵¹ X-ray data collection and crystallographic refinement
4
5 statistics are presented in the Supporting Information (**Table S1**).
6

7
8 The cocrystal structures of all new hybrid compounds were solved
9
10 in the $P2_12_12_1$ space group with one protease homodimer in the
11
12 asymmetric unit and only one orientation of the bound inhibitor in
13
14 the active site, which was crucial for direct comparison of
15
16 inhibitor structures. The cocrystal structure of compound **4** was
17
18 solved in the $P2_1$ space group with two protease homodimers in the
19
20 asymmetric unit, and one inhibitor bound to each dimer in one
21
22 orientation, allowing one $P2_1$ dimer to be directly compared to the
23
24 crystal structures solved in the $P2_12_12_1$ space group. Three
25
26 structures (**10a**, **10b**, **13c**) had significant electron density at the
27
28 flap tips (residues 50–51) indicating that the backbone atoms
29
30 interacted in two conformations, as both hydrogen bond donors and
31
32 acceptors. Figure generation and structural analysis calculations
33
34 (distance differences, vdW, hydrogen bonds) were performed with
35
36 the flap tips modeled in one conformation while the other
37
38 conformation was excluded. Compound **5** formed cocrystals but were
39
40 inadequate for X-ray data collection.
41
42
43
44
45

46 **Intermolecular vdW Contact Analysis of Crystal Structures.** To
47
48 calculate the intermolecular vdW interaction energies the crystal
49
50 structures were prepared using the Schrödinger Protein Preparation
51
52 Wizard.⁵² Hydrogen atoms were added, protonation states were
53
54 determined, and the structures were minimized. The protease active
55
56
57
58
59
60

1
2
3 site was monoprotinated at Asp25. Subsequently, force field
4 parameters were assigned using the OPLS3 force field.⁵³ Interaction
5 energies between the inhibitor and protease were estimated using
6 a simplified Lennard-Jones potential, as previously described in
7 detail.⁵⁴ Briefly, the vdW energy was calculated for pairwise
8 interactions depending on the types of atoms interacting and the
9 distance between them. For each protease residue, the change in
10 vdW interactions relative to a reference complex in the same space
11 group was also calculated for each variant structure.
12
13
14
15
16
17
18
19
20
21
22

23 **Molecular Dynamics Simulations**

24
25
26 **System Preparation.** High resolution crystal structures of LPV and
27 the designed inhibitors bound to WT protease were prepared using
28 the Protein Preparation Wizard using Maestro within the
29 Schrödinger Suite⁵² as previously described.^{55, 56} Briefly, missing
30 atoms were added using Prime⁵⁷ and PROPKA⁵⁸ was used to determine
31 the protonation state of the side chains at pH 7.0. The catalytic
32 aspartic acid with a pKa higher than 7.0 was protonated whereas
33 the one with a pKa less than 7.0 was unprotonated. Co-crystallized
34 fragments such as phosphate ions were removed. Lastly, the
35 structure was minimized to a convergence criterion of 0.3 Å using
36 Impref.⁵⁹
37
38
39
40
41
42
43
44
45
46
47
48
49

50
51 **Molecular Dynamics Simulations.** The prepared systems were placed
52 in a cubic TIP3P implicit water box measuring 12Å on each side.
53 Molecular dynamics simulations were carried out as previously
54
55
56
57
58
59
60

1
2
3 described⁵⁵ using Desmond within Schrödinger Suite.⁵² Briefly,
4 chloride ions were used to neutralize the system and 0.15 M salt
5 were added using sodium and chloride ions. The OPLS3 force field
6 was used to parameterize the ligand and protein. Prior to starting
7 the 100 ns MD simulations, the solvated system was minimized using
8 the stepwise procedure described previously.⁵⁵ Triplicates of 100
9 ns simulations for LPV and the designed inhibitors in complex with
10 WT protease each with a randomized velocity were started using a
11 protocol previously developed.^{55, 56} The root-mean-square deviation
12 (RMSD) and root-mean-square fluctuation (RMSF) were calculated
13 using tools within the Schrodinger Suite.
14
15
16
17
18
19
20
21
22
23
24
25
26
27

