

Structure-Guided Design of Potent Inhibitors of SARS-CoV-2 3CL Protease: Structural, Biochemical, and Cell-based Studies.

Chamandi S. Dampalla,^{1,#} Athri M. Rathnayake,^{1,#} Krishani Dinali Perera,³ Abdul-Rahman M. Jesri,¹ Harry Nhat Nguyen,¹ Matthew J. Miller,¹ Hayden A. Thurman,¹ Jian Zheng,² Maithri M. Kashipathy,⁴ Kevin P. Battaile,⁵ Scott Lovell,⁴ Stanley Perlman,² Yunjeong Kim,^{3*} William C. Groutas,^{1*} Kyeong-Ok Chang^{3*}

¹Department of Chemistry, Wichita State University, Wichita, Kansas 67260, USA

²Department of Microbiology and Immunology, University of Iowa, Iowa City, IA 52242, USA

³Department of Diagnostic Medicine & Pathobiology, College of Veterinary Medicine, Kansas State University, Manhattan, Kansas 66506, USA

⁴Protein Structure Laboratory, The University of Kansas, Lawrence, Kansas 66047, USA

⁵YX, New York Structural Biology Center, Upton, NY 11973, USA

#These authors contributed equally.

*authors to whom correspondence should be addressed.

¹Department of Chemistry, Wichita State University, Wichita, KS 67260
Tel. (316) 978 7374
e-mail: bill.groutas@wichita.edu

³Department of Diagnostic Medicine & Pathobiology, College of Veterinary Medicine, Kansas State University, Manhattan, KS 66506
(K.C.) Phone: (785) 532 3849. E-mail: kchang@vet.ksu.edu
(Y.K.) Phone: (785) 532 4616. E-mail: ykim@ksu.edu

Abstract

The COVID-19 pandemic is having a major impact on public health worldwide and there is an urgent need for the creation of an armamentarium of effective therapeutics, including vaccines, biologics, and small molecule therapeutics, to combat SARS-CoV-2 and emerging variants. Inspection of the virus life cycle reveals multiple viral and host-based choke points that can be exploited to combat the virus. SARS-CoV-2 3C-like protease (3CLpro), an enzyme essential for viral replication, is an attractive target for therapeutic intervention and the design of inhibitors of the protease may lead to the emergence of effective SARS-CoV-2-specific antivirals. We describe herein the results of our studies related to the application of X-ray crystallography, the Thorpe-Ingold effect, deuteration, and stereochemistry in the design of highly potent and non-toxic inhibitors of SARS-CoV-2 3CLpro.

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the etiological agent of coronavirus disease (COVID-19).¹ The severity of the ongoing pandemic is having a major impact on public health worldwide, and is further exacerbated by the emergence of more virulent strains.²⁻³ Intense worldwide efforts to combat the virus have led to the successful development of FDA-approved vaccines, and an array of potential therapeutics, such as monoclonal antibodies, repurposed drugs and others, are currently being evaluated in clinical trials, or are at various stages of clinical development.⁴⁻⁵

The SARS-CoV-2 life cycle encompasses multiple virus and host-based druggable targets that can be exploited, including for example inhibitors that block virus entry and fusion, and replication inhibitors targeting the 3C-like protease (3CLpro) and papain-like protease and the RNA-dependent RNA polymerase, among others. Attractive host-based targets include the proteases transmembrane serine protease 2 (TMPRSS2), cathepsin L, and furin. Thus, the development of small molecule therapeutics that target host or viral targets essential for viral replication is a potentially fruitful avenue of investigation.⁶⁻

9

Our foray in this area has focused on the structure-guided design of inhibitors of SARS-CoV-2 and MERS-CoV 3CLpro,¹⁰⁻¹² as well as feline infectious peritonitis virus (FIPV) protease inhibitors.¹³⁻¹⁴ SARS-CoV-2 3CLpro is a homodimer with a catalytic Cys-His dyad (Cys¹⁴⁵-His⁴¹) and an extended binding cleft. The protease displays a strong preference for a -Y-Z-Leu-Gln-X sequence, corresponding to the residues -P₄-P₃-P₂-P₁-P₁'-,¹⁵ where X is a small amino acid (Ser, Ala, Gly), Y is small hydrophobic amino acid,

and Z is solvent exposed and can tolerate polar or nonpolar amino acid chains. SARS-CoV-2 3CLpro is therefore an attractive target for drug development.¹⁶⁻²³

We recently described the structure-guided design of a dipeptidyl series of MERS-CoV and SARS-CoV-2 3CLpro inhibitors incorporating in their structure a piperidine¹¹ or cyclohexyl¹⁰ moiety capable of engaging in favorable binding interactions with the S₄ pocket. We furthermore demonstrated that members of the cyclohexyl series of compounds improve survival in a mouse model of MERS-CoV infection.¹⁰ In this report, we established a cell-based assay to screen inhibitors against SARS-CoV-2 3CLpro, which is safe (BSL2) and fast (takes less than 24 h). Furthermore, we report the results of structure-guided studies intended to interrogate the effects of stereochemistry, conformation, and structure, including the systematic introduction of fluorine (F-walk)²⁴⁻²⁵ around the structure of GC376¹²⁻¹⁴ and the synthesis of deuterated inhibitors,²⁶⁻²⁸ to modulate pharmacological activity, pharmacokinetic properties, and oral bioavailability.

RESULTS AND DISCUSSION

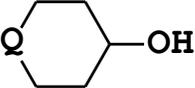
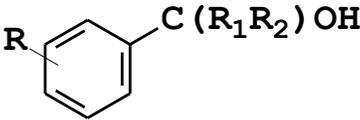
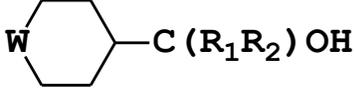
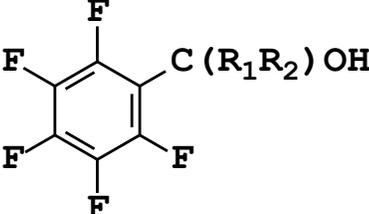
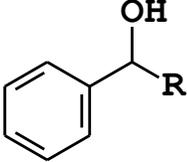
Chemistry

The synthesis of compounds **1-24b/c** entailed the use of a structurally diverse set of precursor alcohols (Table 1), some of which were commercially available. Alcohols **12-16** were readily synthesized from 4,4-difluorocyclohexane carboxylic acid via reduction to the corresponding alcohol by treatment with carbonyl diimidazole and sodium borohydride,²⁹ followed by oxidation with Dess-Martin periodinane reagent to yield the aldehyde. Subsequent treatment with an array of Grignard reagents generated alcohols **12, 14-16** (Scheme /panel A). Alcohol **13** was synthesized by reacting the methyl ester of

4,4-difluorocyclohexane carboxylic acid with excess methyl magnesium iodide, followed by acidic work up (Scheme /panel A). Deuterated alcohols **9**, **11**, **20**, and **22** were obtained by treatment of the precursor carboxylic acid with carbonyl diimidazole followed by the addition of sodium borodeuteride. All trans-substituted alcohols were synthesized by reducing the precursor 4-substituted cyclohexanone with sodium borohydride/CeCl₃.³⁰

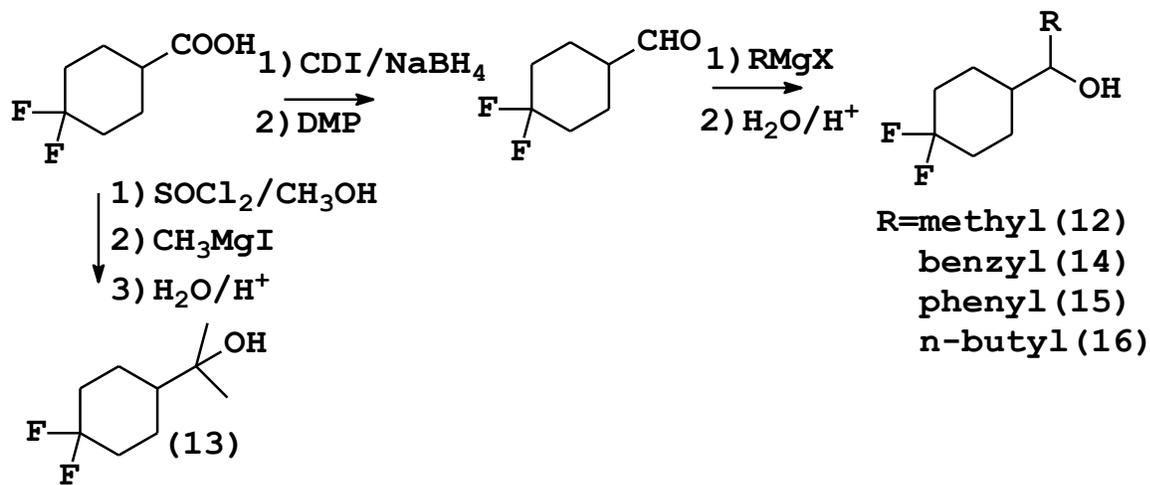
Compounds **1-24b/c** were readily obtained by reacting each precursor alcohol with disuccinimidyl carbonate,³¹ followed by coupling with amino alcohol **A**. The resulting product was treated with Dess-Martin periodinane to yield aldehydes **1-24b** which were converted to the corresponding bisulfite adducts **1-24c** upon treatment with sodium bisulfite (Scheme/panel B).³² An alternative synthesis was used in the case of compounds **6-8**, **10-16**, **23,24** which involved the reaction of the precursor alcohol with (L) leucine methyl ester isocyanate, as described in detail previously.³³ The synthesis of precursor amino alcohol **A** was readily accomplished by coupling (L) Z-Leu with a glutamine surrogate, followed by sequential reduction with LiBH₄ and removal of the protective group (H₂/Pd) (Scheme/panel C).

Table 1. Alcohol inputs

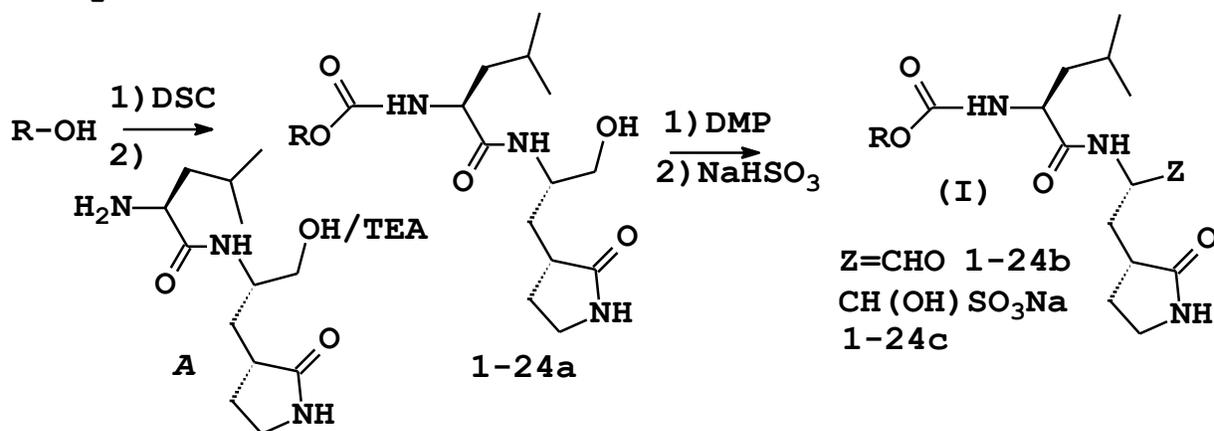
| | |
|---|--|
|  |  |
| $Q = C(CH_3)_2$ <u>1</u> | $R = o-F, R_1 = R_2 = H$ <u>17</u> |
| $iso-Pr$ <u>2</u> (trans) | $m-F, R_1 = R_2 = H$ <u>18</u> |
| $n-Pr$ <u>3</u> (trans) | $p-F, R_1 = R_2 = H$ <u>19</u> |
| $n-Bu$ <u>4</u> (trans) | $p-F, R_1 = R_2 = D$ <u>20</u> |
| $phenyl$ <u>5</u> (trans) | |
|  |  |
| $W = CF_3, R_1 = R_2 = H$ (trans) <u>6</u> | $R_1 = R_2 = H$ <u>21</u> or |
| $CF_3, R_1 = R_2 = H$ (cis) <u>7</u> | D <u>22</u> |
| $W = n-Pr, R_1 = R_2 = H$ (trans) <u>8</u> | |
| $n-Pr, R_1 = R_2 = D$ (trans) <u>9</u> | |
| $W = CF_2, R_1 = R_2 = H$ <u>10</u> | |
| $CF_2, R_1 = R_2 = D$ <u>11</u> |  |
| $CF_2, R_1 = Me, R_2 = H$ <u>12</u> | $R = n-Pr$ <u>23</u> |
| $CF_2, R_1 = R_2 = methyl$ <u>13</u> | $benzyl$ <u>24</u> |
| $CF_2, R_1 = benzyl, R_2 = H$ <u>14</u> | $H (GC376)$ |
| $CF_2, R_1 = phenyl, R_2 = H$ <u>15</u> | |
| $CF_2, R_1 = n-Bu, R_2 = H$ <u>16</u> | |

Scheme

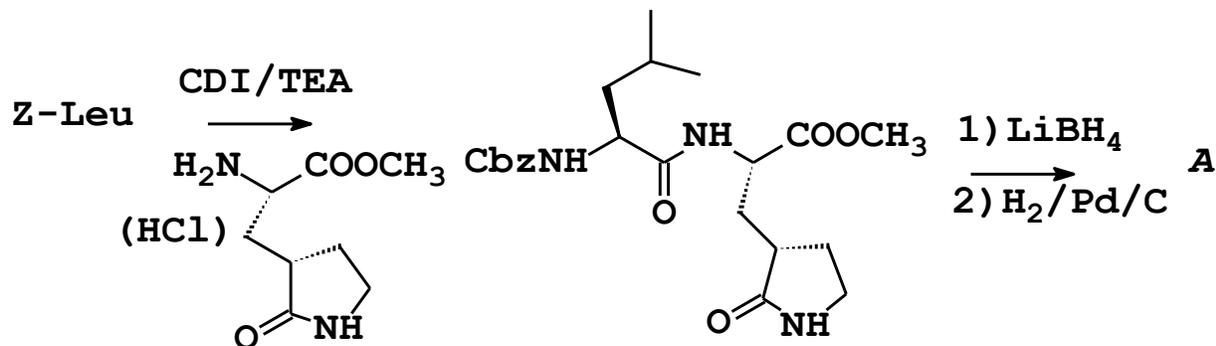
A Synthesis of precursor alcohols 12-16



B Synthesis of inhibitors 1-24b-c



C Synthesis of amino alcohol A



Biochemical Studies

Enzyme assays. The inhibitory activity of compounds **1-24b/c** against SARS-CoV-2 3CLpro in biochemical assays was determined as described in the experimental section. The IC₅₀ values against SARS-CoV-2 and MERS-CoV-2 in the enzyme assays are summarized in Table 2 and they are the average of at least two determinations. Most of the compounds potently inhibited SARS-CoV-2 3CLpro and displayed IC₅₀ values that ranged between 0.13 to 1.25 μM. Compounds **15b** and **15c** were the most effective against SARS-CoV-2 3CLpro with IC₅₀ values 0.13 and 0.17 μM, respectively. The inhibitory activity of a select number of compounds against MERS-CoV 3CLpro was also investigated. The compounds were found to be 3-5-fold more potent against MERS-CoV-3CLpro, with IC₅₀ values in the 40 to 150 nM range (Table 2). Interestingly, compounds **15b** and **15c** were the most effective against MERS-CoV-3CLpro as well with IC₅₀ values 0.04 and 0.05 μM, respectively. The broad spectrum of inhibitory activity displayed by these compounds enhances their therapeutic potential.

Table 2. IC₅₀ values of compounds **1-24b-c** against SARS-CoV-2 3CLpro and selected compounds against MERS-CoV 3CLpro.

| Compound | IC ₅₀ (μM) | | Compound | IC ₅₀ (μM) | |
|-------------|-----------------------|-----------------|------------|-----------------------|-----------|
| | SARS-CoV-2 | MERS-CoV | | SARS-CoV-2 | MERS-CoV |
| 1b | 0.25±0.05 | 0.12±0.02 | 13b | 1.20±0.57 | NT |
| 1c | 0.24±0.01 | 0.12±0.04 | 13c | 1.25±0.49 | NT |
| 2b | 0.24±0.01 | NT [#] | 14b | 0.31±0.11 | 0.10±0.02 |
| 2c | 0.24±0.02 | NT | 14c | 0.29±0.09 | 0.14±0.03 |
| 3b | 0.25±0.02 | NT | 15b | 0.13±0.04 | 0.04±0.01 |
| 3c | 0.19±0.02 | NT | 15c | 0.17±0.01 | 0.05±0.01 |
| 4b | 0.17±0.01 | 0.12±0.01 | 16b | 0.19±0.04 | NT |
| 4c | 0.18±0.01 | 0.15±0.01 | 16c | 0.27±0.03 | NT |
| 5b | 0.22±0.03 | NT | 17b | 0.34±0.06 | NT |
| 5c | 0.23±0.04 | NT | 17c | 0.38±0.04 | NT |
| 6b | 0.25±0.07 | NT | 18b | 0.35±0.04 | 0.07±0.01 |
| 6c | 0.19±0.03 | NT | 18c | 0.36±0.06 | 0.07±0.01 |
| 7b | 0.23±0.04 | NT | 19b | 0.39±0.04 | NT |
| 7c | 0.20±0.03 | NT | 19c | 0.40±0.03 | NT |
| 8b | 0.20±0.03 | 0.06±0.01 | 20b | 0.39±0.06 | NT |
| 8c | 0.20±0.01 | 0.11±0.01 | 20c | 0.32±0.06 | NT |
| 9b | 0.33±0.11 | NT | 21b | 0.33±0.01 | 0.05±0.01 |
| 9c | 0.33±0.01 | NT | 21c | 0.36±0.01 | 0.08±0.01 |
| 10b* | 0.43±0.10 | 0.08±0.01 | 22b | 0.47±0.23 | NT |
| 10c* | 0.41±0.10 | 0.08±0.01 | 22c | 0.60±0.21 | NT |
| 11b | 0.24±0.06 | NT | 23b | 0.33±0.25 | 0.08±0.01 |
| 11c | 0.21±0.13 | NT | 23c | 0.21±0.04 | 0.13±0.01 |
| 12b | 0.26±0.07 | NT | 24b | 0.50±0.16 | NT |
| 12c | 0.26±0.08 | NT | 24c | 0.31±0.01 | NT |

* IC₅₀s were from Rathnayake et. al.¹⁰ # NT: not tested.

Establishment of the cell-based assay for SARS-CoV-2 3CLpro inhibitors. We have previously reported EC₅₀ values determined by incubating SARS-CoV-2 3CLpro inhibitors and Vero E6 cells that were inoculated with SARS-CoV-2 at 50-100 plaque forming units/well.^{10,33} This cell-based assay requires a BSL3 facility and takes at least 2-3 days.

