Received: 19 March 2022 | Revised: 12 June 2022 | Accepted: 3 July 2022

DOI: 10.1111/cbdd.14115

RESEARCH ARTICLE

Hydrogen bonding penalty used for virtual screening to discover potent inhibitors for Papain-Like cysteine proteases of SARS-CoV-2

Guangjian Zhao[1,2,3] | **Xiaochun Liu**[1,2,3] | **Suyun Wang**[4] | **Zhongyue Bai**[1,2,3] | **Siyu Zhang**[2,3,5] | **Yifan Wang**[2,3,6] | **Haibo Yu**[7] | **Ximing Xu**[1,2,3]

[1]Key Laboratory of Marine Drugs of Ministry of Education, School of Medicine and Pharmacy, Ocean University of China, Qingdao, China

[2]Pilot National Laboratory for Marine Science and Technology, Center for Innovation Marine Drug Screening & Evaluation, Qingdao, China

[3]Marine Biomedical Research Institute of Qingdao, Qingdao, China

[4]Key Laboratory of Special Pathogens and Biosafety, Center for Biosafety Mega-Science, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, China

[5]School of Life Science, Lanzhou University, Lanzhou, China

[6]College of Food Science and Engineering, Ocean University of China, Qingdao, China

[7]School of Chemistry and Molecular Bioscience, University of Wollongong, Wollongong, New South Wales, Australia

Correspondence
Ximing Xu, Key Laboratory of Marine Drugs of Ministry of Education, School of Medicine and Pharmacy, Ocean University of China, Qingdao, Shandong 266003, China.
Email: xuximing@ouc.edu.cn

Haibo Yu, School of Chemistry and Molecular Bioscience, University of Wollongong, Wollongong, New South Wales, Australia.
Email: hyu@uow.edu.au

Funding information
Benefiting People Policy-guided science and technology program of Qingdao, Grant/Award Number: 20-4-1-4-nsh; Shandong Provincial Major Science and Technology Innovation Project, Grant/Award Number: 2018SDKJ0402 and 2020CXGC010503; Project supported by the Young Scientists Fund of the National Natural Science Foundation of China, Grant/Award Number: 31900910

Abstract

The Papain-Like proteases (PLpro) of SARS-CoV-2 play a crucial role in viral replication and the formation of nonstructural proteins. To find available inhibitors, the 3D structure of PLpro of SARS2 was obtained by homologous modelling, and we used this structure as a target to search for inhibitors through molecular docking and MM/GBSA binding free energy rescoring. A novel hydrogen bonding penalty was applied to the screening process, which meanwhile took desolvation into account. Finally, 61 compounds were acquired and 4 of them with IC_{50} at micromolar level tested in vitro enzyme activity assay, which includes clinical drugs tegaserod. Considering the importance of crystal water molecules, the 4 compounds were re-docked and considered bound waters in the active site as a part of PLpro. The binding modes of these 4 compounds were further explored with metadynamics simulations. The hits will provide a starting point for future key interactions identified and lead optimization targetting PLpro.

KEYWORDS

hydrogen bonding penalty, molecular dynamics simulation, papain-like cysteine proteases, SARS-CoV-2, virtual screening

1 | INTRODUCTION

There is a new public health crisis threatening the world with the emergence and spread of SARS-CoV-2 (SARS2) (Shin et al., 2020; Singhal, 2020). An early and essential process of SARS2 is the cleavage of polyprotein into 16 mature components termed nonstructural proteins (NSPs). Papain-Like protease (PLpro) catalyzes itself and

Guangjian Zhao, Xiaochun Liu and Suyun Wang are the co-first authors.

 © 2022 John Wiley & Sons Ltd.

other NSPs release from the polyprotein, thereby initiating virus-mediated RNA replication (Klemm et al., 2020; Ratia et al., 2008). The same as SARS-CoV (SARS) PLpro, SARS2 PLpro has deubiquitinating and deISGylating enzyme activities, and evades the innate immune response of host cells by deubiquitinating and deISGylating structural and nonstructural proteins of SARS2 (Klemm et al., 2020; Mielech, Chen, et al., 2014; Mielech, Kilianski, et al., 2014). PLpro uses the thiol group of cysteine as a nucleophile to attack the carbonyl group of the scissile peptide bond (Chou et al., 2014). In the cytosol, the membrane-associated PLpro domain found in the boundaries of nsp1/2, nsp2/3, and nsp3/4, recognizes the P4–P1 consensus cleavage sequence LXGG, and proteolytic of the peptide bond occurred after the glycine at the position, which is an essential process for viral replication (Gao et al., 2021; Han et al., 2005; Rut et al., 2020). The numerous functions and requisite roles of PLpro in viral replication and pathogenesis suggest that PLpro is essential for the viral life cycle, thereby it is a vital target for drug discovery studies against the recent epidemics caused by SARS2 (Klemm et al., 2020).

GRL0617 is a potent, selective, and competitive noncovalent inhibitor of SARS PLpro/deubiquitinase, with an IC_{50} of 0.6 μM and K_i of 0.49 μM (Ratia et al., 2008). Aiming at the GRL0617 binding site, structure-based inhibitors designs have been reported (Freitas et al., 2020; Fu et al., 2021; Rut et al., 2020). The main body of this site is populated with polar amino acids, with only one end consisting of two prolines forming a hydrophobic region (Freitas et al., 2020; Fu et al., 2021; Gao et al., 2021; Rut et al., 2020). Water-mediated hydrogen bonds have been proposed to play important roles in PLpro–ligand complexes. In crystal structure 7JIT (C111S mutant), a water molecule mediated hydrogen bond between Lys157 and carbamylurea of Y95, a similar phenomenon is also seen in crystal structure 7JIW (wild-type). Not only that, the sidechain of Glu167 showed minor migration due to water molecule extrusion, which resulted in Glu167 forming a new hydrogen bond with ligand (Shen et al., 2021). Water molecules in the binding pocket can not only occupy a certain space but also stabilize the residues around them by complexation (Osipiuk et al., 2021; Trujillo et al., 2012). Improper treatment of water molecules leads to a dramatic decrease in inhibitor activity, in contrast, the reasonable design of molecules to replace the water molecules in the binding pocket can improve the activity (Fell et al., 2020; Trujillo et al., 2012). The process which drugs exert their effects involves competition between ligands and water molecules in the binding pocket of the target protein. In other words, if a ligand wants to drive away water molecules and occupy the active pocket, it must break hydrogen bonds between water molecules and amino acid residues and form new hydrogen bonds (Abel et al., 2008). Hence, it is very important to correctly handle the hydrogen bond interaction between protein and ligand. However, in virtual screening methods, the scoring function stipulates the solvent as a continuous medium model that is poorly suited for bound water and therefore tends to give poor results in predicting binding affinities (Feig et al., 2004; Friesner et al., 2006; Schneider, 2010). To this end, there have also been attempts to consider bound water as part of the receptor for molecular docking, with statistical results showing improved precision (Thilagavathi & Mancera, 2010). Unfortunately, MMPB/GBSA does not deal well with water molecules in the receptor structure (Hayes et al., 2011), usually the water molecules in the crystal structure need to be removed for subsequent calculation, which reduced the accuracy of the scoring function.

To address these concerns, we applied the hydrogen bonding penalty (HBP) to the calculation of binding free energy (Zhao & Huang, 2011). Hydrogen bond penalty is based on a theory that hydrogen bonds formed between receptor and water molecules are considered optimal status, and after a water molecule is replaced by a ligand, energy loss will occur if the new hydrogen bonds energies cannot well compensate for the original hydrogen bonds formed between water molecular and receptor. Due to the large degree of freedom of water molecules, the entropy gain caused by the substitution of ligands is beneficial to the enhancement of the binding affinity between receptors and ligands. In contrast, usually, the enthalpy gain cannot be fully complemented because the newly formed hydrogen bonds do not conform to reasonable geometric constraints or cannot form new hydrogen bonds after desolvation, regardless of ligands or receptors (Mahmoud et al., 2020; Young et al., 2007). Therefore, we use HBP to characterize the enthalpic penalties. *Lewater*, a program for hydrogen bond penalty, is designed to predict the position of crystal water and polar interactions at the protein–ligand interface. The program calculates the enthalpy by detecting both the solvation and hydrogen bonding state at the protein–ligand interface. In recent years, hydrogen bond energies have been introduced into modified binding free energy calculations. Hydrogen bond energy usually consists of the distance function and angle function between hydrogen bond donors and hydrogen bond acceptors (Bao et al., 2020; Zhao & Huang, 2011). Solvent exposed moieties of ligand may also form hydrogen bonds with surrounding residues, but these interaction sites are competed by the solvent, consequently, estimating whether these moieties contribute to protein–ligand binding or given a weight value is necessary, which is commonly characterized by the ratio of solvent accessible surface area (SASA) (Mahmoud et al., 2020). Indeed, SASA is a very important

parameter term when calculating hydrogen bond energy and hydrogen bond penalty because it directly affects the angle of hydrogen bonds and the binding "gesture" of ligands. *Lewater* takes into account both the polar and hydrophobic character of the binding site residues and ligands to predict the possible water molecules' binding points in the structure and makes a reasonable treatment for the sites where hydrogen bonds formed. Additionally, if the hydrogen bonds between receptors and donors are formed within solvent inaccessible regions, nonhydrogen bond penalty is created (Bao et al., 2020; Wang et al., 2002; Zhao & Huang, 2011).