28 **ASSOCIATED CONTENT**

31 **Supporting Information**

32
33
34 The Supporting Information is available free of charge on the ACS
35 Publications website at DOI:
36
37

38
39 Synthesis details and analytical data of Phe-Phe isostere **7**,
40 amino acid derivatives **11a-h**, and inhibitors **4-5**; Figures S1-S9;
41
42 Tables S1-S3 (PDF)
43
44

45
46
47 Molecular formula strings (CSV)
48
49

50 **Accessions Codes**

51
52
53 The PDB accession codes for X-ray cocrystal structures of wild-
54 type HIV-1 protease with DRV, LPV, **4**, **12a**, **12b**, **12c**, **12d**, **12e**,
55
56
57
58
59
60

1
2
3 **12f, 12g, 12h, 15a, 15b,** and **15c** are 6DGX, 6PJB, 6PJC, 6PJD, 6PJE,
4
5 6PJG, 6PJH, 6PJI, 6PJK, 6PJL, 6PJM, 6PJN, and 6PJO, respectively.
6
7 Authors will release the atomic coordinates and experimental data
8
9 upon article publication.
10
11
12
13

14 **AUTHOR INFORMATION**

15 16 17 **Corresponding Authors:**

18
19 *Akbar Ali: Phone: +1 508 856 8873; Fax: +1 508 856 6464; E-mail:
20
21 Akbar.Ali@umassmed.edu
22
23

24
25 *Celia A. Schiffer: Phone: +1 508 856 8008; Fax: +1 508 856 6464;
26
27 E-mail: Celia.Schiffer@umassmed.edu
28
29

30 **ORCID**

31
32
33 Akbar Ali: 0000-0003-3491-791X
34

35
36 Celia A. Schiffer: 0000-0003-2270-6613
37

38 **Author Contributions**

39
40
41 [§]These authors contributed equally to this work.
42
43

44 **Present Addresses**

45
46
47 For LNR: Raybow Pharmaceutical USA, 158 McLean Road, Brevard,
48
49 North Carolina 28712, United States
50
51

52 **Notes**

53
54
55 The authors declare no competing financial interest.
56
57
58
59
60

Funding Sources

This work was supported by a grant from the National Institute of General Medical Sciences of the NIH (P01 GM109767).

ACKNOWLEDGMENT

This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. GM/CA@APS has been funded in whole or in part with Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). The Eiger 16M detector was funded by an NIH-Office of Research Infrastructure Programs, High-End Instrumentation Grant (1S10OD012289-01A1). We thank the beamline specialists at 23-ID-D for their help in data collection. We thank Ellen A. Nalivaika for providing the HIV-1 protease variants, the AIDS Research and Reference Reagent Program (NIAD, NIH) for reference protease inhibitors, and Drs. William Royer, Brian Kelch and members of the Schiffer and Miller laboratories for helpful discussions. This work was supported by a grant from the National Institute of General Medical Sciences of the NIH (P01 GM109767).

ABBREVIATIONS

1
2
3 ATV, atazanavir; HIV-1, human immunodeficiency virus type 1; *bis*-THF, *bis*-tetrahydrofuran;
4 Boc, *tert*-butoxycarbonyl; DIEA, *N,N*-diisopropylethylamine; DRV, darunavir; HATU, 1-
5
6 [bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate;
7
8 FRET, fluorescence resonance energy transfer; LPV, lopinavir; MDR, multidrug-resistant; PIs,
9
10 protease inhibitors; SAR, structure-activity relationships; TFA, trifluoroacetic acid; THF,
11
12 tetrahydrofuran.
13
14
15
16
17

