Taylor & Francis
Taylor & Francis Group

Check for updates

A multi-stage virtual screening of FDA-approved drugs reveals potential inhibitors of SARS-CoV-2 main protease

Yasmine M. Mandour[a,b] (iD), Darius P. Zlotos[a] and M. Alaraby Salem[c]

[a]Pharmaceutical Chemistry Department, Faculty of Pharmacy and Biotechnology, The German University in Cairo, Cairo, Egypt; [b]School of Life and Medical Sciences, University of Hertfordshire hosted by Global Academic Foundation, New Administrative Capital, Cairo, Egypt; [c]Department of Pharmaceutical Chemistry, Faculty of Pharmacy, October University of Modern Sciences and Arts (MSA), Giza, Egypt

ABSTRACT

Coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an ongoing global health emergency. Repurposing of approved pharmaceutical drugs for COVID-19 treatment represents an attractive approach to quickly identify promising drug candidates. SARS-CoV-2 main protease (M^{pro}) is responsible for the maturation of viral functional proteins making it a key antiviral target. Based on the recently revealed crystal structures of SARS-CoV-2 M^{pro}, we herein describe a multi-stage virtual screening protocol including pharmacophore screening, molecular docking and protein-ligand interaction fingerprints (PLIF) post-docking filtration for efficient enrichment of potent SARS-CoV-2 M^{pro} inhibitors. Potential hits, along with a cocrystallized control were further studied via molecular dynamics. A 150-ns production trajectory was followed by RMSD, free energy calculation, and H-bond analysis for each compound. The applied virtual screening protocol led to identification of five FDA-approved drugs with promising binding modes to key subsites of the substrate-binding pocket of SARS-CoV-2 M^{pro}. The identified compounds belong to different pharmaceutical classes, including several protease inhibitors, antineoplastic agents and a natural flavonoid. The drug candidates discovered in this study present a potential extension of the recently reported SARS-CoV-2 M^{pro} inhibitors that have been identified using other virtual screening protocols and may be repurposed for COVID-19 treatment.

1. Introduction

Coronaviruses (CoV) have been described for more than five decades since the isolation of the prototype murine strain, JHM, which was reported in 1947 (Bailey et al., 1949). Although numerous studies of their replication mechanism and pathogenesis have been reported since the 1970s, the family of CoVs received much attention when a new human CoV was recognized to be responsible for the highly contagious and fatal disease, severe acute respiratory syndrome (SARS) (Drosten et al., 2003; Rota et al., 2003).

Coronaviruses belong to the family Coronaviridae, and are composed of an enveloped single-stranded, positive-sense RNA with a genome ranging from 26 to 32 kb in length (Payne, 2017). They can be classified into four genera: alpha, beta, delta, and gamma, out of which alpha and beta CoVs are known to infect humans (Payne, 2017). The recent pandemic COVID-19 is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) which belongs to beta CoVs.(Li et al., 2020; Xu et al., 2020). The genomic sequence of SARS-CoV-2 shows 76.5% identity to that of SARS-CoV (Y. Z. Zhang & Holmes, 2020). COVID-19 is a highly contagious disease resulting in an unprecedented number of infected

people. World Health Organization has declared COVID-19 as a public health emergency of international concern (WHO, 2020). Despite the efforts to discover potential therapeutics, there are currently no approved COVID-19 drugs or vaccines. The only exception is remdesvir, an RNA dependent RNA polymerase inhibitor, for which Gilead Science was granted Emergency Use Authorization by the FDA on May, 1st, 2020 (FDA, 2020). Repurposing of approved drugs for COVID-19 treatment is an attractive approach to quickly identify promising drug candidates that have been already optimized in terms of pharmacokinetics and limited toxic side effects. Indeed, a recent study demonstrated that each drug from the DrugBank database has on average three different drug targets (J. Wang et al., 2019).

The COVID-19 virus main protease (M^{pro}) is a cysteine protease responsible for proteolysis of replicase polyproteins resulting in the formation of various functional proteins which play a pivotal role in mediating viral replication and transcription (Anand et al., 2003; Hilgenfeld, 2014).The functional importance of M^{pro} in the viral life cycle, together with the absence of closely related homologues in humans identified M^{pro} as an attractive target for antiviral drugs (Bartlam

et al., 2007). Although numerous reports had been published on SARS-CoV M[pro], no protease inhibitor has yet successfully completed a preclinical development program (Anand et al., 2003; Bartlam et al., 2007; Hilgenfeld, 2014). Recently, several X-ray crystal structures of SARS-CoV-2 M[pro] in complex with covalent inhibitors were determined showing high structural similarity to that of SARS-CoV M[pro], as expected from the 96% sequence identity (Jin et al., 2020; L. Zhang et al., 2020). These crystal structures have been recently used in several computational drug repurposing studies (Ancy et al., 2020; Aouidate et al., 2020; Arun et al., 2020; Chatterjee et al., 2020; Elmezayen et al., 2020; Khan et al., 2020; Kumar et al., 2020; J. Wang, 2020). Additional X-ray structure of SARS-CoV-2 M[pro] in complex with reversible dipeptide inhibitor X77 has been recently published (Mesecar, 2020). However, to the best of our knowledge, this latter crystal structure has not been used for virtual screening of drug databases.

The present study describes an alternative drug repurposing approach to identify promising inhibitors of SARS-CoV-2 M[pro]. Based on the crystal structures of SARS-CoV-2 M[pro], a highly selective pharmacophore model was built and used along with docking and PLIF post-docking filtration to screen the currently FDA approved drugs. The binding mode of potential inhibitors was determined and examined for the proper orientation of the pharmacophores in the binding site. Binding modes were further analyzed and compared to a cocrystallized non-covalent ligand by molecular dynamics and molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) free energy calculations. The identified molecules provide new chemotypes for the development of SARS-CoV-2 M[pro] inhibitors and/or may be considered for COVID-19 treatment.

2. Methodology

2.1. Compounds preparation

All molecules were prepared in Molecular Operating Environment (MOE) (MOE, 2019) version 2019.01 by washing, partial charges calculation and energy minimization using the MMFF94x force field and a gradient of 0.0001 kcal/(mol·Å). Multi-conformations of compounds were generated by the stochastic conformational search algorithm implemented in MOE (MOE, 2019) using the following settings: energy window (7 kcal/mol), elimination of duplicate conformer threshold (RMSD, 0.25 Å), total number of iterations (10000 steps), rejection limit (100 steps), MM iteration limit (10000 steps) and maximum conformation limit (10000 conformers).

2.2. Pharmacophore model generation

A structure-based pharmacophore model was constructed based on the two X-ray crystallographic structures of the SARS-CoV-2 M[pro] complexed with N3 (PDB code: 6LU7) (Jin et al., 2020) and X77 (PDB code: 6W63) (Mesecar, 2020). In order to get more information about the active site hotspots, the most favorable energetic properties for ligand binding were determined using the interaction potential map in MOE (MOE, 2019). Interaction energies were calculated using the following chemical probes: aromatic CH group, amide NH group (hydrogen bond donor), carbonyl oxygen atom (hydrogen bond acceptor) and dry (hydrophobic) probe. Regions of favorable interactions were used to define the pharmacophore features. The coordinates of the cocrystallized ligands were used to define the grid box subjected to energy calculation and for pharmacophore feature placement. The generated pharmacophore model was validated by retrospective screening of a test set database composed of 166 molecules. The test set included 12 SARS-CoV M[pro] inhibitors collected from literature (Barretto et al., 2005; Jacobs et al., 2013; Jain et al., 2004; Shie et al., 2005; Shimamoto et al., 2015; Thanigaimalai, Konno, Yamamoto, Koiwai, Taguchi, Takayama, Yakushiji, Akaji, Chen, et al., 2013; Thanigaimalai, Konno, Yamamoto, Koiwai, Taguchi, Takayama, Yakushiji, Akaji, Kiso, et al., 2013; Turlington et al., 2013; Xue et al., 2008; H. Z. Zhang et al., 2006; Zhu et al., 2011) which were labelled as actives (Table S1 in Supporting Information). The remaining 154 molecules were labelled as inactives, comprising 10 weak inhibitors of SARS-CoV M[pro] (Bégué & Bonnet-Delpon, 1991; Gelb et al., 1985; Jain et al., 2004; Shao et al., 2008; Shie et al., 2005) and 144 DUD-E (Mysinger et al., 2012) generated decoys. All molecules were prepared as described in the ligand preparation Section 2.1. The test set compounds were mapped to the pharmacophore model using the pharmacophore search protocol available in MOE. The following metrics were used to evaluate the model performance: sensitivity (Se), specificity (Sp) and enrichment factor (EF).