2 | METHODS

2.1 | Sequence alignment and homologous modelling

Before the crystal structure of SARS2 PLpro was solved, we build the structural model using the homology modelling method. The sequence information between SARS2 and SARS, including secondary structure distribution, disulfide bond position, highly conserved residues, and partially conserved residues alignment were analysed by *ESPript* (Robert & Gouet, 2014). The 3D model was obtained by homologous modelling with SWISS-MODEL (Waterhouse et al., 2018) using the crystal structure of PLpro of SARS (PDBID: 3E9S) as a template.

2.2 | Screening process

The receptor model (model_01.pdb) and ligands were optimized by correcting the bond order, adding hydrogen atoms, distributing charges, and predicting the protonation states (pH 7.0). The ligand library is taken from "TargetMol, USA" (ChemDiv and "Bioactive Compound Library"). From 9175 compounds, 656 compounds with ligand efficiency (LE) <−0.3 and score <−8.0 kcal/mol were obtained by molecule docking using the professional version *ledock* (*ledock_go*) (Wang et al., 2016), then we separate the first pose of docking results of each ligand and apply *prime-mmgbsa* (Jacobson et al., 2004) to each complex to calculate binding free energy (ΔG, kcal/mol). After integrating HBP into binding free energy calculation, a value of 1 in enthalpy calculated by *Lewater* corresponds to about 1.7 kcal/mol in binding free energy (Zhao & Huang, 2011) which is consistent with the experimental value (Fersht et al., 1985), 61 compounds were obtained as active molecules, the screening process can be seen in Figure 1.

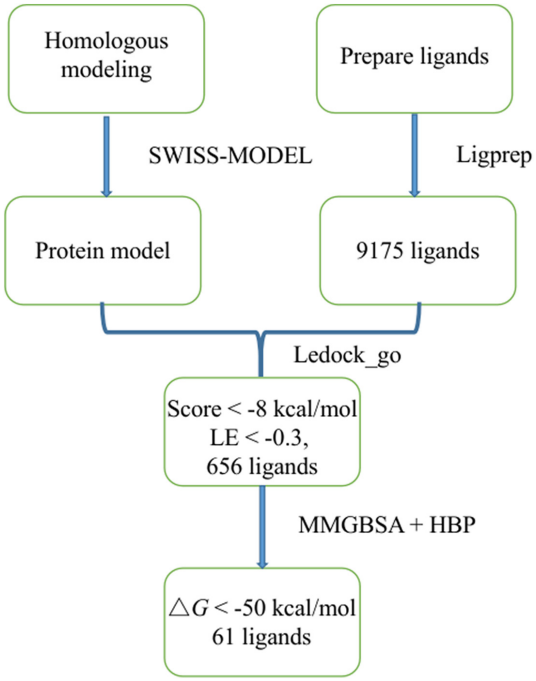


FIGURE 1 Virtual screening workflow

2.3 | Protein expression and purification and inhibitor activity detection

The catalytic domain of SARS2 PLpro used for in vitro compound adduct formation study was sub-cloned into pET-28a vector. This bacterial-expression construct encoded a fusion protein with N-terminal 6×His, protein expression was performed in *Escherichia coli* BL21 (DE3) by induction with 0.1 mM IPTG at 18°C overnight, and purification was achieved using Ni-NTA chromatography. Cells were harvested after 16 h of incubation at 20°C and lysed by sonication in 20 mM phosphate buffer containing 500 mM NaCl, 200 μg/ml of lysozyme, and 0.1 mM PMSF, pH 7.4. Cell lysates were clarified by centrifugation at 13,201 g, 4°C, for 15 min and applied to a Ni-NTA resin. Recombinant protein PLpro was eluted with binding buffer containing 5 mM imidazole and extraction buffer containing 250 mM imidazole, respectively. PLpro was collected in 4 fractions, 5 ml each. The OD280 of the 4 fractions were detected by NanoDrop Spectrophotometer and the fractions with concentration below 0.1 mg/ml were discharged. To further concentrate the PLpro, the protein elution was transferred to a Millipore Amicon Ultra-15 centrifugal filter tube immediately to centrifuge at 4500 *g* at 4°C until there was 5 ml protein left. The protein was diluted with assay buffer containing 200 mM NaH_2PO_4, 200 mM Na_2HPO_4, and pH 6.8.

The purified protease ran at around 36.5 kDa on sodium dodecyl sulfate-polyacrylamide gel electrophoresis, with over 90% purity (data not shown). To find an effective inhibitor, 200 nM PLpro and 20 μM fluorogenic substrate (MCA-Dnp-Lys) in buffer (20 mM Na_2HPO_4–NaH_2PO_4 at pH 6.8) were mixed, and different concentrations of the compounds were added at the incubation temperature of 25°C, shock incubation at room temperature for 15 min, then quickly add 20 μmol/L fluorogenic substrate1 (MCA-Dnp-Lys) which sequence was derived from the amino acid sequence of the viral structural fragment NSP2/3 identified by PLpro. The enhanced fluorescence emission upon substrate cleavage was monitored at the excitation and emission wavelengths of 320 and 425 nm, using SpectraMax iD5, respectively, in the presence or absence of the compounds assayed. Because the active site of PLpro is similar to that of ubiquitin, Cbz-RLRGG-AMC which sequence is derived from five c-terminal residues of ubiquitin was used as the fluorogenic substrate2, 40 nM PLpro and 2.4 μM fluorogenic substrate in buffer (20 mM Na_2HPO_4–NaH_2PO_4 at pH 6.8) were mixed, different concentrations of the compounds were added at the incubation temperature of 25°C, shock incubation at room temperature for 15 min, then quickly add 20 μmol/L fluorogenic substrate2, the enhanced fluorescence emission upon substrate cleavage was monitored at the excitation and emission wavelengths of 360 and 460 nm, using SpectraMax iD5, respectively, in the presence or absence of the compounds assayed. The logarithm of compound concentration was taken as the abscissa, and the corresponding inhibitory rate was taken as the ordinate. The IC_{50} value of the compound was calculated by the four-parameter method.

2.4 | Docking with explicit water molecules

To evaluate the role of ordered water molecules, the crystal water molecules in the binding pocket were included in the docking. We modelled the ordered water molecules based on the template X-ray structure (PDB ID: 3E9S). Five crystal water molecules were included and hydrogen atoms were generated with Chimera (Pettersen et al., 2004). All other parameters were consistent with the docking method in the screening process. The top 3 docked poses for each compound were obtained for further analysis.

2.5 | Conformational verification

To hunt for the correct binding poses, we used metadynamics (Cutrona et al., 2020) to evaluate the docking results. The receptor structure containing 5 crystal water molecules was prepared using *Protein Preparation Wizard* panel (Sastry et al., 2013), including adding hydrogen atoms, removing heteroatoms, and protonation state prediction (pH 7.0). The docked poses were evaluated with the *Binding Pose Metadynamics*. The stability of the binding pose was evaluated with the RMSD over the course of the simulation and the persistence of key interactions between the ligand and the receptor. The docked poses from *ledock_go* for each ligand were evaluated with 10 trials each. To facilitate differentiation, we created an input file for each ligand.

2.6 | Molecular dynamics simulation

We employed the *Desmond* (Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, 2020-2) program to run molecular dynamics simulation based on the OPLS3e force filed (Harder et al., 2016). The protonation states of residues in PLpro and ligands were assigned at pH 7.0. The complexes were solvated in a cubic box of TIP3P water models and NaCl was added to neutralize the system. The system was subjected to 100 ps minimization (10,000 steps with Brownian motion simulation) to adequate equilibrated complexes and solvent molecules. The harmonic position restraints were applied on the backbone of PLpro with a force constant of 5 $kcal.mol^{-1}.Å^{-2}$. A 5 ns constant volume and constant temperature (NVT ensemble) simulation at 300 K was carried out to allow adequate equilibration of the water molecules with harmonic position restraints were applied on the backbone (force constant 10 $kcal.mol^{-1}.Å^{-2}$) and side chains (force constant 5 $kcal.mol^{-1}.Å^{-2}$). The 20 ns constant pressure and constant temperature (NPT ensemble) simulations at 300 K and 1.01325 bar were carried out with 5 $kcal.mol^{-1}.Å^{-2}$ force constant restraint for backbone, and 100 ns equilibrium MD simulation was performed with NPT ensemble at 300 K with a Nose–Hoover chain thermostat and 1 atm with an isotropic Martyna–Tobias–Klein barostat. The time step was set to 2.0 fs. Short-range electrostatics was calculated with a cutoff of 9.0 Å and long-range electrostatics was calculated with QuadS. The trajectories were saved every 20 ps for further analyses.

