RSC Medicinal Chemistry

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: X. Jiang, J. Li, A. Viayna, F. J. Luque, M. Woodson, L. Jing, S. Gao, F. Zhao, M. Xie, K. Toth, J. Tavis, Ann. E. Tollefson, X. Liu and P. Zhan, *RSC Med. Chem.*, 2023, DOI: 10.1039/D3MD00306J.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/medchem

View Article Online
DOI: 10.1039/D3MD00306J

Identification of Novel 1,2,3-triazole Isatin Derivatives as Potent SARS-CoV-2 3CLpro Inhibitors via Click-chemistry-based Rapid Screening

Xiangyi Jiang[a], Jing Li[a], Antonio Viayna[b,c], F. Javier Luque[b,c,d], Molly Woodson[e,f], Lanlan Jing[a], Shenghua Gao[a], Fabao Zhao[a], Minghui Xie[a], Karoly Toth[e,f], John Tavis[e,f], Ann E. Tollefson[e,f,*], Xinyong Liu[a,*] and Peng Zhan[a,*]

[a]*Department of Medicinal Chemistry, Key Laboratory of Chemical Biology (Ministry of Education), School of Pharmaceutical Sciences, Cheeloo College of Medicine, Shandong University, 44 West Culture Road, 250012 Jinan, Shandong, PR China*

[b]*Departament de Nutrició, Ciències de l'Alimentació i Gastronomia, Facultat de Farmàcia i Ciències de l'Alimentació, Universitat de Barcelona (UB), Av. Prat de la Riba 171, 08921 Santa Coloma de Gramenet, Spain.*

[c]*Institut de Biomedicina (IBUB), Universitat de Barcelona (UB), Barcelona, Spain.*

[d]*Institut de Química Teòrica i Computacional (IQTCUB), Universitat de Barcelona (UB), Barcelona, Spain.*

[e]*Department of Molecular Microbiology and Immunology, Saint Louis University School of Medicine, St. Louis, Missouri 63104, United States.*

[f]*Saint Louis University Institute for Drug and Biotherapeutic Innovation, St. Louis, Missouri 63104, United States.*

[*]Corresponding authors.

ann.tollefson@health.slu.edu (Ann E. Tollefson)

xinyongl@sdu.edu.cn (Xinyong Liu);

zhanpeng1982@sdu.edu.cn (Peng Zhan).

View Article Online
DOI: 10.1039/D3MD00306J

ABSTRACT

The SARS-CoV-2 3-chymotrypsin-like protease ($3CL^{pro}$) is considered an attractive target for the development of anti-COVID-19 agents due to its vital function. The *N*-substituted isatin derivative L-26 is a potential SARS-CoV-2 $3CL^{pro}$ inhibitor, but it has poor cell-based antiviral activity and high cytotoxicity. With L-26 as the lead compound, 58 isatin derivatives were prepared using click-chemistry-based miniaturized synthesis and their $3CL^{pro}$ inhibitory activities was determined by a fluorescence resonance energy transfer-based enzymatic assay. Compounds **D1N8** (IC_{50} = 0.44 ± 0.12 μM) and **D1N52** (IC_{50} = 0.53 ± 0.21 μM) displayed excellent inhibitory potency against SARS-CoV-2 $3CL^{pro}$, being equivalent to that of L-26 (IC_{50} = 0.30 ± 0.14 μM). In addition, the cytotoxicity of **D1N8** (CC_{50} > 20 μM) and **D1N52** (CC_{50} > 20 μM) decreased significantly compared with L-26 (CC_{50} < 2.6 μM). Further molecular dynamics simulations revealed the potential binding interactions between **D1N52** and SARS-CoV-2 $3CL^{pro}$. These efforts lay a solid foundation for the research of novel anti-SARS-CoV-2 agents targeting $3CL^{pro}$.

Keywords: SARS-CoV-2, 3-chymotrypsin-like protease, isatin derivatives, click-chemistry.

View Article Online
DOI: 10.1039/D3MD00306J

1. Introduction

The coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), remains the most serious global infectious disease at this stage [1]. As a single-stranded RNA virus, the extreme mutagenicity of SARS-CoV-2 results in dramatically reduced preventive efficacy of existing vaccines [2-4]. Worse yet, the effective antiviral agents used to treat COVID-19 are fairly limited [5-7]. Therefore, there is an urgent need to develop potent anti-SARS-CoV-2 drugs to alleviate patient symptoms and mortality. The 3-chymotrypsin-like protease ($3CL^{pro}$) is a main protease (M^{pro}) that cleaves viral polyproteins into crucial non-structural proteins in cooperation with papain-like protease (PL^{pro}) [8]. $3CL^{pro}$ provides an indispensable role in viral replication, and it is considered as one of the most preferred targets for the development of broad-spectrum anti-SARS-CoV-2 agents due to its high conservation and lack of human homologs [9, 10].

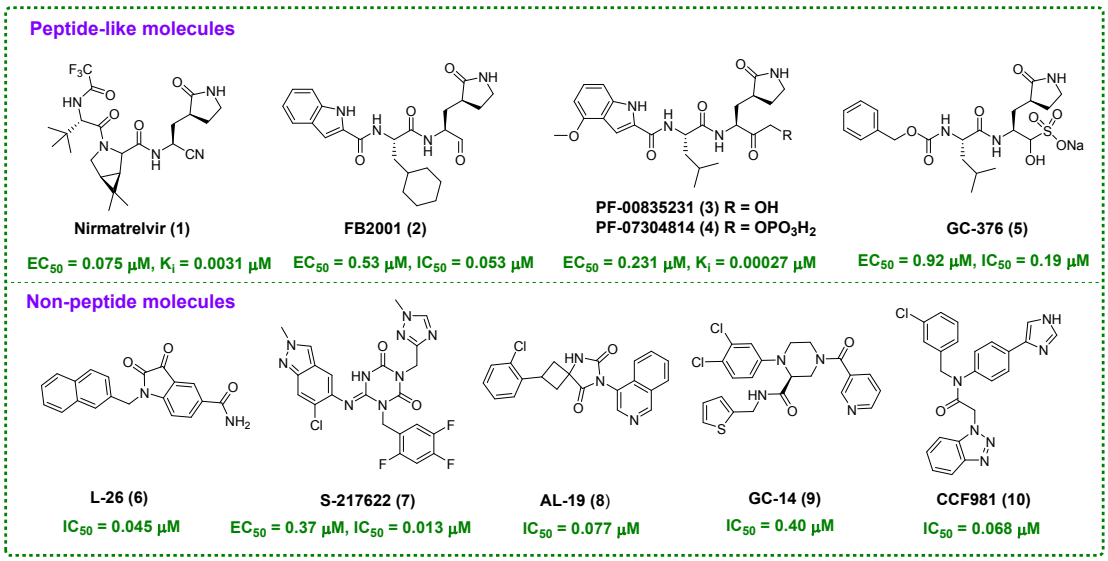
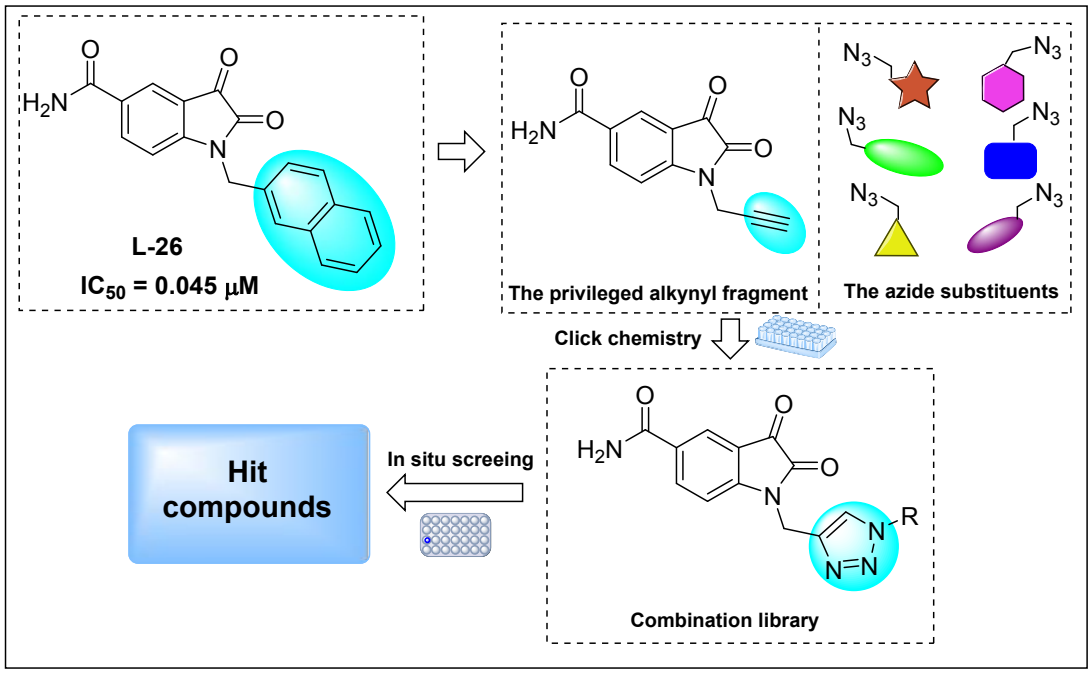


