

## Article

# Search of Novel Small Molecule Inhibitors for the Main Protease of SARS-CoV-2

Wenfa Zhang and Sheng-Xiang Lin \* 

Axe Molecular Endocrinology and Nephrology, CHU Research Center and Laval University, Québec City, QC G1 V 4G2, Canada

\* Correspondence: sxlin@crchul.ulaval.ca

**Abstract:** The current outbreak of coronavirus disease 2019 (COVID-19) has prompted the necessity of efficient treatment strategies. The COVID-19 pandemic was caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Main protease (Mpro), also called 3-chymotrypsin-like protease (3CL protease), plays an essential role in cleaving virus polyproteins for the functional replication complex. Therefore, Mpro is a promising drug target for COVID-19 therapy. Through molecular modelling, docking and a protease activity assay, we found four novel inhibitors targeting Mpro with the half maximal inhibitory concentration (IC<sub>50</sub>) and their binding affinities shown by the dissociation constants (K<sub>D</sub>s). Our new inhibitors CB-21, CB-25, CP-1 and LC24-20 have IC<sub>50</sub>s at 14.88 μM (95% Confidence Interval (95% CI): 10.35 μM to 20.48 μM), 22.74 μM (95% CI: 13.01 μM to 38.16 μM), 18.54 μM (95% CI: 6.54 μM to 36.30 μM) and 32.87 μM (95% CI: 18.37 μM to 54.80 μM), respectively. The evaluation of interactions suggested that each inhibitor has a hydrogen bond or hydrophobic interactions with important residues, including the most essential catalytic residues: His<sup>41</sup> and Cys<sup>145</sup>. All the four inhibitors have a much higher 50% lethal dose (LD<sub>50</sub>) compared with the well-known Mpro inhibitor GC376, demonstrating its low toxicity. These four inhibitors can be potential drug candidates for further in vitro and in vivo studies against COVID-19.

**Keywords:** SARS-CoV-2 main protease; molecular docking; protease kinetics inhibition; ligand/protein interaction; toxicity; inhibitor efficacy



**Citation:** Zhang, W.; Lin, S.-X. Search of Novel Small Molecule Inhibitors for the Main Protease of SARS-CoV-2. *Viruses* **2023**, *15*, 580. <https://doi.org/10.3390/v15020580>

Academic Editor: Graciela Andrei

Received: 1 January 2023

Revised: 13 February 2023

Accepted: 17 February 2023

Published: 20 February 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The recent outbreak of COVID-19 caused by SARS-CoV-2 has rapidly become a pandemic since its first report in December, 2019 [1]. More than 640 million total cases have been reported, with 6.6 million deaths worldwide by 25 November 2022 (Johns Hopkins Coronavirus Resource Center, <https://coronavirus.jhu.edu/map.html> (accessed on 25 November 2022)).

SARS-CoV-2 is a single strand positive-sense RNA virus that is highly contagious in humans [2]. It has 10 open reading frames (ORFs): ORF1a, ORF1b, spike protein (S), ORF3a, Envelope protein (E), Membrane (M), ORF6, ORF7a, ORF7b, ORF8, Nucleocapsid (N) and ORF10 [3]. Structural proteins S, E and M together can create the viral envelope, and N is used to hold the RNA genome. ORF1a and ORF1b embed 16 non-structural proteins (nsp1 to nsp16) including the RNA-dependent RNA polymerase (RdRp) used for RNA genome replication and Mpro, and papain-like protease (PLpro) used for polyprotein cleavage, to generate a functional active viral replication complex [4]. The Mpro corresponds to nsp5. PLpro can cleave three conserved sites for functional nsp1, nsp2 and nsp3 and Mpro can cleave the eleven conserved sites for functional nsp4 to nsp16 [5]. Therefore, Mpro is a promising target for drug design.

Several inhibitors targeting SARS-CoV-2 Mpro have been tested and reported to have a low IC<sub>50</sub> or EC<sub>50</sub>. For example, 11a, 11b, 11r, 13b, or Ebselen have a tested IC<sub>50</sub> lower than 1 μM and 13a; Disulfiram, Tideglusib and Carmofur have IC<sub>50</sub>s between 1 μM to

10  $\mu\text{M}$ , Shikonin has an  $\text{IC}_{50}$  at 15.75  $\mu\text{M}$ , PX-12 has an  $\text{IC}_{50}$  at 21.39  $\mu\text{M}$  and Cinanserin has an  $\text{IC}_{50}$  at 124  $\mu\text{M}$  [6–8]. All these inhibitors could be candidate drugs through pre-clinical and clinical tests. According to the data from [ClinicalTrials.gov](https://clinicaltrials.gov) in the USA, several drug candidates targeting Mpro are under evaluation for clinical trial: PBI-0451, Ebselen, Ritonavir, S-217622, Nirmatrelvir, Lopinavir, Apixaban and Azithromycin. However, no drugs have yet been approved by the Food and Drug Administration (FDA). Therefore, searching for a new drug candidate for high efficacy and low toxicity is still necessary.

In this study, we found four novel Mpro inhibitors through docking and validated by in vitro experiments. All the four inhibitors have a relatively high bioavailability score and low toxicity. They thus have low  $\text{IC}_{50}$ s and low  $K_{\text{D}}$ s, demonstrating a high affinity for Mpro binding. We also evaluated the hydrogen and hydrophobic interactions between each inhibitor and Mpro, demonstrating that each drug has interactions with important amino acid residues from Mpro. These four inhibitors can be potential drug candidates for COVID-19 therapy through preclinical and clinical improvements.