As an alternative method, we report herein a relatively fast and safe cell-based assay system to screen SARS-CoV-2 3CLpro inhibitors using two-plasmids. A similar cell-based assay has been reported;³⁴ however, in contrast, the present system utilizes the replication units of porcine respiratory and reproductive syndrome virus (PRRSV)³⁵ to express SARS-CoV-2 3CLpro. In this system plasmid 1, pR-SARS-CoV-2 3CLpro, was used to express SARS-CoV-2 3CLpro whereas plasmid 2, pGlo-VRLQS, was used to express luciferase-VRLQS in HEK293T cells (**Figure 1/ panel A**). The expressed inactive luciferase is activated by the catalytic mechanism of SARS-CoV-2 3CLpro in HEK293T cells. Hence, the inhibition of SARS-CoV-2 3CLpro was measured as a function of firefly luciferase activity (**Figure 1/ panel B**). The EC₅₀ values of a select number of compounds, including **1b/1c**, **4b/4c**, **8b/8c**, **14b/14c**, **15b/15c**, **18b/18c**, **21b/21c**, and **23b/23c**, were determined (Table 3). Inhibition curves by each compound were consistent with a dose-dependent mode and R² > 0.8 (**Figure 1/ panel C**). The IC₅₀ and EC₅₀ values of **15b** and **15c** were the lowest among tested compounds listed in Table 2 and Table 3.

Figure 1. Generation of a cell-based assay for screening SARS-CoV-2 3CLpro inhibitors in HEK293T cells. **Panel A.** Plasmid 1; pR-SA2-3CLpro, encodes SARS-CoV-2 3CLpro from the PRRSV reverse genetics system. Plasmid 2; pGlo-VRLQS encodes firefly luciferase with coronavirus 3CLpro recognition sequences VRLQS. Active luciferase is generated by the cleavage with CoV 3CLpro. **Panel B.** Semi-confluent HEK293T cells were transfected with two plasmids, and after overnight, various concentrations of each compound are applied to the cells. The inhibition of SARS-CoV-2 3CLpro is determined by measuring luciferase activity. **Panel C.** Inhibition curves of selected compounds, **1c**,

4c, 8c, 14c, 15c, 18c, 21c, and 23c using the cell-based assay with pR-SA2-3CLpro and pGlo-VRLQS.

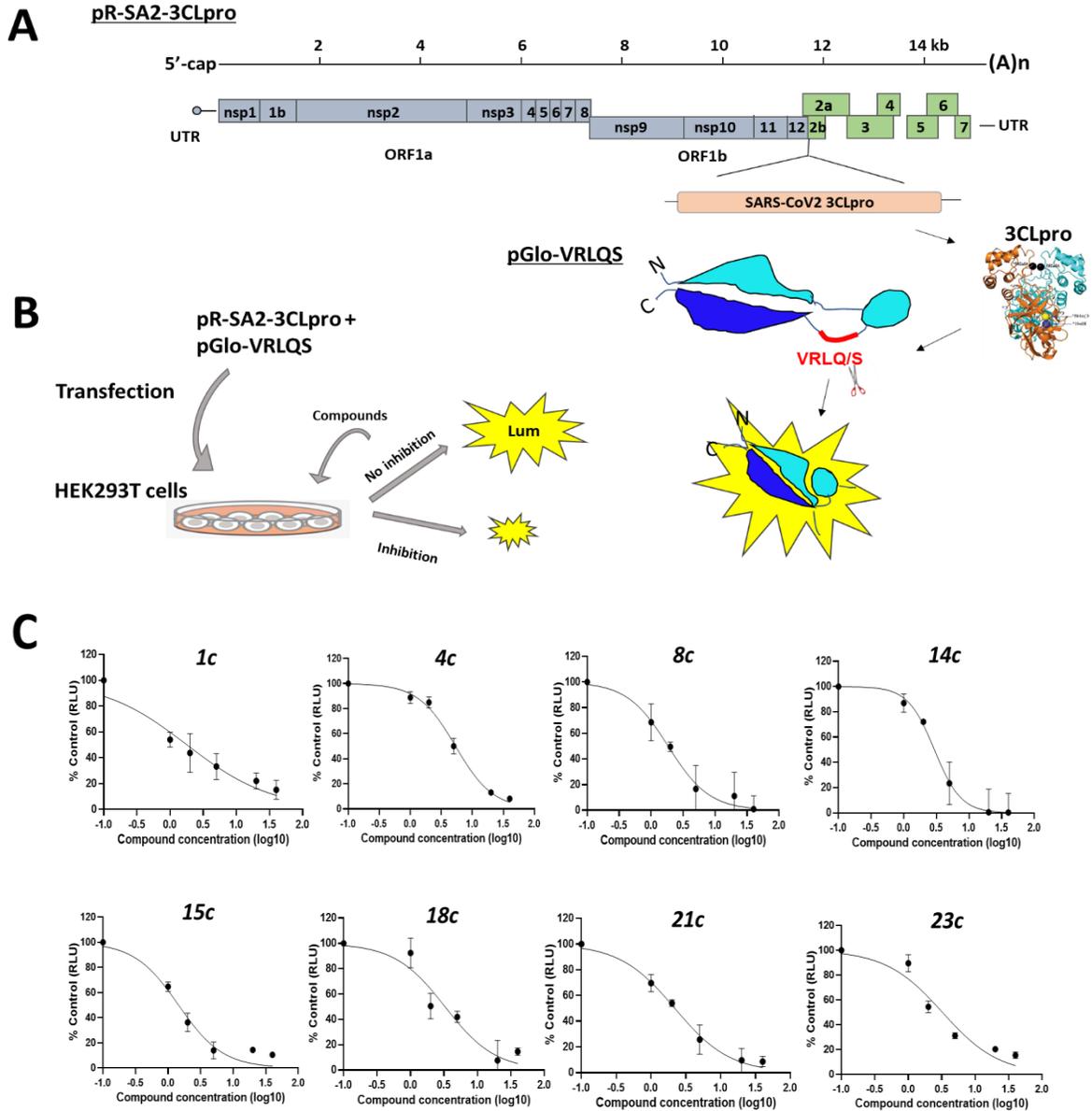


Table 3. EC₅₀ values of selected compounds against SARS-CoV-2 3CLpro from two-plasmid system and CC₅₀ values.

| Compound | EC ₅₀ (μM) | CC ₅₀ (μM) |
|---|-----------------------|-----------------------|
| 1b | 3.48±0.05 | > 100 |
| 1c | 1.67±0.45 | > 100 |
| 4b | 4.84±0.64 | 55±5 |
| 4c* | 4.23±1.54 | 58±8 |
| 8b | 1.03±0.49 | 43±9 |
| 8c | 4.43±0.35 | 44±4 |
| 14b | 2.95±0.94 | > 100 |
| 14c | 3.50±1.20 | > 100 |
| 15b | 1.03±0.47 | > 100 |
| 15c* | 1.45±0.42 | > 100 |
| 18b | 5.28±2.62 | > 100 |
| 18c | 3.40±0.59 | > 100 |
| 21b | 3.27±1.03 | > 100 |
| 21c | 4.84±0.64 | > 100 |
| 23b | 2.39±1.51 | > 100 |
| 23c | 3.72±1.60 | > 100 |
| * EC ₅₀ : 0.85±0.1 and 0.70±0.08 μM for 4c and 15c , respectively from live SARS-CoV-2 in Vero E6 cells. | | |

For comparative purposes, we determined EC₅₀ values using live SARS-CoV-2 in Vero E6 cells and the established two-plasmid system of two compounds (**4c**, **15c**) from this work, and a previously published 3CLpro inhibitor, GC376. The EC₅₀ of GC376 was found to be 0.23±0.01 μM in live SARS-CoV-2 in Vero E6 cells.³³ In the two-plasmid system, the EC₅₀ of GC376 was determined to be 3.15 ± 0.67 μM (14-fold higher), and the R² values of the inhibition curves were > 0.9. When the antiviral effects of compounds **4c** and **15c** were examined from live SARS-CoV-2 in Vero E6 cells, the EC₅₀ values were 0.85±0.1 and 0.70±0.08 μM, respectively (Table 3). The results show that while compounds in this series are cell permeable, the EC₅₀ values were higher from the two-plasmid system (2-fold for **15c** and 5-fold for **4c**) than those by live SARS-CoV-2 in Vero

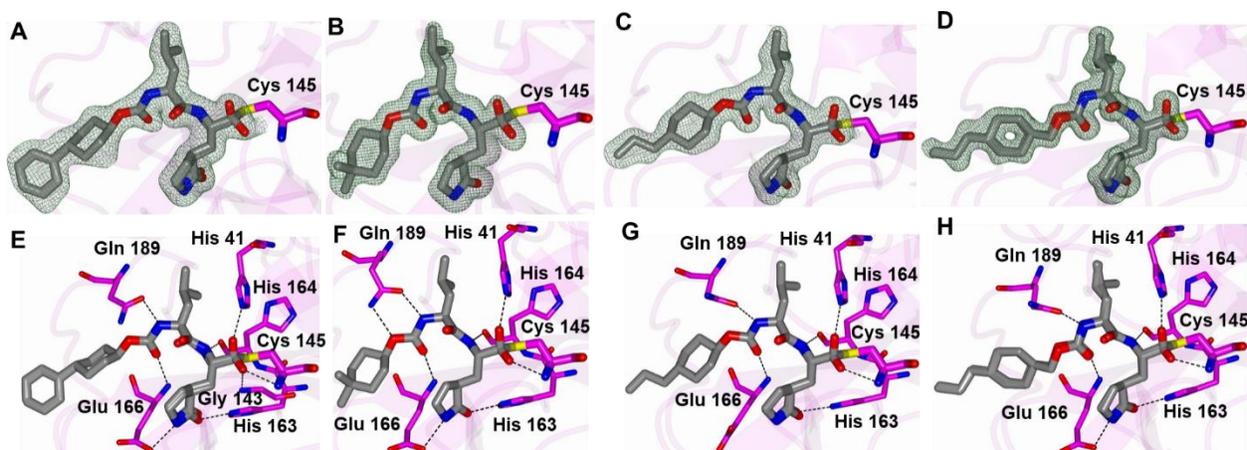
E6 cells. The higher EC₅₀s may be due to various reasons including different cell types, presence of transfection reagent, and higher and synchronized expression of the 3CLpro in HEK293T cells by the transfection of pR-SA2-3CLpro in the two-plasmid system. Most of the examined compounds showed minimal toxicity up to 100 μM, however, the CC₅₀ values for **4b/4c** and **8b/8c** were in the 40-60 μM range (Table 3). Although the EC₅₀ values obtained from the two-plasmid system are higher than those obtained from live SARS-CoV-2 in Vero E6 cells, considering the feasibility of conducting the experiments under BSL2 laboratory conditions and the relatively short amount of time required (24h), the cell based two-plasmid method could be a useful initial screening tool for 3CLpro inhibitors against SARS-CoV-2.

X-Ray Crystallographic Studies

A series of high resolution cocrystal structures were determined to elucidate the interaction of the inhibitors with the active site of SARS-CoV-2 3CLpro. Specifically, we sought to confirm the mechanism of action, identify the structural determinants associated with the binding of the inhibitors to the active site of the protease and ultimately harness the accumulated structural information and insights gained to further optimize pharmacological activity and PK parameters. Three groups of inhibitor types were analyzed with respect to their functional groups that interact within the S₄ subsite which are 1) non-polar substituents, 2) 4,4-difluorocyclohexyl groups that are connected to a stereocenter and, 3) fluorinated aryl compounds based on the structure of GC376.

For all structures described below, the active sites contained prominent difference electron density consistent with inhibitors covalently bound to Cys 145. Additionally, the

electron density was consistent with both the R and S enantiomers at the stereocenter formed by covalent attachment of the S_γ atom of Cys 145 and were therefore, modeled as each enantiomer with 0.5 occupancy. The γ-lactam ring of the inhibitor forms direct hydrogen bonds with Glu166 and His163, and Glu166 and Gln189 form additional H-bonds with the C=O and NH of the carbamate moiety in the inhibitor. The inhibitor engages in hydrophobic interactions with the leucine side chain, which is snugly accommodated in the S₂ pocket. The cocrystal structure confirms that the reaction of



Cys145 with the aldehyde warhead results in the formation of a tetrahedral hemithioacetal that is stabilized by a H-bond to His164.

Figure 2. Binding mode of inhibitors containing non-polar substituents. **5c (A/E)**, **1c (B/F)**, **3c (C/G)** and **8b (D/H)** with SARS-CoV-2 3CLpro. Fo-Fc Polder omit map (**A-D**) contoured at 3 σ . Hydrogen bond interactions (**E-H**) are drawn as dashed lines.

Non-polar substituents. The structures of **5c**, **1c**, **3c** and **8b** displayed well-defined electron density and similar hydrogen bond interactions as shown in **Figure 2**. For all structures, the non-polar groups are mainly positioned within the S₄ subsite near a hydrophobic ridge formed by residues Leu 167, Pro 168, Gly 170 and Ala 191 (**Figure 3**).

However, the dimethyl cyclohexyl ring in **1c** is too short to fully engage the hydrophobic ridge in the S₄ subsite (**Figure 3B**). The addition of an n-propyl group in **3c** permits further engagement with the hydrophobic cleft and the extra carbon atom in **8b** allows the propyl group to extend even further (**Figures 3C** and **3D**). Superposition of **3c** and **1c** (**Figure 3E**) shows that the 4,4-dimethylcyclohexyl ring is moved slightly out of the S₄ subsite relative to the n-propyl group in **3c**. Additionally, superposition of **3c** and **8b** revealed quite similar binding modes although the n-propyl group of **8b** is positioned deeper within the S₄ subsite (**Figure 3F**). Overall, the similar binding modes and attendant high potency of the inhibitors are reflected in their low IC₅₀ values and similar potencies (Table 2, compounds **1-5b/c**). With respect to compound **8**, it was envisaged that the corresponding deuterated compound **9**, found to be nearly equipotent to non-deuterated compound **8** (Table 2), would likely display improved PK properties.²⁶⁻²⁸

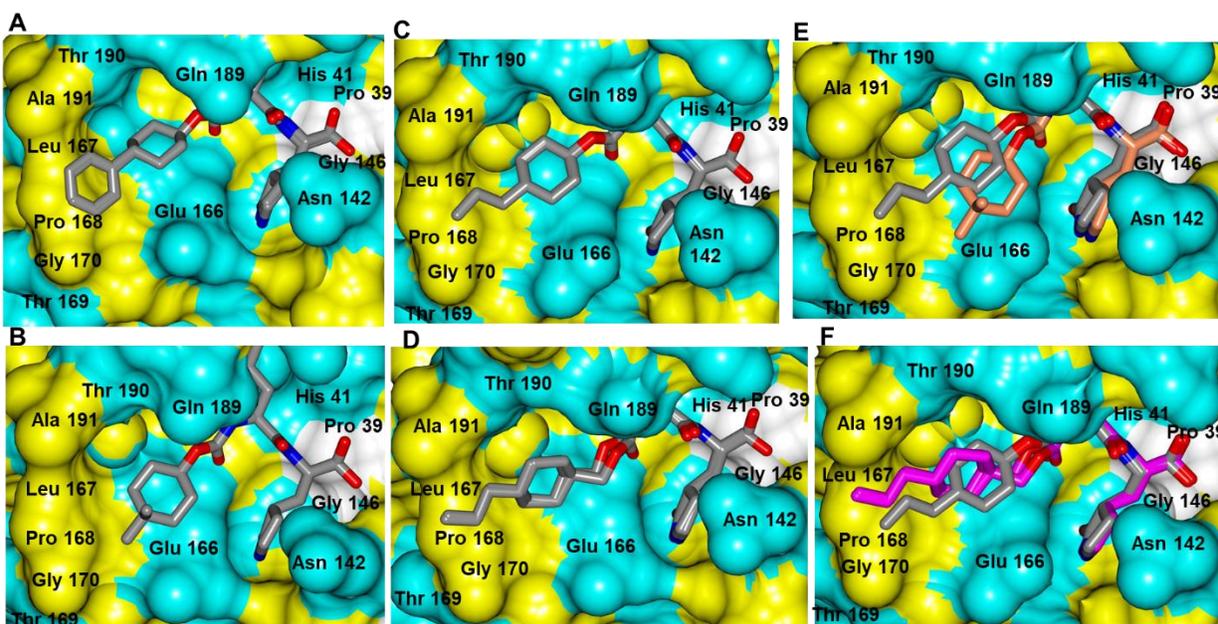


Figure 3. Surface representation showing the orientation of non-polar groups near the S₄ subsite of SARS-CoV-2 3CLpro with neighboring residues colored yellow (nonpolar),

cyan (polar), and white (weakly polar). **5c (A)**, **1c (B)**, **3c (C)**, **8b (D)**. Superposition of **3c** (gray) and **1c** (coral) (**E**). Superposition of **3c** (gray) and **8b** (magenta) (**F**).

4,4-Difluorocyclohexyl compounds. In previous studies related to norovirus 3CLpro inhibitors, the strategic introduction of a *gem*-dimethyl group into the inhibitor structure resulted in enhanced potency by restricting rotation around the nearby single bonds and lowering the entropic penalty associated with binding.³⁶ Thus, we sought to capitalize on this by synthesizing *gem*-dimethyl-substituted compound **13c** and, additionally, achieve the same end by introducing a stereocenter (**12c**). The structures of **12b**, **13c** and **14c** with SARS-CoV 3CLpro displayed well-defined electron density and the typically observed hydrogen bond interactions (**Figure 4**). The 4,4-difluorocyclohexyl rings for all structures are positioned near the hydrophobic cleft in the S₄ subsite as shown in **Figure 5 A-C**. Superposition of these structures revealed a nearly identical binding mode for **12b** and **13c** in which the 4,4-difluorocyclohexyl groups are positioned in the same region within the S₄ subsite (**Figure 5D**). For **14c**, the benzyl ring is oriented in a wide cleft formed by Asn 142 and Gln 189. However, the 4,4-difluorocyclohexyl ring of **13c** contacts residues Thr 190 and Ala 191 (3.0-3.2 Å) and forms new hydrogen bond interactions with the backbone oxygen and nitrogen atoms respectively (**Figure 4E**). This positions the 4,4-difluorocyclohexyl ring of **13c** deeper into the S₄ pocket and results in a conformational change in the loop spanning Gln 189 to Gly 195 and Glu 166 to Gly 170, in order to accommodate the new interactions and avoid steric clash. This results in the loss of the typical hydrogen bond between the side chain of Glu 166 and the glutamine surrogate of **13c** (**Figures 4E** and **5D**) which may explain why the IC₅₀ of **13c** is ~4-fold higher than those of **12b** and **14c**.

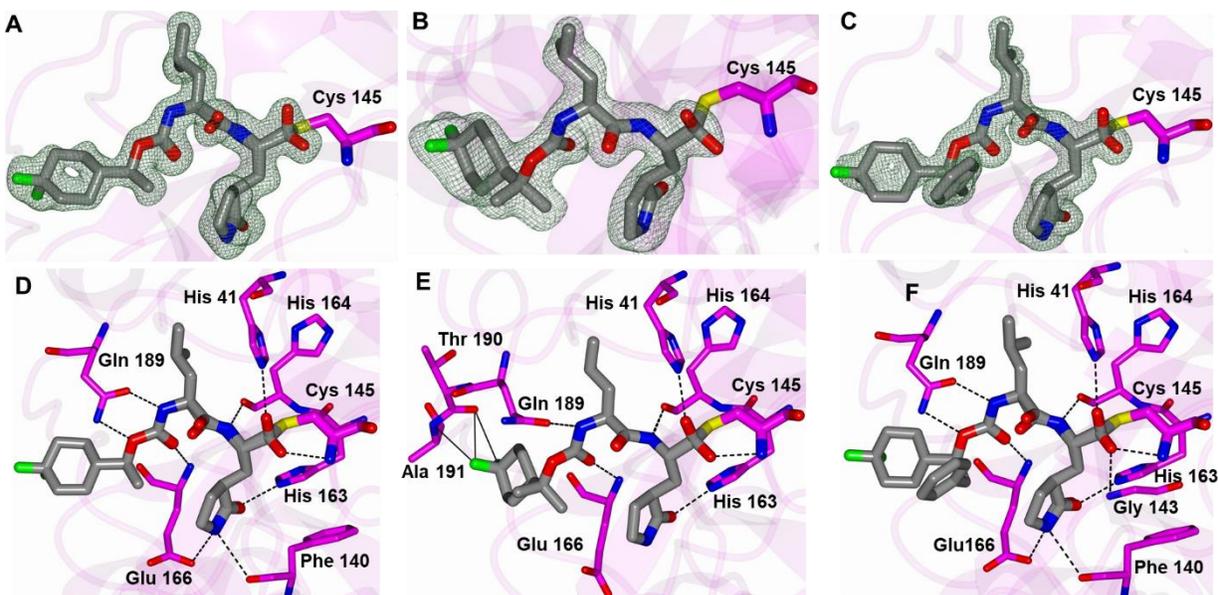


Figure 4. Binding mode of inhibitors containing a 4,4-difluorocyclohexyl group. **12b** (A/D), **13c** (B/E) and **14c** (C/F) with SARS-CoV-2 3CLpro. Fo-Fc Polder omit map (A-C) contoured at 3σ . Hydrogen bond interactions (D-F) are drawn as dashed lines.