Fig.1 Representative peptide-like and non-peptide SARS-CoV-2 $3CL^{pro}$ inhibitors.

To date, a variety of chemotypes have been identified as SARS-CoV-2 $3CL^{pro}$

View Article Online
DOI: 10.1039/D3MD00306J

inhibitors and can be structurally divided into peptide-like and non-peptide molecules [11]. Peptidomimetics consist of a peptide-based scaffold mimicking natural substrates and a covalent warhead interacting with Cys145, such as nirmatrelvir (**1**), FB2001 (**2**), PF-00835231 (**3**), PF-07304814 (**4**) and GC-376 (**5**) [9, 12, 13]. However, the undesirable pharmacokinetic (PK) properties caused by the peptide-based scaffold severely limit their clinical applications. For example, nirmatrelvir, approved by the U.S. FDA, needs to be used in combination with ritonavir (cytochrome P450 enzymatic inhibitor) to improve metabolic stability *in vivo*[7]. Consequently, some potential non-peptide $3CL^{pro}$ inhibitors have attracted much attention, such as L-26 (**6**), S-217622 (**7**), AL-19 (**8**), GC-14 (**9**) and CCF981 (**10**) [8, 14-17]. Among these, the *N*-substituted isatin derivative L-26 is a promising non-covalent $3CL^{pro}$ inhibitor with an IC_{50} value of 0.045 μM [14]. However, the relatively high cytotoxicity interfered with the quantitative determination of the antiviral activity against SARS-CoV-2, indicating that further structural optimization of L-26 is necessary.



View Article Online
DOI: 10.1039/D3MD00306J

Fig.2 Design of novel 1,2,3-triazole isatin SARS-CoV-2 3CL^{pro} inhibitors *via* click-chemistry-based rapid screening

Unlike previous studies, the click-chemistry-based rapid screening technique was first applied to the identification of potent SARS-CoV-2 3CL^{pro} inhibitors in this work. The "click chemistry" is an efficacious fragment-based assembly reaction for the construction of high-quality compound libraries owing to its high yields, remarkable selectivity, gentle reaction conditions and exceptional functional group compatibility [18]. In particular, copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC), the most generic click chemical reaction, can rapidly prepare 1,2,3-triazole-derived compounds libraries [19]. It is widely known that 1,2,3-triazole is a preferred fragment in drug design with unique properties such as hydrogen bonding capability, positive solubility, rigidity and stability *in vivo* [20]. The previous structure-activity relationships (SARs) indicated that the amide group at the 5-position of isatin is the preferred pharmacophore for maintaining activity against SARS-CoV-2 3CL^{pro}, while the *N*-substituted part could well accommodate various groups [14]. Based on this, the methylene alkynyl was introduced into the 1-position of 2,3-dioxoindoline-5-carboxamide to prepare the privileged alkynyl fragment, that was reacted with azide substituents in 96-well plates using CuAAC click chemistry. A total of 58 triazole-containing isatin derivatives were obtained by miniaturized synthesis, and then were directly tested for inhibitory potency against SARS-CoV-2 3CL^{pro} without purification. Finally, compounds **D1N8** and **D1N52** were identified as potent SARS-CoV-2 3CL^{pro} inhibitors by further biological evaluation *in vitro* and

View Article Online
DOI: 10.1039/D3MD00306J

molecular dynamics (MD) simulation.

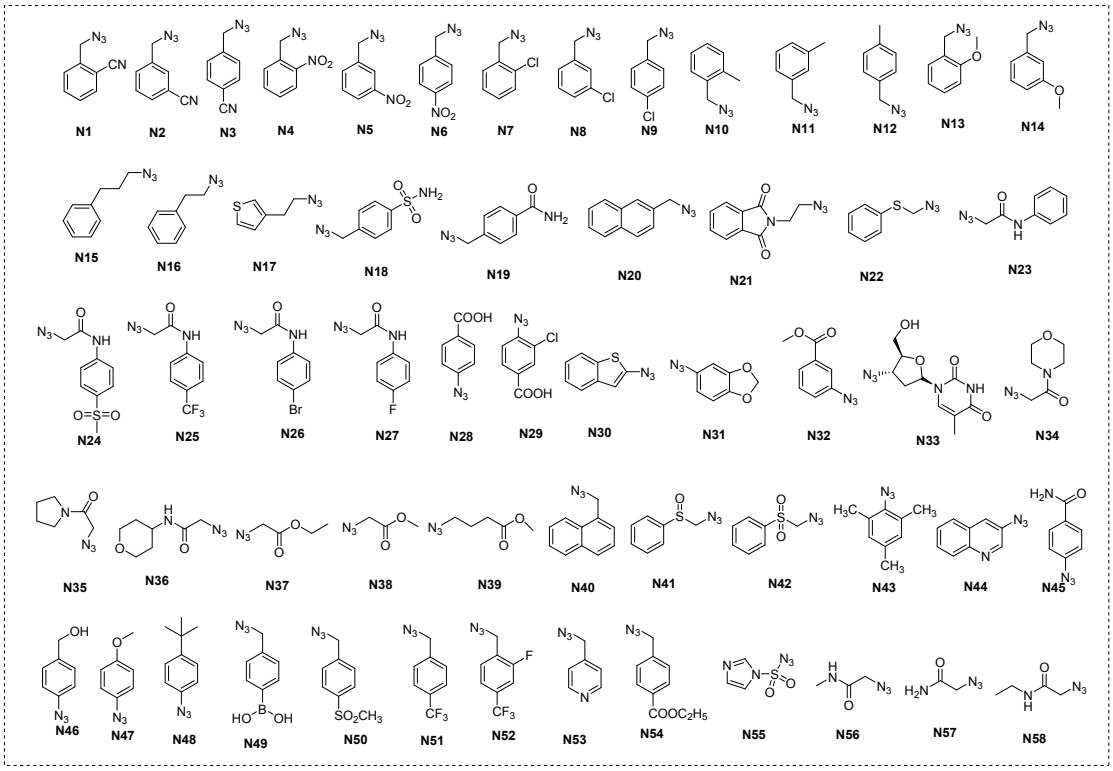
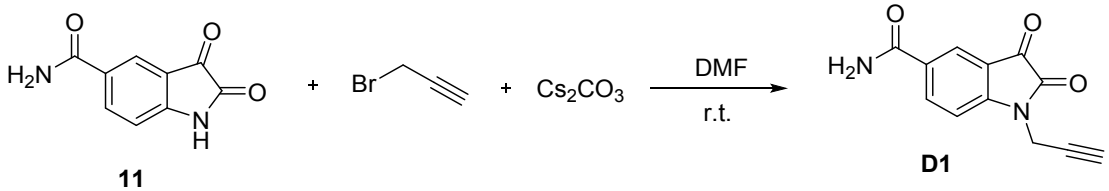


Fig 3. Structures of various azide substituents used in this study.