## 2. Materials and Methods

### 2.1. Drug Preparation and Docking Analysis

A total of 26,832 compounds from Life Chemicals (<https://lifechemicals.com/> (accessed on 1 October 2021)), 29,524 compounds from Chembridge (<https://chembridge.com/> (accessed on 1 January 2022)) and 107,341 compounds from Chempspace (<https://chemspace.com/> (accessed on 1 January 2022)) were downloaded for docking analysis. All drugs were converted from a structure data file (.sdf) to mol2 format. Hydrogens were added, and 3D coordinates were also generated at the same time.

The crystal structure of the main protease (6WTT and 5RGI) was downloaded from the Protein Data Bank (<https://www.rcsb.org/> (accessed on 1 October 2021)). The ligand, ions and waters of the crystal structure were removed before docking. GOLD software was used for docking analysis for all compounds. An amount of 10 Å around the original ligand was set as the docking site. For each ligand, a maximum of 10 structures were generated, and the ChemPLP fitness was computed.

### 2.2. $\text{IC}_{50}$ Test

SARS-CoV-2 Mpro and Mpro protease fluorogenic substrate peptide were purchased from Millipore Sigma (Cat#: SAE0172 and SAE0180, respectively). Proteolytic cleavage by Mpro releases the fluorescent AFC group (sequence TSAVLQ) which can be detected by a fluorescence microscope.

$\text{IC}_{50}$  was tested with an 8  $\mu\text{g}/\text{mL}$  main protease and a 15  $\mu\text{g}/\text{mL}$  substrate in 25 mM HEPES (PH 7.5) in a 96-well plate. The reaction volume was 50  $\mu\text{L}$ . Then 0.5  $\mu\text{L}$  of a different concentration of inhibitors dissolved in 100% DMSO was added to each well and reacted for 90 min. The increase in fluorescence was detected with a microscope (excitation = 400 nm, emission = 505 nm).

$\text{IC}_{50}$  and the 95% confidence interval (95% CI) was calculated by GraphPad Prism (version 8.0.1).

### 2.3. Dissociation Constant ( $K_{\text{D}}$ ) Test

The  $K_{\text{D}}$  was determined by titration as described previously [9]. A total of 350  $\mu\text{L}$  of mixture with 8  $\mu\text{g}/\text{mL}$  main protease, 25 mM HEPES (PH7.5) and 2% DMSO was added to the cuvette. An amount of 2  $\mu\text{L}$  200  $\mu\text{M}$  of each compound with 25mM HEPES (PH7.5) and 2% DMSO was added into each cuvette and left to wait for 90 s. The data was collected with a fluorometer every 90 s (excitation = 285nm and slit = 2 nm; emission = 480 nm and slit = 8 nm). The affinity of each ligand ( $K_{\text{D}}$ ) was derived by a linear relation using the Scatchard plot [10]:

$$\Delta F = -K_{\text{D}} \Delta F / [S] + \Delta F_{\infty}$$

where  $\Delta F$  and  $\Delta F_{\infty}$  stands for the fluorescence change at the substrate concentration  $[S]$ , and the fluorescence change when all enzyme molecules are complexed with the substrate.

The inner-filter correction was made as below [11]:

$$f = \frac{I_0}{I} = \frac{P_0 + \Delta A}{P_0} \frac{1 - 10^{-P_0}}{1 - 10^{-(P_0 + \Delta A)}}$$

where  $f$ ,  $I_0$  and  $I$  refer to the correction factor, corrected and observed fluorescence intensity, respectively.  $P_0$  and  $\Delta A$  stands for the sample absorption before titration and absorption change with the addition of substrates.

#### 2.4. Compound Pharmacokinetic, Drug Likeness and Toxicity Predictions

SwissADME (<http://www.swissadme.ch/index.php#> (accessed on 1 September 2022)) allows one to compute the ADME (Absorption, Distribution, Metabolism, Excretion) parameters, pharmacokinetic properties and drug likeness and medicinal chemistry friendliness [12]. ProTox-II ([https://tox-new.charite.de/prottox\\_II/](https://tox-new.charite.de/prottox_II/) (accessed on 1 September 2022)) is also a free web server based on a total of 33 models for the prediction of various toxicity endpoints like acute toxicity, cytotoxicity and hepatotoxicity. All compounds were subjected to the SwissADME and ProTox-II for evaluation [13].

#### 2.5. Hydrogen and Hydrophobic Interactions Analysis

Interactions between each compound and Mpro were analyzed by LigPlus (version 2.2.5) by default parameters. The maximum distance of the hydrogen–acceptor distance is 2.7 Å, and the donor–acceptor distance is 3.35 Å. For hydrophobic contacts, the minimum contact distance is 2.90 Å, and the maximum contact distance is 3.90 Å.