2.3. Molecular docking

A genetic algorithm based on Cambridge Crystallographic Data Center (CCDC) Genetic Optimization for Ligand Docking (GOLD version 5.8) (Jones et al., 1995; 1997) was employed for molecular docking based on the crystal structure of SARS-CoV-2 M[pro] complexed with the reversible inhibitor "X77" (PDB code: 6W63) (Mesecar, 2020). Binding site residues were defined by specifying the crystal structure ligand coordinates and using the default cutoff radius of 6 Å, with the "detect cavity" option enabled. The docking experiments were performed using the goldscore scoring function. The search efficiency of the genetic algorithm was at 200% setting with the receptor kept rigid. Water molecules were kept in the pocket while allowing the ligand to displace them during the docking experiment. For each compound, 50 complexes were generated and clustered based on their RMSD with the threshold set at 1.5 Å using the complete linkage method. The quality of pose prediction was assessed by calculating the heavy atom RMSD between the docked poses and the original PDB coordinates of X77. The resultant docked poses were rescored using DrugScore (DSX) (version 0.9) (Neudert & Klebe, 2011) utilizing the DrugScore[PDB] potential. All the docked poses were imported into MOE database for calculating their Protein-Ligand Interaction Fingerprints (PLIF) rows which were utilized for generating

amino acids interaction fingerprints using eight types of interactions (sidechain hydrogen bonds (donor or acceptor), backbone hydrogen bonds (donor or acceptor), solvent hydrogen bonds (donor or acceptor), ionic interactions, and π interactions). The cavity used for the PLIF analysis consisted of the same set of residues used in the docking experiments. Finally, the resultant docked poses were filtered using a set of reference PLIFs and the docking protocol validated by retrospective screening of the same test set used in pharmacophore validation. Figures were prepared using Pymol (Schrodinger, LLC, 2015).

2.4. Virtual screening

The validated pharmacophore model was utilized as a 3 D query for screening the FDA approved drug dataset retrieved from Selleckchem Inc. (WA, USA, http://www.selleckchem.com). The dataset comprised 2684 compounds which were prepared as described in the ligand preparation section and mapped to the pharmacophore model such that hit molecules match all the query features. Using GOLD, the identified hit compounds were docked into the active site of SARS-CoV-2 M^{pro} (PDB code: 6W63) (Mesecar, 2020) using goldscore scoring function. The search efficiency of the genetic algorithm was at 200% setting with the receptor kept rigid. Finally, the docked poses were filtered using the predefined PLIF and their binding modes were visually examined to check their similarity to those of SARS-CoV-2 M^{pro} cocrystallized ligands. The selected protein-drug complexes were further subjected to MD simulations for understanding the structural stability of protein-drug complexes at the long-interval.

2.5. Molecular dynamics

2.5.1. Preparation of initial box for MD simulations

Molecular dynamics were performed on 5 of the six most promising hits from our virtual screening in complex with the SARS-CoV-2 M^{pro}; namely: boceprevir, epirubicin, nelfinavir, rutin and thymopentin. Bortezomib was not considered due to the inherent difficulty of deriving parameters for the boron atom. Topology and coordinate files were prepared via Ambertools 14 (D.A. Case, 2020). One extra dynamics trajectory was run for the SARS-CoV-2 M^{pro} in complex with the co-crystallized ligand named X77 (PDB code: 6W63) (Mesecar, 2020), which serves as a control reference. In all trajectories, the dry MOE-prepared protein was used as initial protein coordinates. The starting coordinates for all ligands were provided from the best docked poses. The pdb4amber and reduce programs (Word et al., 1999) implemented in Ambertools (D.A. Case, 2020) where used to convert the files to an AMBER-friendly format. The all-atom force field AMBER ff14SB (Roe & Cheatham, 2013; Turner, 2005) was universally used. As the force field does not contain parameters for the ligand, all ligands were parameterized using ANTECHAMBER (J. Wang et al., 2006) to generate parameters that are consistent with the General Amber Force Field (GAFF) (J. Wang et al., 2004). The AM1-BCC method was used to assign

charges. Crystallographic waters were retained and each protein was solvated with approximately 15192 TIP3P water molecules in a $92.6 \times 69.7 \times 95.4$ Å box which extends 12 Å beyond solute in each direction. Sodium ions were added to neutralize the negatively-charged protein, given the positive charge on nelfinavir and thymopentin, and additional 27 NaCl ions were added to reach a salt concentration of 100 mM to resemble cellular conditions.

2.5.2. MD Simulations and analysis

MD simulations were carried out using the PMEMD.cuda code of the AMBER Molecular Dynamics package (D.A. Case, 2020) following a standard protocol adopted of minimization, heating, density equilibration and production. The AMBER input files are similar to those in the supplementary information of the previous work of Salem and Brown (Alaraby Salem & Brown, 2015) while using 306 as the number of residues (305 protein residues in addition to the ligand) whenever restraints are applied. The trajectory lengths for heating, density equilibration and production were 20 ps, 50 ps and 150 ns, respectively. Extra 100 ns of production where simulated for epirubicin, as explained in the discussion. Prior to heating, we applied two cycles of minimization to each complex. We applied strong restraints on the non-solvent residues, with a force constant of 500 kcal/mol.Å2, in the first minimization run only to relieve unfavorable hydrogen contacts. Heating was employed at 300 K using the Langevin thermostat. Density equilibration and production were conducted at constant pressure (1 atm). Langevin dynamics (Loncharich et al., 1992) were generally employed and the Particle-Mesh Ewald method (Darden et al., 1993) was used to treat long-range electrostatics under periodic boundary conditions. The trajectories were analyzed using CPPTraj (Roe & Cheatham, 2013). Plots and visual inspection of the trajectories was done using XMgrace (Turner, 2005) and VMD (Humphrey et al., 1996), respectively.

2.6. Free energy calculation

A total of 1200 snapshots were evenly sampled from the 150-ns production phase of each trajectory; that is, every 125 ps. In the case of epirubicin, the sampling was repeated in the extra 100 ns production phase, at similar intervals. The molecular mechanics Poisson−Boltzmann surface area (MM-PBSA) (Homeyer & Gohlke, 2012) was used to estimate the binding free energy, as implemented in the MMPBSA.py script available with AMBER (Miller et al., 2012). The entropy contribution was not calculated as the method is used to mainly compare the ligands binding to the same protein, rather than attempting to predict absolute affinities (Genheden & Ryde, 2015). Parameters for internal and external dielectric constants were set to 4 and 80, respectively. For comparison, we also calculated free energy using the generalized-Born surface area method (Genheden & Ryde, 2015). In general, the free energy (G) of the ligand, L, or the protein, P, is computed according to the following equation:

$$G = E_{bind} + E_{el} + E_{vdW} + G_{pol} + G_{np} - TS \qquad (1)$$
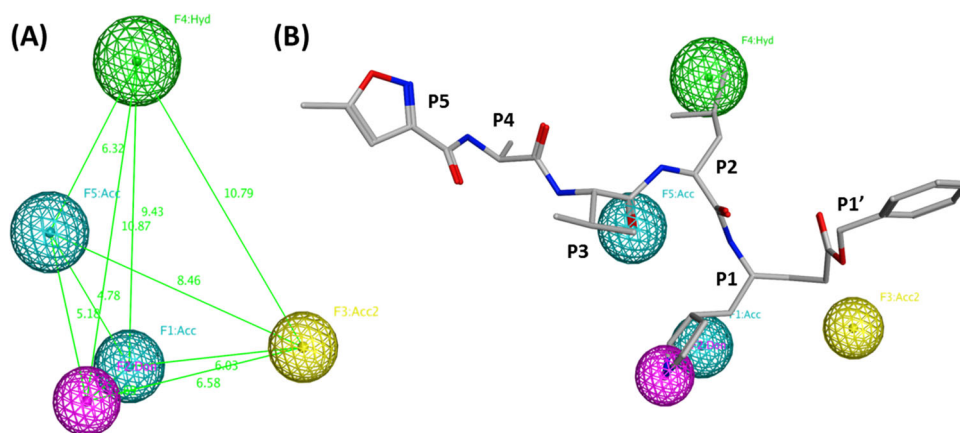
Figure 1. The structure-based pharmacophore model **(A)** 3 D spatial arrangement and distance constraints between the chemical features of the pharmacophore model represented by cyano (hydrogen bond acceptor, Acc), yellow (projection of hydrogen bond acceptor: Acc2), magenta (hydrogen bond donor, Don) and green (hydrophobic, Hyd) spheres. **(B)** Overlay of the crystal coordinates (PDB code: 6LU7) (L. Zhang et al., 2020) of the peptide-like inhibitor; N3; on the pharmacophore model.

where E_{bind}, E_{el} and E_{vdW} are the bonding, electrostatic and van der Waals terms, respectively, as in standard molecular mechanics. The polar and non-polar contributions to solvation energies are G_{pol} and G_{np}, respectively. In MMPBSA, G_{pol} is obtained by solving the PB equation, while in the case of GBSA, the Generalized-Born model is used. T is the absolute temperature and S is the entropy which was excluded from our calculations as mentioned above. We adopted the one-trajectory approach by computing the trajectory for a solvated complex and then computing the difference in free energy:

$$\Delta G_{bind} = \langle G_{PL} - G_P - G_L \rangle \qquad (2)$$

where G_{PL} refers to the protein-ligand complex. This difference must lead to the cancellation of E_{bind} from Equation (1).