2. Results and discussion

2.1. *Generation of combinatorial library by click chemistry*

First, we synthesized the alkynyl fragment and a total of 58 azide substituents (Fig. 3). As shown in Scheme 1, the commercially available 2,3-dioxoindoline-5-carboxamide (**11**) was treated with 3-bromopropyne under the presence of cesium carbonate to prepare the privileged alkynyl fragment (**D1**) [21]. The synthesis of azide substituents **N1-N58** is described at length in our previous work [22].



Scheme 1. Synthetic route of **D1**.

View Article Online
DOI: 10.1039/D3MD00306J

A 58-member isatin derivatives library was built in microtiter plates *via* click-chemistry-based miniaturized synthesis. The various azide substituents **N1-58** (1.4 eq) was dissolved in dimethyl sulfoxide, and alkynyl fragment **D1** (1.0 eq) was added, followed by tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA, 0.2 eq), $CuSO_4 \cdot 5H_2O$ (0.2 eq) and sodium ascorbate (5.0 eq). The minute reaction conditions are shown in Table 1. The reactions were stirred on the constant temperature oscillation incubator at room temperature for 24 h, and then were monitored on thin-layer chromatography (TLC) using iodine vapors and UV light (λ = 254, 365 nm).

Table 1. The detailed reaction conditions.

Reagents	Concentration	Volume	Final Concentration
Azide-unit	35 mM DMSO	20 μL	7 mM
Alkyne-unit	25 mM DMSO	20 μL	5 mM
TBTA	10 mM DMSO	10 μL	20 mol%
$CuSO_4 \cdot 5H_2O$	4 mM MilliQ	25 μL	20 mol%
Sodium Ascorbate	20 mM MilliQ	25 μL	100 mol%

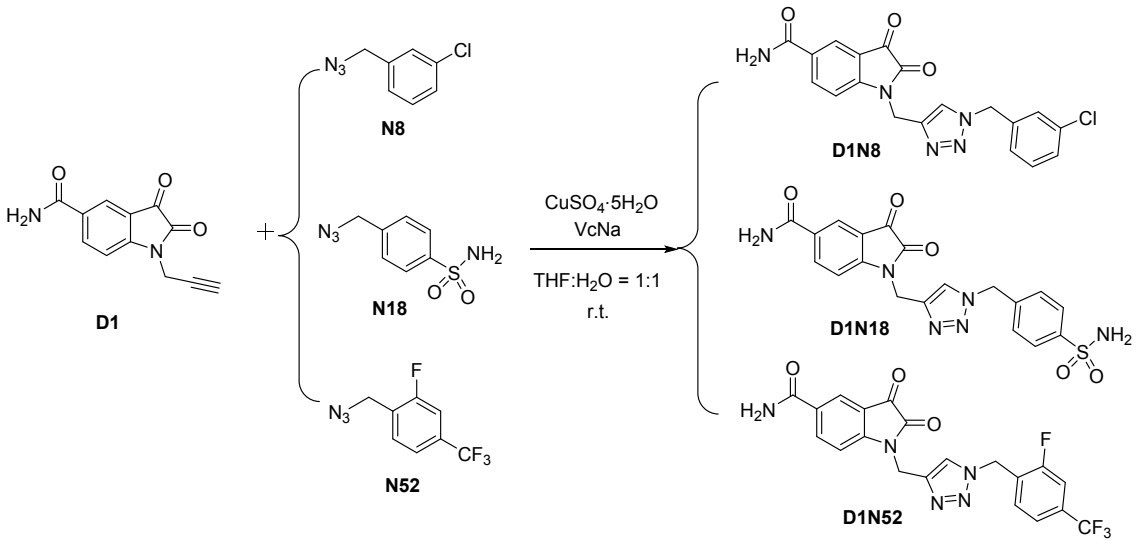
2.2. In situ screening against SARS-CoV-2 3CL^{pro}

The prepared compounds in the library were tested for their inhibitory potency against SARS-CoV-2 3CL^{pro} using a fluorescence resonance energy transfer (FRET)-based enzymatic assay in order to identify potent 3CL^{pro} inhibitors quickly, along with lead compound L-26 as the positive control. The biological results of tested compounds were expressed as inhibition percentage at a concentration of 1 μM.

View Article Online
DOI: 10.1039/D3MD00306J

As shown in Fig. 4, the target compounds showed from weak to excellent inhibitory potency against SARS-CoV-2 $3CL^{pro}$, indicating various azide substituents have an obvious impact on the activity. Notably, compounds **D1N8** (76.41%), **D1N18** (73.34%) and **D1N52** (78.12%) turned out to be the most active inhibitors, which were equivalent to lead compound L-26 (87.30%). Subsequently, they were synthesized (Scheme 2) on a milligram scale to quantitatively measure the $3CL^{pro}$ inhibitory activity.

Fig 4. SARS-CoV-2 $3CL^{pro}$ inhibitory activity in the presence of 1.0 μM **D1NX**. The blank is reaction buffer without azide substituent and alkynyl fragment.



Scheme 2. Synthesis of **D1N8**, **D1N18** and **D1N52**.

2.3. Biological evaluation

Inhibitory activity towards SARS-CoV-2 $3CL^{pro}$ under gradient concentrations (0.1 μM, 0.5 μM, 1.0 μM, 5.0 μM) was determined for the newly prepared compounds **D1N8**, **D1N18**, **D1N52** and the lead compound L-26. The 50% inhibition concentration (IC_{50}) of compounds are outlined in Table 2. Compounds **D1N8** bearing

View Article Online
DOI: 10.1039/D3MD00306J

3-chlorobenzyl and **D1N52** bearing 2-fluoro-4-(trifluoromethyl)benzyl showed outstanding anti-3CL^{pro} potency with the IC_{50} values of 0.44 ± 0.12 and 0.53 ± 0.21 μM, respectively, which are similar to L-26 (IC_{50} = 0.30 ± 0.14 μM), while **D1N18** (IC_{50} = 1.12 ± 0.54 μM) containing 4-sulfamoylbenzyl is slightly weaker than other compounds.

Table 2. Inhibitory activity against SARS-CoV-2 3CL^{pro}.

Compounds	D1N8	D1N18	D1N52	L-26
IC_{50} (μM) [a]	0.44 ± 0.12	1.12 ± 0.54	0.53 ± 0.21	0.30 ± 0.14

[a] IC_{50}: 50% concentration required to inhibit SARS-CoV-2 3CL^{pro}. The numbers in the table are the average of three independent tests.

Subsequently, the three compounds and L-26 were further evaluated for their anti-SARS-CoV-2 activity and cytotoxicity in Vero E6 cells. As shown in Table S1, the tested compounds did not display significant anti-SARS-CoV-2 activity under the concentration of 20 μM, as well as lead L-26. Still, three compounds (CC_{50} > 20 μM) displayed remarkable reduction in cytotoxicity compared with L-26 (CC_{50} < 2.2 μM). This study also confirms the conclusion that the drug discovery process is full of challenges and uncertainties, therefore a new round of iterative optimization is necessary [23].

2.4. Molecular dynamics (MD) simulation studies

Docking of compound **D1N52** led to similar poses in the binding pocket of both 7EN8 and 7V1T, where the isatin ring of **D1N52** partially overlaps the 2-(methylcarbamoyl)-4-bromanyl-6-nitro-phenyl moiety of the ligand bound to

View Article Online
DOI: 10.1039/D3MD00306J

3CL^{pro} in PDB ID 7EN8, and the 2-fluoro-4-trifluromethyl-benzene moiety is close to the hydrophobic area shaped by the methyl group of Thr25 and the side chain of Leu27. Compared with the docking performed using 7M8P as template, the main difference is the orientation of the 2-fluoro-4-trifluromethyl-benzene unit, which lies close to Thr190 and Ala191 due to the enlarged cavity generated by the shift of the Asp187-Ala193 region in 7M8P (Fig. S1).