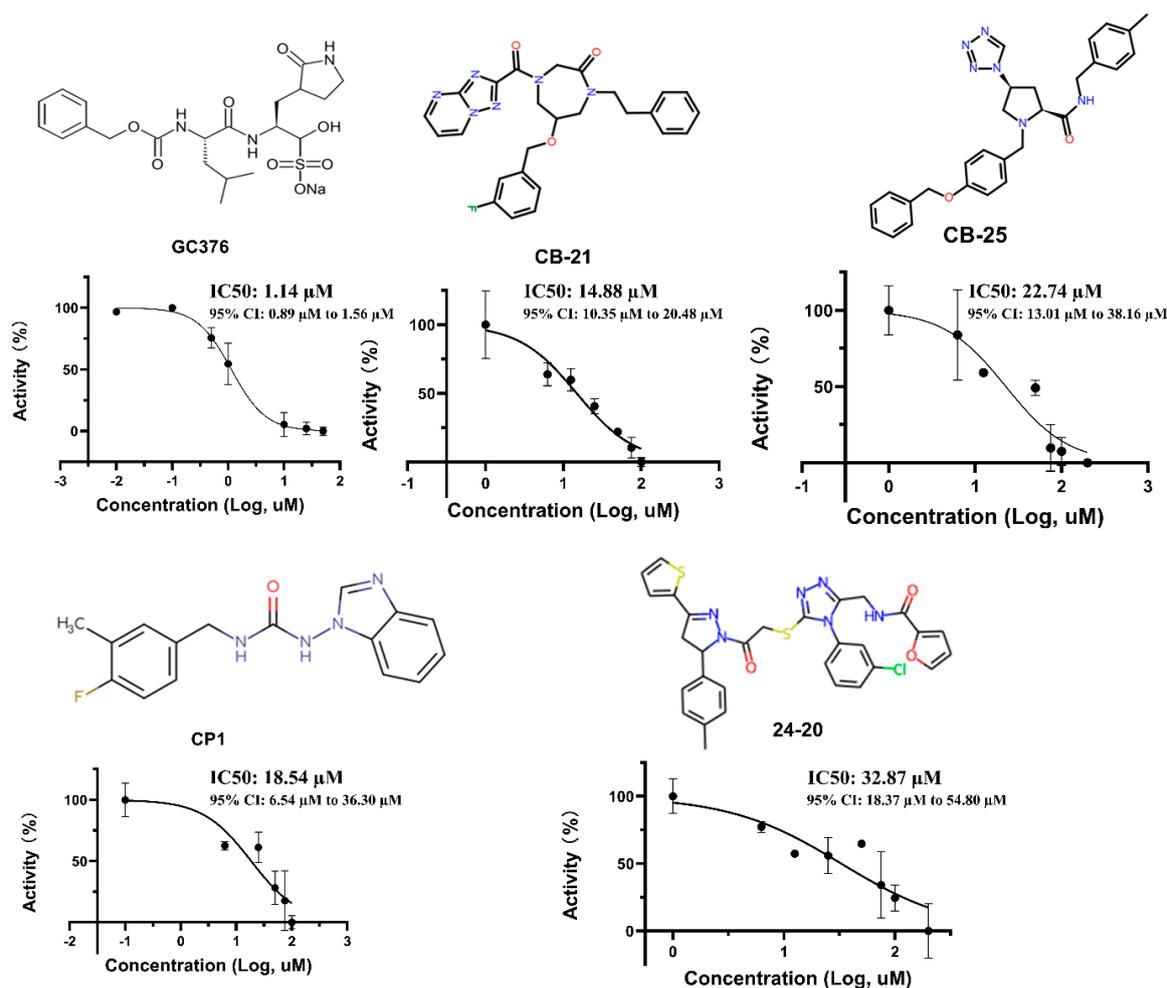
### 3. Results

#### 3.1. Molecular Docking

An in silico molecular docking method was used to investigate potential drug candidates. A total of 26,832 compounds from Life Chemicals, 29,524 compounds from Chembridge and 107,341 compounds from Chemspace were downloaded for docking analysis. Docking analysis was performed by Genetic Optimization for Ligand Docking (GOLD), which is based on a genetic algorithm for docking flexible ligands. The ChemPLP was utilized as the main scoring function. ChemPLP has been validated as the most effective scoring function using the ChemScore hydrogen bonding term as well as multiple linear potentials to model van der Waals and repulsive terms. The Mpro crystal structure 5RGI was downloaded from the Protein Data Bank, and the ligand Z369936976 was removed for docking. Top ranking drugs were purchased and further tested, and several selected drugs are shown in Table 1 and Figure 1. We found that the docking ChemPLP fitness scores of CB-21 and CB-25 from Chembridge, CP-1 from Chemspace and LC24-20 from Life Chemicals are 98.08, 92.36, 81.19 and 90.29, respectively. These scores demonstrate that the four inhibitors have good potential to inhibit Mpro.

**Table 1.** Docking score of selected compounds.

Name	Company	ID	Molecular Mass (Da)	ChemPLP Score
CB-21	Chembridge	25917626	489	98.08
CB-25	Chembridge	55464888	483	92.36
CP-1	ChemSpace	CSCS00026316954	298	81.19
LC24-20	Life Chemicals	F0514-4479	617	90.29



**Figure 1.** Structure and IC<sub>50</sub> test of each compound. 95% CI: 95% Confidence Interval.

### 3.2. IC<sub>50</sub> and Inhibitor Binding

The half maximal inhibitory concentrations (IC<sub>50</sub>s) of five inhibitors were tested by an enzymatic assay (Figure 1). As the IC<sub>50</sub> of the inhibitor Z369936976 in 5RGI was not tested, the well-known SARS-CoV-2 Mpro inhibitor GC376 was used as the positive control in the IC<sub>50</sub> test. The GC376 has already been shown to be an effective Mpro inhibitor, with an IC<sub>50</sub> ranging from 0.15  $\mu\text{M}$  to 29.4  $\mu\text{M}$  from different reports [14–18]. According to our experiment, the IC<sub>50</sub> of GC376 is 1.14  $\mu\text{M}$  (95% CI: 0.89  $\mu\text{M}$  to 1.56  $\mu\text{M}$ ), similar to previously reported results. Our newly found inhibitor CB-21 has an IC<sub>50</sub> at 14.88  $\mu\text{M}$  (95% CI: 10.35  $\mu\text{M}$  to 20.48  $\mu\text{M}$ ) and CB-25 has an IC<sub>50</sub> at 22.74  $\mu\text{M}$  (95% CI: 13.01  $\mu\text{M}$  to 38.16  $\mu\text{M}$ ). CP-1 has an IC<sub>50</sub> at 18.54  $\mu\text{M}$  (95% CI: 6.54  $\mu\text{M}$  to 36.30  $\mu\text{M}$ ) and LC24-20 has a higher IC<sub>50</sub> at 32.87  $\mu\text{M}$  (95% CI: 18.37  $\mu\text{M}$  to 54.80  $\mu\text{M}$ ).