3. Results and discussion

The two recently resolved X-ray crystallographic structures of SARS-CoV-2 M^{pro} complexed with N3; an irreversible substrate-like inhibitor (PDB code: 6LU7, resolution 2.16 Å) (L. Zhang et al., 2020) and X77; a reversible dipeptide inhibitor (PDB code: 6W63, resolution 2.1 Å) (Mesecar, 2020) were used in this study for inferring chemical information on inhibitors' binding to SARS-CoV-2 main protease. Overlay of the two structures showed high structural similarity with an average all-atom RMSD of 1.29 Å and 0.693 Å for all protein residues and binding site residues; respectively. The substrate-binding site is located in a cleft formed between domains I and III, hosting a catalytic dyad composed of Cys145 and His41. In the peptdiomimetic inhibitor N3 the residues P1'-P5 bind to complementary subsites S1'-S5 in the substrate-binding pocket of SARS-CoV-2 M^{pro}. Moreover, as the main protease proteins of SARS-CoV (PDB code: 1UK4) (Yang et al., 2003) and SARS-CoV-2 (PDB code:6W63) (Mesecar, 2020) show high structural similarity with an average all-atom RMSD of 1.26 Å and 0.698 Å for all protein residues and binding site residues; respectively. Structural information of previously reported SARS-CoV M^{pro} inhibitors

was used to guide the discovery of novel SARS-CoV-2 M^{pro} inhibitors in this study

3.1. Pharmacophore model

To obtain comprehensive chemical information on inhibitor binding to SARS-CoV-2 M^{pro}, a structure-based pharmacophore model was constructed based on the two X-ray crystallographic structures of M^{pro} complexed with N3 (L. Zhang et al., 2020) and X77 (Mesecar, 2020). The most favorable energetic properties for ligand binding were determined using the interaction potential map in MOE (MOE, 2019). The interaction potential is a grid-based approach developed for scanning the protein surface to identify hot spots. This method is based on determination of the chemical nature of the protein's active site by calculating interaction energies with chemical probes of different electronic properties. Regions of favorable interactions represent hotspots and were used as a basis to define the pharmacophore features. The resultant pharmacophore model comprised features representing mainly S1'- S3 subsites. The generated structure-based model included five features, namely two hydrogen bond acceptors (Acc), one hydrogen bond acceptor projection (Acc2), one hydrogen bond donor (Don) and a hydrophobic center (Figure 1(A)). Overlay of the crystallized coordinates of the substrate-like inhibitor "N3" on the pharmacophore model showed its peptide P1 orientation defined by two features (Figure 1(B)), namely the hydrogen bond interactions of the side chain of P1 with the imidazole ring of His163 and backbone carbonyl of Phe140 represented by F1:Acc and F2:Don; respectively. This complies with the S1 subsite specificity to P1-Gln, similarly to other previously reported M^{pro} inhibitors where anchoring the ligand to this site appeared essential to block the enzyme's catalytic activity (F. Wang et al., 2016; Xue et al., 2008). The hydrogen bond acceptor projection feature (F3:Acc2) represents the hydrogen bond interaction of P1' residue of N3 with the backbone amine of Gly163 in S1' subsite. A hydrophobic feature (F4: Hyd) reflects the side chain of P2 fitting in the S2 subsite lined by side chains of His41 and Met49. The S2

Table 1. Performance evaluation metrics of screening methods.

Screening method	TH	AH	TPR	TNR	EF
Pharmacophore filtration	53	9	0.75	0.71	2.35
PLIF filtration	60	9	0.75	0.67	2.35
Pharmacophore + PLIF filtration	19	9	0.75	0.99	6.55

*Total number of hits (TH), number of active hits (AH), true positive rate (TPR), true negative rate (TNR) and enrichment factor (EF).

subsite in beta coronaviruses is mainly hydrophobic with both aromatic and aliphatic amino acids reported binding to this site (F. Wang et al., 2016). This was also reflected by the high interaction potentials of the dry probe observed in this site. The last feature is a hydrogen bond acceptor (F5: Acc) accounting for the H-bond between the backbone carbonyl of P3 and the backbone amine of Glu166. Retrospective screening was conducted to evaluate the ability of the pharmacophore model to separate active from inactive compounds. To this end, a small-molecule test set was generated for model validation. Thus, 12 peptide-like M^pro inhibitors of diverse structures, that showed high affinity to the protease enzyme and shared a similar binding mode were selected and labeled as actives (Jacobs et al., 2013; Jain et al., 2004; Shie et al., 2005; Shimamoto et al., 2015; Thanigaimalai, Konno, Yamamoto, Koiwai, Taguchi, Takayama, Yakushiji, Akaji, Chen, et al., 2013; Thanigaimalai, Konno, Yamamoto, Koiwai, Taguchi, Takayama, Yakushiji, Akaji, Kiso, et al., 2013; Turlington et al., 2013; Xue et al., 2008; H. Z. Zhang et al., 2006; Zhu et al., 2011). In addition, 10 weak inhibitors of M^pro enzyme were collected from literature (Bégué & Bonnet-Delpon, 1991; Gelb et al., 1985; Jain et al., 2004; Shao et al., 2008; Shie et al., 2005) and labeled as inactives along with 144 DUD-E (Mysinger et al., 2012) generated decoys (see Table S1 in Supporting Information). The multi-conformation test set compounds were mapped to the pharmacophore model. As shown in Table 1, the model could retrieve 9 out of 12 known M^pro inhibitors (sensitivity = 0.75), exclude 110 out of 154 inactives (specificity = 0.71) and showed a mapping EF of 2.35.

3.2. Molecular docking

Molecular docking simulation studies were performed using GOLD 5.8 (Cambridge Crystallographic Data Centre, Cambridge, UK) (Jones et al., 1995; 1997). Only one crystal structure of SARS-CoV-2 M^pro–ligand complex was employed for docking as the binding site did not undergo conformational changes upon binding of different inhibitors and thus the crystal structure of SARS-CoV-2 M^pro complexed with the reversible inhibitor "X77" (PDB code: 6W63) (Mesecar, 2020) was used. To validate the docking protocol, the cocrystallized ligand was docked into its binding pocket of M^pro enzyme. All the resultant poses converged to a binding mode similar to that of the experimentally determined position of X77 with the best ranking pose having a root-mean square deviation (RMSD) value of 0.52 Å (Table S2, Supporting Information). An overlay of the docked pose and crystal structure of X77 is shown in the Supporting Information (Figure S1). The binding orientation of X77 is overall similar to known covalent substrate-like inhibitors (F. Wang et al.,

2016; L. Zhang et al., 2020), where it appeared preferentially occupying the S1'-S3 subsites of SARS-CoV-2 M^pro enzyme. In this orientation, the cyclohexyl amide occupied S3 subsite, the tert-butyl phenyl group occupied a deep S2 subsite and the imidazole amide and the 3-pyridyl group occupied S1' and S1 subsites; respectively.

Docking of the test set compounds was conducted to further assess the ability of the docking protocol to discriminate active from inactive compounds. Goldscore showed poor enrichment performance with an EF at 1% and 5% of the ranked list of 0 and 1.87, respectively. The docked poses were rescored using DSX^PDB (Neudert & Klebe, 2011) which did not show any improvement in the EF values compared to Goldscore. Accordingly, the docked poses were assessed using protein-ligand interaction fingerprints (PLIFs) which were previously reported to give better results than standard scoring functions in terms of identifying the correct binding modes of the ligand and recovering active compounds in virtual screening (VS) trials (Anighoro & Bajorath, 2016a, 2016b; Da & Kireev, 2014). Interactions with key residues targeting the enzyme catalytic residues (Cys145 and His 41) in addition to Glu166 located at the center of the substrate-binding site ensures a binding mode accessing the key enzyme subsites. Filtering the docked poses of the test set compounds based on these PLIFs correctly identified 9 out of 12 known M^pro inhibitors (sensitivity = 0.75), excluded 103 out of 154 inactives (specificity = 0.67) with a mapping EF of 2.35 (Table 1). These metrics reflect the individual performance of PLIF filtration which appeared similar to those of the pharmacophore model. Cascading the filters; pharmacophore, docking and PLIF in a multi-stage manner; could decrease the false-positive rate. After pharmacophore virtual screening, 9 of 12 actives and 110 out of 154 inactives were correctly predicted. Accordingly, a test set composed of 9 actives and 44 inactives was selected as the initial set for docking and PLIF filtration. The multi-level VS showed better enrichment than either of the screening methods alone retrieving all 9 actives (sensitivity = 0.75), correctly excluding 34 inactives (specificity = 0.06) with a significant increase in the mapping EF value of 6.55 (Table 1). The need for sequential filtration can be explained by the large variation in the chemical space and size of reported SARS-CoV M^pro inhibitors along with the large size of the substrate-binding pocket of M^pro enzyme and accordingly was used for prospective screening.