3 | RESULTS AND DISCUSSION

3.1 | Structural comparison

PLpro of SARS2 and SARS are highly homologous (their identity is 82.9% and similarity is 94.9%) Their active sites

are very similar. The details can be seen in Figure 2. The X-ray structures of SARS and SARS2, and the SARS2-PLpro homology model superimpose very well (with binding site atom RMSD 0.72 Å). The key residues in the active pocket adapt highly similar structure except for Glu167 and Leu162 (Figure 3). It is known that the side chains of Glu167 and Leu162 are highly flexible, were there to fit is typically affected by ligand (Figure S1 and Table S1) (Gavory et al., 2018; Leger et al., 2020; Wang et al., 2020).

3.2 | Virtual screening and enzyme activity experiments

The potential binding sites of water molecules in the active pocket were predicted with *Lewater* (Zhao & Huang, 2011) (for compound **56** shown in Figure 4). Obviously, as the ligands occupy positions of four crystal water, water molecules must be discharged into the solvent, resulting in energy loss. For polar residues around the ligand, all the potential points form a surface, which indicates bound waters distributed at these areas. These regions are exposed to the aqueous solution, and there would be an energy penalty in hydrogen bonds with the ligands if the hydrogen bonds were to exit. Therefore, after binding free energy to rescore by MMGBSA, HBP was used to correct the score values. From 61 hit compounds, 4 of them shown inhibitory activity. The results can be seen in Table 1 and Figures S2 and S3.

3.3 | Poses evaluation

We evaluated the effects of explicit water molecules in docking. With the crystal water molecules included, the docking poses were closer to that of the GRL0617 in the crystal structures 3E9S and 7CJM (Figure 6a–d). Superimposing the poses of 4 compounds with GRL0617

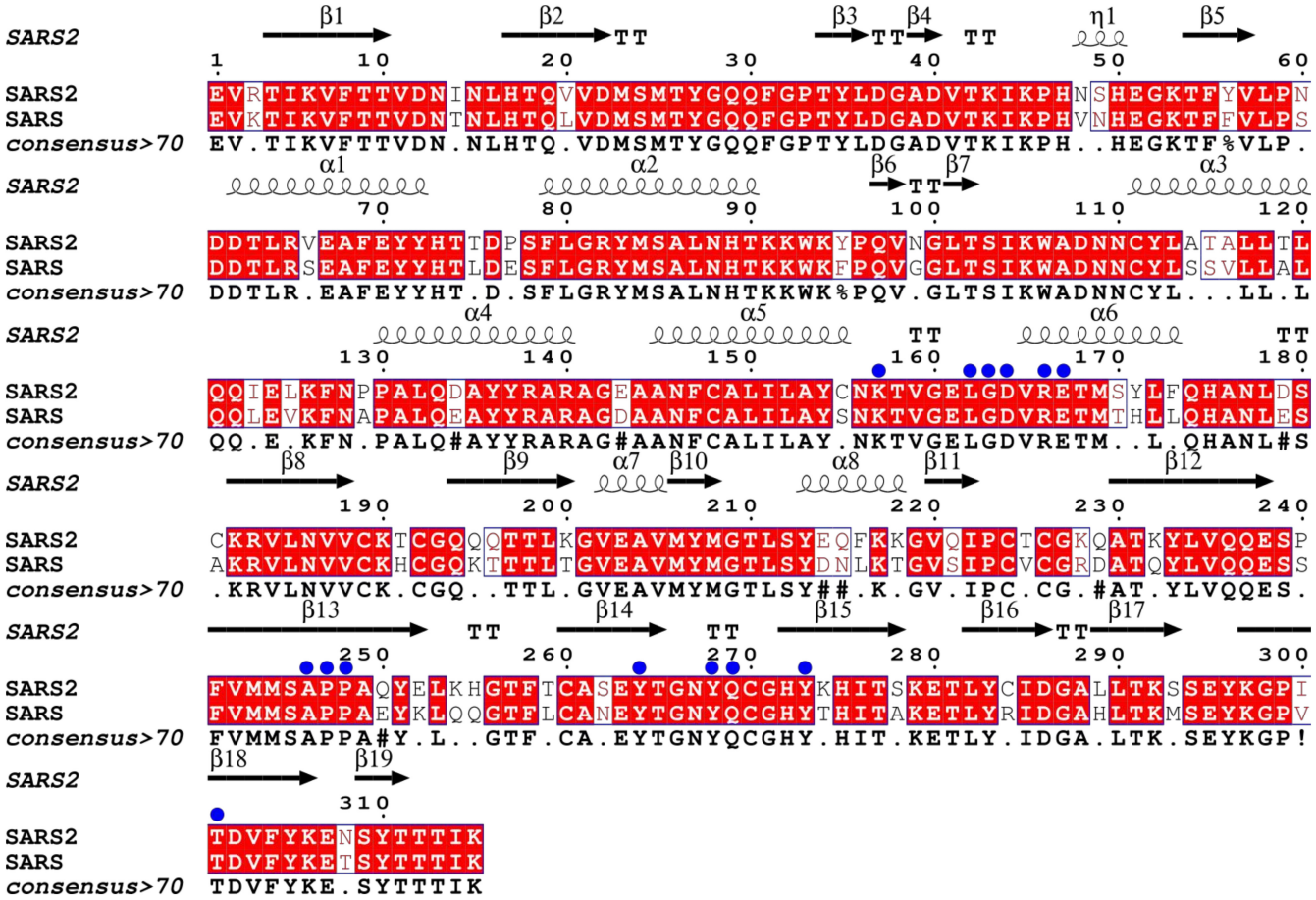


FIGURE 2 PLpro's sequence alignment results of SARS-CoV-2 (SARS2) and SARS-CoV (SARS). α, β, and η are symbols of the secondary structure of SARS2, the arrows represent β-strands, the α represents α-helices, the η represents 310 helices, and "TT" represents strict beta turns. White characters with a red background are the completely conserved amino acids. The red character with white background means the consensus is >70, the black character with white background means the consensus <70. Block uppercase is identity, "!" is anyone of IV, "$" is anyone of LM, "%" is anyone of FY, and "#" is anyone of NDQEBZ. Blue dots represent residues in the active pocket of SARS2 and SARS

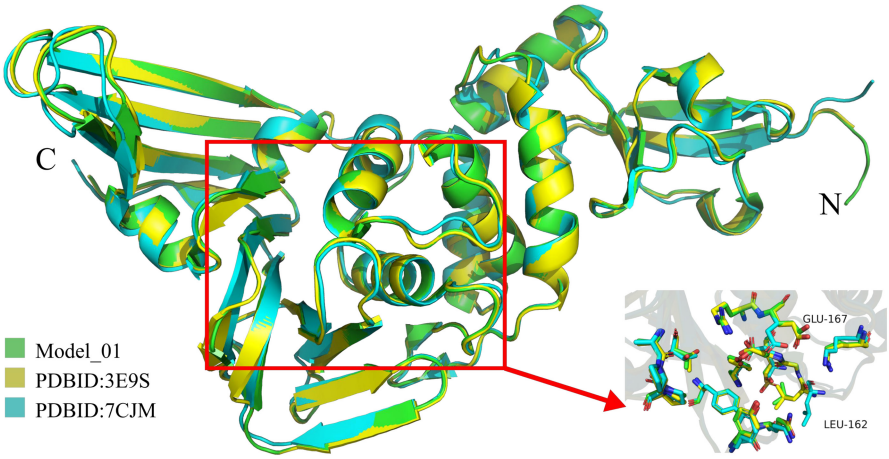


FIGURE 3 Homologous modelling result. Model-01 is the structure model, 3E9S is the crystal structure of SARS and 7CJM is the crystal structure of SARS2. N is the N-terminal and C is the C-terminal. The red box represents the binding site region, and the snapshot on the bottom right shows the distribution of residues in the active pocket