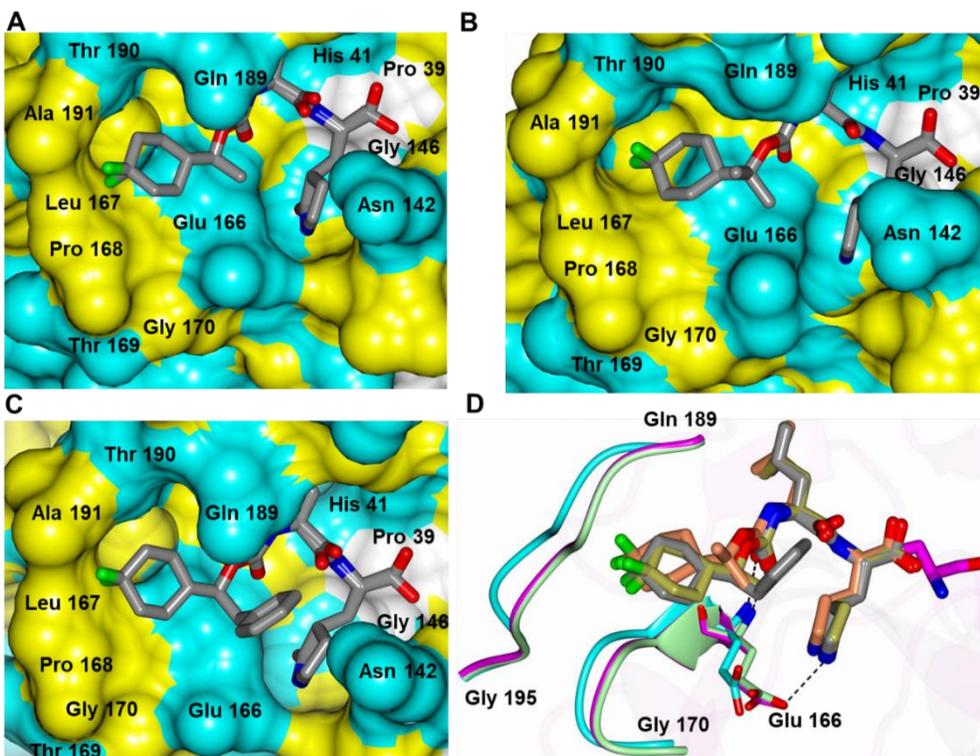


Figure 5. Surface representation showing the orientation of the 4,4-difluorocyclohexyl groups near the S4 subsite of SARS-CoV-2 3CLpro with neighboring residues colored yellow (nonpolar), cyan (polar), and white (weakly polar). **12b** (A), **13c** (B) and **14c** (C). Superposition of **12b** (gold), **13c** (coral) and **14c** (gray) (D). The loop between Gln 189 and Gly 195 is colored cyan for **13c** and magenta for **12b/14c**. Hydrogen bond interactions with Glu 166 are indicated by the dashed lines.

Fluorinated Aryl Compounds. Positional analogue scanning is a widely used strategy for optimizing binding affinity, selectivity, and physicochemical properties of lead compounds containing aromatic or heteroaromatic rings.²⁴ For instance, the introduction of fluorine (F-walk)²⁵ or nitrogen (N-walk)³⁷ is an effective means for multiparameter optimization by leveraging the beneficial impact of fluorine (or nitrogen) and minor

structural changes. In an effort to determine the effect of fluorine on the binding mode in the S₄ subsite of GC376, the structures of the fluorinated benzyl compounds **17c**, **18c**, **19b**, **20b** (deuterated analog of **19b**) and **21c** were determined with SARS-CoV-2 3CLpro. The inhibitor o-fluorobenzyl (**17c**) and m-fluorobenzyl (**18c**) compounds displayed well-defined electron density and similar hydrogen bond interactions as shown in **Figure 6**. Interestingly, the o-fluorobenzyl ring of **17c** adopts a conformation in which the fluorine atom is directed away from Thr 190 and is instead positioned 3.38 Å from the backbone oxygen atom of Glu 166 (**Figure 6C**). Conversely, the fluorine atom in **18c** is positioned between Thr 190/Ala 191 in the S₄ pocket and is 3.10 Å from the backbone nitrogen atom of Ala 191 (**Figure 6D**). The orientations of the fluorine atoms in **17c** and **18c** relative to the hydrophobic ridge in the S₄ pocket is shown in **Figures 6E-F**.

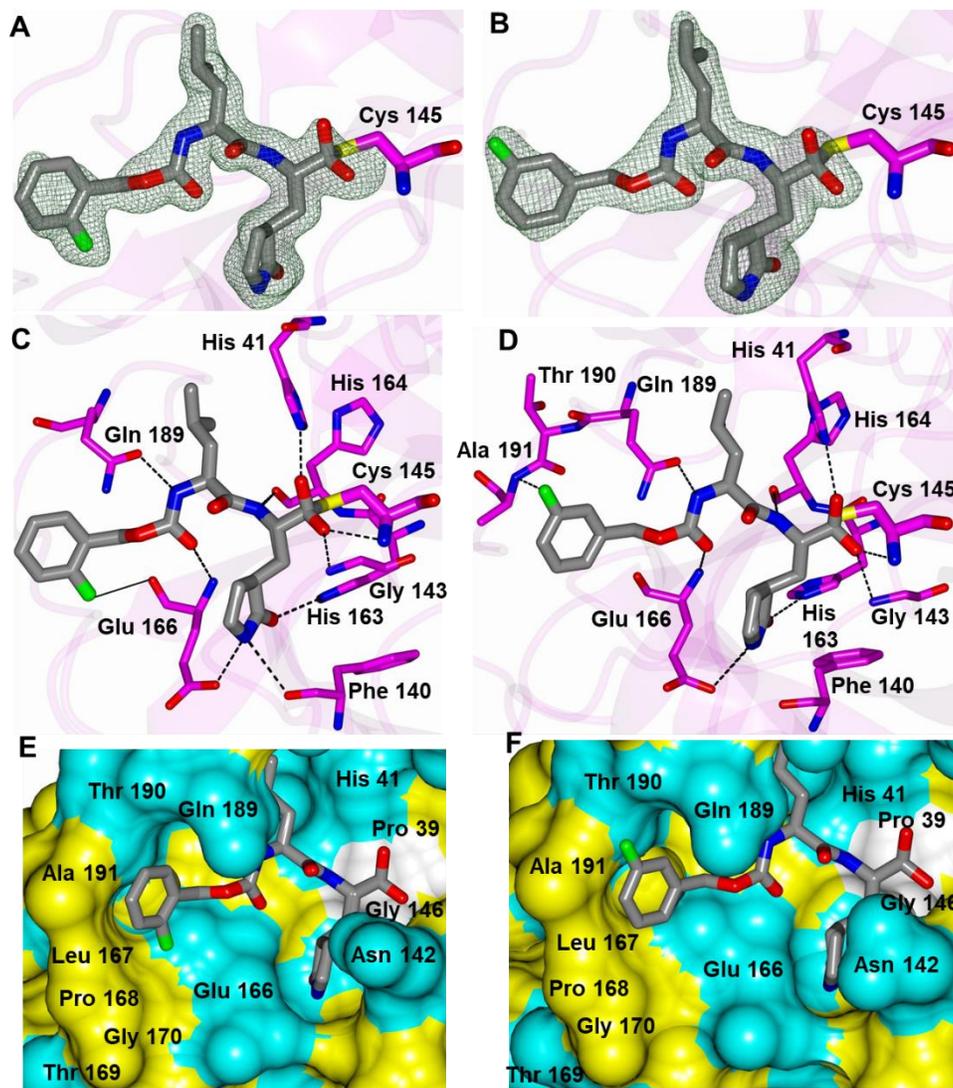


Figure 6. Binding mode of inhibitors containing a fluorinated aromatic group. **17c** (A/C), **18c** (B/D) with SARS-CoV-2 3CLpro. Fo-Fc Polder omit map (A-B) contoured at 3σ . Hydrogen bond interactions (C-D) are drawn as dashed lines. The 3.38 Å contact between the F-atom of **17c** and the backbone O-atom of Glu 166 is drawn as a solid line in panel C. Surface representation of **17c** (E) and **18c** (F) showing the orientation of the 4,4-difluorocyclohexyl groups near the S₄ subsite of SARS-CoV-2 3CLpro with neighboring residues colored yellow (nonpolar), cyan (polar), and white (weakly polar).

The compounds that contain a p-fluorobenzyl group **19b** and its deuterated analog **20b** not surprisingly adopt very similar binding modes and hydrogen bond interactions as shown in **Figures S1** and **S2**. Interestingly, the inhibitor adopts two conformations in which the p-fluorobenzyl ring is projected away from the S₄ subsite in subunit B and is positioned in the S₄ pocket in subunit A. However, the electron density for the p-fluorobenzyl ring is somewhat weaker in subunit A, which suggests that the pose in subunit B is likely the predominant conformation. This may be due to the fact that the fluorine atom does not form any contacts with polar atoms in the S₄ subsite and results in a conformation in which the aryl ring is positioned out of the pocket which is the same conformation observed for the parent compound GC376.

The perfluorinated compound **21c** also displayed well-defined difference electron density consistent with the aryl ring in one conformation (**Figure 7A**). Interestingly, one of the o-fluorine atoms interacts with the backbone oxygen of Glu 166 (3.08 Å) which is shorter than that observed for **17c** described above (3.38 Å). The other o-fluorine atom is positioned 2.92 Å from the backbone N-atom of Thr 190 and 3.12 Å from the side chain N-atom of Gln 189 (**Figure 7B**). Similarly, the m-fluorine atom is positioned near the backbone nitrogen atom of Ala 191 (3.40 Å) which is longer than the distance observed for **18c** (3.10 Å). The pentafluorobenzyl ring is positioned on top of the hydrophobic cleft within the S₄-subsite (**Figure 7C**), unlike GC376 where the phenyl ring undergoes a hydrophobic collapse with the γ -lactam ring and the inhibitor assumes a “paper clip” shape.

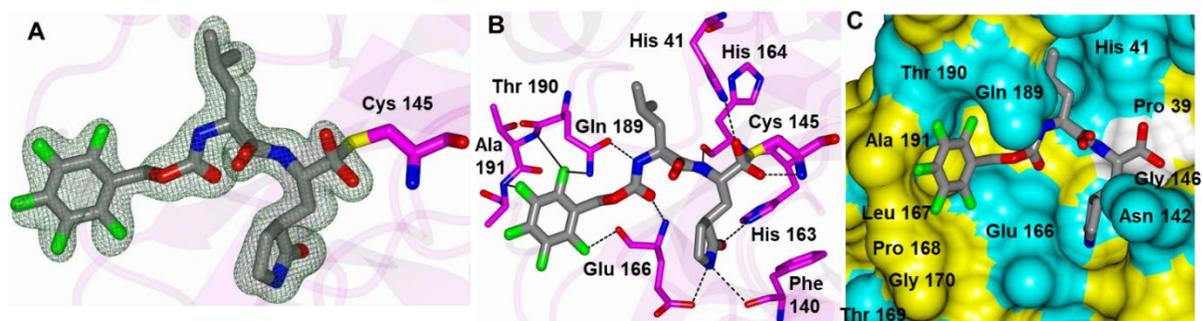


Figure 7. Binding mode of **21c** containing a perfluorinated aromatic group. Fo-Fc Polder omit map contoured at 3σ (**A**). Hydrogen bond interactions are drawn as dashed lines. Close contacts to the perfluorinated ring that are longer than typical polar contact distances are drawn as solid lines (**B**). Surface representation showing the orientation of **21c** near the S4 subsite of SARS-CoV-2 3CLpro with neighboring residues colored yellow (nonpolar), cyan (polar), and white (weakly polar) (**C**).

Finally, GC376 variants **23b/c** and **24b/c** were synthesized and screened as mixtures of epimers. The aldehyde and bisulfite adduct compounds **23b/c** were found to potently inhibit 3CLpro (IC_{50} 0.15 and 0.18 μ M, respectively), and these were 27-fold and 19-fold more potent than the corresponding **24b/c** aldehyde and bisulfite adducts, respectively. These findings provide tentative validation of the design regarding the use of a chiral center to attain directional control and augment binding interactions.

CONCLUSIONS

Effective management of SARS-CoV-2, the causative agent of the COVID-19 pandemic, would require the availability of safe and effective vaccines (already realized), as well as the availability of small-molecule therapeutics and prophylactics that target viral and host-based druggable targets. SARS-CoV-2 3CLpro is an attractive target for the

development of COVID-19 therapeutics because of its vital role in viral replication. An array of approaches was utilized to optimize potency and physicochemical parameters, including conformational and stereochemical control via the introduction of a *gem*-dimethyl group (Thorpe-Ingold effect) or stereocenter, deuteration, and fluorine, into the inhibitors. Virtually all inhibitors were found to display sub-micromolar potency against SARS-CoV-2 and MERS-CoV 3CLpro, and the inhibitory activities were confirmed in a newly established fast and safe cell-based assay. Furthermore, several deuterated inhibitors which are likely to exhibit improved pharmacokinetics were found to be equipotent with the corresponding non-deuterated inhibitors. The fluorine-walk approach was applied to explore bioisosteric replacements for the phenyl ring in GC376 by replacing one or more hydrogen atoms. The effects of these modifications included unanticipated binding modes of the F-substituted phenyl ring and modestly-enhanced potency. The introduction of multiple fluorine atoms resulted in an orientation that allowed the fluorine atoms to engage in H-bonding with residues in the S₄ pocket, although with sub-optimal bond angles. High resolution cocrystal structures with an array of inhibitors unraveled the mechanism of action and provided valuable insights regarding the binding of the inhibitors to the active site and the identity of the structural determinants involved in binding. Collectively, the results of the studies described herein are significant and timely, and provide an effective launching pad for conducting further pre-clinical studies.

EXPERIMENTAL SECTION

General

Reagents and dry solvents were purchased from various chemical suppliers (Sigma-Aldrich, Acros Organics, Chem-Impex, TCI America, Oakwood chemical, APExBIO, Cambridge Isotopes, Alpha Aesar, Fisher and Advanced Chemblocks) and were used as obtained. Silica gel (230-450 mesh) used for flash chromatography was purchased from Sorbent Technologies (Atlanta, GA). Thin layer chromatography was performed using Analtech silica gel plates. Visualization was accomplished using UV light and/or iodine. NMR spectra were recorded in CDCl₃ or DMSO-d₆ using Varian XL-400 spectrometer. Melting points were recorded on a Mel-Temp apparatus and are uncorrected. High resolution mass spectrometry (HRMS) was performed at the Wichita State University Mass Spectrometry lab using Orbitrap Velos Pro mass spectrometer (ThermoFisher, Waltham, MA) equipped with an electrospray ion source. The purity of all final compounds was >95% as evidenced by NMR analysis.

Synthesis of compounds

Preparation of compounds 1-5a, 9a, 17-22a. General procedure. To a solution of alcohol (1 eq) (Table 1) in anhydrous acetonitrile (10 mL/g alcohol) was added N,N'-disuccinimidyl carbonate (1.2 eq) and TEA (3.0 eq) and the reaction mixture was stirred for 4h at room temperature. The solvent was removed *in vacuo* and the residue was dissolved in ethyl acetate (40 mL/g alcohol). The organic phase was washed with saturated aqueous NaHCO₃ (2 x 20 mL/ g alcohol), followed by brine (20 mL/ g alcohol). The organic layers were combined and dried over anhydrous Na₂SO₄, filtered and

concentrated *in vacuo* to yield the mixed carbonate which was used in the next step without further purification.

To a solution of Leu-Gln surrogate amino alcohol (1.0 eq) in dry methylene chloride (10 mL/g of amino alcohol) was added TEA (1.5 eq) and the reaction mixture was stirred for 20 min at room temperature (solution 1). In a separate flask, the mixed carbonate was dissolved in dry methylene chloride (10 mL/g of carbonate) (solution 2). Solution 1 was added to solution 2 and the reaction mixture was stirred 3h at room temperature. Methylene chloride was added to the organic phase (40 mL/g of carbonate) and then washed with saturated aqueous NaHCO₃ (2 x 20 mL/ g alcohol), followed by brine (20 mL/ g alcohol). The organic phase was dried over anhydrous Na₂SO₄, filtered and concentrated *in vacuo*. The resultant crude product was purified by flash chromatography (hexane/ethyl acetate) to yield dipeptidyl alcohol **a** as a white solid.

4,4-Dimethylcyclohexyl ((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**1a**). Yield (36%), ¹H NMR (400 MHz, DMSO-d₆) δ 7.57 (d, *J* = 8.9 Hz, 1H), 7.51 (s, 1H), 7.07 (d, *J* = 8.3 Hz, 1H), 4.67 – 4.63 (m, 1H), 4.47 – 4.40 (m, 1H), 3.98 – 3.88 (m, 1H), 3.78 – 3.74 (m, 1H), 3.36 – 3.28 (m, 1H), 3.27 – 3.17 (m, 1H), 3.13 (t, *J* = 8.9 Hz, 1H), 3.09 – 2.98 (m, 1H), 2.26 – 2.16 (m, 1H), 2.15 – 2.10 (m, 1H), 1.83 – 1.72 (m, 1H), 1.68 – 1.63 (m, 2H), 1.62 – 1.50 (m, 2H), 1.48 – 1.30 (m, 6H), 1.25 – 1.13 (m, 3H), 0.92 – 0.81 (m, 12H).

(1*r*,4*S*)-4-*iso*-Propylcyclohexyl ((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**2a**). Yield (35%), ¹H NMR (400 MHz, DMSO-d₆) δ 7.57 (d, *J* = 8.9 Hz, 1H), 7.51 (s, 1H), 7.07 (d, *J* = 8.2 Hz, 1H), 4.65 (t, *J* = 5.5 Hz, 1H), 4.34 (td, *J* = 11.0, 5.5 Hz, 1H), 3.97 – 3.88 (m, 1H), 3.79 – 3.74 (m, 1H), 3.38

– 3.28 (m, 1H), 3.27 – 3.18 (m, 1H), 3.14 (t, $J = 9.0$ Hz, 1H), 3.05 (q, $J = 8.5$ Hz, 1H), 2.29 – 2.20 (m, 1H), 2.18 – 2.07 (m, 1H), 1.91 (d, $J = 7.9$ Hz, 2H), 1.84 – 1.72 (m, 1H), 1.71 – 1.66 (m, 2H), 1.62 – 1.50 (m, 2H), 1.48 – 1.30 (m, 4H), 1.26 – 1.13 (m, 2H), 1.01 (s, 3H), 0.93 – 0.80 (m, 12H).