Molecular Dynamics was used to examine the stability of the ligand (**D1N52**) pose (Fig. 5a). The root-mean square deviation of the protein backbone was stable along the trajectory, as noted in an average value of 1.9 Å. Regarding the ligand, a structural change occurred after release of the positional restraints used to avoid artefactual changes in the binding pose during the equilibration of the system (at around 25 ns; see section 4.5). Then, **D1N52** adopted a stable arrangement with a slight adjustment at around 130 ns, leading to a stable pose till the end of the MD simulation (Fig. 5b).

The binding mode of **D1N52** involved polar contacts between the isatin-bound amido group with the C=O and NH groups of Gln192, and between the amido NH_2 unit of Asn119 and the trifluoromethyl group of the terminal benzyl unit, as well as hydrophobic contacts between the isatin ring and the side chains of Met165 and Leu167, the benzyl ring and the side chains of Thr25 and Leu27, and the carbon atoms of the triazole ring, which stacks against the imidazole ring of His41, are located close to the side chain of Met49. It can be noticed that the fluorine atom bound to the 4-trifluoromethylbenzene ring is oriented toward Cys44 or Cys144. The

View Article Online
DOI: 10.1039/D3MD00306J

adoption of the two orientations is presumably facilitated by the lack of direct, specific interactions formed by the nitrogen atoms of the triazole ring in the binding pocket. In this regard, the presence of the 2-fluoro-4-(trifluoromethyl)benzyl unit is relevant to counterbalance the polarity of the triazole ring, as noted in a logP value of 2.5 estimated from QM (IEF-PCM/MST) continuum solvation calculations for the corresponding benzylated triazole fragment obtained upon exclusion of the isatin ring from **D1N52**. A similar logP value was estimated for the triazole fragment with chlorobenzyl (logP = 2.4), which may explain the similar inhibitory activity observed for **D1N52** and **D1N8** (IC_{50} = 0.52 μM and 0.44 μM, respectively; Table 2), whereas the presence of the sulfonamidobenzyl unit in **D1N18** led to a more polar triazole fragment (log P = -1.7), which can be expected to favor the solvation in water, thus contributing to the 2-fold reduction in inhibitory potency measured for **D1N18** (IC_{50} = 1.12 μM; Table 2).

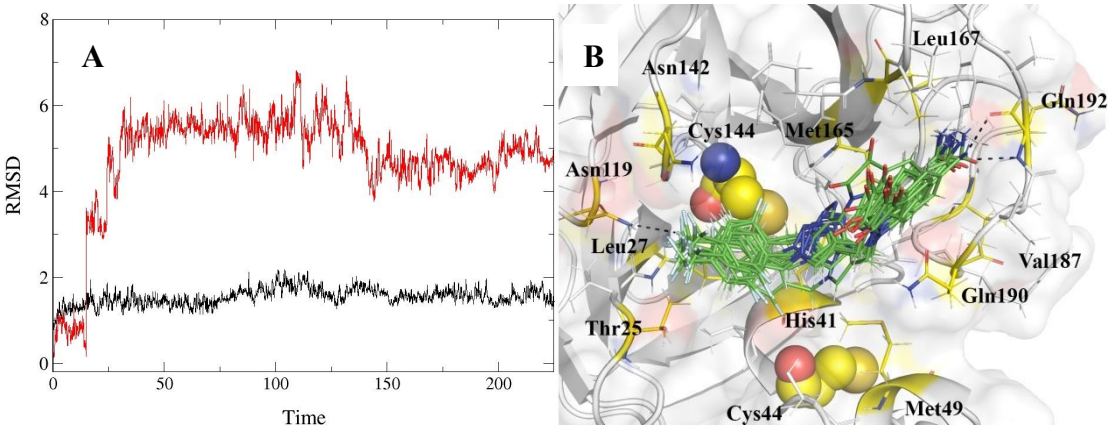


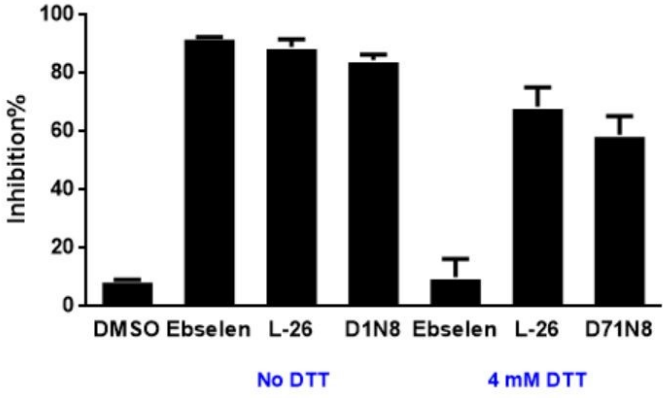
Fig 5. A) Time (ns) evolution of the positional root-mean square deviation (RMSD; Å) of the protein backbone (black) and ligand (**D1N52**; red) bound to the catalytic pocket of $3CL^{pro}$. The profile for the protein backbone was determined by excluding the atoms pertaining to the N-termini (residues 1-8) and loops 188-208 and 280-290 (numbering in the PDB structure 7EN8).

View Article Online
DOI: 10.1039/D3MD00306J

The RMSD was determined relative to the X-ray energy-minimized structure. B) Representation of the binding mode of **D1N52** in the catalytic pocket of 3CLpro (shown as gray surface). The ligands sampled every 5 ns along the last 50 ns of the unrestrained MD simulation are shown as sticks (carbon atoms colored in green). Selected residues in the binding pocket are highlighted as sticks, and Cys44 and Cys144 are shown as spheres (carbon atoms colored in yellow).

2.5. DTT-Dependent assay

To verify the target specificity of these isatin derivatives, compound **D1N8** was measured the anti-3CL^{pro} potency using FRET assay with or without of 4 mM DL-dithiothreitol (DTT) [24]. The non-specific cysteine protease inhibitor ebselen and lead L-26 were selected as controls. At a concentration of 5uM, all compounds exhibited significant inhibitory activity against 3CL^{pro} protease in the absence of DTT (Fig.6). However, ebselen almost completely lost anti-3CL^{pro} activity in the presence of 4 mM DTT. In contrast, isatin derivatives L-26 and **D1N8** maintained high inhibitory activity against 3CL^{pro} protease despite being slightly affected by 4 mM DTT (The reason is that DTT could directly react with isatin derivatives [25]). These results indicated that these isatin derivatives are likely to belong to specific SARS-CoV-2 3CL^{pro} protease inhibitors.



View Article Online
DOI: 10.1039/D3MD00306J

Fig 6. SARS-CoV-2 3CL^{pro} inhibitory activity of compounds (5 μM) in the absence or presence of 4 mM DTT.

3. Conclusion

Herein, taking the promising SARS-CoV-2 3CL^{pro} inhibitor L-26 as the lead compound, a 58-membered isatin derivatives library was constructed *via* click-chemistry-based miniaturized synthesis, and bioactivity against SARS-CoV-2 3CL^{pro} were determined by rapid screening. Among them, compounds **D1N8** (IC_{50} = 0.44 ± 0.12 μM) and **D1N52** (IC_{50} = 0.53 ± 0.21 μM) were identified as the most potent inhibitors towards SARS-CoV-2 3CL^{pro}, which is equivalent to the lead L-26 (IC_{50} = 0.30 ± 0.14 μM). In addition, the cytotoxicity of the tested compounds decreased significantly compared with L-26, but they did not have significant anti-SARS-CoV-2 activity under the concentration of 20 μM, as well as lead L-26. The MD simulation studies allowed us to propose a putative binding mode that balances hydrogen-bond and hydrophobic interactions, providing clues to understand the binding contribution of the different structural fragments present in the isatin derivatives. The DTT-dependent assay revealed that these isatin derivatives are likely to belong to specific 3CL^{pro} protease inhibitors, although slightly affected by reactions with DTT. In our future study, target-specific cellular assays such as the FlipGFP assay [26] will be conducted to further validate this hypothesis. All in all, these efforts provided valuable insights in seeking novel anti-SARS-CoV-2 agents targeting the 3CL^{pro}.