3.3. Virtual screening

Virtual screening was carried out hierarchically using the structure-based pharmacophore model, molecular docking along with poses selection using the predefined PLIFs and finally, visual inspection of the binding modes of the hit compounds. The FDA approved drug dataset comprising 2684 compounds was retrieved from Selleckchem Inc. (WA, USA, http://www.selleckchem.com). An extensive stochastic conformational search was conducted for all compounds with the resultant conformers screened using the pharmacophore model described earlier narrowing down the database to 158 compounds. The hit compounds were docked into
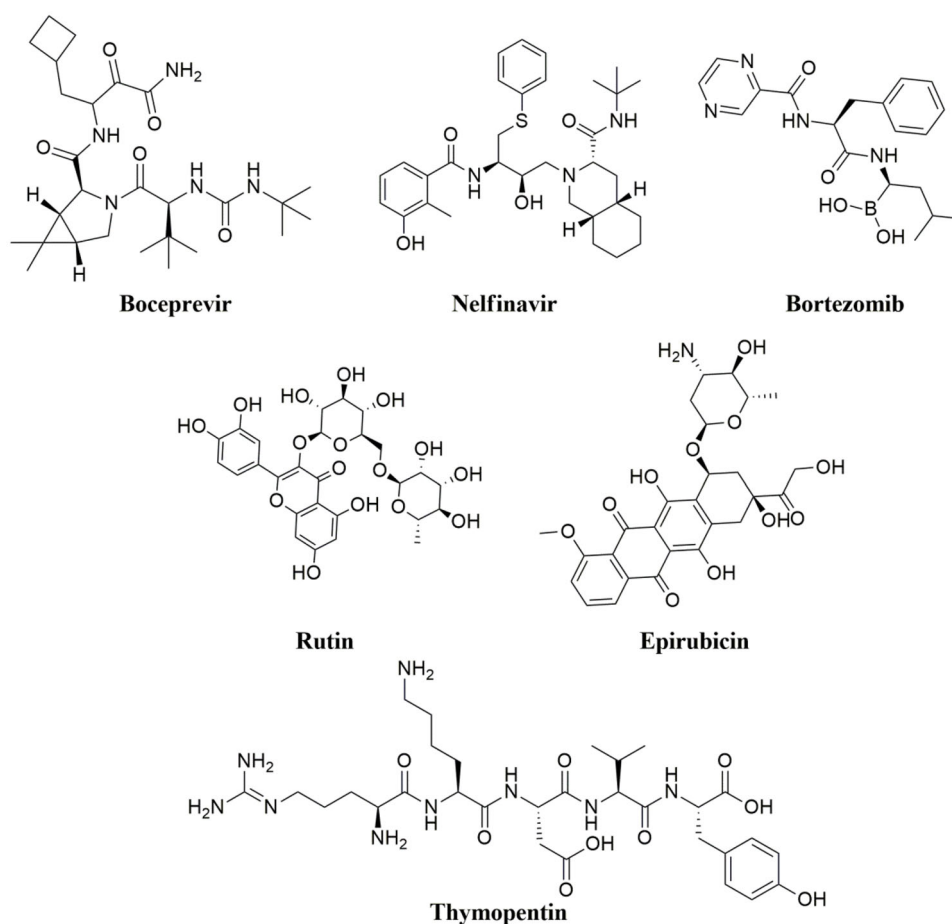
Figure 2. Chemical structures of promising docking hits.

the active site of SARS-CoV-2 M^pro (PDB: 6W63) (Mesecar, 2020) using GOLD (Jones et al., 1995; 1997) with filtration of the docked poses based on the predefined PLIFs reducing the number of hits to 53 compounds. The docked poses of these hits were visually examined using the structural information of the binding modes of previous SARS-CoV M^pro inhibitors for evaluating their binding modes. Finally, 12 compounds revealed promising binding modes to SARS-CoV-2 M^pro enzyme. The selected hits comprised molecules approved by the FDA for one or multiple therapeutic indications as: HIV and HCV infections, antineoplastic agents and natural compounds of multiple indications (Table S3 in the supporting information). Six compounds were particularly interesting because of the high resemblance between their binding modes and those of SARS-CoV-2 M^pro crystallized ligands; X77 and N3. With the exception of bortezomib, molecular dynamics simulations were conducted for these promising docking hits. (Figure 2).

3.4. Molecular dynamics and binding modes analysis

For the promising docking hits, molecular dynamics simulations were conducted using the AMBER software package (D.A. Case, 2020). In total, 6 ligands including the co-crystal X77 ligand were studied and the stability of each MD system was explored. Figure 3(A) showed the RMSD fluctuations of the backbone atoms of the receptor along the MD

simulation time relative to the initial structure. In each complex, it appears that the receptor reached a stable equilibrium after almost 25 ns. To further understand the dynamics of the backbone atoms, the root mean square fluctuation (RMSF) values were calculated for backbone atoms at each time point of the trajectories. Higher RMSF values indicate greater flexibility during the MD simulation. Low RMSF values were observed for all protein complexes indicating their high stabilities during the entire MD simulation with the overall topology similar to the native fold. (Figure S3) The observed RMSD of 3–4 Å is primarily attributed to variations of residues at the termini showing RMSF values > 5 Å. As illustrated in Figure 3(B) for the RMSD of the six ligands, the co-crystallized X77 showed the best stability followed by rutin and boceprevir with RMSDs around 2 Å which is reasonable for large ligands like SARS-CoV-2 main protease inhibitors. These small RMSD fluctuations add confidence to the binding modes of rutin and boceprevir proposed by GOLD docking experiments. On the other hand, thymopentin showed the least stability as shown by its high RMSD. Visual inspection of the trajectory confirmed that thymopnetin completely detached from the binding site after 50 ns and remains in the solution. Accordingly, thymopentin was not further considered in this study. For the other 4 hits, the average structure of the collected snapshots was generated, and the MD snapshot with the highest structural similarity with this average structure was chosen as the representative conformation.
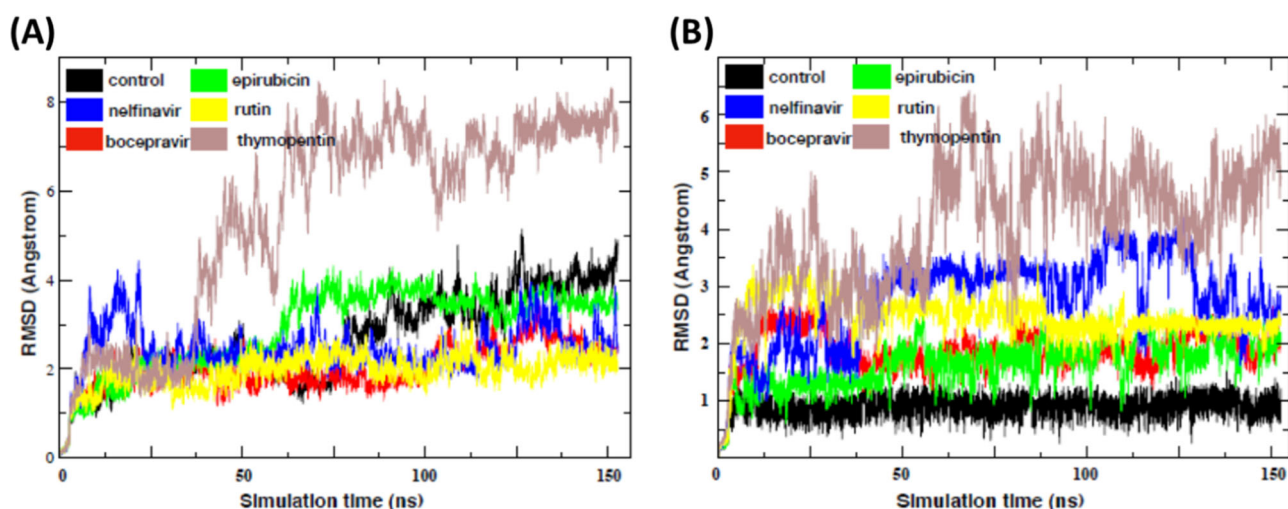
Figure 3. Plots of root-mean-square deviations of **(A)** protein main chain atoms and **(B)** ligand heavy atoms along the MD simulation time.
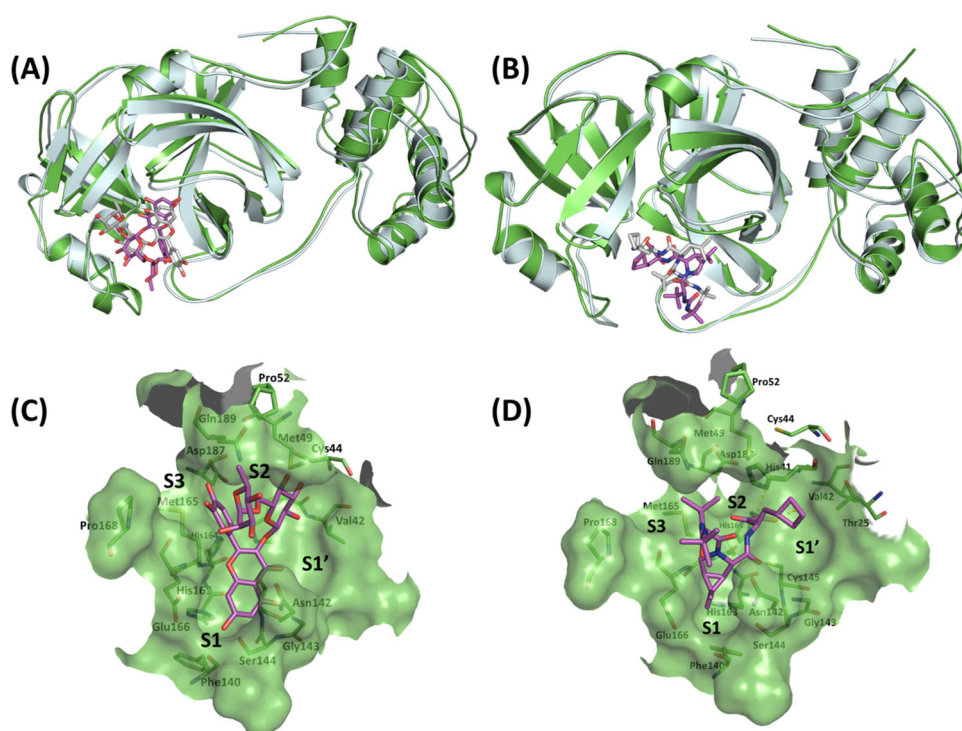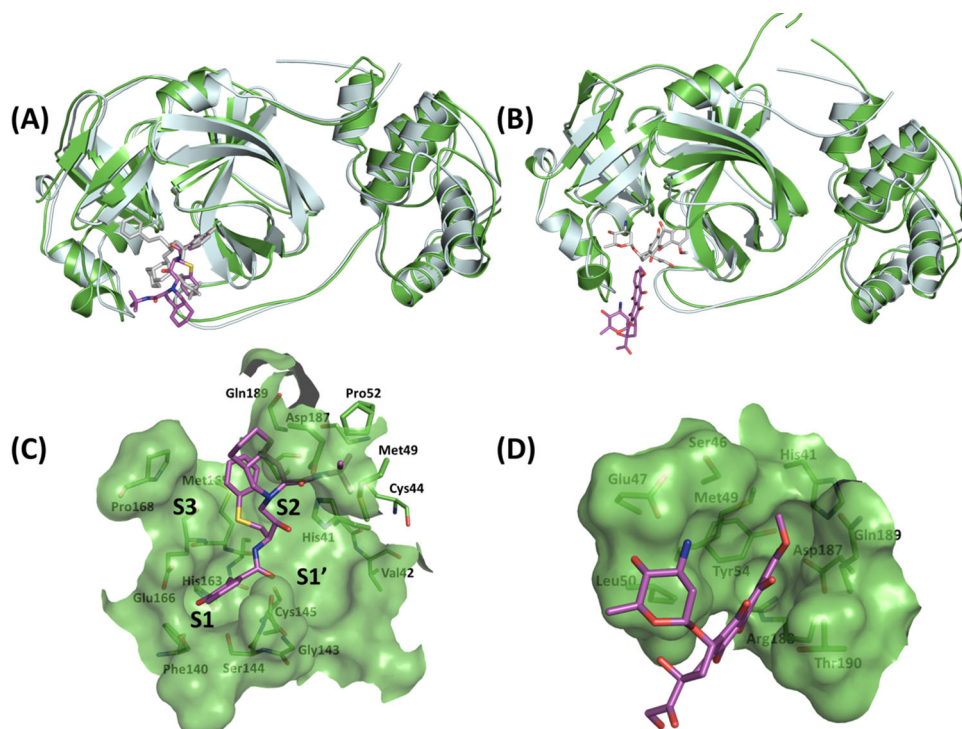