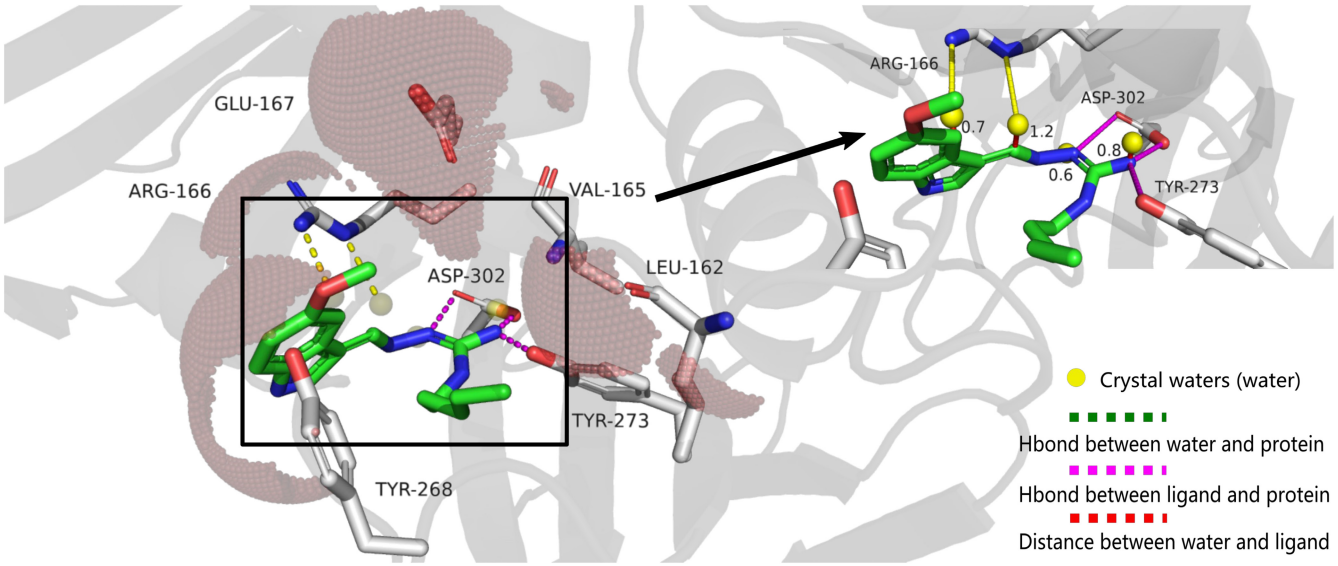


FIGURE 4 The area of bound water is predicted by *Lewater*. The red surface represents the possible location of water molecules. The carbon atoms of the ligand and protein are green and grey, the oxygen atoms are red, and the nitrogen atoms are blue, crystal water molecules were yellow spheres. The ligand is compound **56**

(Figure 6e), a similar trend can be found—they form a "V-shaped" structure, and the turning point occurs at "knee point nitrogen atoms" (Figure 6e, kermesinus oval frame) in the middle of the straight-chain (compound **25** locate at piperazine ring). Interestingly, the coordinates of this knee point are very close, which indicates that the nitrogen atom here is very important for the inhibitory activity of the inhibitor (Fu et al., 2021; Gao et al., 2021).

3.4 | Molecular dynamics simulation analysis

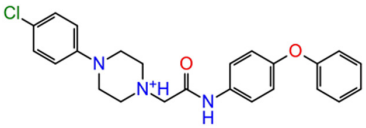
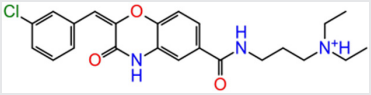
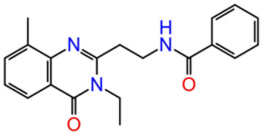
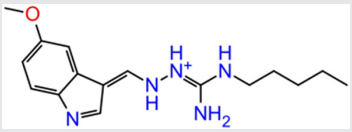
The binding model indicates that the cavity of PLpro is filled with polar residues (Figures 7 and 8). The compound **25** binding with PLpro by π-cation with residue Tyr264. Compound **27** forming salt bridge (with residue Asp164), water bridge (with residue Glu161), π-cation (with residues Tyr264 and Lys157), salt bridge (with residue Asp164), and hydrogen bonds (with residues Asp164 and Gln269) with PLpro. There are hydrophobic interactions (with residues Tyr264 and Pro248, not displayed), hydrogen bonds (with residues Gln269), and Pi–Pi stacking (with residue Tyr268) between compound **38** and PLpro. As for compound **56**, salt bridge (with residue Asp164), π-cation (with residue Tyr264), water bridge (with residue Ala246 and Glu167), and hydrogen bonds (with residues Asp164, Gln269, and Tyr273) in the active pocket of PLpro. From all residues marked in Figures 7 and 8, Tyr264 had the highest frequency between 4 compounds and PLpro, followed by Gln269 (compound **27**, compound **38**, and

compound **56**) and Asp164 (compound **27** and compound **56**), which suggests that they have a relatively important role in maintaining affinity.

In addition to protein-ligand interactions, the internal energy of the ligand is an important factor in assessing binding affinity (Kuhn et al., 2016; Pennington & Moustakas, 2017; Senger et al., 2007). Torsion energy analysis shows that these compounds have 7, 9, 6, and 9 rotational bonds, respectively (Figure S4). The diphenyl ether of compound **25** (Figure S4A) in an extremely flexible state, with the dihedral angle (red and green) sampling extensively between −180° and 180°, the chlorobenzene (deep purple), diethylamine (cyan and pink), and aliphatic chain (wheat, purple, and lime green) of compound **27** (Figure S4B) and aliphatic chain (pink, cyan, and lime green) and methoxy (deep purple) of compound **56** (Figure S4D) sampled much low energy poses, which result in higher RMSD (Figure S5) than compound **38** (Figure S4C). Rotatable bond, linking the "knee point nitrogen atoms" with others, centrally distributed at high energy dihedral angle states for compound **25** (lime green, about 3.68 kcal/mol) and compound **56** (purple, about 4.40 kcal/mol).

The rotatable bonds of compound **25** revealed high entropy (red and green, Figure S4A) and torsional potential energy (lime green, Figure S4A) after entering the binding site. Trajectory analysis shows that compound **38** and compound **56** have lower RMSD (Figure S5) and RMSF

TABLE 1 Chemical structures and inhibitory activity of the SARS-CoV-2 PLpro inhibitors. MCA-Dnp-Lys and Cbz-RLRGG-AMC are different substrates for SARS-CoV-2 PLpro

Compounds	Structure	Score	LE	MMGBSA	HBP	IC_{50} (μM)	IC_{50} (μM)
25		−9.0	−0.30	−65.6	2.7	45.90	60.28
27		−9.5	−0.31	−59.9	2.4	31.03	59.09
38		−8.1	−0.32	−59.4	2.0	19.67	>100
56		−8.7	−0.39	−62.1	3.1	9.25	1.42

Note: Score was calculated by *ledock_go*, the unit of Score, MMGBSA, and HBP is kcal mol^{-1}. LE is the abbreviation of Ligand efficiency. Column 7 is the IC_{50} measured with MCA DNP Lys as the substrate and Column 8 is the IC_{50} measured with Cbz-RLRGG-AMC as the substrate.

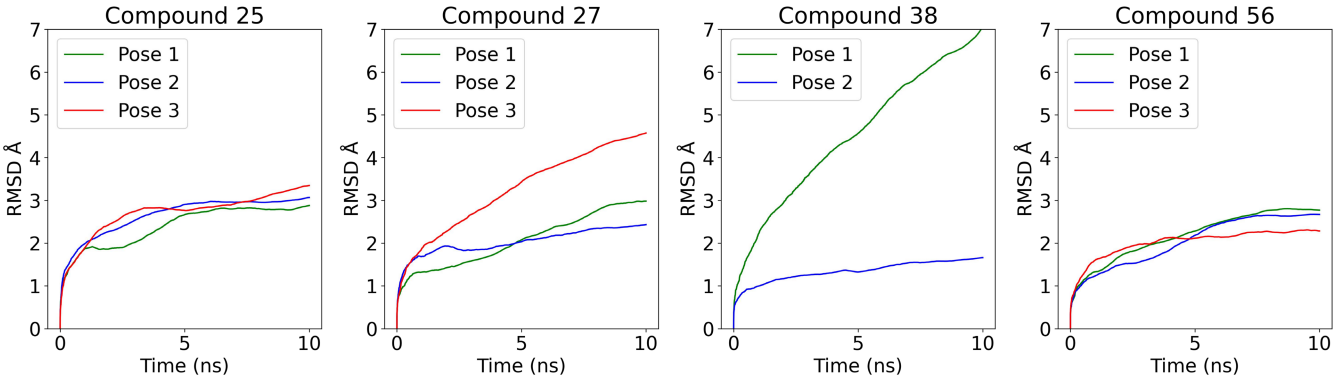


FIGURE 5 Comparison of metadynamics results among 4 compounds. The abscissa is the simulation time of molecular dynamics, and the ordinate is the mean value of RMSD in 10 trials of the compound

TABLE 2 Metadynamics analysis of docking poses for 4 compounds. The persistence of Pi–Pi stacking was zero (not shown)

Number	PoseScore	Persistence	Persistence HBond	CompScore
25-1	2.81	0.36	0.36	1.01
25-2	2.99	0.57	0.57	0.13
25-3	3.21	0.53	0.53	0.55
27-1	2.91	0.36	0.36	1.12
27-2	2.38	0.36	0.36	0.56
27-3	4.41	0	0	4.41
38-1	6.55	0.05	0.05	6.30
38-2	1.61	0.52	0.52	−0.99
56-1	2.79	0.08	0.08	2.38
56-2	2.65	0.24	0.24	1.47
56-3	2.28	0.63	0.63	−0.86

Note: PoseScore was the expectation of the RMSD of the pose over the course of the metadynamics. Persistence was the average persistence of contacts over the course of the metadynamics trajectories. Persistence HBond was the average persistence of HBonds. CompScore was linearly combining the PoseScore and Persistence scores.