The binding of inhibitors to the main protease was parallelly tested by the direct inhibitor titration, yielding the intrinsic fluorescence modification. We found that the  $K_D$  of CB-21, CB-25, CP-1 and LC24-20 are  $12.85 \pm 2.05 \mu\text{M}$ ,  $10.77 \pm 1.32 \mu\text{M}$ ,  $5.08 \pm 0.38 \mu\text{M}$  and  $20.05 \pm 1.27 \mu\text{M}$ , respectively. Like the IC<sub>50</sub> test experiment, the  $K_D$  of CB-21, CB-25 and CP-1 is lower than the  $K_D$  of LC24-20.

### 3.3. Drug Likelihood, Pharmacokinetic and Oral Toxicity Evaluations of Selected Compounds

Drug likelihood, pharmacokinetic and oral toxicity were evaluated for the four inhibitors, as well as the GC376 (Table 2) by SwissADME and ProTox-II. Results show that compared with GC376, all four drugs have high lipophilicity. LC24-20 has the highest lipophilicity, and it is insoluble in water whereas GC376 is soluble in water. CB-21, CB-25

and CP-1 have high gastrointestinal absorption but GC376 and LC24-20 have low gastrointestinal absorption. For the five molecules, only CP-1 has the blood brain barrier permeability. P-glycoprotein is an important protein of cell membrane, primarily expressed in the liver, pancreas, kidney, colon and jejunum [19], and it plays an important role in compound transport. GC376, CB-21 and CB-25 may behave better in their transport. The bioavailability score which refers to the drug becoming completely available to its intended biological destinations of the four drugs is 0.55 and much higher compared with GC376, whose bioavailability score is only 0.17. Similarly, the LD<sub>50</sub>s of CB-21, CB-25, CP-1 and LC24-20 are also much higher than GC376. The LD<sub>50</sub> of GC376 is 300 mg/kg whereas the LD<sub>50</sub>s of CB-21, CB-25, CP-1 and LC24-20 are 1600 mg/kg, 576 mg/kg, 1000 mg/kg and 3000 mg/kg, respectively (Table 2).

**Table 2.** Drug likeliness, pharmacokinetic and oral toxicity evaluations of selected compounds.

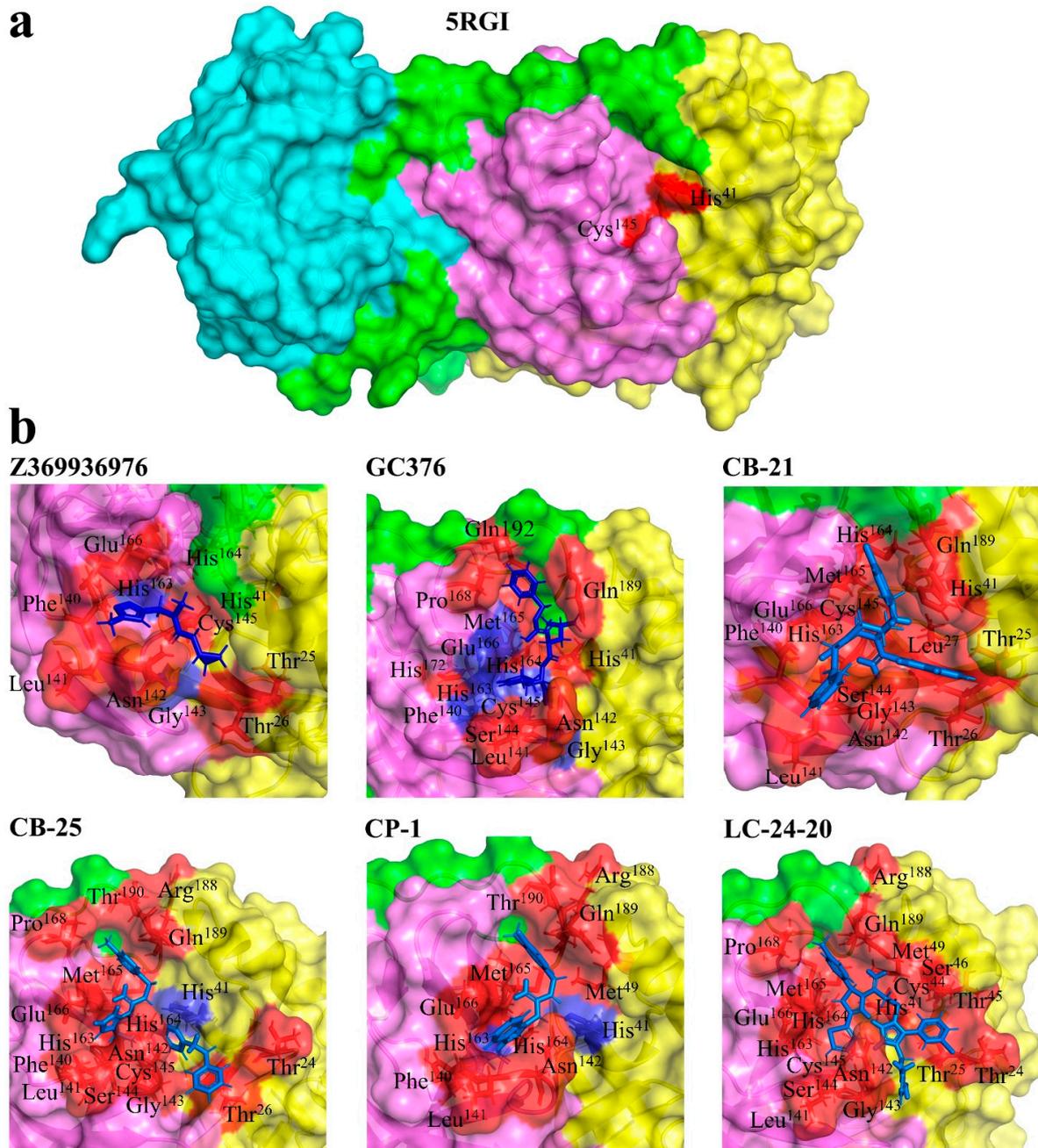
ID	Lipophilicity Consensus Log Po/w	Water Solubility (mol/L)	GI Absorption	BBB Permeability	P-gp Substrate	Bioavailability Score	LD <sub>50</sub> (mg/kg)
GC376	−2.18	Soluble (8.47 × 10 <sup>−4</sup> )	Low	No	Yes	0.17	300
CB-21	2.45	Poorly Soluble (3.00 × 10 <sup>−5</sup> )	High	No	Yes	0.55	1600
CB-25	3.25	Poorly Soluble (4.00 × 10 <sup>−6</sup> )	High	No	Yes	0.55	576
CP-1	2.75	Moderately soluble (1.36 × 10 <sup>−4</sup> )	High	Yes	No	0.55	1000
LC24-20	5.02	Insoluble (3.59 × 10 <sup>−8</sup> )	Low	No	No	0.55	3000