Figure 4. Structural comparison between the crystal structure and a representative MD structure of SARS-CoV-2 main protease bound to Rutin and Boceprevir. The crystal structure is shown as light cyano cartoon with the docked ligand shown as grey sticks, while the representative MD structure is shown in green cartoon and the ligand as magenta sticks. **(A)** Rutin and **(B)** Boceprevir. Surface representation of the substrate-binding pocket of SARS-CoV-2 M^{pro} and ligand − receptor interactions are shown for **(C)** Rutin and **(D)** Boceprevir.

The comparisons between the crystal structure and the representative MD conformations are shown in Figure 4 for Rutin and Boceprevir and in Figure 5 for Nelfinavir and Epirubicin.

Below, we describe in detail the binding poses of the promising hits from our virtual screening, based on collective evidence from docking and molecular dynamics. H-bond analysis has been performed on the production MD trajectory and the results are provided in Table 2. Flavonoids were previously identified as potential inhibitors of SARS-CoV M^{pro} showing inhibitory activity within the micromolar range (L. Chen et al., 2006; Yi et al., 2004). The natural flavonoid rutin was among the identified hits maintaining almost the same

binding mode as its docked pose throughout the entire simulation (Figure 4(A)). Its chromone ring binds mainly to S1 and partially occupies S3 forming interactions with Asn142 and Glu166. The rutinose moiety had one of its saccharides binding to S2 and forming H-bonds with Asp187 and His41, which appeared persistent in 93% and 74% of the production MD trajectory snapshots, respectively (see Table 2 and Figure S2 in the supporting information). Finally, the dihydroxy phenyl ring appeared highly flexible throughout the simulation adopting orientations distinct than that observed in the initial docked pose (Figure 4(C)).

A recent enzymatic study showed boceprevir to inhibit SARS-CoV-2 viral replication by targeting its main protease

Figure 5. Structural comparison between the crystal structure and a representative MD structure of SARS-CoV-2 main protease bound to Nelfinavir and Epirubicin. The crystal structure is shown as light cyano cartoon with the docked ligand shown as grey sticks, while the representative MD structure is shown in green cartoon and the ligand as magenta sticks. **(A)** Nelfinavir and **(B)** Epirubicin. Surface representation of the substrate-binding pocket of SARS-CoV-2 M^pro and ligand − receptor interactions are shown for **(C)** Nelfinavir and **(D)** Epirubicin.

Table 2. H-bond analysis of the 150 ns of the MD trajectory for selected hits identified by our VS protocol. Only H-bonds that are persistent in more than 10% of the snapshots are included. For thymopentin and epirubicin, no H-bonds were persistent more than 10%. For each H-bond, the acceptor residue and atom name are indicated in column 1. The name of the H-atom and the electronegative atom attached to it on the donor residue are given in columns 2 and 3, respectively. The major atoms contributing to H-bonds are mapped in Figure S2.

HB-Acceptor	HB-Donor	Donor	Percentage of persistence along the MD trajectory	Average H-bond Distance (Å)	Average H-bond Angle
Boceprevir					
LIG_306@O8	GLY_143@H	GLY_143@N	50%	2.84	154.5
LIG_306@O16	ASN_142@HD22	ASN_142@ND2	27%	2.82	158.2
HIS_164@O	LIG_306@H45	LIG_306@N37	10%	2.89	149.2
Nelfinavir					
GLU_166@OE2	LIG_306@H41	LIG_306@O33	50%	2.66	162.1
GLN_189@OE1	LIG_306@H17	LIG_306@N12	49%	2.86	162.0
GLU_166@OE1	LIG_306@H41	LIG_306@O33	30%	2.66	162.5
Rutin					
ASP_187@O	LIG_306@H13	LIG_306@O27	93%	2.71	166.0
HIS_41@ND1	LIG_306@H12	LIG_306@O26	74%	2.80	163.9
GLU_166@OE2	LIG_306@H14	LIG_306@O28	35%	2.65	164.5
GLU_166@OE1	LIG_306@H14	LIG_306@O28	31%	2.64	164.5
HIS_41@O	LIG_306@H18	LIG_306@O31	27%	2.79	152.6
GLN_189@OE1	LIG_306@H8	LIG_306@O20	23%	2.76	148.9
LIG_306@O23	GLY_143@H	GLY_143@N	12%	2.92	143.5
ASN_142@OD1	LIG_306@H30	LIG_306@O43	11%	2.75	161.2
Control					
LIG_306@O13	GLU_166@H	GLU_166@N	67%	2.87	162.6
LIG_306@O01	GLY_143@H	GLY_143@N	48%	2.88	161.4
LIG_306@O01	ASN_142@HD22	ASN_142@ND2	32%	2.85	159.4
LIG_306@N18	HIS_163@HE2	HIS_163@NE2	24%	2.91	154.5

(Ma et al., 2020). Interestingly, the binding mode of the synthetic tripeptide HCV NS3/4A protease inhibitor boceprevir mimicked that of X77; occupying the three subsites S1′-S2. Boceprevir's first peptide residue (P1) with its cyclobutyl side chain occupied S1′ while its second peptide residue having a cyclopropyl-fused prolyl ring occupied S1 subsite forming H-bonds with the backbone of Gly143 and side chain of

Asn142, which appeared persistent in 50% and 27% of the 150-ns MD trajectory snapshots, respectively (Figure 4(D) and Table 2). The 3-methyl-L-valyl group of P3 partially occupied S2 while the tertiary butyl group appeared relatively flexible throughout the simulation (Figure 4(B)).

Among the promising hits was the HIV protease inhibitor nelfinavir. Several HIV protease inhibitors had been
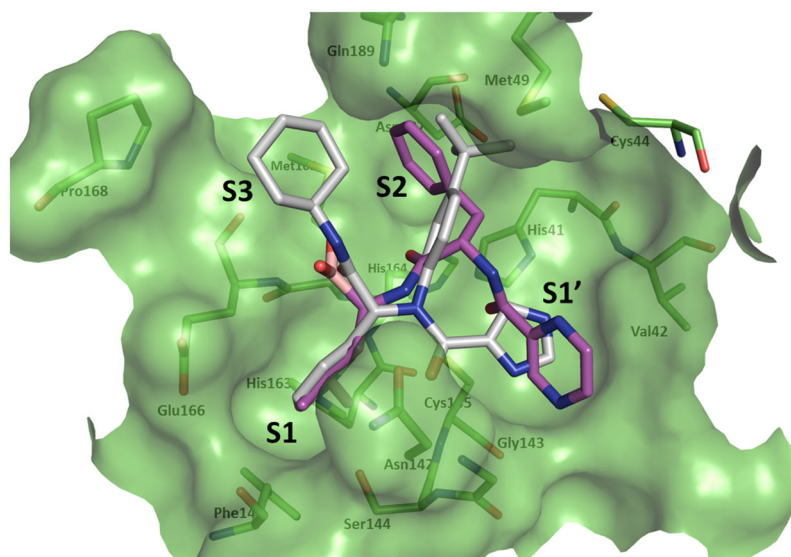
Figure 6. Surface representation of the substrate-binding pocket of SARS-CoV-2 M^{pro} (PDB code: 6W63) (Mesecar, 2020) indicating S1′ –S3 subsites. Overlay of the docked poses of Bortezomib (magenta) and the crystallized coordinates of X77 (grey) (Mesecar, 2020).