(1*s*,4*S*)-4-propylcyclohexyl ((*S*)-1-(((*S*)-1-hydroxy-3-((*S*)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**3a**). Yield (38%), $^1\text{H NMR}$ (400 MHz, DMSO- d_6) δ 7.57 (d, $J = 9.0$ Hz, 1H), 7.51 (s, 1H), 7.07 (d, $J = 8.2$ Hz, 1H), 4.65 (s, 1H), 4.36 (td, $J = 11.0, 5.5$ Hz, 1H), 3.98 – 3.88 (m, 1H), 3.78 – 3.74 (m, 1H), 3.36 – 3.29 (m, 1H), 3.28 – 3.18 (m, 1H), 3.14 (t, $J = 8.9$ Hz, 1H), 3.10 – 2.99 (m, 1H), 2.30 – 2.19 (m, 1H), 2.17 – 2.07 (m, 1H), 1.93 – 1.84 (m, 2H), 1.77 – 1.67 (m, 3H), 1.62 – 1.48 (m, 2H), 1.48 – 1.30 (m, 3H), 1.33 – 1.21 (m, 3H), 1.21 – 1.08 (m, 4H), 0.97 – 0.90 (m, 2H), 0.90 – 0.81 (m, 9H).

(1*s*,4*S*)-4-butylcyclohexyl ((*S*)-1-(((*S*)-1-hydroxy-3-((*S*)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**4a**). Yield (35%), $^1\text{H NMR}$ (400 MHz, DMSO- d_6) δ 7.57 (d, $J = 9.0$ Hz, 1H), 7.51 (s, 1H), 7.07 (d, $J = 8.2$ Hz, 1H), 4.65 (t, $J = 5.5$ Hz, 1H), 4.36 (tt, $J = 10.9, 4.2$ Hz, 1H), 3.98 – 3.88 (m, 1H), 3.79 – 3.74 (m, 1H), 3.38 – 3.28 (m, 1H), 3.28 – 3.18 (m, 1H), 3.14 (t, $J = 8.9$ Hz, 1H), 3.10 – 2.99 (m, 1H), 2.30 – 2.18 (m, 1H), 2.18 – 2.07 (m, 1H), 1.93 – 1.84 (m, 2H), 1.79 – 1.65 (m, 3H), 1.62 – 1.50 (m, 2H), 1.48 – 1.30 (m, 3H), 1.29 – 1.14 (m, 9H), 0.97 – 0.90 (m, 2H), 0.90 – 0.81 (m, 9H).

(1*r*,4*S*)-4-phenylcyclohexyl ((*S*)-1-(((*S*)-1-hydroxy-3-((*S*)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**5a**). Yield (51%), $^1\text{H NMR}$ (400 MHz,

DMSO-d₆) δ 7.59 (d, *J* = 9.0 Hz, 1H), 7.51 (s, 1H), 7.32 – 7.10 (m, 6H), 4.66 (t, *J* = 5.5 Hz, 1H), 4.52 (ddd, *J* = 15.2, 10.8, 4.3 Hz, 1H), 3.99 – 3.90 (m, 1H), 3.80 – 3.75 (m, 1H), 3.41 – 3.28 (m, 1H), 3.28 – 3.18 (m, 1H), 3.14 (t, *J* = 8.9 Hz, 1H), 3.10 – 3.00 (m, 1H), 2.29 – 2.20 (m, 1H), 2.20 – 2.09 (m, 1H), 2.03 – 1.98 (m, 3H), 1.81 (d, *J* = 13.0 Hz, 3H), 1.69 – 1.50 (m, 4H), 1.49 – 1.31 (m, 5H), 0.86 (dd, *J* = 8.8, 6.5 Hz, 6H).

(4-(Trifluoromethyl) cyclohexyl) methyl ((2*S*)-1-(((2*S*)-1-hydroxy-3-(2-oxopyrrolidin-3-yl) propan-2-yl) amino)-4-methyl-1-oxopentan-2-yl) carbamate (**6a**). Yield (83%). ¹H NMR (400 MHz, DMSO-d₆) δ 7.60 (d, *J* = 8.9 Hz, 1H), 7.52 (s, 1H), 7.16 (d, *J* = 8.3 Hz, 1H), 4.66 (s, 1H), 3.99 – 3.89 (m, 1H), 3.83 – 3.71 (m, 2H), 3.27 – 3.19 (m, 2H), 3.15 (t, 2H), 3.10 – 3.02 (m, 1H), 2.26 – 2.08 (m, 3H), 1.91 – 1.73 (m, 4H), 1.64 – 1.52 (m, 4H), 1.49 – 1.31 (m, 1H), 1.30 – 1.15 (m, 1H), 1.08 – 0.94 (m, 4H), 0.90 – 0.81 (m, 6H).

((1*r*,4*S*)-4-(Trifluoromethyl) cyclohexyl) methyl ((2*S*)-1-(((2*S*)-1-hydroxy-3-(2-oxopyrrolidin-3-yl) propan-2-yl) amino)-4-methyl-1-oxopentan-2-yl) carbamate (**7a**). Yield (80%). ¹H NMR (400 MHz, DMSO-d₆) δ 7.61 (d, *J* = 9.0 Hz, 1H), 7.53 (s, 1H), 7.19 (d, *J* = 8.2 Hz, 1H), 4.67 (s, 2H), 4.07 – 3.86 (m, 2H), 3.76 (s, 1H), 3.26 – 3.18 (m, 1H), 3.17 – 3.11 (m, 2H), 3.09 – 3.01 (m, 2H), 2.33 – 2.19 (m, 3H), 2.18 – 2.09 (m, 1H), 1.89 – 1.85 (m, 2H), 1.84 – 1.72 (m, 2H), 1.64 – 1.32 (m, 9H), 0.90 – 0.81 (m, 6H).

((1*r*,4*R*)-4-propylcyclohexyl) methyl ((2*S*)-1-(((2*S*)-1-hydroxy-3-(2-oxopyrrolidin-3-yl) propan-2-yl) amino)-4-methyl-1-oxopentan-2-yl) carbamate (**8a**). Yield (86%). ¹H NMR (400 MHz, DMSO-d₆) δ 7.62 – 7.57 (m, 1H), 7.52 (s, 1H), 7.12 (d, *J* = 8.3 Hz, 1H), 4.65 (s, 1H), 3.99 – 3.89 (m, 1H), 3.82 – 3.68 (m, 3H), 3.37 – 3.19 (m, 2H), 3.18 – 3.02 (m, 2H), 2.26 – 2.07 (m, 4H), 1.76 – 1.65 (m, 4H), 1.62 – 1.51 (m, 8H), 1.49 – 1.32 (m, 2H), 1.32 – 1.24 (m, 2H), 1.21 – 1.10 (m, 2H), 0.99 – 0.78 (m, 9H).

((1*s*,4*S*)-4-propylcyclohexyl) methyl-*d*₂ ((2*S*)-1-(((2*S*)-1-hydroxy-3-(2-oxopyrrolidin-3-yl)propan-2-yl) amino)-4-methyl-1-oxopentan-2-yl) carbamate (**9a**). Yield (43%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.59 (d, *J* = 8.9 Hz, 1H), 7.52 (s, 1H), 7.12 (d, *J* = 8.3 Hz, 1H), 4.65 (s, 1H), 3.98 – 3.89 (m, 1H), 3.80 – 3.70 (m, 1H), 3.27 – 3.10 (m, 2H), 3.10 – 3.00 (m, 2H), 2.28 – 2.06 (m, 4H), 1.70 (s, 4H), 1.63 – 1.51 (m, 2H), 1.49 – 1.34 (m, 8H), 1.32 – 1.23 (m, 2H), 1.20 – 1.10 (m, 2H), 0.90 – 0.81 (m, 9H).

(4,4-Difluorocyclohexyl)methyl ((*S*)-1-(((*S*)-1-hydroxy-3-((*S*)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**10a**). Yield (90%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.60 (d, *J* = 8.9 Hz, 1H), 7.52 (s, 1H), 7.20 (d, *J* = 8.3 Hz, 1H), 4.66 (s, 1H), 3.95 (td, *J* = 8.9, 5.4 Hz, 1H), 3.90 – 3.69 (m, 3H), 3.23 (d, *J* = 5.7 Hz, 1H), 3.18 – 3.00 (m, 2H), 2.28 – 2.07 (m, 2H), 2.06 – 1.93 (m, 2H), 1.90 – 1.64 (m, 6H), 1.56 (dq, *J* = 11.9, 8.8 Hz, 2H), 1.51 – 1.30 (m, 3H), 1.30 – 1.12 (m, 2H), 0.86 (dd, *J* = 10.5, 6.6 Hz, 6H).

(4,4-Difluorocyclohexyl) methyl-*d*₂ ((2*S*)-1-(((2*S*)-1-hydroxy-3-(2-oxopyrrolidin-3-yl)propan-2-yl) amino)-4-methyl-1-oxopentan-2-yl) carbamate (**11a**). Yield (81%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.60 (d, *J* = 8.9 Hz, 1H), 7.52 (s, 1H), 7.20 (d, *J* = 8.2 Hz, 1H), 3.99 – 3.90 (m, 1H), 3.77 (s, 1H), 3.37 – 3.19 (m, 2H), 3.18 – 3.00 (m, 2H), 2.27 – 2.07 (m, 2H), 2.03 – 1.94 (m, 2H), 1.87 – 1.63 (m, 6H), 1.63 – 1.49 (m, 2H), 1.47 – 1.29 (m, 3H), 1.29 – 1.13 (m, 2H), 0.90 – 0.81 (m, 6H).

1-(4,4-Difluorocyclohexyl)ethyl ((*S*)-1-(((*S*)-1-hydroxy-3-((*S*)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**12a**). Yield (76%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.61 – 7.50 (m, 2H), 7.11 (d, *J* = 8.2 Hz, 1H), 4.66 (t, *J* = 5.3 Hz, 1H), 4.59 – 4.50 (m, 1H), 3.98 – 3.90 (m, 1H), 3.79 – 3.75 (m, 1H), 3.39 – 3.29 (m, 1H), 3.28

– 3.18 (m, 1H), 3.18 – 3.10 (m, 1H), 3.10 – 2.99 (m, 1H), 2.27 – 2.18 (m, 1H), 2.16 – 2.09 (m, 1H), 2.02 – 1.97 (m, 2H), 1.83 – 1.63 (m, 5H), 1.63 – 1.51 (m, 3H), 1.47 – 1.31 (m, 3H), 1.27 – 1.22 (m, 2H), 1.12 (dd, 2H), 0.90 – 0.81 (m, 6H).

2-(4,4-Difluorocyclohexyl)propan-2-yl ((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**13a**). Yield (72%). ¹H NMR (400 MHz, DMSO₆) δ 7.59 – 7.47 (m, 2H), 6.94 (d, *J* = 8.2 Hz, 1H), 4.66 (t, *J* = 5.5 Hz, 1H), 3.93 – 3.83 (m, 1H), 3.79 – 3.75 (m, 1H), 3.42 – 3.29 (m, 1H), 3.27 – 3.18 (m, 1H), 3.14 (t, *J* = 8.9 Hz, 1H), 3.10 – 2.99 (m, 1H), 2.27 – 2.09 (m, 2H), 2.06 – 1.91 (m, 3H), 1.85 – 1.61 (m, 5H), 1.61 – 1.51 (m, 2H), 1.46 – 1.34 (m, 3H), 1.33 (s, 6H), 1.31 – 1.18 (m, 2H), 0.90 – 0.81 (m, 6H).

1-(4,4-Difluorocyclohexyl)-2-phenylethyl ((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**14a**). Yield (79%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.61 – 7.54 (m, 1H), 7.52 (s, 1H), 7.11 (d, *J* = 8.2 Hz, 1H), 4.66 (t, *J* = 5.3 Hz, 1H), 4.59 – 4.50 (m, 1H), 3.98 – 3.90 (m, 1H), 3.79 – 3.75 (m, 1H), 3.39 – 3.29 (m, 1H), 3.28 – 3.18 (m, 1H), 3.18 – 3.10 (m, 1H), 3.10 – 2.99 (m, 1H), 2.27 – 2.18 (m, 1H), 2.16 – 2.09 (m, 1H), 2.02 – 1.97 (m, 2H), 1.83 – 1.63 (m, 5H), 1.63 – 1.51 (m, 3H), 1.47 – 1.31 (m, 3H), 1.27 – 1.22 (m, 2H), 1.12 (dd, 3H), 0.90 – 0.81 (m, 6H).

(4,4-Difluorocyclohexyl)(phenyl)methyl ((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**15a**). Yield (86%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.64 – 7.50 (m, 2H), 7.39 – 7.23 (m, 6H), 4.68 – 4.64 (m, 1H), 4.00 – 3.89 (m, 1H), 3.87 – 3.73 (m, 2H), 3.39 – 3.29 (m, 1H), 3.27 – 3.20 (m, 1H), 3.19 – 3.10 (m, 1H), 3.10 – 3.02 (m, 1H), 2.25 – 2.09 (m, 2H), 2.06 – 1.92 (m, 2H), 1.87 – 1.66

(m, 5H), 1.65 – 1.51 (m, 2H), 1.49 – 1.32 (m, 4H), 1.32 – 1.18 (m, 2H), 0.93 – 0.71 (m, 6H).

1-(4,4-Difluorocyclohexyl)pentyl ((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**16a**). Yield (91%). ¹H NMR (400 MHz, DMSO-d₆) δ 7.60 (d, *J* = 9.0 Hz, 1H), 7.52 (s, 1H), 7.20 (d, *J* = 8.3 Hz, 1H), 4.66 (t, *J* = 5.6 Hz, 1H), 4.00 – 3.90 (m, 1H), 3.90 – 3.77 (m, 1H), 3.77 – 3.74 (m, 1H), 3.40 – 3.31 (m, 1H), 3.29 – 3.19 (m, 1H), 3.14 (t, *J* = 9.0 Hz, 1H), 3.11 – 3.00 (m, 1H), 2.25 – 2.15 (m, 1H), 2.15 – 2.09 (m, 1H), 2.02 – 1.97 (m, 3H), 1.88 – 1.80 (m, 1H), 1.79 – 1.65 (m, 8H), 1.63 – 1.51 (m, 2H), 1.49 – 1.31 (m, 4H), 1.26 – 1.18 (m, 3H), 0.93 – 0.81 (m, 9H).

2-Fluorobenzyl ((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**17a**). Yield (33%). ¹H NMR (400 MHz, CDCl₃) δ 7.76 (d, *J* = 7.2 Hz, 1H), 7.42 – 7.34 (m, 1H), 7.34 – 7.27 (m, 1H), 7.12 (td, *J* = 7.5, 1.2 Hz, 1H), 7.09 – 7.00 (m, 1H), 6.23 (s, 1H), 5.56 (d, *J* = 8.3 Hz, 1H), 5.19 – 5.14 (m, 2H), 4.27 – 4.22 (m, 1H), 4.01 – 3.96 (m, 1H), 3.77 – 3.46 (m, 2H), 3.37 – 3.24 (m, 2H), 2.46 – 2.33 (m, 2H), 2.11 – 1.87 (m, 2H), 1.85 – 1.74 (m, 1H), 1.74 – 1.58 (m, 2H), 1.57 – 1.46 (m, 1H), 0.94 (d, *J* = 5.8 Hz, 6H).

3-Fluorobenzyl ((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**18a**). Yield (41%). ¹H NMR (400 MHz, CDCl₃) δ 7.82 (d, *J* = 7.2 Hz, 1H), 7.35 – 7.28 (m, 1H), 7.12 – 7.02 (m, 2H), 7.02 – 6.94 (m, 1H), 6.41 (s, 1H), 5.70 (d, *J* = 8.3 Hz, 1H), 5.16 – 5.00 (m, 2H), 4.29 – 4.22 (m, 1H), 4.02 –

3.93 (m, 1H), 3.74 – 3.53 (m, 2H), 3.36 – 3.22 (m, 2H), 2.45 – 2.30 (m, 2H), 2.07 – 1.95 (m, 1H), 1.79 (td, $J = 9.4, 2.8$ Hz, 1H), 1.73 – 1.43 (m, 4H), 0.94 (d, $J = 6.1$ Hz, 6H).

4-Fluorobenzyl ((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**19a**). Yield (37%). ^1H NMR (400 MHz, CDCl_3) δ 7.73 (d, $J = 7.4$ Hz, 1H), 7.37 – 7.28 (m, 2H), 7.08 – 6.97 (m, 2H), 6.37 (s, 1H), 5.56 (d, $J = 8.1$ Hz, 1H), 5.12 – 5.02 (m, 2H), 4.31 – 4.13 (m, 1H), 4.05 – 3.90 (m, 1H), 3.66 – 3.53 (m, 2H), 3.39 – 3.25 (m, 2H), 2.46 – 2.31 (m, 2H), 2.05 – 1.93 (m, 1H), 1.88 – 1.76 (m, 1H), 1.74 – 1.57 (m, 3H), 1.57 – 1.45 (m, 1H), 0.93 (d, $J = 6.1$ Hz, 6H).

(4-Fluorophenyl)methyl- d_2 ((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**20a**). Yield (37%). ^1H NMR (400 MHz, CDCl_3) δ 7.76 (d, $J = 7.3$ Hz, 1H), 7.36 – 7.29 (m, 2H), 7.07 – 6.98 (m, 2H), 6.32 (s, 1H), 5.55 (d, $J = 8.3$ Hz, 1H), 4.27 – 4.20 (m, 1H), 4.00 – 3.96 (m, 1H), 3.66 – 3.53 (m, 2H), 3.37 – 3.23 (m, 2H), 2.48 – 2.32 (m, 2H), 2.03 – 1.93 (m, 1H), 1.87 – 1.76 (m, 1H), 1.73 – 1.57 (m, 3H), 1.55 – 1.45 (m, 1H), 0.93 (d, $J = 6.1$ Hz, 6H).

(Perfluorophenyl)methyl ((2S)-1-(((2S)-1-hydroxy-3-(2-oxopyrrolidin-3-yl) propan-2-yl) amino)-4-methyl-1-oxopentan-2-yl) carbamate (**21a**). Yield (21%). ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 7.66 (d, $J = 8.9$ Hz, 1H), 7.52 (s, 1H), 7.46 (d, $J = 8.0$ Hz, 1H), 5.18 – 5.07 (m, 2H), 4.64 (s, 1H), 3.99 – 3.90 (m, 1H), 3.75 (s, 1H), 3.25 – 3.02 (m, 4H), 2.27 – 2.16 (m, 1H), 2.15 – 2.05 (m, 2H), 1.81 – 1.71 (m, 1H), 1.62 – 1.50 (m, 2H), 1.47 – 1.31 (m, 2H), 0.90 – 0.78 (m, 6H).

(Perfluorophenyl)methyl- d_2 ((2S)-1-(((2S)-1-hydroxy-3-(2-oxopyrrolidin-3-yl) propan-2-yl) amino)-4-methyl-1-oxopentan-2-yl) carbamate (**22a**). Yield (11%). ^1H NMR (400 MHz,

DMSO-d₆) δ 7.65 (d, *J* = 8.9 Hz, 1H), 7.52 (s, 1H), 7.46 (d, *J* = 8.0 Hz, 1H), 4.64 (s, 1H), 3.99 – 3.90 (m, 1H), 3.75 (s, 1H), 3.25 – 3.02 (m, 4H), 2.26 – 2.16 (m, 1H), 2.15 – 2.05 (m, 2H), 1.81 – 1.71 (m, 1H), 1.60 – 1.49 (m, 2H), 1.46 – 1.31 (m, 2H), 0.90 – 0.77 (m, 6H).