Po/w: octanol/water partition coefficient, GI: Gastrointestinal, BBB: Blood Brain Barrier, 50%, P-gp: P-glycoprotein, LD<sub>50</sub>: 50% Lethal Dose.

### 3.4. Interactions

The main protease has three domains: domain I (residues 8–101), domain II (residues 102–184) and domain III (residues 201–303). Domain I and domain II have an antiparallel β-barrel structure similar to structures of the trypsin-like serine proteases, and domain III consists of five α-helices, forming an antiparallel conglomerate connected to the domain by a long loop region composed of residues 185–200 [20,21]. The active site of SARS-CoV-2 Mpro is located in a cleft between domains I and II, holding a histidine/cysteine catalytic dyad (His<sup>41</sup> and Cys<sup>145</sup>).

Elaborating on the interactions between the ligand and the enzyme is important for understanding the enzyme mechanism and developing novel chemical agents or inhibitors [22]. For each inhibitor, the hydrogen bond interactions and hydrophobic interactions between each inhibitor and Mpro were analyzed (Figure 2, Table 3). The amino acid alignment reveals that all residues potentially interacting with the ligand in the active site of SARS-CoV-1 and SARS-CoV-2 are conserved. These residues include Thr<sup>24</sup>, Thr<sup>25</sup>, His<sup>41</sup>, Cys<sup>44</sup>, Met<sup>49</sup>, Tyr<sup>54</sup>, Phe<sup>140</sup>, Asn<sup>142</sup>, Gly<sup>143</sup>, Cys<sup>145</sup>, His<sup>163</sup>, His<sup>164</sup>, Met<sup>165</sup>, Glu<sup>166</sup>, Leu<sup>167</sup>, Pro<sup>168</sup>, Asp<sup>187</sup>, Arg<sup>188</sup>, Gln<sup>189</sup> and Thr<sup>190</sup> [21]. All the inhibitors have interactions with the most important catalytic residues His<sup>41</sup> and Cys<sup>145</sup>. Our results showed that Z369936976 has two hydrogen bond interactions with Mpro: Gly<sup>143</sup> and His<sup>163</sup>, while it has nine hydrophobic interactions with Mpro: Thr<sup>25</sup>, Thr<sup>26</sup>, His<sup>41</sup>, Phe<sup>140</sup>, Leu<sup>141</sup>, Asn<sup>142</sup>, Cys<sup>145</sup>, His<sup>164</sup> and Glu<sup>166</sup>. GC376 has hydrogen bond interactions with five amino acids: Phe<sup>140</sup>, Gly<sup>143</sup>, Cys<sup>145</sup>, His<sup>163</sup> and Glu<sup>166</sup> and hydrophobic interactions with another ten amino acids including His<sup>41</sup>. CB-21 has interactions with all residues that interact with Z369936976, and it also interacts with several other residues such as Leu<sup>27</sup>, Ser<sup>144</sup>, Met<sup>165</sup> and Gln<sup>189</sup>. CB-25 has nine of the same interactions with Z369936976, LC24-20 has eight of the same interactions with Z369936976 and CP-1 has seven of the same interactions with Z369936976. CB-25 only has one hydrogen bond interaction with His<sup>41</sup> and CP-1 has two hydrogen bond interactions with His<sup>41</sup> and His<sup>163</sup>, and all the remaining interactions are hydrophobic interactions.



**Figure 2.** Structure of main protease and interactions with each compound. Main protease has three domains. Domain I (residues 8–101) is colored in yellow, domain II (residues 102–184) is colored in violet and domain III (residues 201–303) is colored in cyan. (a) Surface plot of crystal structure 5RGI. Essential catalytic residues His<sup>41</sup> and Cys<sup>145</sup> are colored in red. (b) Interactions between each inhibitor and main protease. Crystal structure of Mpro with inhibitor GC376 was retrieved from 6WTT and others are retrieved from 5RGI. Residues which have hydrogen bond interactions with ligands are colored in blue while residues which have hydrophobic interactions with ligands are colored in red and shown by sticks.