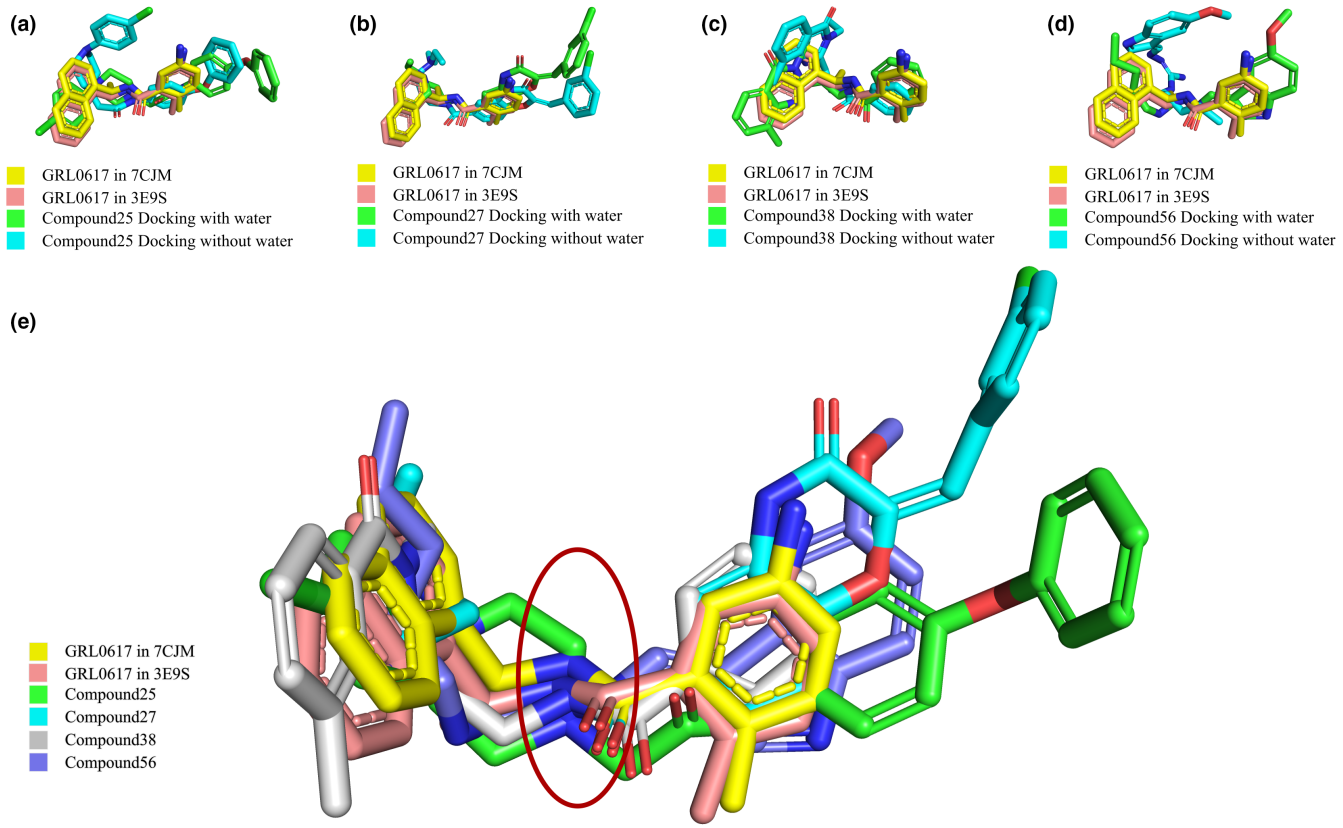


FIGURE 6 (a) Compound **25**. (b) Compound **27**. (c) Compound **38**. (d) Compound **56**. Superimposing poses of regarded crystal water molecules as a part of the protein or not with GRL0617 in 3E9S and 7CJM. (e) Superimposing re-docking poses of 4 compounds with GRL0617 in 3E9S and 7CJM

(Figure S6). During the simulations, compound **25** caused large fluctuations in the protein backbone (Figures S7 and S8), compounds **27** and **56** had comparable effects, and compound **38** caused the smallest fluctuation.

In the statistics of metadynamics interactions, compound **25** had many interactions with PLpro (Table S2), however, an unstable binding pose leads to a low proportion of many interactions in a long time MD process. For

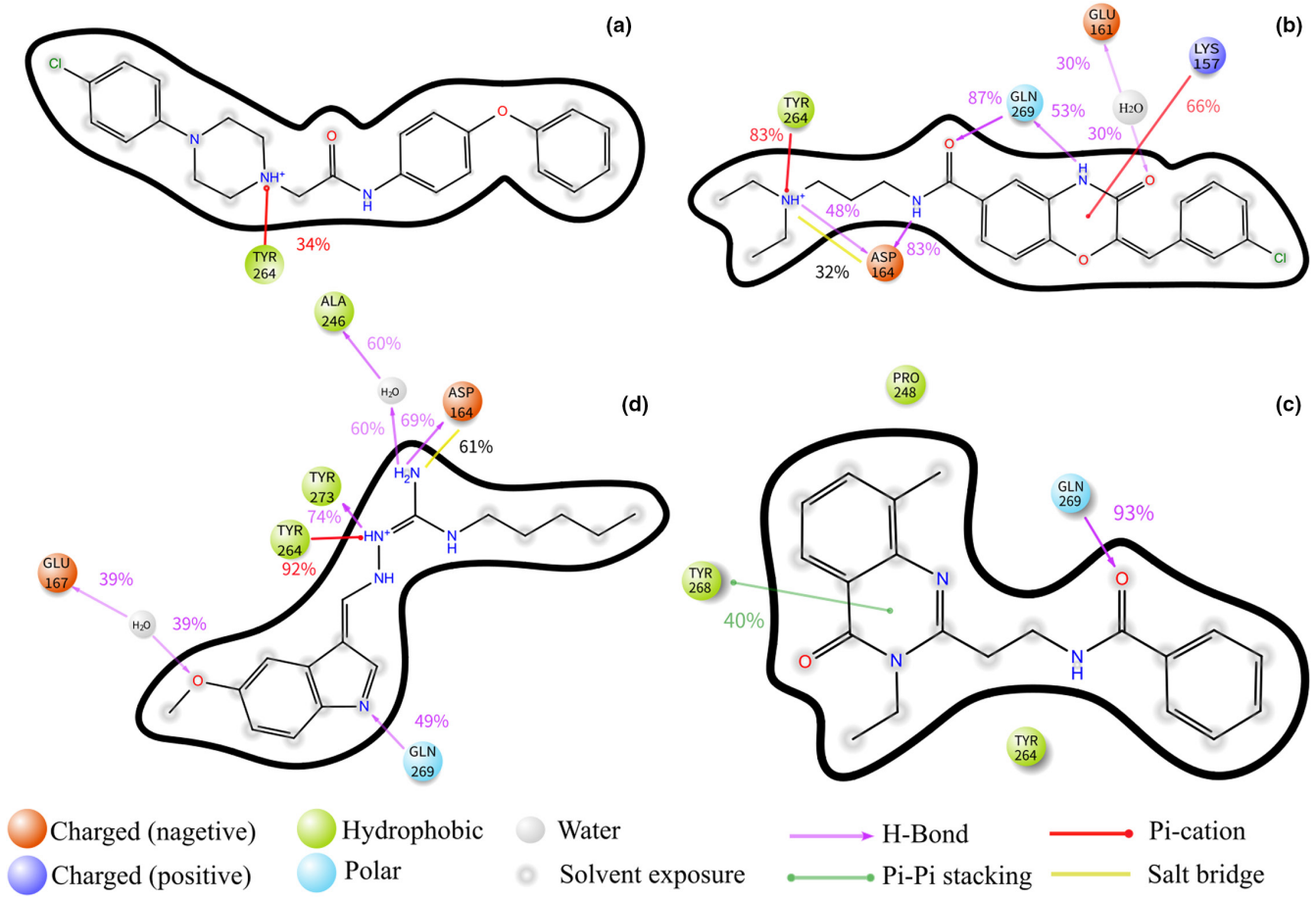


FIGURE 7 Interactions between 4 compounds and PLpro analysed by MD (more than 30% rate is displayed). The percentage represents the probability of total MD process. (a) compound **25**, (b) compound **27**, (c) compound **38**, and (d) compound **56**