4. Experimental section

View Article Online
DOI: 10.1039/D3MD00306J

4.1. Chemistry

The melting point data of the prepared compounds was obtained on a micro melting point apparatus and were uncorrected. ^{1}H NMR and ^{13}C NMR spectra were recorded on a Bruker Avance-400 NMR spectrometer or a Bruker Avance-600 NMR spectrometer in DMSO-d_6 with TMS as an internal reference. Chemical shifts and *J* values were expressed in δ units (ppm) and in hertz (Hz), respectively. Mass spectra were measured *via* an electrospray ionization on an LC Autosampler Device: Standard G1313A instrument. The reactions were monitored on thin-layer chromatography (TLC) using iodine vapors and UV light ($\lambda = 254$, 365 nm). The mixtures were separated on column packed with Silica Gel 60 (200 - 300 mesh) by flash column chromatography. Reagent-grade solvents were applied, and purified *via* standard methods if necessary. Rotary evaporator (EYELA N-1300D) was used to concentrate the reaction solutions under reduced pressure condition.

2,3-dioxo-1-(prop-2-yn-1-yl)indoline-5-carboxamide (**D1**). A solution of 2,3-dioxoindoline-5-carboxamide (**11**, 0.1 g, 0.53 mmol) in 5 mL *N*,*N*-dimethylformamide was added to cesium carbonate (0.43 g, 1.33 mmol) and 3-bromopropyne (0.08 g, 0.64 mmol), and the mixture stirred at room temperature for 0.5 h. After monitored by TLC, the resulting mixture was treated with 20 mL water, and extracted with ethyl acetate (3 × 10 mL). Then, the organic phase was washed with saturated sodium bicarbonate (3 × 20 mL), dried over anhydrous sodium sulfate, filtered, concentrated under reduced pressure, and finally was purified by flash column chromatography (ethyl acetate: petroleum ether = 1:2) to afford intermediate

View Article Online
DOI: 10.1039/D3MD00306J

D1 as a yellow solid with a yield of 82.0%. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.25 (dd, J = 8.3, 1.8 Hz, 1H, Ph-H), 8.11 (d, J = 1.6 Hz, 1H, Ph-H), 8.09 (s, 1H, 1/2 NH_2), 7.44 (s, 1H, 1/2 NH_2), 7.32 (d, J = 8.3 Hz, 1H, Ph-H), 4.59 (s, 2H, CH_2), 3.37 (d, J = 2.4 Hz, 1H, alkynyl-H). ESI-MS: m/z 229.09 $[M+H]^+$, $C_{12}H_8N_2O_3$ [228.05].

General procedure for the synthesis of target compounds **D1N8**, **D1N18** and **D1N52**. The prepared compound **D1** (1.0 eq), azide substituents (**N8**, **N18** or **N52**, 1.2 eq), ascorbic acid sodium (0.6 eq) and $CuSO_4 \cdot 5H_2O$ (0.3 eq) were dissolved in the solution of tetrahydrofuran/water (8 mL, v:v = 1:1). The mixture was stirred at room temperature for 8 h, and then monitored by TLC. The reaction mixture was treated with 20 mL water, and extracted with ethyl acetate (3 × 10 mL). The combined organic phase was washed with saturated salt water (3 × 15 mL), dried over anhydrous sodium sulphate, filtered, and concentrated under reduced pressure to give crude products, which were purified by flash column chromatography and recrystallized using ethyl acetate and petroleum ether to afford corresponding target compounds **D1N8**, **D1N18** and **D1N52**.

*1-((1-(3-chlorobenzyl)-1H-1,2,3-triazol-4-yl)methyl)-2,3-dioxoindoline-5-carbox amide (**D1N8**)*. Yellow solid, yield: 14.5%, mp: 218-220℃. ^{1}H NMR (600 MHz, DMSO-d_6) δ 8.24 (s, 1H, triazoyl-H), 8.15 (dd, J = 8.3, 1.8 Hz, 1H, Ph-H), 8.08 (d, J = 1.6 Hz, 1H, Ph-H), 8.04 (s, 1H, 1/2 NH_2), 7.42 (s, 1H, 1/2 NH_2), 7.40 – 7.37 (m, 2H, Ph-H), 7.34 (s, 1H, Ph-H), 7.25 – 7.20 (m, 2H, Ph-H), 5.59 (s, 2H, CH_2), 5.00 (s, 2H, CH_2). ^{13}C NMR (100 MHz, DMSO-d_6) δ 183.01, 166.72, 158.71, 152.61, 142.22, 138.75, 137.84, 133.76, 131.15, 129.63, 128.62, 128.25, 127.09, 124.45, 123.75,

117.92, 111.27, 52.55, 35.74. ESI-MS: m/z 396.20 $[M+H]^+$, $C_{19}H_{14}ClN_5O_3$ [395.80].

2,3-dioxo-1-((1-(4-sulfamoylbenzyl)-1H-1,2,3-triazol-4-yl)methyl)indoline-5-carboxamide ***(D1N18)***. Yellow solid, yield: 12.0%, mp: 208-210℃. ^{1}H NMR (600 MHz, DMSO-d_6) δ 8.25 (s, 1H, triazoyl-H), 8.16 (d, J = 7.8 Hz, 1H, Ph-H), 8.08 (s, 1H, Ph-H), 8.05 (s, 1H, 1/2 NH_2), 7.80 (d, J = 7.7 Hz, 2H, Ph-H), 7.44 (d, J = 7.9 Hz, 3H, 1/2 NH_2 + 2 Ph-H), 7.36 (s, 2H, NH_2), 7.25 (d, J = 8.1 Hz, 1H), 5.66 (s, 2H, CH_2), 5.00 (s, 2H, CH_2). ^{13}C NMR (100 MHz, DMSO-d_6) δ 183.01, 166.73, 158.72, 152.61, 144.33, 142.25, 140.09, 137.84, 129.63, 128.89, 126.58, 124.43, 123.74, 117.94, 111.27, 52.73, 35.74. ESI-MS: m/z 441.4$[M+H]^+$, 463.3 $[M+Na]^+$, $C_{19}H_{16}N_6O_5S$ [440.43].

1-((1-(2-fluoro-4-(trifluoromethyl)benzyl)-1H-1,2,3-triazol-4-yl)methyl)-2,3-dioxoindoline-5-carboxamide ***(D1N52)***. Yellow solid, yield: 12.0%, mp: 320-322℃. ^{1}H NMR (600 MHz, DMSO-d_6) δ 8.25 (s, 1H, triazoyl-H), 8.15 (dd, J = 8.3, 1.9 Hz, 1H, Ph-H), 8.08 (d, J = 1.9 Hz, 1H, Ph-H), 8.03 (s, 1H, 1/2 NH_2), 7.74 (d, J = 9.7 Hz, 1H), 7.61 (d, J = 7.9 Hz, 1H, Ph-H), 7.49 (t, J = 7.6 Hz, 1H, Ph-H), 7.41 (s, 1H, 1/2 NH_2), 7.23 (d, J = 8.3 Hz, 1H, Ph-H), 5.73 (s, 2H, CH_2), 5.00 (s, 2H, CH_2). ^{13}C NMR (100 MHz, DMSO) δ 183.01, 166.70, 160.28 (d, J_{CF} = 249.8 Hz), 158.71, 152.59, 142.18, 137.82, 132.23 (d, J_{CF} = 3.9 Hz), 131.46 (q, J_{CF} = 10.3 Hz), 129.64, 128.02 (d, J_{CF} = 14.4 Hz), 124.67, 123.74, 122.45 (q, J_{CF} = 167.2 Hz), 122.27 (q, J_{CF} = 4.0 Hz), 117.93, 113.71 (dq, J_{CF} = 25.1, 3.7 Hz), 111.28, 47.12 (d, J_{CF} = 3.4 Hz), 35.70. ESI-MS: m/z 448.4 $[M+H]^+$, $C_{20}H_{13}F_4N_5O_3$ [447.35].