18 REFERENCES

- 19
20
21 1. Cihlar, T.; Fordyce, M. Current status and prospects of HIV
22 treatment. *Curr. Opin. Virol.* **2016**, *18*, 50–56.
23
24
25 2. Saag, M. S.; Benson, C. A.; Gandhi, R. T.; Hoy, J. F.;
26 Landovitz, R. J.; Mugavero, M. J.; Sax, P. E.; Smith, D. M.;
27 Thompson, M. A.; Buchbinder, S. P.; Del Rio, C.; Eron, J. J., Jr.;
28 Fatkenheuer, G.; Gunthard, H. F.; Molina, J. M.; Jacobsen, D. M.;
29 Volberding, P. A. Antiretroviral drugs for treatment and
30 prevention of HIV infection in adults: 2018 recommendations of the
31 international antiviral society-USA panel. *JAMA* **2018**, *320*, 379–
32 396.
33
34 3. Panel on Antiretroviral Guidelines for Adults and
35 Adolescents. *Guidelines for the Use of Antiretroviral Agents in*
36 *Adults and Adolescents with HIV*; Department of Health and Human
37 Services. Available at
38 <https://files.aidsinfo.nih.gov/contentfiles/lvguidelines/AdultandAdolescentGL.pdf>. (Accessed February 15, 2020).
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 4. Ali, A.; Bandaranayake, R. M.; Cai, Y.; King, N. M.; Kolli,
4
5 M.; Mittal, S.; Murzycki, J. F.; Nalam, M. N.; Nalivaika, E. A.;
6
7 Ozen, A.; Prabu-Jeyabalan, M. M.; Thayer, K.; Schiffer, C. A.
8
9 Molecular basis for drug resistance in HIV-1 protease. *Viruses*
10
11 **2010**, *2*, 2509-2535.
12
13
14 5. Ghosh, A. K.; Osswald, H. L.; Prato, G. Recent progress in
15
16 the development of HIV-1 protease inhibitors for the treatment of
17
18 HIV/AIDS. *J. Med. Chem.* **2016**, *59*, 5172-5208.
19
20
21 6. Surleraux, D. L.; Tahri, A.; Verschueren, W. G.; Pille, G.
22
23 M.; de Kock, H. A.; Jonckers, T. H.; Peeters, A.; De Meyer, S.;
24
25 Azijn, H.; Pauwels, R.; de Bethune, M. P.; King, N. M.; Prabu-
26
27 Jeyabalan, M.; Schiffer, C. A.; Wigerinck, P. B. Discovery and
28
29 selection of TMC114, a next generation HIV-1 protease inhibitor.
30
31 *J. Med. Chem.* **2005**, *48*, 1813-1822.
32
33
34 7. Ghosh, A. K.; Chapsal, B. D.; Weber, I. T.; Mitsuya, H. Design
35
36 of HIV protease inhibitors targeting protein backbone: an
37
38 effective strategy for combating drug resistance. *Acc. Chem. Res.*
39
40 **2008**, *41*, 78-86.
41
42
43 8. Tie, Y.; Boross, P. I.; Wang, Y. F.; Gaddis, L.; Hussain, A.
44
45 K.; Leshchenko, S.; Ghosh, A. K.; Louis, J. M.; Harrison, R. W.;
46
47 Weber, I. T. High resolution crystal structures of HIV-1 protease
48
49 with a potent non-peptide inhibitor (UIC-94017) active against
50
51 multi-drug-resistant clinical strains. *J. Mol. Biol.* **2004**, *338*,
52
53 341-352.
54
55
56
57
58
59
60

1
2
3 9. Ghosh, A. K.; Ramu Sridhar, P.; Kumaragurubaran, N.; Koh, Y.;
4
5 Weber, I. T.; Mitsuya, H. Bis-tetrahydrofuran: a privileged ligand
6
7 for darunavir and a new generation of hiv protease inhibitors that
8
9 combat drug resistance. *ChemMedChem* **2006**, *1*, 939-50.

10
11
12 10. Hazen, R.; Harvey, R.; Ferris, R.; Craig, C.; Yates, P.;
13
14 Griffin, P.; Miller, J.; Kaldor, I.; Ray, J.; Samano, V.; Furfine,
15
16 E.; Spaltenstein, A.; Hale, M.; Tung, R.; St Clair, M.; Hanlon,
17
18 M.; Boone, L. In vitro antiviral activity of the novel, tyrosyl-
19
20 based human immunodeficiency virus (HIV) type 1 protease inhibitor
21
22 brecanavir (GW640385) in combination with other antiretrovirals
23
24 and against a panel of protease inhibitor-resistant HIV.
25
26 *Antimicrob. Agents Chemother.* **2007**, *51*, 3147-3154.