Table 3. List of MM-PBSA binding free energies (in kcal/mol) for potential inhibitors binding to SARS-CoV-2 main protease. The free energy is calculated over the course of 150 ns MD simulation for all ligands except epirubicin, which detaches and reattaches in this trajectory. The quoted free energy for epirubicin is for the extended 100 ns, after it reattaches at a close-by site in the active pocket.

Compound Name	MM-PBSA(Kcal/mol)	Std deviation	Std error of mean
Boceprevir	−13.6	4.2	0.1
Epirubicin	−11.4	3.5	0.1
Nelfinavir	−13.4	4.2	0.1
Rutin	−14.9	4.3	0.1
X77 "control"	−17.0	3.4	0.1

previously reported to inhibit the replication of SARS CoV (Jenwitheesuk & Samudrala, 2003) and have shown potential in reducing the symptoms of the recent COVID-19 (Z. M. Chen et al., 2020). Nelfinavir adopted a different conformation throughout the simulation compared to its initial docked pose with a single anchor part represented by its benzoyl group fitting in S1 subsite and forming H-bonds with Glu166 (Figure 5(A)). Analysis of the MD trajectory reveals the formation of H-bonds with Glu166 and Gln189 in approximately 50% of the snapshots for each amino acid. On the other hand, the phenylsulfanyl group flipped 180°, fitting in S3 subsite instead of S1′ in its initial docked pose. Finally, the octahydroisoquinoline ring appeared rather flexible adopting different orientations throughout the entire simulation (Figure 5(C)).

The anthracycline antineoplastic agent valrubicin has been recently discovered as possible SARS-CoV-2 M^{pro} inhibitor in a computational drug repurposing study (L. Zhang et al., 2020). Our VS protocol identified another anthracycline antineoplastic agent epirubicin. Anthracenes have been previously reported as inhibitors of dengue virus NS2B-NS3 protease, a member of the chymotrypsin family as SARS-CoV-2 M^{pro} (Tomlinson & Watowich, 2011). Interestingly, MD simulations of epirubicin showed it detaching from the protein and reattaching again, adopting a more solvent–exposed binding

mode distinct than that of its initial docked pose (Figure 5(B)). The new binding mode showed epirubicin's main tetracene skeleton residing in a deep groove whose base overlap with S2 subsite and lined by side chains of His41 and Met49 and backbone atoms of Gln189 and Thr190. The oxane ring appeared more solvent exposed with its hydroxyl and basic amino groups forming H-bonds with Glu47 (Figure 5(D)) which appeared in only 8% of the snapshots.

Although bortezomib was not considered in MD simulations due to the inherent difficulty of deriving parameters for its boron atom, the binding mode of its docked pose was investigated. Interestingly, previous reports showed that boronic acid compounds bind reversibly to SARS-CoV M^{pro} and inhibiting its activity (Bacha et al., 2004). Moreover, the proteasome inhibitor carfilzomib was recently discovered by virtual screening as the top candidate among all SARS-CoV-2 M^{pro} inhibitor hits (L. Zhang et al., 2020). Bortezomib represents a dipeptide boronic acid analogue belonging to the same class of anticancer agents. The docked pose of bortezomib revealed its boronic acid group mimicking the backbone carboxylic acid of P3 and forming a H-bond with Glu166. The pyrazine ring occupies the S1′ subsite forming a H-bond with the catalytic Cys145. The deep S2 subsite is completely occupied by bortezomib's phenyl ring forming C–H–π interactions with His41. Finally, the isopropyl group occupied the S1 subsite (Figure 6).

3.5. MMPBSA calculation

The ligand binding affinity was measured using the MM-PBSA method (Homeyer & Gohlke, 2012). Structures were evenly sampled from the 150-ns trajectory of each compound, and, in the case of epirubicin, from the extra 100-ns trajectory. The calculated binding free energies are summarized in Table 3 with its individual terms listed in table S4. The binding affinity of all ligands was compared to that of

the cocrystallized ligand X77. Based on the MM-PBSA values reported in Table 3, the co-crystallized ligand X77 is expected to have the largest binding affinity (-17.0 kcal/mol), closely followed by Rutin (-14.9 kcal/mol), within the range of error of the calculated standard deviation (3.4 and 4.3 kcal/mol, respectively). Rutin is followed by Boceprevir and Nelfinavir, with MM-PBSA values of −13.6 and −13.4 kcal/mol, respectively. Epirubicin is expected to have the lowest affinity (-11.4 kcal/mol). The relative binding affinities of the hit compounds appeared consistent with their observed stabilities throughout the MD simulations. This trend was also observed for free energy calculated by GBSA method, see table S5 in the supplementary information.

4. Conclusion

A sequential virtual screening protocol including pharmacophore modeling, molecular docking with PLIF post-docking filtration was developed for the SARS-CoV-2 M^pro enzyme. Molecular dynamics simulations were performed for 5 promising hits and the cocrystallized compound followed by binding free energy calculations. Several promising known drugs stand out as potential inhibitors of SARS-CoV-2 main protease, including rutin, boceprevir, epirubicin, nelfinavir and bortezomib with binding modes mimicking those of SARS-CoV-2 M^pro cocrystallized ligands N3 and X77. Most of the discovered hits belong to pharmaceutical classes that were previously reported to have inhibitory activity on several viruses including SARS-CoV, a highly homologous virus to SARS-CoV-2 (Bacha et al., 2004; L. Chen et al., 2006; Tomlinson & Watowich, 2011; Yi et al., 2004). The five potential drug candidates discovered in this study are a valuable extension of the recently reported SARS-CoV-2 M^pro inhibitors that have been identified using other virtual screening protocols and may be repurposed for COVID-19 treatment. Moreover, our screening protocol proved to be an efficient tool for a possible repurposing of FDA approved drugs towards SARS-CoV-2 main protease and could be further applied on large compound libraries such as ZINC.

Disclosure statement

The authors declare no conflict of interest.

ORCID

Yasmine M. Mandour 🔟 http://orcid.org/0000-0001-7104-3268

Data availability statement

All data are within the manuscript and the supplementary materials.

References

Alaraby Salem, M., & Brown, A. (2015). Two-photon absorption of fluorescent protein chromophores incorporating non-canonical amino acids: TD-DFT screening and classical dynamics. *Physical Chemistry Chemical Physics: PCCP*, *17*(38), 25563–25571. https://doi.org/10.1039/c5cp03875h

Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R., & Hilgenfeld, R. (2003). Coronavirus main proteinase (3CLpro) structure: Basis for design of anti-SARS drugs. *Science (New York, N.Y.)*, *300*(5626), 1763–1767. https://doi.org/10.1126/science.1085658

Ancy, I., Sivanandam, M., & Kumaradhas, P. (2020). Possibility of HIV-1 protease inhibitors-clinical trial drugs as repurposed drugs for SARS-CoV-2 main protease: A molecular docking, molecular dynamics and binding free energy simulation study. *Journal of Biomolecular Structure and Dynamics*, 1–8. https://doi.org/10.1080/07391102.2020.1786459

Anighoro, A., & Bajorath, J. (2016a). Binding mode similarity measures for ranking of docking poses: A case study on the adenosine A2A receptor. *Journal of Computer-Aided Molecular Design*, *30*(6), 447–456. https://doi.org/10.1007/s10822-016-9918-z

Anighoro, A., & Bajorath, J. (2016b). Three-dimensional similarity in molecular docking: Prioritizing ligand poses on the basis of experimental binding modes. *Journal of Chemical Information and Modeling*, *56*(3), 580–587. https://doi.org/10.1021/acs.jcim.5b00745

Aouidate, A., Ghaleb, A., Chtita, S., Aarjane, M., Ousaa, A., Maghat, H., Sbai, A., Choukrad, M. b., Bouachrine, M., & Lakhlifi, T. (2020). Identification of a novel dual-target scaffold for 3CLpro and RdRp proteins of SARS-CoV-2 using 3D-similarity search, molecular docking, molecular dynamics and ADMET evaluation. *Journal of Biomolecular Structure and Dynamics*, 1–14. https://doi.org/10.1080/07391102.2020.1779130

Arun, K. G., Sharanya, C. S., Abhithaj, J., Francis, D., & Sadasivan, C. (2020). Drug repurposing against SARS-CoV-2 using E-pharmacophore based virtual screening, molecular docking and molecular dynamics with main protease as the target. *Journal of Biomolecular Structure and Dynamics*, 1–12. https://doi.org/10.1080/07391102.2020.1779819

Bacha, U., Barrila, J., Velazquez-Campoy, A., Leavitt, S. A., & Freire, E. (2004). Identification of novel inhibitors of the SARS coronavirus main protease 3CLpro. Biochemistry, *43*(17), 4906–4912. https://doi.org/10.1021/bi0361766

Bailey, O. T., Pappenheimer, A. M., Cheever, F. S., & Daniels, J. B. (1949). A murine virus (Jhm) causing disseminated encephalomyelitis with extensive destruction of myelin: II. Pathology. *The Journal of Experimental Medicine*, *90*(3), 195–212. https://doi.org/10.1084/jem.90.3.195