4.2. 3CL^{pro} Enzymatic Assay

View Article Online
DOI: 10.1039/D3MD00306J

The anti-3CL^{pro} activity of prepared compounds was evaluated using fluorescence resonance energy transfer (FRET) method with the fluorescent substrate MCAAVLQSGFR-Lys(Dnp)-Lys-NH_2 (Beyotime Biotechnology, Shanghai, China) [15]. The solution of 3CL^{pro} and fluorescent substrate were respectively diluted to a concentration of 1.5 μM and 500 μM with the assay buffer (50 mM Tris-HCl, 150 mM NaCl, 20% glycerol, pH = 7.3) and stored at -20°C. The tested compounds were dissolved in DMSO to prepare 5 mM stock solutions, and diluted to desired concentrations with assay buffer. The black 96-well plate was added with 50 μL of assay buffer, 20 μL of diluted enzyme solution, and 20 μL inhibitor solution were added, incubated for 10 minutes at 37℃, and then 10 μL of substrate solution was added to each well. The fluorescence signal, with excitation wavelength of 320 nm and emission wavelength of 405 nm, was detected using SpectraMax iD5 multimode plate reader (Molecular Devices) every 10 seconds, lasting for 10 minutes. The control wells were blank wells with no enzyme or inhibitor and wells with no inhibitor. Inhibition (i%) in each well can be calculated by the following formula: i% = 1- $V_i/V_0 \times 100\%$ (V_i for test well, and V_0 for control well). All tested compounds was measured for the inhibition rate at the concentration of 1 μM in preliminary screening, and the hit compounds was measured at 4 different concentrations (0.1 μM, 0.5 μM, 1 μM, and 5 μM, three independent experiments in each concentration) to determine the IC_{50} values (three independent experiments were performed).

4.3. Cellular Anti-SARS-CoV-2 Activity Assay

Vero E6 cells (from Prof. Ann E. Tollefson's laboratory, Department of

View Article Online
DOI: 10.1039/D3MD00306J

Molecular Microbiology and Immunology, Saint Louis University School of Medicine, United States.) were plated at 1.7×10^4 cells/well into 96-well plates about 24 hours prior to infection. Test compounds were prepared to a 10 mM stock solution in DMSO. The test compounds were diluted to 80 μM in Dulbecco's MEM/5% FBS (fetal bovine serum) and then 1:3 dilutions were done on a dilution plate (each compound was done in duplicate), to form a dilution series. Medium was removed from the cells and 50 μL of the compounds dilutions were added to each well of the cell plate. The plates were then transferred to the BSL-3 facility for infection with SARS-CoV-2. Virus was diluted to approximately 100 focus forming units (infectious particles) per 50 μL (the volume of diluted virus added per well). After the addition of virus to the wells, the plates were incubated at 37℃, 5% CO_2 for 1 hour. After the incubation, medium containing methyl cellulose was added to each well (100 μL). The total final volume in each well was 200 μL. The final compound concentration range was 20 μM to 0.0274 μM. After the plates were incubated for 25-26 hours, the methyl cellulose overlay was removed and plates were washed twice with PBS. 5% paraformaldehyde in PBS was then added to each well and the entire plates were subsequently submerged in 5% formalin for 15 minutes. Plates were then submerged in PBS for 15 minutes and then fresh PBS was added to each well and plates were transferred out of the BSL-3 laboratory. Plates carrying immobilized Vero E6 cells were rinsed with focus forming assay buffer and incubated with a guinea pig anti-SARS-CoV antibody generated against virions that cross-reacts with SARS-CoV-2. The secondary antibody was goat anti-guinea pig IgG (horseradish

View Article Online
DOI: 10.1039/D3MD00306J

peroxidase conjugate). The substrate was TrueBlue and the infected cells/foci appeared as individual or clusters of blue cells. After staining, plates were read on an EliSpot reader. The relative surface area that was stained was used to calculate inhibition activity of the concentration gradient of drug, and to generate the EC_{50} curves. EC_{50} values were generated by curve fitting using Graphpad Prism 8.0.

4.4. Cellular Toxicity Assay

Vero E6 Cells (from Prof. Ann E. Tollefson's laboratory, Department of Molecular Microbiology and Immunology, Saint Louis University School of Medicine, United States.) were plated on 96-well plates (Greiner) at 1.7×10^4 cells per well in DMEM (Dulbecco's Modified Eagle Medium) high glucose with 10% FBS. Cells were incubated for 24 h in 37°C and 5% CO_2 at saturating humidity. Cells were treated with compound in DMEM high glucose with 5% FBS at a final concentration of 1% DMSO for 26 h. Cellular viability was measured using the CellTiter 96 aqueous nonradioactive cell proliferation assay (Promega). Three or more replicate assays were done to calculate the average values. The 50% cytotoxic concentration (CC_{50}) values were determined with GraphPad Prism 8.0 using the four-parameter variable slope algorithm with the bottom set to zero.

4.5. Molecular Dynamics Simulation

A combined strategy involving molecular docking and molecular dynamics simulations was used to investigate the binding mode of isatin derivatives to SARS-CoV-2 $3CL^{pro}$. For the docking stage calculations were performed considering X-ray structures deposited in the Protein Data Bank [27] with PDB codes 7EN8, 7M8P

View Article Online
DOI: 10.1039/D3MD00306J

and 7V1T [28-30] to take into account i) the non-covalent inhibitory binding mode of the enzyme-bound ligands, ii) the distinct arrangement of the side chains of residues Met49 and Gln189, and iii) the different conformation adopted by the flexible loop defined by residues Asp187-Ala193. It is worth noting that the binding mode of the ligands bound to these structures fills the distinct subpockets that can be identified in the binding site (see Supplementary Information Fig. S2). Docking was accomplished using the XP score function of Glide [31-33] in conjunction with a defined inner/outer box of 40/76 Å centered on the catalytic pocket of $3CL^{pro}$.

Molecular Dynamics (MD) simulations were performed using the Amber20 package [34] to explore the structural stability of **D1N52** bound to $3CL^{pro}$. The ff19SB protein force-field [35] was used for the enzyme, while the isatin derivative was parametrized using the gaff2 force field [36,37] considering the RESP atomic charges [38] determined by fitting the HF/6-31G(d) electrostatic potential. The enzyme-ligand complex was solvated with the OPC water model [39], and counterions were used to maintain the system neutrality and the physiological ionic atmosphere. Finally, production calculations were performed using the SHAKE algorithm [40] and treating electrostatic interactions with Particle Mesh Ewald (PME) [41] with a cutoff of 10 Å for non-bonded interactions. Due to the flexibility of $3CL^{pro}$ binding pocket, the restraints were maintained along the first 25 ns, then reduced to 2 kcal mol^{-1} A^{-2} in the next 5 ns, and finally an unrestrained MD simulation (200 ns) was run under constant volume and temperature.

Quantum mechanical (QM) continuum solvation calculations were performed to

View Article Online
DOI: 10.1039/D3MD00306J

estimate the hydrophobicity of the isatin derivatives. To this end, the n-cotanol/water partition coefficient (logP) was calculated using the IEF-PCM/MST method [42-44] as implemented in a local version of Gaussian 16 [45]. Molecular geometries were optimized at the B3LYP/6-31G(d) level of theory [46-48] considering the solvation effects in both water and *n*-octanol. Finally, the logP was estimated by considering the solvation free energies in water and *n*-octanol.

4.6. DTT-Dependent assay

The method is similar to $3CL^{pro}$ enzymatic assay in Section 4.2. The experimental group added 4 mM DTT to the assay buffer, while the control group did not add DTT to the assay buffer.

Acknowledgments

The authors gratefully acknowledge financial support from Major Basic Research Project of Shandong Provincial Natural Science Foundation (no. ZR2021ZD17), Guangdong Basic and Applied Basic Research Foundation (no. 2021A1515110740), China Postdoctoral Science Foundation (no. 2021M702003), Science Foundation for Outstanding Young Scholars of Shandong Province (no. ZR2020JQ31), and Foreign Cultural and Educational Experts Project (no. GXL20200015001). A.V. and F.J.L. acknowledge financial support from the Spanish Ministerio de Ciencia e Innovación (AEI/10.13039/501100011033; grants PID2020-117646RB-I00 and CEX2021-001202-M), the Generalitat de Catalunya (grant 2021SGR00671). The Consorci de Serveis Universitaris de Catalunya (CSUC) is acknowledged for

View Article Online
DOI: 10.1039/D3MD00306J

computational resources (Molecular Recognition project).