27
28
29
30 11. Cihlar, T.; He, G. X.; Liu, X.; Chen, J. M.; Hatada, M.;
31
32 Swaminathan, S.; McDermott, M. J.; Yang, Z. Y.; Mulato, A. S.;
33
34 Chen, X.; Leavitt, S. A.; Stray, K. M.; Lee, W. A. Suppression of
35
36 HIV-1 protease inhibitor resistance by phosphonate-mediated
37
38 solvent anchoring. *J. Mol. Biol.* **2006**, *363*, 635-647.

39
40
41 12. Dandache, S.; Sevigny, G.; Yelle, J.; Stranix, B. R.; Parkin,
42
43 N.; Schapiro, J. M.; Wainberg, M. A.; Wu, J. J. In vitro antiviral
44
45 activity and cross-resistance profile of PL-100, a novel protease
46
47 inhibitor of human immunodeficiency virus type 1. *Antimicrob.*
48
49 *Agents Chemother.* **2007**, *51*, 4036-4043.

50
51
52 13. Kesteleyn, B.; Amsoms, K.; Schepens, W.; Hache, G.;
53
54 Verschueren, W.; Van De Vreken, W.; Rombauts, K.; Meurs, G.;

1
2
3 Sterkens, P.; Stoops, B.; Baert, L.; Austin, N.; Wegner, J.;
4
5 Masungi, C.; Dierynck, I.; Lundgren, S.; Jonsson, D.; Parkes, K.;
6
7 Kalayanov, G.; Wallberg, H.; Rosenquist, A.; Samuelsson, B.; Van
8
9 Emelen, K.; Thuring, J. W. Design and synthesis of HIV-1 protease
10
11 inhibitors for a long-acting injectable drug application. *Bioorg.*
12
13 *Med. Chem. Lett.* **2013**, *23*, 310-317.

14
15
16 14. Bungard, C. J.; Williams, P. D.; Schulz, J.; Wiscount, C. M.;
17
18 Holloway, M. K.; Loughran, H. M.; Manikowski, J. J.; Su, H. P.;
19
20 Bennett, D. J.; Chang, L.; Chu, X. J.; Crespo, A.; Dwyer, M. P.;
21
22 Keertikar, K.; Morriello, G. J.; Stamford, A. W.; Waddell, S. T.;
23
24 Zhong, B.; Hu, B.; Ji, T.; Diamond, T. L.; Bahnck-Teets, C.;
25
26 Carroll, S. S.; Fay, J. F.; Min, X.; Morris, W.; Ballard, J. E.;
27
28 Miller, M. D.; McCauley, J. A. Design and synthesis of piperazine
29
30 sulfonamide cores leading to highly potent HIV-1 protease
31
32 inhibitors. *ACS Med. Chem. Lett.* **2017**, *8*, 1292-1297.

33
34
35 15. Miller, J. F.; Andrews, C. W.; Brieger, M.; Furfine, E. S.;
36
37 Hale, M. R.; Hanlon, M. H.; Hazen, R. J.; Kaldor, I.; McLean, E.
38
39 W.; Reynolds, D.; Sammond, D. M.; Spaltenstein, A.; Tung, R.;
40
41 Turner, E. M.; Xu, R. X.; Sherrill, R. G. Ultra-potent P1 modified
42
43 arylsulfonamide HIV protease inhibitors: the discovery of GW0385.
44
45 *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1788-1794.