Barretto, N., Jukneliene, D., Ratia, K., Chen, Z., Mesecar, A. D., & Baker, S. C. (2005). The papain-like protease of severe acute respiratory syndrome coronavirus has deubiquitinating activity. *Journal of Virology*, *79*(24), 15189–15198.https://doi.org/10.1128/JVI.79.24.15189-15198.2005

Bartlam, M., Xu, Y., & Rao, Z. (2007). Structural proteomics of the SARS coronavirus: A model response to emerging infectious diseases. *Journal of Structural and Functional Genomics*, *8*(2–3), 85–97. https://doi.org/10.1007/s10969-007-9024-5

Bégué, J.-P., & Bonnet-Delpon, D. (1991). Preparation of trifluoromethyl ketones and related fluorinated ketones. Tetrahedron, *47*(20–21), 3207–3258. https://doi.org/10.1016/S0040-4020(01)86391-2

Chatterjee, S., Maity, A., Chowdhury, S., Islam, M. A., Muttinini, R. K., & Sen, D. (2020). In silico analysis and identification of promising hits

against 2019 novel coronavirus 3C-like main protease enzyme. *Journal of Biomolecular Structure and Dynamics*, 1–14. https://doi.org/10.1080/07391102.2020.1787228

Chen, L., Li, J., Luo, C., Liu, H., Xu, W., Chen, G., Liew, O. W., Zhu, W., Puah, C. M., Shen, X., & Jiang, H. (2006). Binding interaction of quercetin-3-beta-galactoside and its synthetic derivatives with SARS-CoV 3CL(pro): Structure-activity relationship studies reveal salient pharmacophore features. *Bioorganic & Medicinal Chemistry*, 14(24), 8295–8306. https://doi.org/10.1016/j.bmc.2006.09.014

Chen, Z. M., Fu, J. F., Shu, Q., Chen, Y. H., Hua, C. Z., Li, F. B., Lin, R., Tang, L. F., Wang, T. L., Wang, W., Wang, Y. S., Xu, W. Z., Yang, Z. H., Ye, S., Yuan, T. M., Zhang, C. M., & Zhang, Y. Y. (2020). Diagnosis and treatment recommendations for pediatric respiratory infection caused by the 2019 novel coronavirus. *World Journal of Pediatrics: WJP*, 16(3), 240–246. https://doi.org/10.1007/s12519-020-00345-5

Case, D. A., Belfon, K., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, T. E., Cruzeiro III, V. W. D., Darden, T. A., Duke, R. E., Giambasu, G., Gilson, M. K., Gohlke, H., Goetz, A. W., Harris, R., Izadi, S., Izmailov, S. A., Kasavajhala, K., Kovalenko, A., Krasny, R., … Kollman, P. A. (2020). *Amber 12*, University of California.

Da, C., & Kireev, D. (2014). Structural protein-ligand interaction fingerprints (SPLIF) for structure-based virtual screening: Method and benchmark study. *Journal of Chemical Information and Modeling*, 54(9), 2555–2561. https://doi.org/10.1021/ci500319f

Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12), 10089–10092. https://doi.org/10.1063/1.464397

Drosten, C., Gunther, S., Preiser, W., van der Werf, S., Brodt, H. R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R. A., Berger, A., Burguiere, A. M., Cinatl, J., Eickmann, M., Escriou, N., Grywna, K., Kramme, S., Manuguerra, J. C., Muller, S., … Doerr, H. W. (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *The New England Journal of Medicine*, 348(20), 1967–1976. https://doi.org/10.1056/NEJMoa030747

Elmezayen, A. D., Al-Obaidi, A., Şahin, A. T., & Yelekçi, K. (2020). Drug repurposing for coronavirus (COVID-19): In silico screening of known drugs against coronavirus 3CL hydrolase and protease enzymes. *Journal of Biomolecular Structure and Dynamics*, 1–13. https://doi.org/10.1080/07391102.2020.1758791

Food and Drug Administration (FDA). (2020). Retrieved May 3, 2020, from https://www.fda.gov/media/137564/download

Gelb, M. H., Svaren, J. P., & Abeles, R. H. (1985). Fluoro ketone inhibitors of hydrolytic enzymes. Biochemistry, 24(8), 1813–1817. https://doi.org/10.1021/bi00329a001

Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery*, 10(5), 449–461. https://doi.org/10.1517/17460441.2015.1032936

Hilgenfeld, R. (2014). From SARS to MERS: Crystallographic studies on coronaviral proteases enable antiviral drug design. *The FEBS Journal*, 281(18), 4085–4096. https://doi.org/10.1111/febs.12936

Homeyer, N., & Gohlke, H. (2012). Free energy calculations by the molecular mechanics poisson-boltzmann surface area method. *Molecular Informatics*, 31(2), 114–122. https://doi.org/10.1002/minf.201100135

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1), 33–38, 27-38. https://doi.org/10.1016/0263-7855(96)00018-5

Jacobs, J., Grum-Tokars, V., Zhou, Y., Turlington, M., Saldanha, S. A., Chase, P., Eggler, A., Dawson, E. S., Baez-Santos, Y. M., Tomar, S., Mielech, A. M., Baker, S. C., Lindsley, C. W., Hodder, P., Mesecar, A., & Stauffer, S. R. (2013). Discovery, synthesis, and structure-based optimization of a series of N-(tert-butyl)-2-(N-arylamido)-2-(pyridin-3-yl) acetamides (ML188) as potent noncovalent small molecule inhibitors of the severe acute respiratory syndrome coronavirus (SARS-CoV) 3CL protease. *Journal of Medicinal Chemistry*, 56(2), 534–546. https://doi.org/10.1021/jm301580n

Jain, R. P., Pettersson, H. I., Zhang, J., Aull, K. D., Fortin, P. D., Huitema, C., Eltis, L. D., Parrish, J. C., James, M. N., Wishart, D. S., & Vederas, J. C. (2004). Synthesis and evaluation of keto-glutamine analogues as

potent inhibitors of severe acute respiratory syndrome 3CLpro. *Journal of Medicinal Chemistry*, 47(25), 6113–6116. https://doi.org/10.1021/jm0494873

Jenwitheesuk, E., & Samudrala, R. (2003). Identifying inhibitors of the SARS coronavirus proteinase. *Bioorganic & Medicinal Chemistry Letters*, 13(22), 3989–3992. https://doi.org/10.1016/j.bmcl.2003.08.066

Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., Duan, Y., Yu, J., Wang, L., Yang, K., Liu, F., Jiang, R., Yang, X., You, T., Liu, X., … Yang, H. (2020). Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. Nature, 582(7811), 289–293. https://doi.org/10.1038/s41586-020-2223-y

Jones, G., Willett, P., & Glen, R. C. (1995). Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *Journal of Molecular Biology*, 245(1), 43–53. https://doi.org/10.1016/S0022-2836(95)80037-9

Jones, G., Willett, P., Glen, R. C., Leach, A. R., & Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology*, 267(3), 727–748. https://doi.org/10.1006/jmbi.1996.0897

Khan, R. J., Jha, R. K., Amera, G. M., Jain, M., Singh, E., Pathak, A., Singh, R. P., Muthukumaran, J., & Singh, A. K. (2020). Targeting SARS-CoV-2: A systematic drug repurposing approach to identify promising inhibitors against 3C-like proteinase and 2′-O-ribose methyltransferase. *Journal of Biomolecular Structure and Dynamics*, 1–14. https://doi.org/10.1080/07391102.2020.1753577

Kumar, D., Kumari, K., Vishvakarma, V. K., Jayaraj, A., Kumar, D., Ramappa, V. K., Patel, R., Kumar, V., Dass, S. K., Chandra, R., & Singh, P. (2020). Promising inhibitors of main protease of novel corona virus to prevent the spread of COVID-19 using docking and molecular dynamics simulation. *Journal of Biomolecular Structure and Dynamics*, 1–15. https://doi.org/10.1080/07391102.2020.1779131

Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K. S. M., Lau, E. H. Y., Wong, J. Y., Xing, X., Xiang, N., Wu, Y., Li, C., Chen, Q., Li, D., Liu, T., Zhao, J., Liu, M., … Feng, Z. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *The New England Journal of Medicine*, 382(13), 1199–1207. https://doi.org/10.1056/NEJMoa2001316

Loncharich, R. J., Brooks, B. R., & Pastor, R. W. (1992). Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanyl-N′-methylamide. Biopolymers, 32(5), 523–535. https://doi.org/10.1002/bip.360320508

Ma, C., Sacco, M. D., Hurst, B., Townsend, J. A., Hu, Y., Szeto, T., Zhang, X., Tarbet, B., Marty, M. T., Chen, Y., & Wang, J. (2020). Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Research*, 30(8), 678–692. https://doi.org/10.1038/s41422-020-0356-z

Mesecar, A. D. (2020). A taxonomically-driven approach to development of potent, broad-spectrum inhibitors of coronavirus main protease including SARS-CoV-2 (COVID-19). https://doi.org/doi:10.2210/pdb6W63/pdb

Miller, B. R., 3rd, McGee, T. D., Jr., Swails, J. M., Homeyer, N., Gohlke, H., & Roitberg, A. E. (2012). MMPBSA.py: An efficient program for end-state free energy calculations. *Journal of Chemical Theory and Computation*, 8(9), 3314–3321. https://doi.org/10.1021/ct300418h

Molecular Operating Environment (MOE). (2019). 2019.01, Chemical Computing Group ULC, 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7.