Conflict of interest

The authors declare no conflict of interest.

References

1. P. Zhou, X.L. Yang, X.G. Wang, B. Hu, L. Zhang, W. Zhang, H.R. Si, Y. Zhu, B. Li, C.L. Huang, H.D. Chen, J. Chen, Y. Luo, H. Guo, R.D. Jiang, M.Q. Liu, Y. Chen, X.R. Shen, X. Wang, X.S. Zheng, K. Zhao, Q.J. Chen, F. Deng, L.L. Liu, B. Yan, F.X. Zhan, Y.Y. Wang, G.F. Xiao, Z.L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin, *Nature*, 2020, **579** 270-273.
2. J. Chen, R. Wang, N.B. Gilby, G.W. Wei, Omicron Variant (B.1.1.529): Infectivity, Vaccine Breakthrough, and Antibody Resistance, *J Chem Inf Model*, 2022, **62**, 412-422.
3. T. Fiolet, Y. Kherabi, C.J. MacDonald, J. Ghosn, N. Peiffer-Smadja, Comparing COVID-19 vaccines for their characteristics, efficacy and effectiveness against SARS-CoV-2 and variants of concern: a narrative review, *Clin Microbiol Infect*, 2022, **28**, 202-221.
4. W.T. Harvey, A.M. Carabelli, B. Jackson, R.K. Gupta, E.C. Thomson, E.M. Harrison, C. Ludden, R. Reeve, A. Rambaut, C.-G.U. Consortium, S.J. Peacock, D.L. Robertson, SARS-CoV-2 variants, spike mutations and immune escape, *Nat Rev Microbiol*, 2021, **19**, 409-424.
5. W. Wen, C. Chen, J.K. Tang, C.Y. Wang, M.Y. Zhou, Y.R. Cheng, X. Zhou, Q. Wu, X.W. Zhang, Z.H. Feng, M.W. Wang, Q. Mao, Efficacy and safety of three new oral antiviral treatment (molnupiravir, fluvoxamine and Paxlovid) for COVID-19: a meta-analysis, *Ann Med*, 2022, **54**,

View Article Online
DOI: 10.1039/D3MD00306J

516-523.

6. S.M.R. Hashemian, M.H. Pourhanifeh, M.R. Hamblin, M.K. Shahrzad, H. Mirzaei, RdRp inhibitors and COVID-19: Is molnupiravir a good option?, *Biomedicine & pharmacotherapy*, 146 (2022) 112517.

7. Y.N. Lamb, Nirmatrelvir Plus Ritonavir: First Approval, *Drugs*, 2022, **82**, 585-591.

8. S.H. Han, C.M. Goins, T. Arya, W.J. Shin, J. Maw, A. Hooper, D.P. Sonawane, M.R. Porter, B.E. Bannister, R.D. Crouch, A.A. Lindsey, G. Lakatos, S.R. Martinez, J. Alvarado, W.S. Akers, N.S. Wang, J.U. Jung, J.D. Macdonald, S.R. Stauffer, Structure-Based Optimization of ML300-Derived, Noncovalent Inhibitors Targeting the Severe Acute Respiratory Syndrome Coronavirus 3CL Protease (SARS-CoV-2 3CL(pro)), *J Med Chem*, 2022, **65**, 2880-2904.

9. S. Gao, T. Huang, L. Song, S. Xu, Y. Cheng, S. Cherukupalli, D. Kang, T. Zhao, L. Sun, J. Zhang, P. Zhan, X. Liu, Medicinal chemistry strategies towards the development of effective SARS-CoV-2 inhibitors. *Acta Pharm Sin B*. 2022, 12, 581-599.

10. R. Banerjee, L. Perera, L.M.V. Tillekeratne, Potential SARS-CoV-2 main protease inhibitors, *Drug Discov Today*, 2021, **26**, 804-816.

11. L. Agost-Beltran, S. de la Hoz-Rodriguez, L. Bou-Iserte, S. Rodriguez, A. Fernandez-de-la-Pradilla, F.V. Gonzalez, Advances in the Development of SARS-CoV-2 Mpro Inhibitors, *Molecules*, 2022, **27**, 2523.

12. W. Vuong, M.B. Khan, C. Fischer, E. Arutyunova, T. Lamer, J. Shields, H.A. Saffran, R.T. McKay, M.J. van Belkum, M.A. Joyce, H.S. Young, D.L. Tyrrell, J.C. Vederas, M.J. Lemieux, Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication, *Nat Commun*, 2020, **11**, 4282.

View Article Online
DOI: 10.1039/D3MD00306J

13. D.R. Owen, C.M.N. Allerton, A.S. Anderson, L. Aschenbrenner, M. Avery, S. Berritt, B. Boras, R.D. Cardin, A. Carlo, K.J. Coffman, A. Dantonio, L. Di, H. Eng, R. Ferre, K.S. Gajiwala, S.A. Gibson, S.E. Greasley, B.L. Hurst, E.P. Kadar, A.S. Kalgutkar, J.C. Lee, J. Lee, W. Liu, S.W. Mason, S. Noell, J.J. Novak, R.S. Obach, K. Ogilvie, N.C. Patel, M. Pettersson, D.K. Rai, M.R. Reese, M.F. Sammons, J.G. Sathish, R.S.P. Singh, C.M. Steppan, A.E. Stewart, J.B. Tuttle, L. Updyke, P.R. Verhoest, L. Wei, Q. Yang, Y. Zhu, An oral SARS-CoV-2 M(pro) inhibitor clinical candidate for the treatment of COVID-19, *Science*, 2021, **374**, 1586-1593.

14. P. Liu, H. Liu, Q. Sun, H. Liang, C. Li, X. Deng, Y. Liu, L. Lai, Potent inhibitors of SARS-CoV-2 3C-like protease derived from N-substituted isatin compounds, *Eur J Med Chem*, 206 (2020) 112702.

15. S.H. Gao, K. Sylvester, L.T. Song, T. Claff, L.L. Jing, M. Woodson, R.H. Weisse, Y.S. Cheng, L. Schakel, M. Petry, M. Gutschow, A.C. Schiedel, N. Strater, D.W. Kang, S.J. Xu, K. Toth, J. Tavis, A.E. Tollefson, C.E. Muller, X.Y. Liu, P. Zhan, Discovery and Crystallographic Studies of Trisubstituted Piperazine Derivatives as Non-Covalent SARS-CoV-2 Main Protease Inhibitors with High Target Specificity and Low Toxicity, *J Med Chem*, 2022, **65**, 13343-13364.

16. Y. Unoh, S. Uehara, K. Nakahara, H. Nobori, Y. Yamatsu, S. Yamamoto, Y. Maruyama, Y. Taoda, K. Kasamatsu, T. Suto, K. Kouki, A. Nakahashi, S. Kawashima, T. Sanaki, S. Toba, K. Uemura, T. Mizutare, S. Ando, M. Sasaki, Y. Orba, H. Sawa, A. Sato, T. Sato, T. Kato, Y. Tachibana, Discovery of S-217622, a Noncovalent Oral SARS-CoV-2 3CL Protease Inhibitor Clinical Candidate for Treating COVID-19, *J Med Chem*, 2022, **65** 6499-6512.

17. A. Luttens, H. Gullberg, E. Abdurakhmanov, D. D. Vo, D. Akaberi, V. O. Talibov, N. Nekhotiaeva, L. Vangeel, S. D. Jonghe, D. Jochmans, J. Krambrich, A. Tas, B. Lundgren, Y.

View Article Online
DOI: 10.1039/D3MD00306J

Gravenfors, A. J. Craig, Y. Atilaw, A. Sandström, L. W. Moodie, Å. Lundkvist, M. J Hemert, J. Neyts, J. Lennerstrand, J. Kihlberg, K. Sandberg, U. H. Danielson, J. Carlsson, Ultralarge Virtual Screening Identifies SARS-CoV-2 Main Protease Inhibitors with Broad-Spectrum Activity against Coronaviruses, *J Am Chem Soc*. 2022, 144, 2905-2920.