1-Phenylbutyl ((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**23a**). Yield (60%). ¹H NMR (400 MHz, CDCl₃) δ 7.59 (dd, *J* = 79.1, 7.3 Hz, 1H), 7.36 – 7.21 (m, 5H), 6.24 (d, *J* = 37.3 Hz, 1H), 5.60 (t, *J* = 7.1, 7.1 Hz, 1H), 5.41 (dd, *J* = 23.4, 8.0 Hz, 1H), 4.21 – 4.11 (m, 1H), 4.04 – 3.89 (m, 1H), 3.69 – 3.47 (m, 2H), 3.36 – 3.19 (m, 2H), 2.53 – 2.17 (m, 2H), 2.03 – 1.80 (m, 3H), 1.79 – 1.42 (m, 3H), 1.41 – 1.21 (m, 2H), 0.98 – 0.83 (m, 11H).

1,2-Diphenylethyl ((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**24a**). Yield (83%). ¹H NMR (400 MHz, CDCl₃) δ 7.58 (dd, *J* = 49.0, 7.4 Hz, 1H), 7.33 – 7.16 (m, 8H), 7.11 – 7.02 (m, 2H), 6.09 (d, *J* = 26.6 Hz, 1H), 5.88 – 5.75 (m, 1H), 5.42 – 5.32 (m, 1H), 4.17 – 4.07 (m, 1H), 4.04 – 3.87 (m, 1H), 3.68 – 3.43 (m, 2H), 3.33 – 2.98 (m, 4H), 2.52 – 2.21 (m, 2H), 2.02 – 1.39 (m, 6H), 0.94 – 0.79 (m, 6H).

Preparation of compounds 1-24b. General procedure. To a solution of dipeptidyl alcohol **a** (1 eq) in anhydrous dichloromethane (300 mL/g dipeptidyl alcohol) kept at 0-5 °C under a N₂ atmosphere was added Dess-Martin periodinane reagent (3.0 eq) and the reaction mixture was stirred for 3 h at 15-20 °C. The organic phase was washed with 10% aq Na₂S₂O₃ (2 x 100 mL/g dipeptidyl alcohol), followed by saturated aqueous NaHCO₃ (2 x 100 mL/g dipeptidyl alcohol), distilled water (2 x 100 mL/g dipeptidyl alcohol), and brine (100 mL/g dipeptidyl alcohol). The organic phase was dried over anhydrous Na₂SO₄,

filtered and concentrated *in vacuo*. The resulting crude product was purified by flash chromatography (hexane/ethyl acetate) to yield aldehyde **b** as a white solid.

4,4-Dimethylcyclohexyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**1b**). Yield (80%). ¹H NMR (400 MHz, DMSO-d₆) δ 9.40 (s, 1H), 8.39 (d, *J* = 7.7 Hz, 1H), 7.62 (s, 1H), 7.19 (d, *J* = 8.0 Hz, 1H), 4.49 – 4.40 (m, 1H), 4.24 – 4.14 (m, 1H), 4.07 – 3.97 (m, 1H), 3.16 (t, *J* = 9.2 Hz, 1H), 3.12 – 3.00 (m, 1H), 2.31 – 2.22 (m, 1H), 2.19 – 2.09 (m, 1H), 1.95 – 1.83 (m, 1H), 1.74 – 1.55 (m, 5H), 1.54 – 1.31 (m, 6H), 1.25 – 1.15 (m, 2H), 0.92 – 0.81 (m, 12H). HRMS *m/z*: [M+Na]⁺ Calculated for C₂₂H₃₇N₃NaO₅ 446.2631; Found 446.2612.

(1*r*,4*S*)-4-isopropylcyclohexyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**2b**). Yield (78%). ¹H NMR (400 MHz, DMSO-d₆) δ 9.40 (s, 1H), 8.40 (d, *J* = 7.6 Hz, 1H), 7.63 (s, 1H), 7.19 (d, *J* = 8.0 Hz, 1H), 4.42 – 4.30 (m, 1H), 4.19 (ddd, *J* = 11.4, 7.6, 4.2 Hz, 1H), 4.07 – 3.96 (m, 1H), 3.28 – 3.01 (m, 2H), 2.35 – 2.19 (m, 1H), 2.19 – 2.08 (m, 1H), 1.95 – 1.84 (m, 3H), 1.73 – 1.57 (m, 5H), 1.53 – 1.36 (m, 3H), 1.27 – 1.20 (m, 2H), 1.02 (s, 3H), 0.97 – 0.80 (m, 12H). HRMS *m/z*: [M+Na]⁺ Calculated for C₂₃H₃₉N₃NaO₅ 460.2788; Found 460.2777.

(1*s*,4*S*)-4-propylcyclohexyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**3b**). Yield (73%). ¹H NMR (400 MHz, DMSO-d₆) δ 9.40 (s, 1H), 8.40 (d, *J* = 7.7 Hz, 1H), 7.63 (s, 1H), 7.19 (d, *J* = 7.9 Hz, 1H), 4.37 (td, *J* = 11.0, 5.5 Hz, 1H), 4.24 – 4.12 (m, 1H), 4.07 – 3.97 (m, 1H), 3.21 – 3.01 (m, 2H), 2.35 – 2.22 (m, 1H), 2.19 – 2.08 (m, 1H), 1.95 – 1.85 (m, 3H), 1.76 – 1.55 (m, 5H), 1.54 – 1.37 (m, 2H), 1.35 – 1.19 (m, 4H), 1.19 – 1.09 (m, 3H), 1.01 – 0.91 (m, 2H), 0.91

– 0.81 (m, 9H). HRMS m/z : $[M+H]^+$ Calculated for $C_{23}H_{40}N_3O_5$ 438.2968; Found 438.2952.

(1*s*,4*S*)-4-Butylcyclohexyl ((*S*)-4-methyl-1-oxo-1-(((*S*)-1-oxo-3-((*S*)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**4b**). Yield (65%). 1H NMR (400 MHz, DMSO- d_6) δ 9.40 (s, 1H), 8.40 (d, $J = 7.6$ Hz, 1H), 7.63 (s, 1H), 7.19 (d, $J = 7.9$ Hz, 1H), 4.37 (td, $J = 11.0, 5.3$ Hz, 1H), 4.19 (ddd, $J = 11.5, 7.7, 4.1$ Hz, 1H), 4.02 (q, $J = 8.5$ Hz, 1H), 3.17 (t, $J = 9.1$ Hz, 1H), 3.13 – 3.01 (m, 1H), 2.32 – 2.23 (m, 1H), 2.20 – 2.08 (m, 1H), 1.96 – 1.84 (m, 3H), 1.76 – 1.72 (m, 3H), 1.72 – 1.55 (m, 3H), 1.53 – 1.37 (m, 2H), 1.29 – 1.13 (m, 8H), 1.01 – 0.91 (m, 2H), 0.91 – 0.80 (m, 9H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{24}H_{41}N_3NaO_5$ 474.2944; Found 474.2931.

(1*r*,4*S*)-4-Phenylcyclohexyl ((*S*)-4-methyl-1-oxo-1-(((*S*)-1-oxo-3-((*S*)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**5b**). Yield (63%). 1H NMR (400 MHz, DMSO- d_6) δ 9.41 (s, 1H), 8.43 (d, $J = 7.7$ Hz, 1H), 7.63 (s, 1H), 7.32 – 7.10 (m, 6H), 4.53 (td, $J = 10.9, 5.4$ Hz, 1H), 4.20 (ddd, $J = 11.4, 7.5, 4.0$ Hz, 1H), 4.07 – 3.92 (m, 1H), 3.29 – 3.02 (m, 2H), 2.36 – 2.23 (m, 1H), 2.21 – 2.10 (m, 1H), 2.07 – 1.96 (m, 2H), 1.96 – 1.85 (m, 1H), 1.81 (d, $J = 12.8$ Hz, 2H), 1.70 – 1.56 (m, 6H), 1.56 – 1.42 (m, 4H), 0.89 (dd, $J = 9.5, 6.6$ Hz, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{26}H_{37}N_3NaO_5$ 494.2631; Found 494.2608.

(4-(Trifluoromethyl) cyclohexyl) methyl ((*S*)-4-methyl-1-oxo-1-(((*S*)-1-oxo-3-(2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**6b**). Yield (88%). 1H NMR (400 MHz, DMSO- d_6) δ 9.40 (s, 1H), 8.43 (d, $J = 7.6$ Hz, 1H), 7.63 (s, 1H), 7.29 (d, $J = 8.0$ Hz, 1H), 4.22 – 4.15 (m, 1H), 4.07 – 3.99 (m, 1H), 3.78 (d, $J = 2.0$ Hz, 2H), 3.20 –

3.05 (m, 3H), 2.31 – 2.08 (m, 3H), 1.95 – 1.76 (m, 4H), 1.70 – 1.40 (m, 5H), 1.29 – 1.16 (m, 1H), 1.02 (q, $J = 13.0$ Hz, 4H), 0.92 – 0.81 (m, 6H). HRMS m/z : $[M+H]^+$ Calculated for $C_{22}H_{35}F_3N_3O_5$ 478.2529; Found 478.2522. HRMS m/z : $[M+Na]^+$ Calculated for $C_{22}H_{34}F_3N_3NaO_5$ 500.2349; Found 500.2328.

((1*r*,4*S*)-4-(Trifluoromethyl) cyclohexyl) methyl ((2*S*)-4-methyl-1-oxo-1-(((2*S*)-1-oxo-3-(2-oxopyrrolidin-3-yl) propan-2-yl) amino) pentan-2-yl) carbamate (**7b**). Yield (91%). 1H NMR (400 MHz, DMSO- d_6) δ 9.40 (s, 1H), 8.43 (d, $J = 7.6$ Hz, 1H), 7.63 (s, 1H), 7.31 (d, $J = 8.0$ Hz, 1H), 4.24 – 4.14 (m, 1H), 4.08 – 3.99 (m, 1H), 3.98 – 3.92 (m, 2H), 3.21 – 3.04 (m, 3H), 2.33 – 2.20 (m, 3H), 2.19 – 2.08 (m, 1H), 1.92 – 1.83 (m, 4H), 1.69 – 1.38 (m, 9H), 0.92 – 0.81 (m, 6H). HRMS m/z : $[M+H]^+$ Calculated for $C_{22}H_{35}F_3N_3O_5$ 478.2529; Found 478.2512. HRMS m/z : $[M+Na]^+$ Calculated for $C_{22}H_{34}F_3N_3NaO_5$ 500.2349; Found 500.2320.

((1*r*,4*R*)-4-Propylcyclohexyl) methyl ((2*S*)-4-methyl-1-oxo-1-(((2*S*)-1-oxo-3-(2-oxopyrrolidin-3-yl) propan-2-yl) amino) pentan-2-yl) carbamate (**8b**). Yield (42%). 1H NMR (400 MHz, DMSO- d_6) δ 9.40 (s, 1H), 8.42 (d, $J = 7.7$ Hz, 1H), 7.63 (s, 1H), 7.25 (d, $J = 8.0$ Hz, 1H), 4.22 – 4.15 (m, 1H), 4.05 – 4.01 (m, 1H), 3.78 – 3.74 (m, 2H), 3.20 – 3.05 (m, 2H), 2.19 – 2.10 (m, 3H), 1.94 – 1.85 (m, 3H), 1.75 – 1.60 (m, 10H), 1.54 – 1.40 (m, 2H), 1.33 – 1.23 (m, 2H), 1.17 – 1.10 (m, 2H), 0.97 – 0.81 (m, 9H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{24}H_{41}N_3NaO_5$ 474.2944; Found 474.2013.

((1*s*,4*S*)-4-Propylcyclohexyl) methyl- d_2 ((2*S*)-4-methyl-1-oxo-1-(((2*S*)-1-oxo-3-(2-oxopyrrolidin-3-yl) propan-2-yl) amino) pentan-2-yl) carbamate (**9b**). Yield (96%). 1H NMR (400 MHz, DMSO- d_6) δ 9.40 (s, 1H), 8.42 (d, $J = 7.6$ Hz, 1H), 7.63 (s, 1H), 7.25 (d, $J =$

8.0 Hz, 1H), 4.24 – 4.14 (m, 1H), 4.08 – 3.97 (m, 1H), 3.21 – 3.03 (m, 4H), 2.31 – 2.10 (m, 4H), 1.75 – 1.67 (m, 4H), 1.55 – 1.35 (m, 8H), 1.35 – 1.22 (m, 2H), 1.18 – 1.10 (m, 2H), 1.00 – 0.76 (m, 9H). HRMS m/z: [M+Na]⁺ Calculated for C₂₄H₃₉D₂N₃NaO₅ 476.3070; Found 476.3051.

(4,4-Difluorocyclohexyl)methyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**10b**). Yield (51%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 9.40 (s, 1H), 8.44 (d, *J* = 7.5 Hz, 1H), 7.95 (s, 1H), 7.36 – 7.28 (m, 1H), 4.19 (ddd, *J* = 11.4, 7.6, 4.2 Hz, 1H), 4.03 (td, *J* = 8.8, 6.2 Hz, 1H), 3.84 (d, *J* = 6.1 Hz, 2H), 3.24 – 3.02 (m, 2H), 2.37 – 2.08 (m, 2H), 2.08 – 1.94 (m, 2H), 1.94 – 1.80 (m, 1H), 1.80 – 1.55 (m, 7H), 1.55 – 1.33 (m, 2H), 1.33 – 1.11 (m, 3H), 0.97 – 0.79 (m, 6H). HRMS m/z: [M+H]⁺ Calculated for C₂₁H₃₄F₂N₃O₅ 446.2467; Found 446.2452, m/z: [M+Na]⁺ Calculated for C₂₁H₃₃F₂N₃O₅Na 468.2286; Found 468.2281.

(4,4-Difluorocyclohexyl) methyl-*d*₂ ((2S)-4-methyl-1-oxo-1-(((2S)-1-oxo-3-(2-oxopyrrolidin-3-yl) propan-2-yl) amino) pentan-2-yl) carbamate (**11b**). Yield (63%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 9.40 (s, 1H), 8.44 (d, *J* = 7.5 Hz, 1H), 7.63 (s, 1H), 7.32 (d, *J* = 8.1 Hz, 1H), 4.23 – 4.15 (m, 1H), 4.07 – 3.99 (m, 1H), 3.21 – 3.05 (m, 2H), 2.32 – 2.20 (m, 1H), 2.19 – 2.10 (m, 2H), 2.05 – 1.95 (m, 2H), 1.79 – 1.58 (m, 8H), 1.55 – 1.40 (m, 2H), 1.30 – 1.15 (m, 2H), 0.92 – 0.83 (m, 6H). HRMS m/z: [M+H]⁺ Calculated for C₂₁H₃₂D₂F₂N₃O₅ 448.2592; Found 448.2576. HRMS m/z: [M+Na]⁺ Calculated for C₂₁H₃₁D₂F₂N₃NaO₅ 470.2412; Found 470.2412.

1-(4,4-Difluorocyclohexyl)ethyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**12b**). Yield (70%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 9.40 (s, 1H), 8.46 – 8.36 (m, 1H), 7.63 (s, 1H), 7.23 (d, *J* = 8.2 Hz, 1H), 4.59

– 4.52 (m, 1H), 4.23 – 4.18 (m, 1H), 4.06 – 3.98 (m, 1H), 3.21 – 3.03 (m, 2H), 2.32 – 2.22 (m, 1H), 2.18 – 2.09 (m, 1H), 2.06 – 1.95 (m, 3H), 1.95 – 1.76 (m, 2H), 1.72 – 1.59 (m, 5H), 1.55 – 1.40 (m, 2H), 1.31 – 1.14 (m, 3H), 1.12 (d, $J = 6.3$ Hz, 3H), 0.92 – 0.81 (m, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{22}H_{35}F_2N_3NaO_5$ 482.2443; Found 482.2433.

2-(4,4-Difluorocyclohexyl)propan-2-yl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**13b**). Yield (42%). 1H NMR (400 MHz, DMSO- d_6) δ 9.40 (s, 1H), 8.35 (d, $J = 7.7$ Hz, 1H), 7.70 – 7.60 (m, 1H), 7.06 (d, $J = 8.1$ Hz, 1H), 4.02 – 3.89 (m, 2H), 3.21 – 3.00 (m, 2H), 2.31 – 2.11 (m, 1H), 2.07 – 1.83 (m, 4H), 1.82 – 1.74 (m, 4H), 1.69 – 1.58 (m, 4H), 1.56 – 1.36 (m, 3H), 1.34 (s, 6H), 1.30 – 1.23 (m, 1H), 0.92 – 0.81 (m, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{23}H_{38}F_2N_3O_5$ 474.2779; Found 474.2771.

1-(4,4-Difluorocyclohexyl)-2-phenylethyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**14b**). Yield (80%). 1H NMR (400 MHz, DMSO- d_6) δ 9.39 (s, 1H), 8.38 – 8.30 (m, 1H), 7.62 (s, 1H), 7.26 – 7.15 (m, 5H), 4.77 – 4.73 (m, 1H), 4.22 – 4.18 (m, 1H), 3.93 – 3.87 (m, 1H), 3.18 – 2.99 (m, 2H), 2.92 – 2.83 (m, 1H), 2.78 – 2.68 (m, 1H), 2.32 – 2.18 (m, 1H), 2.16 – 2.07 (m, 1H), 2.07 – 1.94 (m, 2H), 1.91 – 1.80 (m, 3H), 1.79 – 1.68 (m, 1H), 1.67 – 1.41 (m, 7H), 1.41 – 1.27 (m, 3H), 0.92 – 0.80 (m, 4H), 0.75 (dd, $J = 11.6, 6.5$ Hz, 2H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{28}H_{39}F_2N_3NaO_5$ 558.2756; Found 558.2740.

(4,4-Difluorocyclohexyl)(phenyl)methyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**15b**). Yield (63%). 1H NMR (400 MHz, DMSO- d_6) δ 9.41 (s, 1H), 8.48 – 8.35 (m, 1H), 7.65 (d, $J = 5.4$ Hz, 1H), 7.52 –

7.43 (m, 1H), 7.39 – 7.20 (m, 5H), 4.22 – 4.17 (m, 1H), 4.00 – 3.96 (m, 1H), 3.85 (d, $J = 6.0$ Hz, 1H), 3.23 – 3.05 (m, 2H), 2.32 – 2.22 (m, 1H), 2.17 – 2.13 (m, 1H), 2.05 – 1.87 (m, 3H), 1.84 – 1.72 (m, 2H), 1.70 – 1.59 (m, 4H), 1.55 – 1.38 (m, 3H), 1.33 – 1.16 (m, 3H), 0.97 – 0.74 (m, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{27}H_{37}F_2N_3NaO_5$ 544.2599; Found 544.2586.

1-(4,4-Difluorocyclohexyl)pentyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**16b**). Yield (88%). 1H NMR (400 MHz, DMSO- d_6) δ 9.40 (s, 1H), 8.43 (d, $J = 7.6$ Hz, 1H), 7.63 (s, 1H), 7.32 (d, $J = 8.0$ Hz, 1H), 4.24 – 4.14 (m, 1H), 4.05 – 4.00 (m, 1H), 3.87 – 3.81 (m, 1H), 3.21 – 3.04 (m, 2H), 2.30 – 2.22 (m, 1H), 2.19 – 2.08 (m, 1H), 2.06 – 1.94 (m, 3H), 1.94 – 1.80 (m, 3H), 1.79 – 1.71 (m, 5H), 1.69 – 1.58 (m, 4H), 1.55 – 1.41 (m, 3H), 1.26 – 1.18 (m, 3H), 0.92 – 0.81 (m, 9H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{25}H_{41}F_2N_3NaO_5$ 524.2912; Found 524.2888.