example, in the first 10 ns (0–500 frames) of the MD process, the hydrogen bond formed with Gln269 appeared 365 times, accounting for 73.00%, but 1334 times in 100 ns (5000 frames), accounting for 26.68% (data not shown), which further explains the reason why the low activity of compound **25**. As for compound **27**, π-cation interactions with Lys157 appeared 223 times in the first 500 frames, and 4198 times in the whole MD process (5000 frames). Similarly, Tyr264 appeared 369 times and 4183 times in the first 10 ns and all simulation time (100 ns), respectively (data not shown), which indicates that the contribution of these interactions was relatively weak in the early MD but their importance were reflected along with the simulation system gradually stabilized. The same reason can be invoked to explain the case of hydrogen bonds between compound **56** and Asp164 and Gln269 (data not shown), as the results of metadynamics simulations (Table S2), the average value of these interactions in the 10 evaluation results was very low. Although compound **27** and compound **56** were comparable in terms of interaction, the RMSD (Figure S5) and RMSF (Figure S6) of compound **27** were higher than compound **56**, and metadynamics analysis (Figure 5 and Table 2) also confirmed, in binding poses, compound **56** (**56-3**) being more stable than compound **27** (**27-2**). On the contrary, compound **38** showed the lowest RMSD and RMSF, however, a paucity of interactions resulted in its affinity being much lower than that of compound **56**. It is worth mentioning that the torsional potential energy of ligands is a potential factor to improve binding affinity (Engelhardt et al., 2019; Muegge et al., 2017; Sellers et al., 2019) and the potence of compound **56** will be better if the dihedral angle of rotatable bond (purple, Figure S4D) been optimized.

Tegaserod, a 5-hydroxytryptamine (serotonin) type 4 receptor agonist, is used to treat lower gastrointestinal hypomotility associated with irritable bowel syndrome (IBS-C), and exhibits favourable efficacy for the treatments of colitis (Madia et al., 2020). From Figure 7d, the precursor of pentylaminoguanidine (PAG) undertakes the main interactions between tegaserod and PLpro. In previous research, PAG, a product of acid hydrolysis of tegaserod in the stomach, is the crucial consistency that exerts pharmacodynamic effects for IBS-C. SARS2 can cause a

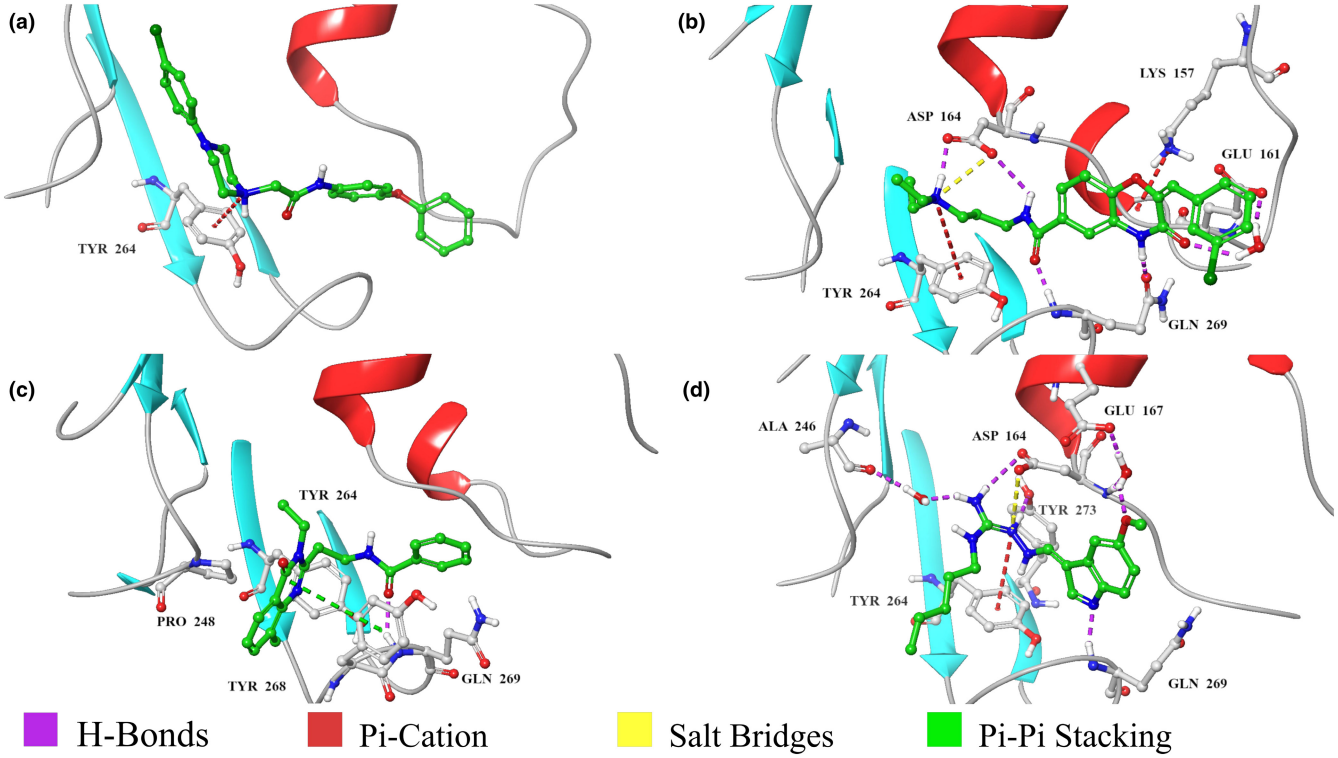


FIGURE 8 3D modes between 4 compounds and PLpro. (a) compound **25**, (b) compound **27**, (c) compound **38**, and (d) compound **56**

severe inflammatory response, which is characterized by the continuous increase of inflammatory cytokines, such as IL-6 and TNF-α (Del Valle et al., 2020). IL-6 can activate JAKs (Janus kinase), which will phosphorylate signal transducer and activator of transcription 3 (STAT3), a crucial transcriptional factor that is overactivated in multiple cancer types and many other diseases. However, tegaserod can down-regulate STAT3 activity by specifically inhibiting the phosphorylation of STAT3 in cancer cells, and this inhibition is independent of the 5-HT4R (Zhang et al., 2020). Thus, tegaserod has potential activity against SARS2 but also has anti-tumor activity, which is good news for cancer patients with SARS2 infection. What's more, inhibition of PLpro activity also contributes to the innate immune function of the host, which is dual effects on the treatment of viral infection (Klemm et al., 2020; Mielech, Chen, et al., 2014; Mielech, Kilianski, et al., 2014). In summary, tegaserod showed a prominent inhibitory effect on PLpro. As a clinical drug, tegaserod may be a good choice for anti-COVID-19.

4 | CONCLUSION

In this work, virtual screening and molecular dynamics simulations were used to discover novel inhibitors for PLpro. Among 61 compounds tested, 4 of them were found to have micromolar inhibitory activity against PLpro. Molecular docking results showed that the existence of bound water molecules had crucial effects on the binding conformation of ligands, and the "knee point nitrogen atoms" is the distinct features of active ligands. The structure–activity relationship suggests that Tyr264 and Gln269 are key residues for PLpro.

ACKNOWLEDGEMENTS

Project supported by the Young Scientists Fund of the National Natural Science Foundation of China (31900910), the Benefiting People Policy-guided science and technology program of Qingdao (Grant No. 20-4-1-4-nsh), Shandong Provincial Major Science and Technology Innovation Project (2018SDKJ0402, 2020CXGC010503).