**Table 3.** Hydrogen bond and hydrophobic interactions between each compound and main protease.

Compound	Hydrogen Bond Interaction	Hydrophobic Interaction
Z369936976	Gly <sup>143</sup> , His <sup>163</sup>	Thr <sup>25</sup> , Thr <sup>26</sup> , His <sup>41</sup> , Phe <sup>140</sup> , Leu <sup>141</sup> , Asn <sup>142</sup> , Cys <sup>145</sup> , His <sup>164</sup> , Glu <sup>166</sup>
GC376	Phe <sup>140</sup> , Gly <sup>143</sup> , Cys <sup>145</sup> , His <sup>163</sup> , Glu <sup>166</sup>	His <sup>41</sup> , Leu <sup>141</sup> , Asn <sup>142</sup> , Ser <sup>144</sup> , His <sup>164</sup> , Met <sup>165</sup> , Pro <sup>168</sup> , His <sup>172</sup> , Gln <sup>189</sup> , Gln <sup>192</sup>
CB-21		Thr <sup>25</sup> , Thr <sup>26</sup> , Leu <sup>27</sup> , His <sup>41</sup> , Phe <sup>140</sup> , Leu <sup>141</sup> , Asn <sup>142</sup> , Gly <sup>143</sup> , Ser <sup>144</sup> , Cys <sup>145</sup> , His <sup>163</sup> , His <sup>164</sup> , Met <sup>165</sup> , Glu <sup>166</sup> , Gln <sup>189</sup>
CB-25	His <sup>41</sup>	Thr <sup>24</sup> , Thr <sup>26</sup> , Phe <sup>140</sup> , Leu <sup>141</sup> , Asn <sup>142</sup> , Gly <sup>143</sup> , Cys <sup>145</sup> , His <sup>163</sup> , His <sup>164</sup> , Met <sup>165</sup> , Glu <sup>166</sup> , Pro <sup>168</sup> , Arg <sup>188</sup> , Gln <sup>189</sup> , Thr <sup>190</sup>
CP-1	His <sup>41</sup> , His <sup>163</sup>	Met <sup>49</sup> , Phe <sup>140</sup> , Leu <sup>141</sup> , Asn <sup>142</sup> , His <sup>164</sup> , Met <sup>165</sup> , Glu <sup>166</sup> , Arg <sup>188</sup> , Gln <sup>189</sup> , Thr <sup>190</sup>
LC24-20		Thr <sup>24</sup> , Thr <sup>25</sup> , His <sup>41</sup> , Cys <sup>44</sup> , Thr <sup>45</sup> , Ser <sup>46</sup> , Met <sup>49</sup> , Leu <sup>141</sup> , Asn <sup>142</sup> , Gly <sup>143</sup> , Ser <sup>144</sup> , Cys <sup>145</sup> , His <sup>163</sup> , His <sup>164</sup> , Met <sup>165</sup> , Glu <sup>166</sup> , Pro <sup>168</sup> , Arg <sup>188</sup> , Gln <sup>189</sup>

#### 4. Discussion

The ongoing COVID-19 pandemic brought disaster to health and economic crises to people all over the world. Despite the usage of various kinds of vaccines, the pandemic is still spreading worldwide due to the fast viral mutation. As the main protease of SARS-CoV-2, Mpro plays an important role in SARS-CoV-2 replication and thus becomes a promising drug target.

A total of 163,697 compounds from Life Chemicals, Chembridge and Chemspace were prepared for the docking analysis. The larger the compound number utilized for docking, the more possibility there was to find efficient inhibitors. Compounds with high PLP fitness scores were then tested by an enzymatic assay for IC50. However, a higher PLP fitness score does not always correlate with a lower IC50 for a compound. For example, for the four inhibitors described here, CP-1 has a PLP fitness score of 81.19 and an IC50 at 18.54  $\mu$ M while LC24-20 has a PLP fitness score of 90.29 and an IC50 at 32.87  $\mu$ M. In fact most compounds we purchased with a high docking score may have low or even no inhibition to the enzyme. Therefore, more validation methods need to be used jointly for a better prediction. Molecular dynamics-based methods, although computationally expensive, are reported to perform better in ranking molecules compared with docking methods including ChemPLP, AutoDock Vina and Glide [23]. For example, top ranking molecules selected by a PLP fitness score could be further screened by molecular dynamics-based methods before inhibitor purchasing for experimental verification.

CB-21 has a PLP fitness score of 98.08 and an IC50 at 14.88  $\mu$ M, seeming to be the best of the four inhibitors found. Compared with the well-known Mpro inhibitor GC376 with a tested IC50 at 1.14  $\mu$ M (95% CI: 0.89  $\mu$ M to 1.56  $\mu$ M) the inhibition efficacy is not high enough. However, CB-21 has a high GI absorption, high bioavailability score and much higher LD<sub>50</sub> (1600 mg/kg) than GC376. The calculated LD<sub>50</sub> of inhibitor 11a, 11b and 11r [8], which has an IC50 at less than 1  $\mu$ M, are 200 mg/kg, 200 mg/kg and 1000 mg/kg, respectively. It is also higher than that of CB-21. Several inhibitors targeting Mpro are under evaluation for clinical trial by the FDA: PBI-0451, Ebselen, Ritonavir, S-217622, Nirmatrelvir, Lopinavir, Apixaban and Azithromycin. Paxlovid (nirmatrelvir and ritonavir) has been issued for emergency use for the treatment of mild-to-moderate cases in certain adults and pediatric patients (Paxlovid LOA 10272022 ([fda.gov](https://www.fda.gov))). However, no drugs targeting Mpro have yet been approved by the FDA. All of our four new inhibitors were directly purchased from the company and tested by an enzymatic assay without any modification. Proper group modification of the molecular may increase their performance.