2-Fluorobenzyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**17b**). Yield (68%). 1H NMR (400 MHz, $CDCl_3$) δ 9.48 (s, 1H), 8.35 (d, $J = 5.9$ Hz, 1H), 7.39 (t, $J = 7.6$ Hz, 1H), 7.35 – 7.27 (m, 1H), 7.18 – 7.00 (m, 2H), 6.02 (s, 1H), 5.38 (d, $J = 8.5$ Hz, 1H), 5.22 – 5.16 (m, 2H), 4.44 – 4.26 (m, 2H), 3.39 – 3.27 (m, 2H), 2.49 – 2.30 (m, 2H), 2.01 – 1.92 (m, 2H), 1.91 – 1.80 (m, 1H), 1.81 – 1.45 (m, 3H), 0.97 (d, $J = 5.8$ Hz, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{21}H_{28}FN_3NaO_5$: 444.1911, Found: 444.1907.

3-Fluorobenzyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**18b**). Yield (68%). 1H NMR (400 MHz, $CDCl_3$) δ 9.48 (s, 1H), 8.45 (d, $J = 5.7$ Hz, 1H), 7.35 – 7.24 (m, 1H), 7.13 – 7.05 (m, 2H), 7.05 – 6.94 (m,

1H), 6.18 (s, 1H), 5.48 (d, $J = 8.6$ Hz, 1H), 5.16 – 5.03 (m, 2H), 4.41 – 4.26 (m, 2H), 3.41 – 3.27 (m, 2H), 2.56 – 2.29 (m, 2H), 2.00 – 1.79 (m, 2H), 1.77 – 1.64 (m, 3H), 1.60 – 1.51 (m, 1H), 0.97 (d, $J = 6.0$ Hz, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{21}H_{28}FN_3NaO_5$: 444.1911, Found: 444.1911.

4-Fluorobenzyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**19b**). Yield (60%). 1H NMR (400 MHz, $CDCl_3$) δ 9.48 (s, 1H), 8.36 (d, $J = 5.7$ Hz, 1H), 7.38 – 7.29 (m, 2H), 7.09 – 6.98 (m, 2H), 6.00 (s, 1H), 5.36 (d, $J = 8.6$ Hz, 1H), 5.11 – 5.04 (m, 2H), 4.35 – 4.28 (m, 2H), 3.41 – 3.28 (m, 2H), 2.50 – 2.31 (m, 2H), 1.97 – 1.93 (m, 1H), 1.90 – 1.76 (m, 1H), 1.74 – 1.65 (m, 3H), 1.59 – 1.51 (m, 1H), 0.96 (d, $J = 6.0$ Hz, 6H). HRMS m/z : $[M+H]^+$ Calculated for $C_{21}H_{29}FN_3O_5$: 422.2091, Found: 422.2085. HRMS m/z : $[M+Na]^+$ Calculated for $C_{21}H_{28}FN_3NaO_5$: 444.1911, Found: 444.1903.

(4-Fluorophenyl)methyl-d2 ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**20b**). Yield (75%). 1H NMR (400 MHz, $CDCl_3$) δ 9.48 (s, 1H), 8.36 (d, $J = 5.7$ Hz, 1H), 7.37 – 7.28 (m, 2H), 7.08 – 6.98 (m, 2H), 6.06 (s, 1H), 5.38 (d, $J = 8.6$ Hz, 1H), 4.39 – 4.14 (m, 2H), 3.41 – 3.30 (m, 2H), 2.48 – 2.29 (m, 2H), 2.01 – 1.92 (m, 1H), 1.92 – 1.80 (m, 1H), 1.80 – 1.62 (m, 3H), 1.59 – 1.46 (m, 1H), 0.96 (d, $J = 5.9$ Hz, 6H). HRMS m/z : $[M+H]^+$ Calculated for $C_{21}H_{27}D_2FN_3O_5$: 424.2217, Found: 424.2210. HRMS m/z : $[M+Na]^+$ Calculated for $C_{21}H_{26}D_2FN_3NaO_5$: 446.2037, Found: 446.2027.

(Perfluorophenyl)methyl ((2S)-4-methyl-1-oxo-1-(((2S)-1-oxo-3-(2-oxopyrrolidin-3-yl)propan-2-yl) amino) pentan-2-yl) carbamate (**21b**). Yield (86%). 1H NMR (400 MHz, $DMSO-d_6$) δ 9.39 (s, 1H), 8.47 (d, $J = 7.5$ Hz, 1H), 7.63 (s, 1H), 7.59 (d, $J = 7.8$ Hz, 1H),

5.14 (s, 2H), 4.23 – 4.14 (m, 1H), 4.07 – 3.98 (m, 1H), 3.22 – 3.03 (m, 2H), 2.33 – 2.20 (m, 1H), 2.17 – 2.06 (m, 2H), 1.93 – 1.83 (m, 1H), 1.70 – 1.56 (m, 2H), 1.51 – 1.39 (m, 2H), 0.91 – 0.80 (m, 6H). HRMS m/z: [M+H]⁺ Calculated for C₂₁H₂₅F₅N₃O₅ 494.1714; Found 494.1715. HRMS m/z: [M+Na]⁺ Calculated for C₂₁H₂₄F₅N₃NaO₅ 516.1534; Found 516.1520.

(Perfluorophenyl)methyl-*d*₂ ((*2S*)-4-methyl-1-oxo-1-(((*2S*)-1-oxo-3-(2-oxopyrrolidin-3-yl)propan-2-yl) amino) pentan-2-yl) carbamate (**22b**). Yield (83%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 9.38 (s, 1H), 8.46 (d, *J* = 7.6 Hz, 1H), 7.63 (s, 1H), 7.58 (d, *J* = 7.8 Hz, 1H), 4.22 – 4.14 (m, 1H), 4.06 – 3.99 (m, 1H), 3.20 – 3.05 (m, 2H), 2.27 (d, *J* = 7.9 Hz, 1H), 2.16 – 2.07 (m, 2H), 1.92 – 1.83 (m, 1H), 1.68 – 1.58 (m, 2H), 1.50 – 1.41 (m, 2H), 0.90 – 0.80 (m, 6H). HRMS m/z: [M+H]⁺ Calculated for C₂₁H₂₃D₂F₅N₃O₅ 496.1840; Found 496.1837. HRMS m/z: [M+Na]⁺ Calculated for C₂₁H₂₂D₂F₅N₃NaO₅ 518.1660; Found 518.1646.

1-Phenylbutyl ((*S*)-4-methyl-1-oxo-1-(((*S*)-1-oxo-3-((*S*)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**23b**). Yield (71%). ¹H NMR (400 MHz, CDCl₃) δ 9.44 (d, *J* = 36.5 Hz, 1H), 8.21 (dd, *J* = 43.6, 6.1 Hz, 1H), 7.41 – 7.17 (m, 5H), 6.28 (d, *J* = 30.3 Hz, 1H), 5.68 – 5.54 (m, 1H), 5.36 (dd, *J* = 26.5, 8.5 Hz, 1H), 4.40 – 4.19 (m, 2H), 3.41 – 3.15 (m, 2H), 2.56 – 2.14 (m, 2H), 2.01 – 1.81 (m, 3H), 1.77 – 1.62 (m, 3H), 1.59 – 1.44 (m, 1H), 1.42 – 1.25 (m, 1H), 1.00 – 0.77 (m, 11H). HRMS m/z: [M+Na]⁺ Calculated for C₂₄H₃₅N₃NaO₅ 468.2475; Found 468.2463.

1,2-Diphenylethyl ((*S*)-4-methyl-1-oxo-1-(((*S*)-1-oxo-3-((*S*)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (**24b**). Yield (82%). ¹H NMR (400 MHz, DMSO-*d*₆) δ

9.35 (dd, $J = 24.5, 7.6$ Hz, 1H), 8.48 – 8.32 (m, 1H), 7.61 (d, $J = 21.6$ Hz, 1H), 7.53 – 7.42 (m, 1H), 7.36 – 7.09 (m, 10H), 5.83 – 5.71 (m, 1H), 4.24 – 4.08 (m, 1H), 4.06 – 3.89 (m, 1H), 3.24 – 2.86 (m, 4H), 2.34 – 2.05 (m, 3H), 1.94 – 1.78 (m, 1H), 1.67 – 1.28 (m, 4H), 0.90 – 0.68 (m, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{28}H_{35}N_3NaO_5$ 516.2475; Found 516.2462.

Preparation of compounds 1-24c. General procedure. To a solution of dipeptidyl aldehyde **b** (1 eq) in ethyl acetate (10 mL/g of dipeptidyl aldehyde) was added absolute ethanol (5 mL/g of dipeptidyl aldehyde) with stirring, followed by a solution of sodium bisulfite (1 eq) in water (1 mL/ g of dipeptidyl aldehyde). The reaction mixture was stirred for 3 h at 50 °C. The reaction mixture was allowed to cool to room temperature and then vacuum filtered. The solid was thoroughly washed with absolute ethanol and the filtrate was dried over anhydrous sodium sulfate, filtered, and concentrated to yield a white solid. The white solid was stirred with dry ethyl ether (3 x 10 mL/ g of dipeptidyl aldehyde), followed by careful removal of the solvent using a pipette and dried using a vacuum pump for 2 h to yield dipeptidyl bisulfite adduct **c** as a white solid.

Sodium (2S)-2-((S)-2-(((4,4-dimethylcyclohexyl)oxy)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**1c**). Yield (39%). 1H NMR (400 MHz, DMSO- d_6) δ 7.50 (dd, $J = 13.4, 9.1$ Hz, 1H), 7.44 (s, 1H), 7.21 – 7.11 (m, 1H), 5.45 – 5.27 (m, 1H), 4.53 – 4.39 (m, 1H), 4.01 – 3.92 (m, 1H), 3.95 – 3.78 (m, 1H), 3.14 – 3.08 (m, 1H), 3.06 – 3.01 (m, 1H), 2.19 – 2.05 (m, 1H), 2.04 – 1.86 (m, 1H), 1.70 – 1.65 (m, 3H), 1.62 – 1.52 (m, 3H), 1.52 – 1.31 (m, 6H), 1.28 – 1.09 (m, 2H), 0.92 – 0.80 (m, 12H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{22}H_{38}N_3Na_2O_8S$ 550.2175; Found 550.2154.

Sodium (2S)-1-hydroxy-2-((S)-2-((((1R,4S)-4-isopropylcyclohexyl)oxy)carbonyl)amino)-4-methylpentanamido)-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**2c**). Yield (36%). ¹H NMR (400 MHz, DMSO-d₆) δ 7.54 – 7.42 (m, 2H), 7.23 – 7.11 (m, 1H), 5.45 – 5.28 (m, 1H), 4.41 – 4.30 (m, 1H), 4.00 – 3.89 (m, 1H), 3.89 – 3.78 (m, 1H), 3.17 – 3.09 (m, 1H), 3.08 – 2.97 (m, 1H), 2.14 – 2.04 (m, 2H), 2.01 – 1.84 (m, 3H), 1.73 – 1.65 (m, 2H), 1.62 – 1.51 (m, 3H), 1.48 – 1.34 (m, 3H), 1.30 – 1.14 (m, 2H), 1.01 (s, 3H), 0.89 – 0.80 (m, 12H). HRMS m/z: [M+Na]⁺ Calculated for C₂₃H₄₀N₃Na₂O₈S 564.2332; Found 564.2450.

Sodium (2S)-1-hydroxy-2-((S)-4-methyl-2-((((1S,4S)-4-propylcyclohexyl)oxy)carbonyl)amino)pentanamido)-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**3c**). Yield (40%). ¹H NMR (400 MHz, DMSO-d₆) δ 7.54 – 7.39 (m, 2H), 7.23 – 7.10 (m, 1H), 5.45 – 5.28 (m, 1H), 4.42 – 4.32 (m, 1H), 4.03 – 3.88 (m, 1H), 3.89 – 3.78 (m, 1H), 3.18 – 3.08 (m, 1H), 3.07 – 2.97 (m, 1H), 2.19 – 2.03 (m, 2H), 1.97 – 1.80 (m, 3H), 1.71 (d, *J* = 13.2 Hz, 2H), 1.59 – 1.56 (m, 3H), 1.48 – 1.37 (m, 2H), 1.34 – 1.02 (m, 7H), 0.97 – 0.89 (m, 2H), 0.92 – 0.80 (m, 9H). HRMS m/z: [M+Na]⁺ Calculated for C₂₃H₄₀N₃Na₂O₈S 564.2332; Found 564.2311.

Sodium (2S)-2-((S)-2-((((1S,4S)-4-butylcyclohexyl)oxy)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**4c**). Yield (33%). ¹H NMR (400 MHz, DMSO-d₆) δ 7.55 – 7.42 (m, 2H), 7.22 – 7.12 (m, 1H), 5.48 – 5.30 (m, 1H), 4.42 – 4.32 (m, 1H), 3.97 – 3.89 (m, 1H), 3.88 – 3.79 (m, 1H), 3.17 – 3.08 (m, 1H), 3.07 – 2.98 (m, 1H), 2.19 – 2.02 (m, 2H), 1.99 – 1.83 (m, 3H), 1.72 (d, *J* = 13.3 Hz, 3H), 1.60 – 1.55 (m, 3H), 1.48 – 1.37 (m, 2H), 1.27 – 1.17 (m, 6H), 1.17 – 1.13 (m, 3H), 0.97 – 0.92 (m, 1H), 0.90 – 0.80 (m, 9H). HRMS m/z: [M+Na]⁺ Calculated for C₂₄H₄₂N₃Na₂O₈S 578.2488; Found 578.2473.

Sodium (2S)-1-hydroxy-2-((S)-4-methyl-2-((((1*r*,4S)-4-phenylcyclohexyl)oxy)carbonyl)amino)pentanamido)-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**5c**). Yield (35%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.57 – 7.47 (m, 1H), 7.44 (s, 1H), 7.32 – 7.10 (m, 6H), 5.52 – 5.22 (m, 1H), 4.55 – 4.50 (m, 1H), 4.06 – 3.89 (m, 1H), 3.89 – 3.77 (m, 1H), 3.15 – 3.10 (m, 1H), 3.09 – 3.00 (m, 1H), 2.14 – 2.09 (m, 2H), 2.05 – 1.97 (m, 2H), 1.81 (d, *J* = 13.1 Hz, 3H), 1.62 – 1.52 (m, 5H), 1.50 – 1.39 (m, 4H), 1.14 – 1.03 (m, 1H), 0.93 – 0.81 (m, 6H). HRMS *m/z*: [M+Na]⁺ Calculated for C₂₆H₃₈N₃Na₂O₈S 598.2175; Found 598.2152.

Sodium (2S)-1-hydroxy-2-((S)-4-methyl-2-(((4-(trifluoromethyl) cyclohexyl) methoxy) carbonyl) amino) pentanamido)-3-(2-oxopyrrolidin-3-yl) propane-1-sulfonate (**6c**). Yield (81%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.62 – 7.49 (m, 1H), 7.46 (s, 1H), 7.25 – 7.13 (m, 1H), 5.50 (d, *J* = 6.3 Hz, 1H), 5.34 (d, *J* = 6.0 Hz, 1H), 4.41 – 4.33 (m, 1H), 4.27 – 4.18 (m, 1H), 3.81 – 3.70 (m, 2H), 3.13 (s, 2H), 3.08 – 2.99 (m, 1H), 2.20 – 2.06 (m, 3H), 1.91 – 1.76 (m, 4H), 1.63 – 1.57 (m, 4H), 1.49 – 1.37 (m, 1H), 1.27 – 1.16 (m, 1H), 1.13 – 0.92 (m, 4H), 0.93 – 0.80 (m, 6H). HRMS *m/z*: [M+Na]⁺ Calculated for C₂₂H₃₅F₃N₃Na₂O₈S 604.1893; Found 604.1871.

Sodium (2S)-1-hydroxy-2-((S)-4-methyl-2-((((1*r*,4S)-4-(trifluoromethyl) cyclohexyl) methoxy) carbonyl) amino) pentanamido)-3-(2-oxopyrrolidin-3-yl) propane-1-sulfonate (**7c**). Yield (76%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.58 – 7.50 (m, 1H), 7.46 (d, *J* = 9.6 Hz, 1H), 7.25 – 7.16 (m, 1H), 5.41 (d, *J* = 6.3 Hz, 1H), 5.28 (d, *J* = 5.9 Hz, 1H), 4.45 – 4.32 (m, 1H), 4.00 – 3.90 (m, 2H), 3.50 – 3.25 (m, 1H), 3.08 – 3.01 (m, 2H), 2.31 – 2.26 (m, 3H), 2.15 – 2.08 (m, 1H), 1.90 – 1.85 (m, 4H), 1.67 – 1.33 (m, 10H), 0.91 – 0.81 (m, 6H). HRMS *m/z*: [M+Na]⁺ Calculated for C₂₂H₃₅F₃N₃Na₂O₈S 604.1893; Found 604.1862.

Sodium (2*S*)-1-hydroxy-2-(((*S*)-4-methyl-2-((((1*r*,4*R*)-4-propylcyclohexyl) methoxy) carbonyl) amino) pentanamido)-3-(2-oxopyrrolidin-3-yl) propane-1-sulfonate (**8c**). Yield (82%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.53 – 7.40 (m, 1H), 7.20 – 7.09 (m, 1H), 6.10 – 6.04 (m, 1H), 5.41 (d, *J* = 6.2 Hz, 1H), 5.29 (d, *J* = 6.0 Hz, 1H), 4.43 – 4.32 (m, 1H), 4.00 – 3.89 (m, 1H), 3.74 (s, 2H), 3.16 – 3.01 (m, 2H), 2.15 – 2.05 (m, 6H), 1.71 (d, *J* = 11.4 Hz, 10H), 1.53 – 1.36 (m, 2H), 1.35 – 1.22 (m, 2H), 1.19 – 1.03 (m, 2H), 0.98 – 0.77 (m, 9H). HRMS *m/z*: [M+Na]⁺ Calculated for C₂₄H₄₂N₃Na₂O₈S 578.2488; Found 578.2460.

Sodium (2*S*)-1-hydroxy-2-(((*S*)-4-methyl-2-((((1*s*,4*S*)-4-propylcyclohexyl) methoxy-*d*₂) carbonyl) amino) pentanamido)-3-(2-oxopyrrolidin-3-yl) propane-1-sulfonate (**9c**). Yield (77%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.57 – 7.42 (m, 1H), 7.27 – 7.09 (m, 1H), 6.07 (d, *J* = 10.0 Hz, 1H), 5.40 (d, *J* = 6.4 Hz, 1H), 5.29 (d, *J* = 6.0 Hz, 1H), 3.99 – 3.89 (m, 1H), 3.78 – 3.69 (m, 1H), 3.17 – 2.98 (m, 4H), 2.20 – 2.06 (m, 4H), 1.71 (d, *J* = 11.6 Hz, 4H), 1.49 – 1.35 (m, 8H), 1.35 – 1.22 (m, 2H), 1.18 – 1.09 (m, 2H), 0.90 – 0.80 (m, 9H). HRMS *m/z*: [M+Na]⁺ Calculated for C₂₄H₄₀D₂N₃Na₂O₈S 580.2614; Found 580.2582.

Sodium (2*S*)-2-(((*S*)-2-((((4,4-difluorocyclohexyl) methoxy) carbonyl) amino)-4-methylpentanamido)-1-hydroxy-3-(((*S*)-2-oxopyrrolidin-3-yl) propane-1-sulfonate (**10c**). Yield (50.5%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.57 (t, *J* = 8.9 Hz, 1H), 7.45 (s, 1H), 7.38 – 7.17 (m, 1H), 4.29 – 4.10 (m, 1H), 4.05 – 3.67 (m, 4H), 3.09 (dt, *J* = 29.8, 8.8 Hz, 2H), 2.33 – 2.05 (m, 2H), 2.05 – 1.88 (m, 4H), 1.88 – 1.64 (m, 5H), 1.64 – 1.48 (m, 2H), 1.43 (q, *J* = 7.3 Hz, 2H), 1.30 – 1.11 (m, 2H), 1.04 – 0.78 (m, 6H).