46
47
48 16. He, G.-X.; Yang, Z.-Y.; Williams, M.; Callebaut, C.; Cihlar,
49
50 T.; Murray, B. P.; Yang, C.; Mitchell, M. L.; Liu, H.; Wang, J.;
51
52 Arimilli, M.; Eisenberg, E.; Stray, K. M.; Tsai, L. K.; Hatada,
53
54
55
56
57
58
59
60

1
2
3 M.; Chen, X.; Chen, J. M.; Wang, Y.; Lee, M. S.; Strickley, R. G.;
4
5 Iwata, Q.; Zheng, X.; Kim, C. U.; Swaminathan, S.; Desai, M. C.;
6
7 Lee, W. A.; Xu, L. Discovery of GS-8374, a potent human
8
9 immunodeficiency virus type 1 protease inhibitor with a superior
10
11 resistance profile. *MedChemComm* **2011**, *2*, 1093-1098.
12
13

14
15 17. Windsor, I. W.; Palte, M. J.; Lukesh, J. C., 3rd; Gold, B.;
16
17 Forest, K. T.; Raines, R. T. Sub-picomolar inhibition of HIV-1
18
19 protease with a boronic acid. *J. Am. Chem. Soc.* **2018**, *140*, 14015-
20
21 14018.
22

23
24 18. Nalam, M. N.; Ali, A.; Reddy, G. S.; Cao, H.; Anjum, S. G.;
25
26 Altman, M. D.; Yilmaz, N. K.; Tidor, B.; Rana, T. M.; Schiffer, C.
27
28 A. Substrate envelope-designed potent HIV-1 protease inhibitors to
29
30 avoid drug resistance. *Chem. Biol.* **2013**, *20*, 1116-1124.
31

32
33 19. Aoki, M.; Hayashi, H.; Rao, K. V.; Das, D.; Higashi-Kuwata,
34
35 N.; Bulut, H.; Aoki-Ogata, H.; Takamatsu, Y.; Yedidi, R. S.; Davis,
36
37 D. A.; Hattori, S. I.; Nishida, N.; Hasegawa, K.; Takamune, N.;
38
39 Nyalapatla, P. R.; Osswald, H. L.; Jono, H.; Saito, H.; Yarchoan,
40
41 R.; Misumi, S.; Ghosh, A. K.; Mitsuya, H. A novel central nervous
42
43 system-penetrating protease inhibitor overcomes human
44
45 immunodeficiency virus 1 resistance with unprecedented aM to pM
46
47 potency. *Elife* **2017**, *6*, e28020.
48
49

50
51 20. Ghosh, A. K.; P, R. N.; Kovala, S.; Rao, K. V.; Brindisi, M.;
52
53 Osswald, H. L.; Amano, M.; Aoki, M.; Agniswamy, J.; Wang, Y. F.;
54
55 Weber, I. T.; Mitsuya, H. Design and synthesis of highly potent
56
57
58
59
60

1
2
3 HIV-1 protease inhibitors containing tricyclic fused ring systems
4
5 as novel P2 ligands: structure-activity studies, biological and X-
6
7 ray structural analysis. *J. Med. Chem.* **2018**, *61*, 4561-4577.

8
9
10 21. Ghosh, A. K.; Rao, K. V.; Nyalapatla, P. R.; Osswald, H. L.;
11
12 Martyr, C. D.; Aoki, M.; Hayashi, H.; Agniswamy, J.; Wang, Y. F.;
13
14 Bulut, H.; Das, D.; Weber, I. T.; Mitsuya, H. Design and
15
16 development of highly potent HIV-1 protease inhibitors with a
17
18 crown-like oxotricyclic core as the P2-ligand to combat multidrug-
19
20 resistant HIV variants. *J. Med. Chem.* **2017**, *60*, 4267-4278.

21
22
23 22. Ghosh, A. K.; Chapsal, B. D.; Baldrige, A.; Steffey, M. P.;
24
25 Walters, D. E.; Koh, Y.; Amano, M.; Mitsuya, H. Design and
26
27 synthesis of potent HIV-1 protease inhibitors incorporating
28
29 hexahydrofuropyranol-derived high affinity P2 ligands: structure-
30
31 activity studies and biological evaluation. *J. Med. Chem.* **2011**,
32
33
34 *54*, 622-634.