Elaborating on the detailed interactions between the enzyme and ligands is the basis for understanding the binding and catalytic process, being important for developing novel inhibitors or therapeutic agents [22]. All of the four inhibitors as well as GC376 and the ligand Z369936976 have hydrophobic or hydrogen bond interactions with the essential catalytic residues His<sup>41</sup> and Cys<sup>145</sup> of Mpro. Inhibitor GC376 has a peptide bond. Similar

to the four selected inhibitors, CB-25, CP-1 and LC24-20 also have a peptide bond in each of them. GC376 and CB21, while having a relatively lower IC<sub>50</sub>, have 15 hydrogen bond or hydrophobic interactions with Mpro while LC24-20, with a higher IC<sub>50</sub>, has 19 hydrogen bond or hydrophobic interactions with Mpro. It seems that having more interactions does not correlate with a lower IC<sub>50</sub>, as each interaction can contribute a different free energy. Compared with GC376, both CB-21 and LC24-20 have interactions with 12 of the same amino acids. Among them, 11 amino acids were the same and only 1 amino acid was different: GC376 and CB-21 had interactions with Phe<sup>140</sup> while GC376 and LC24-20 had interactions with Pro<sup>168</sup>. Besides the 12 same amino acids, GC376 had interactions with Pro<sup>168</sup>, His<sup>172</sup> and Gln<sup>192</sup> with Mpro while CB-21 had interactions with Thr<sup>25</sup>, Thr<sup>26</sup> and Leu<sup>27</sup>, possibility leading to the IC<sub>50</sub> difference.

As SARS-CoV-2 mutates rapidly, to prioritize global monitoring and research, WHO, in collaboration with experts, national authorities, institutions and researchers, characterized specific Variants of Interest (VOIs) and Variants of Concern (VOCs), including the variants alpha, beta, delta and omicron. We searched the data from GISAID (<https://gisaid.org/> (accessed on 23 January 2023)) and Nextstrain (<https://nextstrain.org/> (accessed on 23 January 2023)) with 2910 SARS-CoV-2 genomes samples in the past six months (updated on 23 January 2023) [24]. Normalized Shannon entropy per codon for the whole genome was calculated, including Mpro (supplementary file). Results show that the main protease has 64 codon diversity entropies among the Mpro's 306 amino acids. Not surprisingly, 74.6% of the codon diversity entropies equals to or is lower than 0.01. The two essential catalytic residues His<sup>41</sup> and Cys<sup>145</sup> are also reserved. Only four amino acids (1.3%) have a diversity entropy over 0.05: Gly<sup>15</sup> (entropy 0.159), Leu<sup>89</sup> (entropy 0.119), Lys<sup>90</sup> (entropy 0.1) and Pro<sup>132</sup> (entropy 0.584). However, these four amino acids are more than 10 Å from the catalytic center of Mpro. Therefore, the influence of mutations for Mpro activity and inhibition by our compounds are relatively low. On the contrary, 94 out of 1273 amino acids (7.4%) of SARS-CoV-2 spike protein have a diversity entropy over 0.05 (supplementary file). At the same time, 27 out of 1273 amino acids (2.1%) of spike protein have a diversity entropy over 0.5. Therefore, compared with inhibitors targeting spike protein, inhibitors targeting Mpro should be much more conserved.

## 5. Conclusions

In this study, we evaluated four novel SARS-CoV-2 Mpro inhibitors with low IC<sub>50</sub>s and a high binding affinity, through a combined study of molecular modelling, docking, inhibition of enzyme activity, inhibitor binding, inhibitor toxicity and a ligand/enzyme interaction study. The new inhibitors have relatively higher bioavailabilities and lower toxicity compared with the well-known GC376 with high efficacy. They all have interactions with the essential catalytic residues His<sup>41</sup> and Cys<sup>145</sup>, thus contributing to their inhibition efficacy. These inhibitors have potential to become drug candidates following tests with SARS-CoV-2 and animal models. The established combined computational and experimental study has brought potential drug candidates with good efficacy and low toxicity and will continue to be useful for the fight against COVID-19 as well as for new viral pandemics.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/v15020580/s1>. Table S1: Normalized Shannon entropy per codon for SARS-CoV-2 Mpro and Spike protein.