CONFLICT OF INTEREST

There is no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

ORCID

Haibo Yu https://orcid.org/0000-0002-1099-2803

REFERENCES

Abel, R., Young, T., Farid, R., Berne, B. J., & Friesner, R. A. (2008). The role of the active site solvent in the thermodynamics of

factor Xa-ligand binding. *Journal of the American Chemical Society*, *130*(9), 2817–2831. https://doi.org/10.1021/ja0771033

Bao, J., He, X., & Zhang, J. Z. H. (2020). Development of a new scoring function for virtual screening: APBScore. *Journal of Chemical Information and Modeling*, *60*, 6355–6365. https://doi.org/10.1021/acs.jcim.0c00474

Chou, C. Y., Lai, H. Y., Chen, H. Y., Cheng, S. C., Cheng, K. W., & Chou, Y. W. (2014). Structural basis for catalysis and ubiquitin recognition by the severe acute respiratory syndrome coronavirus papain-like protease. *Acta Crystallographica*, *70*(2), 572–581. https://doi.org/10.1107/S1399004713031040

Cutrona, K. J., Newton, A. S., Krimmer, S. G., Tirado-Rives, J., & Jorgensen, W. L. (2020). Metadynamics as a Postprocessing Method for Virtual Screening with Application to the Pseudokinase Domain of JAK2. *Journal of Chemical Information and Modeling*, *60*(9), 4403–4415. https://doi.org/10.1021/acs.jcim.0c00276

Del Valle, D. M., Kim-Schulze, S., Huang, H.-H., Beckmann, N. D., Nirenberg, S., Wang, B., Lavin, Y., Swartz, T. H., Madduri, D., Stock, A., Marron, T. U., Xie, H., Patel, M., Tuballes, K., Van Oekelen, O., Rahman, A., Kovatch, P., Aberg, J. A., Schadt, E., … Gnjatic, S. (2020). An inflammatory cytokine signature predicts COVID-19 severity and survival. *Nature Medicine*, *26*(10), 1636–1643. https://doi.org/10.1038/s41591-020-1051-9

Engelhardt, H., Böse, D., Petronczki, M., Scharn, D., Bader, G., Baum, A., Bergner, A., Chong, E., Döbel, S., Egger, G., Engelhardt, C., Ettmayer, P., Fuchs, J. E., Gerstberger, T., Gonnella, N., Grimm, A., Grondal, E., Haddad, N., Hopfgartner, B., … McConnell, D. B. (2019). Start selective and rigidify: The discovery path toward a next generation of EGFR tyrosine kinase inhibitors. *Journal of Medicinal Chemistry*, *62*(22), 10272–10293. https://doi.org/10.1021/acs.jmedchem.9b01169

Feig, M., Onufriev, A., Lee, M. S., Im, W., Case, D. A., & Brooks, C. L. (2004). Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. *Journal of Computational Chemistry*, *25*(2), 265–284. https://doi.org/10.1002/jcc.10378

Fell, J. B., Fischer, J. P., Baer, B. R., Blake, J. F., Bouhana, K., Briere, D. M., Brown, K. D., Burgess, L. E., Burns, A. C., Burkard, M. R., Chiang, H., Chicarelli, M. J., Cook, A. W., Gaudino, J. J., Hallin, J., Hanson, L., Hartley, D. P., Hicken, E. J., Hingorani, G. P., … Marx, M. A. (2020). Identification of the clinical development candidate MRTX849, a covalent KRASG12C inhibitor for the treatment of cancer. *Journal of Medicinal Chemistry*, *63*(13), 6679–6693. https://doi.org/10.1021/acs.jmedchem.9b02052

Fersht, A. R., Shi, J. P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M., & Winter, G. (1985). Hydrogen bonding and biological specificity analysed by protein engineering. *Nature*, *314*(6008), 235–238. https://doi.org/10.1038/314235a0

Freitas, B. T., Durie, I. A., Murray, J., Longo, J. E., Miller, H. C., Crich, D., Hogan, R. J., Tripp, R. A., & Pegan, S. D. (2020). Characterization and noncovalent inhibition of the deubiquitinase and deISGylase activity of SARS-CoV-2 papain-like protease. *ACS Infectious Diseases*, *6*(8), 2099–2109. https://doi.org/10.1021/acsinfecdis.0c00168

Friesner, R., Murphy, R., Repasky, M., Frye, L., Greenwood, J., Halgren, T., Sanschagrin, P., & Mainz, D. (2006). Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein—ligand complexes. *Journal of Medicinal Chemistry*, *49*, 6177–6196. https://doi.org/10.1021/jm051256o

Fu, Z., Huang, B., Tang, J., Liu, S., Liu, M., Ye, Y., Liu, Z., Xiong, Y., Zhu, W., Cao, D., Li, J., Niu, X., Zhou, H., Zhao, Y. J., Zhang, G., & Huang, H. (2021). The complex structure of GRL0617 and SARS-CoV-2 PLpro reveals a hot spot for antiviral drug discovery. *Nature Communications*, *12*(1), 488. https://doi.org/10.1038/s41467-020-20718-8

Gao, X., Qin, B., Chen, P., Zhu, K., Hou, P., Wojdyla, J. A., Wang, M., & Cui, S. (2021). Crystal structure of SARS-CoV-2 papain-like protease. *Acta Pharmaceutica Sinica B*, *11*(1), 237–245. https://doi.org/10.1016/j.apsb.2020.08.014

Gavory, G., O'Dowd, C. R., Helm, M. D., Flasz, J., Arkoudis, E., Dossang, A., Hughes, C., Cassidy, E., McClelland, K., Odrzywol, E., Page, N., Barker, O., Miel, H., & Harrison, T. (2018). Discovery and characterization of highly potent and selective allosteric USP7 inhibitors. *Nature Chemical Biology*, *14*(2), 118–125. https://doi.org/10.1038/nchembio.2528

Han, Y. S., Chang, G. G., Juo, C. G., Lee, H. J., Yeh, S. H., Hsu, T. A., & Chen, X. (2005). Papain-like protease 2 (PLP2) from severe acute respiratory syndrome coronavirus (SARS-CoV): Expression, purification, characterization, and inhibition. *Biochemistry*, *44*(30), 10349–10359. https://doi.org/10.1021/bi0504761

Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J. Y., Wang, L., Lupyan, D., Dahlgren, M. K., Knight, J. L., Kaus, J. W., Cerutti, D. S., Krilov, G., Jorgensen, W. L., Abel, R., & Friesner, R. A. (2016). OPLS3: A force field providing broad coverage of drug-like small molecules and proteins. *Journal of Chemical Theory and Computation*, *12*(1), 281–296. https://doi.org/10.1021/acs.jctc.5b00864

Hayes, J. M., Skamnaki, V. T., Archontis, G., Lamprakis, C., Sarrou, J., Bischler, N., Skaltsounis, A.-L., Zographos, S. E., & Oikonomakos, N. G. (2011). Kinetics, in silico docking, molecular dynamics, and MM-GBSA binding studies on prototype indirubins, KT5720, and staurosporine as phosphorylase kinase ATP-binding site inhibitors: The role of water molecules examined. *Proteins*, *79*(3), 703–719. https://doi.org/10.1002/prot.22890

Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J. F., & Friesner, R. A. (2004). A hierarchical approach to all-atom protein loop prediction. *Protns Structure Function & Bioinformatics*, *55*(2), 351–367. https://doi.org/10.1002/prot.10613

Klemm, T., Ebert, G., Calleja, D. J., Allison, C. C., Richardson, L. W., Bernardini, J. P., Lu, B. G., Kuchel, N. W., Grohmann, C., Shibata, Y., Gan, Z. Y., Cooney, J. P., Doerflinger, M., Au, A. E., Blackmore, T. R., Van Der Heden Van Noort, G. J., Geurink, P. P., Ovaa, H., Newman, J., … Komander, D. (2020). Mechanism and inhibition of the papain-like protease, PLpro, of SARS-CoV-2. *The EMBO Journal*, *39*, e106275. https://doi.org/10.15252/embj.2020106275

Kuhn, B., Guba, W., Hert, J., Banner, D., Bissantz, C., Ceccarelli, S., Haap, W., Körner, M., Kuglstatter, A., Lerner, C., Mattei, P., Neidhart, W., Pinard, E., Rudolph, M. G., Schulz-Gasch, T., Woltering, T., & Stahl, M. (2016). A real-world perspective on molecular design. *Journal of Medicinal Chemistry*, *59*(9), 4087–4102. https://doi.org/10.1021/acs.jmedchem.5b01875

Leger, P. R., Hu, D. X., Biannic, B., Bui, M., Han, X., Karbarz, E., Maung, J., Okano, A., Osipov, M., Shibuya, G. M., Young,

K., Higgs, C., Abraham, B., Bradford, D., Cho, C., Colas, C., Jacobson, S., Ohol, Y. M., Pookot, D., ... Wustrow, D. J. (2020). Discovery of potent, selective, and orally bioavailable inhibitors of USP7 with in vivo antitumor activity. *Journal of Medicinal Chemistry*, *63*(10), 5398–5420. https://doi.org/10.1021/acs.jmedchem.0c00245