35
36
37 23. King, N. M.; Prabu-Jeyabalan, M.; Nalivaika, E. A.; Schiffer,
38
39 C. A. Combating susceptibility to drug resistance: lessons from
40
41 HIV-1 protease. *Chem. Biol.* **2004**, *11*, 1333-1338.

42
43
44 24. Gerlits, O.; Wymore, T.; Das, A.; Shen, C. H.; Parks, J. M.;
45
46 Smith, J. C.; Weiss, K. L.; Keen, D. A.; Blakeley, M. P.; Louis,
47
48 J. M.; Langan, P.; Weber, I. T.; Kovalevsky, A. Long-range
49
50 electrostatics-induced two-proton transfer captured by neutron
51
52 crystallography in an enzyme catalytic site. *Angew. Chem. Int. Ed.*
53
54
55 *Engl.* **2016**, *55*, 4924-4927.

- 1
2
3 25. Kovalevsky, A.; Gerlits, O.; Beltran, K.; Weiss, K. L.; Keen,
4 D. A.; Blakeley, M. P.; Louis, J. M.; Weber, I. T. Proton transfer
5 and drug binding details revealed in neutron diffraction studies
6 of wild-type and drug resistant HIV-1 protease. *Methods Enzymol.*
7
8 **2020**, *634*, 257-279.
9
10
11
12
13
14 26. Chen, X.; Kempf, D. J.; Li, L.; Sham, H. L.; Vasavanonda, S.;
15 Wideburg, N. E.; Saldivar, A.; Marsh, K. C.; McDonald, E.; Norbeck,
16 D. W. Synthesis and SAR studies of potent HIV protease inhibitors
17 containing novel dimethylphenoxy acetates as P2 ligands. *Bioorg.*
18
19 *Med. Chem. Lett.* **2003**, *13*, 3657-3660.
20
21
22
23
24
25
26 27. Chen, X.; Li, L.; Kempf, D. J.; Sham, H.; Wideburg, N. E.;
27 Saldivar, A.; Vasavanonda, S.; Marsh, K. C.; McDonald, E.; Norbeck,
28 D. W. Evaluation of furofuran as a P2 ligand for symmetry-based
29 HIV protease inhibitors. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2847-
30
31
32
33
34
35
36 2852.
37
38
39 28. Cannizzaro, C. E.; Yang, Z.-Y.; Desai, M. C.; Mitchell, M.
40 L.; Xu, L.; Swaminathan, S.; Chen, J. M. Antiviral protease
41 inhibitors. PCT Int. App. WO 2008/011117, 2008.
42
43
44 29. Ozen, A.; Haliloglu, T.; Schiffer, C. A. Dynamics of
45 preferential substrate recognition in HIV-1 protease: redefining
46 the substrate envelope. *J. Mol. Biol.* **2011**, *410*, 726-744.
47
48
49
50
51 30. Paulsen, J. L.; Leidner, F.; Ragland, D. A.; Kurt Yilmaz, N.;
52 Schiffer, C. A. Interdependence of inhibitor recognition in HIV-1
53 protease. *J. Chem. Theory Comput.* **2017**, *13*, 2300-2309.
54
55
56
57
58
59
60