Mysinger, M. M., Carchia, M., Irwin, J. J., & Shoichet, B. K. (2012). Directory of useful decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking. *Journal of Medicinal Chemistry*, 55(14), 6582–6594. https://doi.org/10.1021/jm300687e

Neudert, G., & Klebe, G. (2011). DSX: A knowledge-based scoring function for the assessment of protein-ligand complexes. *Journal of Chemical Information and Modeling*, 51(10), 2731–2745. https://doi.org/10.1021/ci200274q

Payne, S. (2017). Chapter 17- Family coronaviridae. In S. Payne (Ed.), *Viruses* (pp.149–158). Academic Press. https://doi.org/10.1016/B978-0-12-803109-4.00017-9

Roe, D. R., & Cheatham, T. E. 3rd. (2013). PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data.

Journal of Chemical Theory and Computation, 9(7), 3084–3095. https://doi.org/10.1021/ct400341p

Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A., Campagnoli, R., Icenogle, J. P., Penaranda, S., Bankamp, B., Maher, K., Chen, M. H., Tong, S., Tamin, A., Lowe, L., Frace, M., DeRisi, J. L., Chen, Q., Wang, D., Erdman, D. D., Peret, T. C., … Bellini, W. J. (2003). Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science (New York, N.Y.)*, 300(5624), 1394–1399. https://doi.org/10.1126/science.1085952

Schrodinger, LLC. (2015). The PyMOL Molecular Graphics System, Version 1.8.

Shao, Y. M., Yang, W. B., Kuo, T. H., Tsai, K. C., Lin, C. H., Yang, A. S., Liang, P. H., & Wong, C. H. (2008). Design, synthesis, and evaluation of trifluoromethyl ketones as inhibitors of SARS-CoV 3CL protease. *Bioorganic & Medicinal Chemistry*, 16(8), 4652–4660. https://doi.org/10.1016/j.bmc.2008.02.040

Shie, J. J., Fang, J. M., Kuo, C. J., Kuo, T. H., Liang, P. H., Huang, H. J., Yang, W. B., Lin, C. H., Chen, J. L., Wu, Y. T., & Wong, C. H. (2005). Discovery of potent anilide inhibitors against the severe acute respiratory syndrome 3CL protease. *Journal of Medicinal Chemistry*, 48(13), 4469–4473. https://doi.org/10.1021/jm050184y

Shimamoto, Y., Hattori, Y., Kobayashi, K., Teruya, K., Sanjoh, A., Nakagawa, A., Yamashita, E., & Akaji, K. (2015). Fused-ring structure of decahydroisoquinolin as a novel scaffold for SARS 3CL protease inhibitors. *Bioorganic & Medicinal Chemistry*, 23(4), 876–890. https://doi.org/10.1016/j.bmc.2014.12.028

Thanigaimalai, P., Konno, S., Yamamoto, T., Koiwai, Y., Taguchi, A., Takayama, K., Yakushiji, F., Akaji, K., Chen, S. E., Naser-Tavakolian, A., Schon, A., Freire, E., & Hayashi, Y. (2013). Development of potent dipeptide-type SARS-CoV 3CL protease inhibitors with novel P3 scaffolds: Design, synthesis, biological evaluation, and docking studies. *European Journal of Medicinal Chemistry*, 68, 372–384. https://doi.org/10.1016/j.ejmech.2013.07.037

Thanigaimalai, P., Konno, S., Yamamoto, T., Koiwai, Y., Taguchi, A., Takayama, K., Yakushiji, F., Akaji, K., Kiso, Y., Kawasaki, Y., Chen, S. E., Naser-Tavakolian, A., Schon, A., Freire, E., & Hayashi, Y. (2013). Design, synthesis, and biological evaluation of novel dipeptide-type SARS-CoV 3CL protease inhibitors: Structure-activity relationship study. *European Journal of Medicinal Chemistry*, 65, 436–447. https://doi.org/10.1016/j.ejmech.2013.05.005

Tomlinson, S. M., & Watowich, S. J. (2011). Anthracene-based inhibitors of dengue virus NS2B-NS3 protease. *Antiviral Research*, 89(2), 127–135. https://doi.org/10.1016/j.antiviral.2010.12.006

Turlington, M., Chun, A., Tomar, S., Eggler, A., Grum-Tokars, V., Jacobs, J., Daniels, J. S., Dawson, E., Saldanha, A., Chase, P., Baez-Santos, Y. M., Lindsley, C. W., Hodder, P., Mesecar, A. D., & Stauffer, S. R. (2013). Discovery of N-(benzo[1,2,3]triazol-1-yl)-N-(benzyl)acetamido)phenyl) carboxamides as severe acute respiratory syndrome coronavirus (SARS-CoV) 3CLpro inhibitors: Identification of ML300 and noncovalent nanomolar inhibitors with an induced-fit binding. *Bioorganic & Medicinal Chemistry Letters*, 23(22), 6172–6177. https://doi.org/10.1016/j.bmcl.2013.08.112

Turner, P. J. (2005). *XMGRACE* (Version 5.1.19). Beaverton, OR: Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology.

Wang, F., Chen, C., Tan, W., Yang, K., & Yang, H. (2016). Structure of main protease from human coronavirus NL63: Insights for wide spectrum anti-coronavirus drug design. *Scientific Reports*, 6, 22677 https://doi.org/10.1038/srep22677

Wang, J. (2020). Fast Identification of possible drug treatment of coronavirus disease-19 (COVID-19) through computational drug repurposing study. *Journal of Chemical Information and Modeling*, 60(6), 3277–3286. https://doi.org/10.1021/acs.jcim.0c00179

Wang, J., Ge, Y., & Xie, X. Q. (2019). Development and testing of druglike screening libraries. *Journal of Chemical Information and Modeling*, 59(1), 53–65. https://doi.org/10.1021/acs.jcim.8b00537

Wang, J., Wang, W., Kollman, P. A., & Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics & Modelling*, 25(2), 247–260. https://doi.org/10.1016/j.jmgm.2005.12.005

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9), 1157–1174. https://doi.org/10.1002/jcc.20035

World Health Organization (WHO). (2020). Timeline of WHO's response to COVID-19. Retrieved July 24, 2020, from http://www.who.int/

Word, J. M., Lovell, S. C., Richardson, J. S., & Richardson, D. C. (1999). Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. *Journal of Molecular Biology*, 285(4), 1735–1747. https://doi.org/10.1006/jmbi.1998.2401

Xu, X., Chen, P., Wang, J., Feng, J., Zhou, H., Li, X., Zhong, W., & Hao, P. (2020). Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Science China. Life Sciences*, 63(3), 457–460. https://doi.org/10.1007/s11427-020-1637-5

Xue, X., Yu, H., Yang, H., Xue, F., Wu, Z., Shen, W., Li, J., Zhou, Z., Ding, Y., Zhao, Q., Zhang, X. C., Liao, M., Bartlam, M., & Rao, Z. (2008). Structures of two coronavirus main proteases: implications for substrate binding and antiviral drug design. *Journal of Virology*, 82(5), 2515–2527. https://doi.org/10.1128/JVI.02114-07

Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Mo, L., Ye, S., Pang, H., Gao, G. F., Anand, K., Bartlam, M., Hilgenfeld, R., & Rao, Z. (2003). The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proceedings of the National Academy of Sciences of the United States of America*, 100(23), 13190–13195. https://doi.org/10.1073/pnas.1835675100

Yi, L., Li, Z., Yuan, K., Qu, X., Chen, J., Wang, G., Zhang, H., Luo, H., Zhu, L., Jiang, P., Chen, L., Shen, Y., Luo, M., Zuo, G., Hu, J., Duan, D., Nie, Y., Shi, X., Wang, W., … Xu, X. (2004). Small molecules blocking the entry of severe acute respiratory syndrome coronavirus into host cells. *Journal of Virology*, 78(20), 11334–11339. https://doi.org/10.1128/JVI.78.20.11334-11339.2004

Zhang, H. Z., Zhang, H., Kemnitzer, W., Tseng, B., Cinatl, J., Jr., Michaelis, M., Doerr, H. W., & Cai, S. X. (2006). Design and synthesis of dipeptidyl glutaminyl fluoromethyl ketones as potent severe acute respiratory syndrome coronovirus (SARS-CoV) inhibitors. *Journal of Medicinal Chemistry*, 49(3), 1198–1201. https://doi.org/10.1021/jm0507678

Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K., & Hilgenfeld, R. (2020). Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved alpha-ketoamide inhibitors. *Science*, 368(6489), 409–412. https://doi.org/10.1126/science.abb3405

Zhang, Y. Z., & Holmes, E. C. (2020). A genomic perspective on the origin and emergence of SARS-CoV-2. *Cell*, 181(2), 223–227. https://doi.org/10.1016/j.cell.2020.03.035

Zhu, L., George, S., Schmidt, M. F., Al-Gharabli, S. I., Rademann, J., & Hilgenfeld, R. (2011). Peptide aldehyde inhibitors challenge the substrate specificity of the SARS-coronavirus main protease. *Antiviral Research*, 92(2), 204–212. https://doi.org/10.1016/j.antiviral.2011.08.001 [InsertedFromOnline[pubmedMismatch]]