18. X. Jiang, X. Hao, L. Jing, G. Wu, D. Kang, X. Liu, P. Zhan, Recent applications of click chemistry in drug discovery, *Expert Opin Drug Discov*, 2019, **14**, 779-789.

19. D.W. Kang, D. Feng, L.L. Jing, Y.Y. Sun, F.J. Wei, X.Y. Jiang, G.C. Wu, E. De Clercq, C. Pannecouque, P. Zhan, X.Y. Liu, In situ click chemistry-based rapid discovery of novel HIV-1 NNRTIs by exploiting the hydrophobic channel and tolerant regions of NNIBP, *Eur J Med Chem*, 2020, **193**, 112237.

20. X.Y. Jiang, G.C. Wu, W.A. Zalloum, M.E. Meuser, A. Dick, L. Sun, C.H. Chen, D.W. Kang, L.L. Jing, R.F. Jia, S. Cocklin, K.H. Lee, X.Y. Liu, P. Zhan, Discovery of novel 1,4-disubstituted 1,2,3-triazole phenylalanine derivatives as HIV-1 capsid inhibitors, *Rsc Adv*, 2019, **9**, 28961-28986.

21. L. Zhou, Y. Liu, W.L. Zhang, P. Wei, C.K. Huang, J.F. Pei, Y.X. Yuan, L.H. Lai, Isatin compounds as noncovalent SARS coronavirus 3C-like protease inhibitors, *J Med Chem*, 2006, **49**, 3440-3443.

22. L. Jing, G. Wu, X. Hao, F.A. Olotu, D. Kang, C.H. Chen, K.H. Lee, M.E.S. Soliman, X. Liu, Y. Song, P. Zhan, Identification of highly potent and selective Cdc25 protein phosphatases inhibitors from miniaturization click-chemistry-based combinatorial libraries, *Eur J Med Chem*, 2019, **183**, 111696.

23. H Li, W Wei, H. Xu, Drug discovery is an eternal challenge for the biomedical sciences. *Acta*

View Article Online
DOI: 10.1039/D3MD00306J

Materia Medica. 2022, **1**, 1-3.

24. C. Ma, Y. Hu, J. A. Townsend, P. I. Lagarias, M. T. Marty, A. Kolocouris, J. Wang, Ebselen, Disulfiram, Carmofur, PX-12, Tideglusib, and Shikonin Are Nonspecific Promiscuous SARS-CoV-2 Main Protease Inhibitors. *ACS Pharmacol Transl Sci*. 2020, **3**, 1265-1277.
25. L. Zhou, Y. Liu, W. Zhang, P. Wei, C. Huang, J. Pei, Y. Yuan, L. Lai, Isatin compounds as noncovalent SARS coronavirus 3C-like protease inhibitors, *J Med Chem*. 2006, **49**, 3440-3.
26. Z. Xia, M. Sacco, Y. Hu, C. Ma, X. Meng, F. Zhang, T. Szeto, Y. Xiang, Y. Chen, J. Wang, Rational Design of Hybrid SARS-CoV-2 Main Protease Inhibitors Guided by the Superimposed Cocrystal Structures with the Peptidomimetic Inhibitors GC-376, Telaprevir, and Boceprevir. *ACS Pharmacol Transl Sci*. 2021, **4**, 1408-1421.
27. H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, *Nucleic Acids Res*, 2000, **28**, 235-242.
28. N.K. Hou, L. Shuai, L.J. Zhang, X.P. Xie, K.M. Tang, Y.K. Zhu, Y. Yu, W.Y. Zhang, Q.Z. Tan, G.X. Zhong, Z.Y. Wen, C. Wang, X.J. He, H. Huo, H.S. Gao, Y. Xu, J. Xue, C. Peng, J. Zou, C. Schindewolf, V. Menachery, W.J. Su, Y.L. Yuan, Z.Y. Shen, R. Zhang, S.F. Yuan, H.T. Yu, P.Y. Shi, Z.G. Bu, J. Huang, Q. Hu, Development of Highly Potent Noncovalent Inhibitors of SARS-CoV-2 3CLpro, *Acs Central Sci*, 2023, **9**, 217–227.
29. M.G. Deshmukh, J.A. Ippolito, C.H. Zhang, E.A. Stone, R.A. Reilly, S.J. Miller, W.L. Jorgensen, K.S. Anderson, Structure-guided design of a perampanel-derived pharmacophore targeting the SARS-CoV-2 main protease, *Structure*, 2021, **29**, 823.
30. https://doi.org/10.2210/pdb7V1T/pdb
31. R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky,

View Article Online
DOI: 10.1039/D3MD00306J

E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, P.S. Shenkin, Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, *J Med Chem*, 2004, **47**, 1739-1749.

32. T.A. Halgren, R.B. Murphy, R.A. Friesner, H.S. Beard, L.L. Frye, W.T. Pollard, J.L. Banks, Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening, *J Med Chem*, 2004, **47**, 1750-1759.

33. R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, *J Med Chem*, 2006, **49**, 6177-6196.

34. Case DA, Belfon K, Ben-Shalom IY. AMBER 2020, University of California, San Francisco.

35. C. Tian, K. Kasavajhala, K.A.A. Belfon, L. Raguette, H. Huang, A.N. Migues, J. Bickel, Y. Wang, J. Pincay, Q. Wu, C. Simmerling, ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution, *J Chem Theory Comput*, 2020, **16**, 528-552.

36. J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, *J Comput Chem*, 2004, **25**, 1157-1174.

37. J. Wang, W. Wang, P.A. Kollman, D.A. Case, Automatic atom type and bond type perception in molecular mechanical calculations, *J Mol Graph Model*, 2006, **25**, 247-260.

38. C.I. Bayly, P. Cieplak, W.D. Cornell, P.A. Kollman, A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model, *J Phys Chem-Us*, 1993, **97**, 10269-10280.

39. S. Izadi, R. Anandakrishnan, A.V. Onufriev, Building Water Models: A Different Approach,

View Article Online
DOI: 10.1039/D3MD00306J

Journal of Physical Chemistry Letters, 2014, **5**, 3863-3871.

40. J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J Comput Phys*, 1977, **23**, 327–341.
41. T. Darden, D. York, L. Pedersen, Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems, *J Chem Phys*, 1993, **98**, 10089-10092.
42. C. Curutchet, M. Orozco, F.J. Luque, Solvation in octanol: Parametrization of the continuum MST model, *J Comput Chem*, 2001, **22**, 1180-1193.
43. C. Curutchet, A. Bidon-Chanal, I. Soteras, M. Orozco, F.J. Luque, MST continuum study of the hydration free energies of monovalent ionic species, *J Phys Chem B*, 2005, **109**, 3565-3574.
44. I. Soteras, C. Curutchet, A. Bidon-Chanal, M. Orozco, F.J. Luque, Extension of the MST model to the IEF formalism: HF and B3LYP parametrizations, *J Mol Struc-Theochem*, 2005, **727**, 29-40.
45. M.J. Frisch, G.W. Trucks, H.B. Schlegel HB, Gaussian 16, Revision C.01, Gaussian, Inc., Wallingford CT, 2016.
46. C.T. Lee, W.T. Yang, R.G. Parr, Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density, *Phys Rev B*, 1988, **37**, 785-789.
47. K. Raghavachari, Perspective on "Density functional thermochemistry. III. The role of exact exchange" - Becke AD (1993) J Chem Phys 98:5648-52, *Theor Chem Acc*, 2000, **103**, 361-363.
48. P.J. Stephens, F.J. Devlin, C.F. Chabalowski, M.J. Frisch, Ab-Initio Calculation of Vibrational Absorption and Circular-Dichroism Spectra Using Density-Functional Force-Fields, *J Phys Chem-Us*, 1994, **98**, 11623-11627.

View Article Online
DOI: 10.1039/D3MD00306J