Sodium (2*S*)-2-(((*S*)-2-((((4,4-difluorocyclohexyl) methoxy-*d*₂) carbonyl) amino)-4-methylpentanamido)-1-hydroxy-3-(2-oxopyrrolidin-3-yl) propane-1-sulfonate (**11c**). Yield (81%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.60 – 7.50 (m, 1H), 7.45 (s, 1H), 7.29 – 7.18 (m,

1H), 5.41 (d, $J = 6.3$ Hz, 1H), 5.24 (d, $J = 6.0$ Hz, 1H), 4.37 – 4.32 (m, 1H), 3.98 – 3.89 (m, 1H), 3.15 – 3.02 (m, 2H), 2.14 – 2.05 (m, 3H), 2.01 – 1.96 (m, 1H), 1.84 – 1.70 (m, 8H), 1.62 – 1.53 (m, 4H), 1.45 – 1.38 (m, 1H), 0.90 – 0.80 (m, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{21}H_{32}D_2F_2N_3Na_2O_8S$ 574.1956; Found 574.1931.

Sodium (2S)-2-((2S)-2-(((1-(4,4-difluorocyclohexyl)ethoxy)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**12c**). Yield (48%). 1H NMR (400 MHz, DMSO- d_6) δ 7.52 (d, $J = 9.9$ Hz, 1H), 7.43 (s, 1H), 7.28 – 7.13 (m, 1H), 5.36 – 5.17 (m, 1H), 4.58 – 4.51 (m, 1H), 3.96 – 3.91 (m, 1H), 3.83 – 3.78 (m, 1H), 3.18 – 3.09 (m, 1H), 3.06 – 3.01 (m, 1H), 2.17 – 1.88 (m, 3H), 1.87 – 1.77 (m, 4H), 1.74 – 1.66 (m, 2H), 1.65 – 1.51 (m, 3H), 1.48 – 1.35 (m, 2H), 1.30 – 1.16 (m, 3H), 1.16 – 1.04 (m, 3H), 0.89 – 0.80 (m, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{22}H_{36}F_2N_3Na_2O_8S$ 586.1987; Found 586.1978.

Sodium (2S)-2-((S)-2-(((2-(4,4-difluorocyclohexyl)propan-2-yl)oxy)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**13c**). Yield (39%). 1H NMR (400 MHz, DMSO- d_6) δ 7.49 (d, $J = 10.6$ Hz, 1H), 7.43 (s, 1H), 7.12 – 6.95 (m, 1H), 5.47 – 5.26 (m, 1H), 4.06 – 3.71 (m, 2H), 3.17 – 3.08 (m, 1H), 3.08 – 2.97 (m, 1H), 2.14 – 1.93 (m, 6H), 1.86 – 1.65 (m, 5H), 1.64 – 1.49 (m, 2H), 1.44 – 1.36 (m, 2H), 1.33 (s, 6H), 1.28 – 1.23 (m, 2H), 0.84 (ddd, $J = 11.6, 6.5, 3.0$ Hz, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{23}H_{38}F_2N_3Na_2O_8S$ 600.2143; Found 600.2131.

Sodium (2S)-2-((2S)-2-(((1-(4,4-difluorocyclohexyl)-2-phenylethoxy)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**14c**). Yield (41%). 1H NMR (400 MHz, DMSO- d_6) δ 7.53 – 7.47 (m, 1H), 7.43 (s, 1H), 7.39 –

7.33 (m, 1H), 7.28 – 7.15 (m, 5H), 5.38 – 5.15 (m, 1H), 4.75 – 4.71 (m, 1H), 3.96 – 3.91 (m, 1H), 3.87 – 3.68 (m, 1H), 3.13 – 3.09 (m, 2H), 3.07 – 2.96 (m, 1H), 2.90 – 2.82 (m, 1H), 2.00 (s, 4H), 1.92 – 1.80 (m, 3H), 1.76 (s, 3H), 1.59 – 1.51 (m, 4H), 1.49 – 1.33 (m, 3H), 0.89 – 0.79 (m, 4H), 0.79 – 0.67 (m, 2H). HRMS m/z: [M+Na]⁺ Calculated for C₂₈H₄₀F₂N₃NaO₈S 662.2300; Found 662.2288.

Sodium (2S)-2-((2S)-2-(((4,4-difluorocyclohexyl)(phenyl)methoxy)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**15c**). Yield (66%). ¹H NMR (400 MHz, DMSO-d₆) δ 7.64 – 7.50 (m, 1H), 7.49 – 7.44 (m, 1H), 7.42 – 7.19 (m, 6H), 5.48 – 5.35 (m, 1H), 4.01 – 3.75 (m, 3H), 3.18 – 3.10 (m, 1H), 3.08 – 3.01 (m, 1H), 2.22 – 2.06 (m, 1H), 2.06 – 1.91 (m, 3H), 1.87 – 1.68 (m, 4H), 1.61 – 1.54 (m, 3H), 1.49 – 1.40 (m, 3H), 1.35 – 1.14 (m, 3H), 0.93 – 0.71 (m, 6H). HRMS m/z: [M+Na]⁺ Calculated for C₂₇H₃₈F₂N₃Na₂O₈S 648.2143; Found 648.2121.

Sodium (2S)-2-((2S)-2-(((1-(4,4-difluorocyclohexyl)pentyl)oxy)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**16c**). Yield (58%). ¹H NMR (400 MHz, DMSO-d₆) δ 7.60 – 7.51 (m, 1H), 7.44 (s, 1H), 7.38 – 7.23 (m, 1H), 5.46 – 5.21 (m, 1H), 3.98 – 3.74 (m, 3H), 3.14 – 3.07 (m, 1H), 3.06 – 2.98 (m, 1H), 2.16 – 2.04 (m, 2H), 2.04 – 1.92 (m, 4H), 1.86 – 1.66 (m, 8H), 1.64 – 1.52 (m, 3H), 1.49 – 1.36 (m, 3H), 1.21 – 1.17 (m, 3H), 0.83 (ddd, J = 11.9, 6.5, 3.1 Hz, 9H). HRMS m/z: [M+Na]⁺ Calculated for C₂₅H₄₂F₂N₃Na₂O₈S 628.2456; Found 628.2419.

Sodium (2S)-2-((S)-2-(((2-fluorobenzyl)oxy)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**17c**). Yield (71%). ¹H NMR (400 MHz, DMSO-d₆) δ 7.77 – 7.70 (m, 1H), 7.70 – 7.59 (m, 1H), 7.59 – 7.33 (m, 3H), 7.26 –

7.12 (m, 2H), 5.65 (d, $J = 78.6$ Hz, 1H), 5.16 – 5.01 (m, 2H), 4.11 – 3.84 (m, 2H), 3.17 – 2.97 (m, 2H), 2.39 – 2.07 (m, 2H), 2.07 – 1.85 (m, 1H), 1.70 – 1.51 (m, 3H), 1.51 – 1.33 (m, 2H), 0.92 – 0.77 (m, 6H). HRMS m/z : $[M]^-$ Calculated for $C_{21}H_{29}FN_3O_8S$: 502.1659, Found: 502.1650. HRMS m/z : $[M+Na]^+$ Calculated for $C_{21}H_{29}FN_3Na_2O_8S$: 548.1455, Found: 548.1446.

Sodium (2S)-2-((S)-2-(((3-fluorobenzyl)oxy)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**18c**). Yield (89%). 1H NMR (400 MHz, DMSO- d_6) δ 7.75 (d, $J = 9.0$ Hz, 1H), 7.64 (d, $J = 9.3$ Hz, 1H), 7.57 (d, $J = 7.9$ Hz, 1H), 7.54 – 7.35 (m, 2H), 7.25 – 7.08 (m, 2H), 5.75 – 5.46 (m, 1H), 5.15 – 4.98 (m, 2H), 4.13 – 3.87 (m, 2H), 3.17 – 2.89 (m, 2H), 2.23 – 2.04 (m, 2H), 2.04 – 1.91 (m, 1H), 1.89 – 1.74 (m, 1H), 1.70 – 1.31 (m, 4H), 0.91 – 0.77 (m, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{21}H_{29}FN_3Na_2O_8S$: 548.1455, Found: 548.1450. HRMS m/z : $[M]^-$ Calculated for $C_{21}H_{29}FN_3O_8S$: 502.1659, Found: 502.1655.

Sodium (2S)-2-((S)-2-(((4-fluorobenzyl)oxy)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**19c**). Yield (68%). 1H NMR (400 MHz, DMSO- d_6) δ 7.70 – 7.57 (m, 1H), 7.57 – 7.49 (m, 1H), 7.48 – 7.37 (m, 3H), 7.23 – 7.14 (m, 2H), 5.43 (d, $J = 85.6$ Hz, 1H), 5.09 – 4.92 (m, 2H), 4.12 – 3.81 (m, 2H), 3.14 – 2.94 (m, 2H), 2.23 – 1.91 (m, 2H), 1.62 – 1.50 (m, 4H), 1.50 – 1.38 (m, 2H), 0.90 – 0.80 (m, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{21}H_{29}FN_3Na_2O_8S$: 548.1455, Found: 548.1448. HRMS m/z : $[M]^-$ Calculated for $C_{21}H_{29}FN_3O_8S$: 502.1659, Found: 502.1645.

Sodium (2S)-2-((S)-2-(((4-fluorophenyl)methoxy- d_2)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**20c**). Yield (90%). 1H NMR (400 MHz, DMSO- d_6) δ 7.68 (d, $J = 9.2$ Hz, 1H), 7.63 (d, $J = 9.2$

Hz, 1H), 7.53 (d, $J = 7.8$ Hz, 1H), 7.48 – 7.35 (m, 2H), 7.23 – 7.13 (m, 2H), 5.52 (dd, $J = 85.5, 6.2$ Hz, 1H), 4.04 – 3.84 (m, 2H), 3.18 – 2.97 (m, 2H), 2.27 – 1.90 (m, 2H), 1.63 – 1.50 (m, 3H), 1.53 – 1.38 (m, 3H), 0.92 – 0.78 (m, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{21}H_{27}D_2FN_3Na_2O_8S$: 550.1581, Found: 550.1573. HRMS m/z : $[M+Na]^+$ Calculated for $C_{21}H_{27}D_2FN_3O_8S$: 504.1785, Found: 504.1769.

Sodium (2S)-1-hydroxy-2-((S)-4-methyl-2-(((perfluorophenyl) methoxy) carbonyl) amino) pentanamido)-3-(2-oxopyrrolidin-3-yl) propane-1-sulfonate (**21c**). Yield (81%). 1H NMR (400 MHz, DMSO- d_6) δ 7.68 – 7.42 (m, 2H), 6.06 (s, 1H), 5.41 (d, $J = 6.3$ Hz, 1H), 5.24 (d, $J = 5.9$ Hz, 1H), 5.13 (s, 2H), 4.36 (d, $J = 7.2$ Hz, 1H), 4.01 – 3.89 (m, 1H), 3.48 – 3.41 (m, 3H), 2.20 – 2.04 (m, 3H), 1.63 – 1.33 (m, 4H), 0.89 – 0.78 (m, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{21}H_{25}F_5N_3Na_2O_8S$ 620.1078; Found 620.1069.

Sodium (2S)-1-hydroxy-2-((S)-4-methyl-2-(((perfluorophenyl)methoxy- d_2) carbonyl) amino) pentanamido)-3-(2-oxopyrrolidin-3-yl) propane-1-sulfonate (**22c**). Yield (89%). 1H NMR (400 MHz, DMSO- d_6) δ 7.67 (d, $J = 9.1$ Hz, 1H), 7.58 (s, 1H), 7.50 (d, $J = 8.8$ Hz, 1H), 5.12 (d, $J = 12.4$ Hz, 1H), 4.42 – 4.36 (m, 1H), 3.99 – 3.90 (m, 2H), 3.16 – 3.00 (m, 2H), 2.20 – 2.05 (m, 3H), 1.77 (s, 1H), 1.62 – 1.50 (m, 2H), 1.47 – 1.37 (m, 2H), 0.88 – 0.78 (m, 6H). HRMS m/z : $[M+Na]^+$ Calculated for $C_{21}H_{23}D_2F_5N_3Na_2O_8S$ 622.1204; Found 622.1193.

Sodium(2S)-1-hydroxy-2-((2S)-4-methyl-2-(((1-phenylbutoxy)carbonyl)amino)pentanamido)-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**23c**). Yield (63%). 1H NMR (400 MHz, DMSO- d_6) δ 7.60 (d, $J = 9.8$ Hz, 1H), 7.47 (d, $J = 5.0$ Hz, 1H), 7.41 (d, $J = 7.4$ Hz, 1H), 7.38 – 7.22 (m, 5H), 5.54 (t, $J = 6.6, 6.6$

Hz, 1H), 5.37 – 5.25 (m, 1H), 4.12 – 3.78 (m, 2H), 3.17 – 2.95 (m, 2H), 2.31 – 1.88 (m, 3H), 1.85 – 1.38 (m, 4H), 1.36 – 1.18 (m, 2H), 0.93 – 0.82 (m, 10H), 0.79 – 0.70 (m, 2H). HRMS m/z: [M+Na]⁺ Calculated for C₂₄H₃₆N₃Na₂O₈S 572.2019; Found 572.2005.

Sodium (2S)-2-((2S)-2-(((1,2-diphenylethoxy)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (**24c**). Yield (69%). ¹H NMR (400 MHz, DMSO-d₆) δ 7.69 – 7.36 (m, 2H), 7.36 – 7.07 (m, 11H), 5.82 – 5.71 (m, 1H), 5.59 – 5.32 (m, 1H), 4.15 – 3.81 (m, 2H), 3.18 – 2.76 (m, 4H), 2.32 – 1.72 (m, 4H), 1.71 – 1.30 (m, 4H), 0.90 – 0.60 (m, 6H). HRMS m/z: [M+Na]⁺ Calculated for C₂₈H₃₆N₃Na₂O₈S 620.2019; Found 620.2007.

Biochemical Studies

Enzyme assays and inhibition studies. *Cloning and expression of the 3CLpro of SARS-CoV-2 and FRET enzyme assays.* The codon-optimized cDNA of full length of 3CLpro of SARS-CoV-2 (GenBank number MN908947.3) fused with sequences encoding 6 histidine at the N-terminal was synthesized by Integrated DNA (Coralville, IA). The synthesized gene was subcloned into the pET-28a(+) vector. The expression and purification of SARS-CoV-2 3CLpro were conducted following a standard procedure described previously.¹⁰ Briefly, a stock solution of an inhibitor was prepared in DMSO and diluted in assay buffer comprised of 20 mM HEPES buffer, pH 8, containing NaCl (200 mM), EDTA (0.4 mM), glycerol (60%), and 6 mM dithiothreitol (DTT). The SARS-CoV-2 3CLpro was mixed with serial dilutions of compound or with DMSO in 25 μL of assay buffer and incubated at 37°C for 1 h, followed by the addition of 25 μL of assay buffer containing substrate (FAM-SAVLQ/SG-QXL[®]520, AnaSpec, Fremont, CA). The substrate

was derived from the cleavage sites on the viral polyproteins of SARS-CoV. Fluorescence readings were obtained using an excitation wavelength of 480 nm and an emission wavelength of 520 nm on a fluorescence microplate reader (FLx800; Biotec, Winoosk, VT) 1 h following the addition of substrate. Relative fluorescence units (RFU) were determined by subtracting background values (substrate-containing well without protease) from the raw fluorescence values, as described previously. The dose-dependent FRET inhibition curves were fitted with a variable slope by using GraphPad Prism software (GraphPad, La Jolla, CA) in order to determine the IC₅₀ values of the compounds. The expression and purification of the 3CLpro of MERS-CoV, as well as the FRET enzyme assays were performed as described previously.^{10-12,33}

Cell based assay to screen SARS-CoV-2 3CLpro inhibitors. Two plasmids, pR-SARS-CoV-2 3CLpro and pGlo-VRLQS were used for the system (**Figure 1/panel A**). First, the open reading frame of SARS-CoV-2 3CLpro was cloned to the reverse genetics system of PRRSV with GFP³⁵ and designated as pR-SARS-CoV-2 3CLpro. The GFP gene was replaced with SARS-CoV-2 3CLpro gene with AflII and MluI enzyme sites. As a control, inactive form of the 3CLpro was introduced to pR-SARS-CoV-2 3CLpro by the mutagenesis, and resulting plasmid was designated as pR-SARS-CoV-2 3CLpro C145A. Second plasmid was utilized with the pGloSensor caspase 3/7 biosensor (Promega, Madison, WI) which contains a caspase-3/7 cleavage site engineered in the Firefly luciferase gene. This plasmid also contains intact Renilla luciferase gene as an expression control. After transfection of pGloSensor caspase 3/7 to cells, the Firefly luciferase is expressed as an inactive form, and in the presence of caspase 3/7, it is activated after the cleavage. The caspase-3/7 cleavage site in this plasmid was replaced

with CoV 3CLpro recognition sequences, VRLQS, designated as pGlo-VRLQS.³⁴ As a result, the expressed inactive luciferase is activated by the cleavage with CoV 3CLpro in the cells (**Figure 1/panel A**). HEK293T cells were used for transfection and the compound screening (**Figure 1/panel B**). One-day old HEK293T cells in 48-well plates were co-transfected with two plasmids, pR-SARS-CoV-2 3CLpro and pGlo-VRLQS, or pR-SARS-CoV-2 3CLpro C145C and pGlo-VRLQS. Following morning, medium containing Mock-DMSO or serial concentrations of each compound were replaced to the transfected cells and incubate at 37 C for 6 h. Cell lysates were prepared for testing the levels of Firefly and Renilla luciferases (Dual Luciferase assay kit, Promega) in a luminometer (Promega). The expression levels of Firefly luciferase were normalized with Renilla luciferase levels. The transfection of pR-SARS-CoV-2 3CLpro C145A and pGlo-VRLQS resulted in minimal levels of Firefly luciferase, and this was applied to adjust all Firefly expression levels. The inhibition curve (**Figure 1/panel C**) for each compound was prepared, and the 50% effective concentration (EC_{50}) values were determined by GraphPad Prism software using a variable slope (GraphPad, La Jolla, CA). Compounds **4c** and **15c** were selected and examined for the antiviral effects with live SARS-CoV-2 in Vero E6 cells as described before³³, and the EC_{50} was calculated by the same method described above. When GC376 and the selected compounds were examined if they inhibited the replication of PRRSV, none of them resulted in the reduction of viral replication.

Nonspecific cytotoxic effects/In vitro cytotoxicity. HEK293T cells grown in 96-well plates were incubated with various concentrations (1 to 100 μ M) of each compound for 72 h. Cell cytotoxicity was measured by a CytoTox 96 nonradioactive cytotoxicity assay

kit (Promega), and the CC_{50} values were calculated using a variable slope by GraphPad Prism software.