Madia, V. N., Messore, A., Saccoliti, F., Tudino, V., De Leo, A., De Vita, D., Bortolami, M., Scipione, L., Pindinello, I., Costi, R., & Di Santo, R. (2020). Tegaserod for the treatment of irritable bowel syndrome. *Anti-Inflammatory & Anti-Allergy Agents in Medicinal Chemistry*, *19*(4), 342–369. https://doi.org/10.2174/1871523018666190911121306

Mahmoud, A. H., Masters, M. R., Yang, Y., & Lill, M. A. (2020). Elucidating the multiple roles of hydration for accurate protein-ligand binding prediction via deep learning. *Communications Chemistry*, *3*(19), 1–19. https://doi.org/10.1038/s42004-020-0261-x

Mielech, A. M., Chen, Y., Mesecar, A. D., & Baker, S. C. (2014). Nidovirus papain-like proteases: Multifunctional enzymes with protease, deubiquitinating and deISGylating activities. *Virus Research*, *194*, 184–190. https://doi.org/10.1016/j.virusres.2014.01.025

Mielech, A. M., Kilianski, A., Baez-Santos, Y. M., Mesecar, A. D., & Baker, S. C. (2014). MERS-CoV papain-like protease has deISGylating and deubiquitinating activities. *Virology*, *450–451*, 64–70. https://doi.org/10.1016/j.virol.2013.11.040

Muegge, I., Bergner, A., & Kriegl, J. M. (2017). Computer-aided drug design at Boehringer Ingelheim. *Journal of Computer-Aided Molecular Design*, *31*(3), 275–285. https://doi.org/10.1007/s10822-016-9975-3

Osipiuk, J., Azizi, S.-A., Dvorkin, S., Endres, M., Jedrzejczak, R., Jones, K. A., Kang, S., Kathayat, R. S., Kim, Y., Lisnyak, V. G., Maki, S. L., Nicolaescu, V., Taylor, C. A., Tesar, C., Zhang, Y.-A., Zhou, Z., Randall, G., Michalska, K., Snyder, S. A., ... Joachimiak, A. (2021). Structure of papain-like protease from SARS-CoV-2 and its complexes with non-covalent inhibitors. *Nature Communications*, *12*(1), 743. https://doi.org/10.1038/s41467-021-21060-3

Pennington, L. D., & Moustakas, D. T. (2017). The necessary nitrogen atom: A versatile high-impact design element for multiparameter optimization. *Journal of Medicinal Chemistry*, *60*(9), 3552–3579. https://doi.org/10.1021/acs.jmedchem.6b01807

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Ratia, K., Pegan, S., Takayama, J., Sleeman, K., Coughlin, M., Baliji, S., Chaudhuri, R., Fu, W., Prabhakar, B. S., & Johnson, M. E. (2008). A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(42), 16119–16124. https://doi.org/10.1073/pnas.0805240105

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, *42*(W1), W320–W324. https://doi.org/10.1093/nar/gku316

Rut, W., Lv, Z., Zmudzinski, M., Patchett, S., Nayak, D., Snipas, S. J., El Oualid, F., Huang, T. T., Bekes, M., Drag, M., & Olsen, S. K. (2020). Activity profiling and crystal structures of inhibitor-bound SARS-CoV-2 papain-like protease: A framework for anti-COVID-19 drug design. *Science Advances*, *6*(42), eabd4596. https://doi.org/10.1126/sciadv.abd4596

Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design*, *27*(3), 221–234. https://doi.org/10.1007/s10822-013-9644-8

Schneider, G. (2010). Virtual screening: An endless staircase? *Nature Reviews. Drug Discovery*, *9*(4), 273–276. https://doi.org/10.1038/nrd3139

Sellers, B. D., James, N. C., & Gobbi, A. (2019). Correction to "A comparison of quantum and molecular mechanical methods to estimate strain energy in druglike fragments". *Journal of Chemical Information and Modeling*, *59*(4), 1680. https://doi.org/10.1021/acs.jcim.9b00216

Senger, S., Chan, C., Convery, M. A., Hubbard, J. A., Shah, G. P., Watson, N. S., & Young, R. J. (2007). Sulfonamide-related conformational effects and their importance in structure-based design. *Bioorganic & Medicinal Chemistry Letters*, *17*(10), 2931–2934. https://doi.org/10.1016/j.bmcl.2007.02.034

Shen, Z., Ratia, K., Cooper, L., Kong, D., Lee, H., Kwon, Y., Li, Y., Alqarni, S., Huang, F., Dubrovskyi, O., Rong, L., Thatcher, G. R., & Xiong, R. (2021). Potent, Novel SARS-CoV-2 PLpro inhibitors block viral replication in monkey and human cell cultures. *BioRxiv: The Preprint Server for Biology*. https://doi.org/10.1101/2021.02.13.431008

Shin, D., Mukherjee, R., Grewe, D., Bojkova, D., Baek, K., Bhattacharya, A., Schulz, L., Widera, M., Mehdipour, A. R., Tascher, G., Geurink, P. P., Wilhelm, A., van der Heden van Noort, G. J., Ovaa, H., Müller, S., Knobeloch, K.-P., Rajalingam, K., Schulman, B. A., Cinatl, J., ... Dikic, I. (2020). Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature*, *587*(7835), 657–662. https://doi.org/10.1038/s41586-020-2601-5

Singhal, T. (2020). A review of Coronavirus Disease-2019 (COVID-19). *Indian Journal of Pediatrics*, *87*(10223), 281–286. https://doi.org/10.1007/s12098-020-03263-6

Thilagavathi, R., & Mancera, R. L. (2010). Ligand-protein cross-docking with water molecules. *Journal of Chemical Information and Modeling*, *50*(3), 415–421. https://doi.org/10.1021/ci900345h

Trujillo, J. I., Kiefer, J. R., Huang, W., Day, J. E., Moon, J., Jerome, G. M., Bono, C. P., Kornmeier, C. M., Williams, M. L., Kuhn, C., Rennie, G. R., Wynn, T. A., Carron, C. P., & Thorarensen, A. (2012). Investigation of the binding pocket of human hematopoietic prostaglandin (PG) D2 synthase (hH-PGDS): A tale of two waters. *Bioorganic & Medicinal Chemistry Letters*, *22*(11), 3795–3799. https://doi.org/10.1016/j.bmcl.2012.04.004

Wang, R., Lai, L., & Wang, S. (2002). Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *Journal of Computer-Aided Molecular Design*, *16*(1), 11–26. https://doi.org/10.1023/a:1016357811882

Wang, Z., Sun, H., Yao, X., Li, D., & Hou, T. (2016). Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: The prediction accuracy of sampling power and scoring power. *Physical Chemistry Chemical Physics Pccp*, *18*(18), 12964–12975. https://doi.org/10.1039/c6cp01555g

Wang, Z., Wang, X., Kang, Y., Zhong, H., Shen, C., Yao, X., Cao, D., & Hou, T. (2020). Binding affinity and dissociation pathway

predictions for a series of USP7 inhibitors with pyrimidinone scaffold by multiple computational methods. *Physical Chemistry Chemical Physics*, *22*(10), 5487–5499. https://doi.org/10.1039/D0CP00370K

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*(W1), W296–W303. https://doi.org/10.1093/nar/gky427

Young, T., Abel, R., Kim, B., Berne, B. J., & Friesner, R. A. (2007). Motifs for molecular recognition exploiting hydrophobic enclosure in protein–ligand binding. *Proceedings of the National Academy of Sciences*, *104*(3), 808–813. https://doi.org/10.1073/pnas.0610202104

Zhang, L., Song, Q., Zhang, X., Li, L., Xu, X., Xu, X., Li, X., Wang, Z., Lin, Y., Li, X., Li, M., Su, F., Wang, X., Qiu, P., Guan, H., Tang, Y., Xu, W., Yang, J., & Zhao, C. (2020). Zelnorm, an agonist of 5-Hydroxytryptamine 4-receptor, acts as a potential antitumor drug by targeting JAK/STAT3 signaling. *Investigational New Drugs*, *38*(2), 311–320. https://doi.org/10.1007/s10637-019-00790-8

Zhao, H., & Huang, D. (2011). Hydrogen bonding penalty upon ligand binding. *PLoS One*, *6*, e19923. https://doi.org/10.1371/journal.pone.0019923

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zhao, G., Liu, X., Wang, S., Bai, Z., Zhang, S., Wang, Y., Yu, H., & Xu, X. (2022). Hydrogen bonding penalty used for virtual screening to discover potent inhibitors for Papain-Like cysteine proteases of SARS-CoV-2. *Chemical Biology & Drug Design*, *00*, 1–13. https://doi.org/10.1111/cbdd.14115