**Author Contributions:** Conceptualization, S.-X.L. & W.Z.; methodology, W.Z.; software, W.Z.; validation, W.Z. and S.-X.L.; formal analysis, W.Z.; investigation, W.Z. & S.-X.L.; resources, S.-X.L.; data curation, W.Z.; writing—original draft preparation, W.Z.; writing—review and editing, S.-X.L.; visualization, W.Z.; supervision, S.-X.L.; project administration, S.-X.L.; funding acquisition, S.-X.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Cao, B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **2020**, *395*, 497–506. [[CrossRef](#)]
2. Machhi, J.; Herskovitz, J.; Senan, A.M.; Dutta, D.; Nath, B.; Oleynikov, M.D.; Kevadiya, B.D. The natural history, pathobiology, and clinical manifestations of SARS-CoV-2 infections. *J. Neuroimmune Pharmacol.* **2020**, *15*, 359–386. [[CrossRef](#)] [[PubMed](#)]
3. Rastogi, M.; Pandey, N.; Shukla, A.; Singh, S.K. SARS coronavirus 2: From genome to infectome. *Respir. Res.* **2020**, *21*, 318. [[CrossRef](#)] [[PubMed](#)]
4. Romano, M.; Ruggiero, A.; Squeglia, F.; Maga, G.; Berisio, R. A structural view of SARS-CoV-2 RNA replication machinery: RNA synthesis, proofreading and final capping. *Cells* **2020**, *9*, 1267. [[CrossRef](#)] [[PubMed](#)]
5. Fan, K.; Wei, P.; Feng, Q.; Chen, S.; Huang, C.; Ma, L.; Lai, B.; Pei, J.; Liu, Y.; Chen, J.; et al. Biosynthesis, Purification, and Substrate Specificity of Severe Acute Respiratory Syndrome Coronavirus 3C-like Proteinase. *J. Biol. Chem.* **2004**, *279*, 1637–1642. [[CrossRef](#)] [[PubMed](#)]
6. Ullrich, S.; Nitsche, C. The SARS-CoV-2 main protease as drug target. *Bioorganic Med. Chem. Lett.* **2020**, *30*, 127377. [[CrossRef](#)] [[PubMed](#)]
7. Bhardwaj, V.K.; Singh, R.; Das, P.; Purohit, R. Evaluation of acridinedione analogs as potential SARS-CoV-2 main protease inhibitors and their comparison with repurposed anti-viral drugs. *Comput. Biol. Med.* **2021**, *128*, 104117. [[CrossRef](#)]
8. Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Hilgenfeld, R. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved  $\alpha$ -ketoamide inhibitors. *Science* **2020**, *368*, 409–412. [[CrossRef](#)]
9. Lin, S.X.; Wang, Q.; Wang, Y.L. Interactions between Escherichia coli arginyl-tRNA synthetase and its substrates. *Biochemistry* **1988**, *27*, 6348–6353. [[CrossRef](#)]
10. Holler, E.; Bennett, E.L.; Calvin, M. 2-p-Toluidinylnaphthalene-6-sulfonate, a fluorescent reporter group for L-isoleucyl-tRNA synthetase. *Biochem. Biophys. Res. Commun.* **1971**, *45*, 409–415. [[CrossRef](#)]
11. Burshtein, E.A. Quenching of protein fluorescence. I. Principles of the method. Solutions of tryptophan, tyrosine, and denaturated proteins. *Biofizika* **1968**, *13*, 433–442.
12. Daina, A.; Michielin, O.; Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **2017**, *7*, 42717. [[CrossRef](#)] [[PubMed](#)]
13. Banerjee, P.; Eckert, A.O.; Schrey, A.K.; Preissner, R. ProTox-II: A webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res.* **2018**, *46*, W257–W263. [[CrossRef](#)]
14. Lu, J.; Chen, S.A.; Khan, M.B.; Brassard, R.; Arutyunova, E.; Lamer, T.; Lemieux, M.J. Crystallization of feline coronavirus Mpro with GC376 reveals mechanism of inhibition. *Front. Chem.* **2022**, *10*, 852210. [[CrossRef](#)]
15. Hung, H.C.; Ke, Y.Y.; Huang, S.Y.; Huang, P.N.; Kung, Y.A.; Chang, T.Y.; Hsu, J.T.A. Discovery of M protease inhibitors encoded by SARS-CoV-2. *Antimicrob. Agents Chemother.* **2020**, *64*, e00872-e20. [[CrossRef](#)] [[PubMed](#)]
16. Vuong, W.; Fischer, C.; Khan, M.B.; van Belkum, M.J.; Lamer, T.; Willoughby, K.D.; Vederas, J.C. Improved SARS-CoV-2 Mpro inhibitors based on feline antiviral drug GC376: Structural enhancements, increased solubility, and micellar studies. *Eur. J. Med. Chem.* **2021**, *222*, 113584. [[CrossRef](#)]
17. Fu, L.; Ye, F.; Feng, Y.; Yu, F.; Wang, Q.; Wu, Y.; Gao, G.F. Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. *Nat. Commun.* **2020**, *11*, 4417. [[CrossRef](#)]
18. Dampalla, C.S.; Zheng, J.; Perera, K.D.; Wong, L.Y.R.; Meyerholz, D.K.; Nguyen, H.N.; Chang, K.O. Postinfection treatment with a protease inhibitor increases survival of mice with a fatal SARS-CoV-2 infection. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2101555118. [[CrossRef](#)] [[PubMed](#)]
19. Thiebaut, F.; Tsuruo, T.; Hamada, H.; Gottesman, M.M.; Pastan, I.; Willingham, M.C. Cellular localization of the multidrug-resistance gene product P-glycoprotein in normal human tissues. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 7735–7738. [[CrossRef](#)]
20. Li, Q.; Kang, C. Progress in developing inhibitors of SARS-CoV-2 3C-like protease. *Microorganisms* **2020**, *8*, 1250. [[CrossRef](#)] [[PubMed](#)]
21. Banerjee, R.; Perera, L.; Tillekeratne, L.V. Potential SARS-CoV-2 main protease inhibitors. *Drug Discov. Today* **2021**, *26*, 804–816. [[CrossRef](#)] [[PubMed](#)]
22. Kim, S.-K.; Park, K.D.; Lee, D.-W. Editorial: Interactions Between Small Molecule Ligands and Target Enzymes. *Front. Mol. Biosci.* **2021**, *8*, 649450. [[CrossRef](#)] [[PubMed](#)]
23. Breznik, M.; Ge, Y.; Bluck, J.P.; Briem, H.; Hahn, D.F.; Christ, C.D.; Mortier, J.; Mobley, D.L.; Meier, K. Prioritizing Small Sets of Molecules for Synthesis through *in-silico* Tools: A Comparison of Common Ranking Methods. *Chemmedchem* **2022**, *18*, 202200425. [[CrossRef](#)] [[PubMed](#)]
24. Hadfield, J.; Megill, C.; Bell, S.M.; Huddleston, J.; Potter, B.; Callender, C.; Sagulenko, P.; Bedford, T.; Neher, R.A. Nextstrain: Real-time tracking of pathogen evolution. *Bioinformatics* **2018**, *34*, 4121–4123. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.