- 1
2
3 31. Haight, A. R.; Stuk, T. L.; Allen, M. S.; Bhagavatula, L.;
4
5 Fitzgerald, M.; Hannick, S. M.; Kerdesky, F. A. J.; Menzia, J. A.;
6
7 Parekh, S. I.; Robbins, T. A.; Scarpetti, D.; Tien, J.-H. J.
8
9 Reduction of an enaminone: synthesis of the diamino alcohol core
10
11 of ritonavir. *Org. Process Res. Dev.* **1999**, *3*, 94-100.
12
13
14 32. Engle, K. M.; Wang, D. H.; Yu, J. Q. Ligand-accelerated C-H
15
16 activation reactions: evidence for a switch of mechanism. *J. Am.*
17
18 *Chem. Soc.* **2010**, *132*, 14137-14151.
19
20
21 33. Windsor, I. W.; Raines, R. T. Fluorogenic Assay for Inhibitors
22
23 of HIV-1 Protease with Sub-picomolar Affinity. *Sci. Rep.* **2015**, *5*,
24
25 11286.
26
27
28 34. Scott, W. R.; Schiffer, C. A. Curling of flap tips in HIV-1
29
30 protease as a mechanism for substrate entry and tolerance of drug
31
32 resistance. *Structure* **2000**, *8*, 1259-1265.
33
34
35 35. Yu, Y.; Wang, J.; Chen, Z.; Wang, G.; Shao, Q.; Shi, J.; Zhu,
36
37 W. Structural insights into HIV-1 protease flap opening processes
38
39 and key intermediates. *RSC Adv.* **2017**, *7*, 45121-45128.
40
41
42 36. Ozen, A.; Lin, K. H.; Kurt Yilmaz, N.; Schiffer, C. A.
43
44 Structural basis and distal effects of Gag substrate coevolution
45
46 in drug resistance to HIV-1 protease. *Proc. Natl. Acad. Sci. U. S.*
47
48 *A.* **2014**, *111*, 15993-15998.
49
50
51 37. King, N. M.; Melnick, L.; Prabu-Jeyabalan, M.; Nalivaika, E.
52
53 A.; Yang, S. S.; Gao, Y.; Nie, X.; Zepp, C.; Heefner, D. L.;

1
2
3 Schiffer, C. A. Lack of synergy for inhibitors targeting a multi-
4 drug-resistant HIV-1 protease. *Protein Sci.* **2002**, *11*, 418-429.

5
6
7 38. Rose, J. R.; Salto, R.; Craik, C. S. Regulation of
8 autoproteolysis of the HIV-1 and HIV-2 proteases with engineered
9 amino acid substitutions. *J. Biol. Chem.* **1993**, *268*, 11939-11945.

10
11
12 39. Hui, J. O.; Tomasselli, A. G.; Reardon, I. M.; Lull, J. M.;
13 Brunner, D. P.; Tomich, C. S.; Heinrikson, R. L. Large scale
14 purification and refolding of HIV-1 protease from *Escherichia coli*
15 inclusion bodies. *J. Protein Chem.* **1993**, *12*, 323-327.

16
17
18 40. Rusere, L. N.; Lockbaum, G. J.; Lee, S. K.; Henes, M.;
19 Kosovrasti, K.; Spielvogel, E.; Nalivaika, E. A.; Swanstrom, R.;
20 Yilmaz, N. K.; Schiffer, C. A.; Ali, A. HIV-1 protease inhibitors
21 incorporating stereochemically defined P2' ligands to optimize
22 hydrogen bonding in the substrate envelope. *J. Med. Chem.* **2019**,
23 *62*, 8062-8079.

24
25
26 41. Wei, X.; Decker, J. M.; Liu, H.; Zhang, Z.; Arani, R. B.;
27 Kilby, J. M.; Saag, M. S.; Wu, X.; Shaw, G. M.; Kappes, J. C.
28 Emergence of resistant human immunodeficiency virus type 1 in
29 patients receiving fusion inhibitor (T-20) monotherapy.
30 *Antimicrob. Agents Chemother.* **2002**, *46*, 1896-1905.

31
32
33 42. Lee, S. K.; Harris, J.; Swanstrom, R. A strongly transdominant
34 mutation in the human immunodeficiency virus type 1 gag gene
35 defines an Achilles heel in the virus life cycle. *J. Virol.* **2009**,
36 *83*, 8536-8543.