X-ray Crystallographic Studies

Crystallization and Data Collection. Purified SARS-CoV-2 3CLpro¹⁰ in 100mM NaCl, 20mM Tris pH 8.0 was concentrated to 9.6 mg/mL (0.28 mM) for crystallization screening. All crystallization experiments were setup using an NT8 drop-setting robot (Formulatrix Inc.) and UVXPO MRC (Molecular Dimensions) sitting drop vapor diffusion plates at 18 °C. 100 nL of protein and 100 nL crystallization solution were dispensed and equilibrated against 50 μ L of the latter. A stock solution of 100 mM compound was prepared in DMSO and the SARS-CoV-2 3CLpro:compound complex was prepared by mixing 1 μ L of the ligand (2 mM) with 49 μ L (0.28 mM) of SARS2 3CLpro and incubating on ice for 1 hour. Crystals were obtained in 1-2 days from various conditions for the following complexes. **8b:** Proplex HT screen (Molecular Dimensions) condition F7 (0.5 M ammonium sulfate, 100 mM MES pH 6.5). **12b, 13c, 14c** and **21c:** Index HT screen (Hampton Research) condition D10 (20% (w/v) PEG 5000 MME, 100 mM Bis-Tris pH 6.5). **19b** and **20b:** Proplex HT screen (Rigaku Reagents) condition C5 (20% (w/v) PEG 4000, 100 mM Tris pH 8.0). **1c:** Index HT screen (Hampton Research) condition F2 (20% (w/v) PEG 2000 MME, 100 mM Tris pH 8.5, 200 mM Trimethylamine N-oxide dihydrate). **3c:** Index HT screen (Hampton Research) condition F5 (17% (w/v) PEG 10,000, 100 mM Bis-Tris pH 5.5, 100 mM ammonium acetate). **5c:** Index HT screen (Hampton Research) condition F1 (10% (w/v) PEG 3350, 100 mM Hepes pH 7.5, 200 L-proline). **17c** and **18c:** Index HT screen (Rigaku Reagents) condition H11 (30% (w/v) PEG 2000 MME, 100 mM potassium thiocyanate). Samples were transferred to a fresh drop composed of 80% crystallization

solution and 20% (v/v) PEG 200 and stored in liquid nitrogen. Crystals of SARS-CoV-2 3CLpro with **8b** were transferred to a cryoprotectant solution containing 80% crystallant and 20% (v/v) glycerol prior to freezing. X-ray diffraction data were collected at the Advanced Photon Source beamline except for the SARS-CoV-2 3CLpro complex with **14c** which were collected at the National Synchrotron Light Source II (NSLS-II) AMX beamline 17-ID-1. All diffraction data were collected using a Dectris Eiger2 X 9M pixel array detector.

Structure Solution and Refinement. Intensities were integrated using XDS³⁸⁻³⁹ via Autoproc⁴⁰ and the Laue class analysis and data scaling were performed with Aimless⁴¹. Structure solution was conducted by molecular replacement with Phaser⁴² using a previously determined structure of SARS2 3CLpro (PDB 6XMK¹⁰) as the search model. Structure refinement and manual model building were conducted with Phenix⁴³ and Coot⁴⁴ respectively. Disordered side chains were truncated to the point for which electron density could be observed. Structure validation was conducted with Molprobit⁴⁵ and figures were prepared using the CCP4MG⁴⁶ package.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Supplemental information (.docx)

Molecular formula strings – SMILES codes (.csv)

AUTHOR INFORMATION

Corresponding authors

*(Y.K.) Phone: 785-532-4616. E-mail: ykim@ksu.edu

*(W.C.G.) Phone: 316-978-7374. E-mail: bill.groutas@wichita.edu.

*(K-O.C.) Phone: 785-532-3849. E-mail: kchang@vet.ksu.edu

Notes

The authors declare no competing financial interests.

ACKNOWLEDGEMENTS

This research was supported in part by grants from the National Institutes of Health (NIH) (R01 AI109039 to K.O.C). Use of the University of Kansas Protein Structure Laboratory was supported by a grant from the National Institute of General Medical Sciences (P30GM110761) of the NIH. Use of the IMCA-CAT beamline 17-ID at the Advanced Photon Source was supported by the companies of the Industrial Macromolecular Crystallography Association through a contract with Hauptman-Woodward Medical Research Institute. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under contract no. DE-AC02-06CH11357. This research used the AMX beamline of the National Synchrotron Light Source II, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704. The Center for BioMolecular Structure (CBMS) is primarily supported by the National Institutes of Health, National Institute of General Medical Sciences (NIGMS) through a Center Core P30 Grant (P30GM133893), and by the DOE Office of Biological and Environmental Research (KP1605010).

Abbreviations used

ORF, open reading frame; DSC, N,N'-disuccinimidyl carbonate; TEA, Triethyl amine; DMP, Dess-Martin periodinane; CDI, carbonyl diimidazole; DTT, dithiothreitol; DMSO, dimethyl sulfoxide; MNV, murine norovirus; MOI, multiplicity of infection; CPE, cytopathic effects; TCID₅₀, the 50% tissue culture infectious dose; IC₅₀, the 50% inhibitory concentration in the enzyme assay; EC₅₀, the 50% effective concentration in cell culture; CC₅₀, 50% cytotoxic concentration in cell-based assays; GESAMT, general efficient *structural* alignment of macromolecular targets; RMSD, root mean square deviation; XDS, X-ray detector software; MME, monomethyl ether; PK, pharmacokinetics.

Accession Codes

Coordinates and structure factors for the following SARS2 3CLpro complexes with inhibitors were deposited to the Worldwide Protein Databank (wwPDB) with the accession codes: **8b** (7LZT), **12b** (7LZU), **19b** (7LZV), **20b** (7LZW), **1c** (7LZX), **3c** (7LZY), **5c** (7LZZ), **13c** (7M00), **14c** (7M01), **17c** (7M02), **18c** (7M03) and **21c** (7M04). Authors will release the atomic coordinates upon article publication.

References

1. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.-M.; Wang, W.; Song, Z.-G.; Hu, Y.; Tao, Z.-W.; Tian, J.-H.; Pei, Y.-Y.; Yuan, M.-L.; Zhang, Y.-L.; Dai, F.-H.; Liu, Y.; Wang, Q.-M.; Zheng, J.-J.; Xu, L.; Holmes, E. C.; Zhang, Y.-Z. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, *579*, 265-269.
2. Fontanet, A.; Autran, B.; Lina, B.; Kieny, M. P.; Abdool Karim, S. S.; Sridhar, D. SARS-CoV-2 variants and ending the COVID-19 pandemic. *The Lancet* **2021**, *397*, 952-954.
3. Yurkovetskiy, L.; Wang, X.; Pascal, K. E.; Tomkins-Tinch, C.; Nyalile, T. P.; Wang, Y.; Baum, A.; Diehl, W. E.; Dauphin, A.; Carbone, C.; Veinotte, K.; Egri, S. B.; Schaffner, S. F.; Lemieux, J.; Munro, J. B.; Rafique, A.; Barve, A.; Sabeti, P.C.; Kyratsous, C. A.; Dudkina, N. V.; Shen, K.; Luban, J. Structural and functional analysis of the D614G SARS-CoV-2 spike protein. *CellPress* **2020**, *183*, 739-751.
4. Wu, D.; Koganti, R.; Lambe, U. P.; Yadavalli, T.; Nandi, S. S.; Shukla, D. Vaccines and therapies in development for SARS-CoV-2 infections. *J. Clin. Med.* **2020**, *9*, 1885.
5. Wang, C.; Li, W.; Drabek, D.; Okba, N. M. A.; Haperen, R. V.; Osterhaus, A. D. M. E.; Kuppeveld, F. J. M. V; Haagmans, B. L.; Grosveld, F.; Bosch, B.-J. A human monoclonal antibody blocking SARS-CoV-2. *Nature Comm.* **2020**, *11* 2251.

6. Ghosh, A. K.; Brindisi, M.; Shahabi, D.; Chapman, M. E.; Mesecar, A. D. Drug development and medicinal chemistry efforts toward SARS-Coronavirus and Covid-19 therapeutics. *Chem. Med. Chem.* **2020**, *15*, 907-932.
7. Gil, C.; Ginex, T.; Maestro, I.; Nozal, V.; Barrado-Gil, L.; Cuesta-Geijo, M. A.; Urquiza, J.; Ramírez, D.; Alonso, C.; Campillo, N. E., Martinez A., COVID-19: Drug targets and potential treatments. *J. Med. Chem.* **2020**, *63*, 12359-12386.
8. Sharma, A.; Tiwari, S.; Deb, M. K.; MARTY, J. L., Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): A global pandemic and treatments strategies. *Intl J. Antimicrob. Agents* **2020**, *56*, 106054.
9. Cannalire, R.; Cerchia, C.; Beccari, A. R.; Di Leva, F. S.; Summa, V., Targeting SARS-CoV-2 proteases and polymerase for COVID-19 treatment: State of the art and future opportunities. *J. Med. Chem.* [online early access]. DOI: 10.1021/acs.jmedchem.0c01140. Published online: Nov 13, 2020. <https://doi.org/10.1021/acs.jmedchem.0c01140> (accessed Nov 13, 2020).
10. Rathnayake, A. D.; Zheng, J.; Kim, Y.; Perera, K. D.; Mackin, S.; Meyerholz, D. K.; Kashipathy, M. M.; Battaile, K. P.; Lovell, S.; Perlman, S., 3C-like protease inhibitors block coronavirus replication in vitro and improve survival in MERS-CoV–infected mice. *Science Transl. Med.* **2020**, *12* :eabc5332.
11. Galasiti Kankanamalage, A. C.; Kim, Y.; Damalanka, V. C.; Rathnayake, A. D.; Fehr, A. R.; Mehzabeen, N.; Battaille, K. P.; Lovell, S.; Lushington, G. H.; Perlman, S.; Chang, K. O.; Groutas, W. C. Structure-guided design of potent and permeable inhibitors of MERS coronavirus 3CL protease that utilize a piperidine moiety as a novel design element. *Eur. J. Med. Chem.* **2018**, *150*, 334-346.

12. Kim, Y.; Lovell, S.; Tiew, K.-C.; Mandadapu, S. R.; Alliston, K. R.; Battaile, K. P.; Groutas, W. C.; Chang, K.-O., Broad-spectrum antivirals against 3C or 3C-like proteases of picornaviruses, noroviruses, and coronaviruses. *J. Virol.* **2012**, *86*, 11754-11762.
13. Kim, Y.; Liu, H.; Galasiti Kankanamalage, A. C.; Weerasekara, S.; Hua, D. H.; Groutas, W. C.; Chang, K.-O.; Pedersen, N. C., Reversal of the progression of fatal coronavirus infection in cats by a broad-spectrum coronavirus protease inhibitor. *PLoS Pathogens* **2016**, *12*, e1005531.
14. Pedersen, N. C.; Kim, Y.; Liu, H.; Galasiti Kankanamalage, A. C.; Eckstrand, C.; Groutas, W. C.; Bannasch, M.; Meadows, J. M.; Chang, K.-O., Efficacy of a 3C-like protease inhibitor in treating various forms of acquired feline infectious peritonitis. *J. Feline Med. Surg.* **2018**, *20*, 378-392.
15. Schechter, I., Reprint of "on the size of the active site in proteases. I. Papain". *Biochem. Biophys. Res. Com.* **2012**, *425*, 497-502. The nomenclature used is that of Schechter, I. and Berger, A., where the residues on the N-terminus side of the peptide bond that is cleaved are designated as P₁-P_n and those on the C-terminus side are designated P₁'-P_n'. The corresponding active site subsites are designated S₁-S_n and S₁'-S_n'. S₁ is the primary substrate specificity subsite and P₁-P₁' is the scissile bond.
16. Ullrich, S.; Nitsche, C., The SARS-CoV-2 main protease as drug target. *Bioorg. Med. Chem. Lett.* **2020**, *30*, 127377.

17. Konwar, M.; Sarma, D., Advances in developing small molecule SARS 3CLpro inhibitors as potential remedy for Corona Virus infection. *Tetrahedron* **2021**, *77*, 131761.
18. He, J.; Hu, L.; Huang, X.; Wang, C.; Zhang, Z.; Wang, Y.; Zhang, D.; Ye, W., Potential of coronavirus 3C-like protease inhibitors for the development of new anti-SARS-CoV-2 drugs: Insights from structures of protease and inhibitors. *Intl J. Antimicrob. Agents* **2020**, *56(2)*, 106055.
19. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C., Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L.W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H., Structure of M pro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020**, *582*, 289-293.
20. Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R., Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science* **2020**, *368*, 409-412.
21. Dai, W.; Zhang, B.; Jiang, X.-M.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Peng, J.; Liu, F., Li, C.; Li, Y.; Bai, F.; Wang, H.; Cheng, X.; Cen, X.; Hu, S.; Yang, X.; Wang, J.; Liu, X.; Xiao, G.; Jiang, H.; Rao, Z.; Zhang, L.-K.; Xu, Y.; Yang, H.; Liu, H. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* **2020**, *368*, 1331-1335.
22. Qiao, J.; Li, Y-S.; Zeng, R.; Liu, F-L.; Luo, R-H.; Huang C.; Wang, Y-F.; Zhang, J.; Quan, B.; Shen, C.; Mao, X.; Liu, X.; Sun, W.; Yang, W.; Ni, X.; Wang, K.; Xu, L.;

- Duan, Z-L.; Zou, Q-C.; Zhang, H-L.; Qu, W.; Long, Y-H-P.; Li, M-H.; Yang, R-C.; Liu, X.; You, J.; Zhou, Y.; Yao, R.; Li, W-P.; Liu, J-M.; Chen, P.; Liu, Y.; Lin, G-F.; Yang, X.; Zou, J.; Li, L.; Hu, Y.; Lu, G-W.; Li, W-M.; Wei, Y-Q.; Zheng, Y-T.; Lei, J.; Yang, S., SARS-CoV-2 M^{pro} inhibitors with antiviral activity in a transgenic mouse model. *Science* **2021**, *371*, 1374-1378.
23. Kreutzer, A. G.; Krumberger, M.; Parrocha, C. M. T.; Morris, M. A.; Guaglianone, G.; Nowick, J. S. Structure-based design of a cyclic peptide inhibitor of the SARS-CoV-2 Main protease. *bioRxiv* [online early access]. DOI: 10.1101/2020.08.03.234872. Published online: Aug 03, 2020. <https://doi.org/10.1021/acs.jmedchem.0c01140> (accessed Aug 03, 2020).
24. Pennington, L. D.; Aquila, B. M.; Choi, Y.; Valiulin, R. A.; Muegge, I. Positional analog scanning: an effective strategy for multiparameter optimization in drug design. *J. Med. Chem.* **2020**, *63*, 8956-8976.
25. Vulpetti, A.; Hommel, U.; Landrum, G.; Lewis, R.; Dalvit, C. Design and NMR-based screening of LEF, a library of chemical fragments with different local environments of fluorine. *J. Am. Chem. Soc.* **2009**, *131*, 12949-12959.
26. Pirali, T.; Serafini, M.; Cargnin, S.; Genazzani, A. A., Applications of deuterium in medicinal chemistry. *J. Med. Chem.* **2019**, *62*, 5276-5297.
27. Timmins, G. S. Deuterated drugs: updates and obviousness analysis. *Expert Opin. Ther. Pat.* **2017**, *27*, 1353-1361.
28. Liu, J. F. A decade of deuteration in medicinal chemistry. *Ann. Rep. Med. Chem.* **2017**, *50*, 519-542.

29. Kim, H-O.; Kahn, M. The synthesis of aminoazole analogs of lysine and arginine: the Mitsunobu reaction with lysinol and arginol. *Synlett*. **1999**, *8*, 1239-1240.
30. Maity, P.; Gujjar, M.; Vellingiri, R.; Lakshminarasimhan, T.; DelMonte, A. J.; Young, I. S.; Eastgate, M. D.; Vaidyanathan, R. Cerium(III) chloride-mediated stereoselective reduction of a 4-substituted cyclohexanone using NaBH₄. *Org. Process Res. Dev.* **2019**, *23*, 2754-2757.
31. Ghosh, A. K.; Duong, T. T.; McKee, S. P.; Thompson, W. J., N, N'-Disuccinimidyl carbonate: A useful reagent for alkoxycarbonylation of amines. *Tetrahedron Lett.* **1992**, *33*, 2781-2784.
32. Kjell, D. P.; Slattery, B. J.; Semo, M. J., A novel, nonaqueous method for regeneration of aldehydes from bisulfite adducts. *J. Org. Chem.* **1999**, *64*, 5722-5724.
33. Dampalla, C. S.; Zheng, J.; Perera, K. D.; Wong, L-Y. R.; Meyerholz, D. K.; Nguyen, H. N.; Kashipathy, M. M.; Battaile, K. P.; Lovell, S.; Kim, Y.; Perlman, S.; Groutas, W. C.; Chang, K-O. Post-infection treatment with a protease inhibitor increases survival of mice with a fatal SARS-CoV-2 infection. *bioRxiv* [online early access]. DOI: 10.1101/ 2021.02.05.429937. Published online: Feb 05, 2021. <https://doi.org/10.1101/ 2021.02.05.429937> (accessed Feb 05, 2021).
34. O'Brien, A.; Chen, D-Y.; Hackbart, M.; Close, B. J.; O'Brien, T.E.; Saeed, M.; Baker S.C.; Detecting SARS-CoV-2 3CLpro expression and activity using a polyclonal antiserum and a luciferase-based biosensor. *Virology*, **2021**, *556*: 73-78.
35. Lawson, S. R.; Li, Y.; Patton, J. B.; Langenhorst, R. J.; Sun, Z.; Jiang, Z.; Christopher-Hennings, J.; Nelson, E. A.; Knudsen, D.; Fang, Y.; Chang, K-O.

- Interleukin-1 β expression by a recombinant PRRSV enhanced viral specific host immunity. *Virus Research*. **2012**, *163*, 461-468.
36. Talele, T. T. Natural products-inspired use of the *gem*-dimethyl group in medicinal chemistry. *J. Med. Chem.* **2018**, *61*, 2166-2210.
37. Pennington, L. D.; Moustakas, D. T. The necessary nitrogen atom: a versatile high-impact design element for multiparameter optimization. *J. Med. Chem.* **2017**, *60*, 3552-3579.
38. Kabsch, W., Automatic indexing of rotation diffraction patterns. *J. Applied Crystallogr.* **1988**, *21*, 67-72.
39. Kabsch, W., XDS. *Acta Crystallogr. D* **2010**, *66*, 125-132.
40. Vonrhein, C.; Flensburg, C.; Keller, P.; Sharff, A.; Smart, O.; Paciorek, W.; Womack, T.; Bricogne, G., Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. Section D: Biolog. Crystallogr.* **2011**, *67*, 293-302.
41. Evans, P. R., An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. Section D: Biolog. Crystallogr.* **2011**, *67*, 282-292.
42. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *J. Applied Crystallogr.* **2007**, *40*, 658-674.
43. Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Section D: Biolog. Crystallogr.* **2010**, *66*, 213-221.

44. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr. Section D: Biolog. Crystallogr.* **2010**, *66*, 486-501.
45. Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Section D: Biolog. Crystallogr.* **2010**, *66*, 12-21.
46. Potterton, L.; McNicholas, S.; Krissinel, E.; Gruber, J.; Cowtan, K.; Emsley, P.; Murshudov, G. N.; Cohen, S.; Perrakis, A.; Noble, M., Developments in the CCP4 molecular-graphics project. *Acta Crystallogr. Section D: Biolog. Crystallogr.* **2004**, *60*, 2288-2294.
47. Evans, P. Scaling and assessment of data quality, *Acta Crystallogr. Sect D: Biol. Crystallogr.* **2006**, *62*, 72-82.
48. Diederichs, K.; Karplus, P. A. Improved R-factors for diffraction data analysis in macromolecular crystallography, *Nat Struct Biol.* **1997**, *4*, 269-275.
49. Weiss, M. S. Global indicators of X-ray data quality, *J. Appl. Crystallogr.* **2001**, *34*, 130-135.
50. Karplus, P. A.; Diederichs, K. Linking crystallographic model and data quality, *Science* **2012**, *336*, 1030-1033.
51. Evans, P. Biochemistry. Resolving some old problems in protein crystallography, *Science* **2012**, *336*, 986-987.

Table of content of graphic