- 1
2
3 43. Lockbaum, G. J.; Leidner, F.; Rusere, L. N.; Henes, M.;
4
5 Kosovrasti, K.; Nachum, G. S.; Nalivaika, E. A.; Ali, A.; Yilmaz,
6
7 N. K.; Schiffer, C. A. Structural Adaptation of Darunavir Analogues
8
9 against Primary Mutations in HIV-1 Protease. *ACS Infect. Dis.* **2019**,
10
11 5, 316-325.
12
13
14 44. Otwinowski, Z.; Minor, W. Processing of X-ray diffraction
15
16 data collected in oscillation mode. *Methods Enzymol.* **1997**, 276,
17
18 307-326.
19
20
21 45. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M.
22
23 D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software.
24
25 *J. Appl. Crystallogr.* **2007**, 40, 658-674.
26
27
28 46. Emsley, P.; Cowtan, K. Coot: model-building tools for
29
30 molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.*
31
32 **2004**, 60, 2126-2132.
33
34
35 47. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.;
36
37 Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G.
38
39 J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.;
40
41 Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.;
42
43 Terwilliger, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-
44
45 based system for macromolecular structure solution. *Acta*
46
47 *Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, 66, 213-221.
48
49
50 48. Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D.
51
52 electronic Ligand Builder and Optimization Workbench (eLBOW): a
53
54
55
56
57
58
59
60

1
2
3 tool for ligand coordinate and restraint generation. *Acta*
4 *Crystallogr., Sect. D: Biol. Crystallogr.* **2009**, *65*, 1074-1080.

5
6
7
8 49. Brunger, A. T. Free R value: a novel statistical quantity for
9 assessing the accuracy of crystal structures. *Nature* **1992**, *355*,
10 472-475.

11
12
13
14 50. Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.;
15 Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B., 3rd;
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Snoeyink, J.; Richardson, J. S.; Richardson, D. C. MolProbity:
all-atom contacts and structure validation for proteins and
nucleic acids. *Nucleic Acids Res.* **2007**, *35*, W375-383.

51. PyMOL. *The PyMOL Molecular Graphics System, Version 2.3*;
Schrödinger, LLC.

52. Schrödinger. *Schrödinger Release 2019-2*; Schrödinger, LLC,
New York, NY, United States, 2019.

53. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang,
J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus,
J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.;
Friesner, R. A. OPLS3: a force field providing broad coverage of
drug-like small molecules and proteins. *J. Chem. Theory Comput.*
2016, *12*, 281-296.

54. Nalam, M. N.; Ali, A.; Altman, M. D.; Reddy, G. S.;
Chellappan, S.; Kairys, V.; Ozen, A.; Cao, H.; Gilson, M. K.;
Tidor, B.; Rana, T. M.; Schiffer, C. A. Evaluating the substrate-
envelope hypothesis: structural analysis of novel HIV-1 protease

1
2
3 inhibitors designed to be robust against drug resistance. *J. Virol.*
4
5 **2010**, *84*, 5368–5378.

7 55. Leidner, F.; Kurt Yilmaz, N.; Paulsen, J.; Muller, Y. A.;
8
9 Schiffer, C. A. Hydration structure and dynamics of inhibitor-
10
11 bound HIV-1 protease. *J. Chem. Theory Comput.* **2018**, *14*, 2784–2796.

14 56. Henes, M.; Lockbaum, G. J.; Kosovrasti, K.; Leidner, F.;
15
16 Nachum, G. S.; Nalivaika, E. A.; Lee, S. K.; Spielvogel, E.; Zhou,
17
18 S.; Swanstrom, R.; Bolon, D. N. A.; Kurt Yilmaz, N.; Schiffer, C.
19
20 A. Picomolar to micromolar: elucidating the role of distal
21
22 mutations in HIV-1 protease in conferring drug resistance. *ACS*
23
24 *Chem. Biol.* **2019**, *14*, 2441–2452.

27 57. Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B. On the
28
29 role of the crystal environment in determining protein side-chain
30
31 conformations. *J. Mol. Biol.* **2002**, *320*, 597–608.

34 58. Olsson, M. H.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J.
35
36 H. PROPKA3: consistent treatment of internal and surface residues
37
38 in empirical pKa predictions. *J. Chem. Theory Comput.* **2011**, *7*,
39
40 525–537.

43 59. Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.;
44
45 Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J.
46
47 R.; Murphy, R.; Philipp, D. M.; Repasky, M. P.; Zhang, L. Y.;
48
49 Berne, B. J.; Friesner, R. A.; Gallicchio, E.; Levy, R. M.
50
51 Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J.*
52
53 *Comput. Chem.* **2005**, *26*, 1752–1780.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

TOC Graphic

