Taylor & Francis
Taylor & Francis Group

Check for updates

Drug repurposing for coronavirus (COVID-19): *in silico* screening of known drugs against coronavirus 3CL hydrolase and protease enzymes

Ammar D. Elmezayen 📵, Anas Al-Obaidi 📵, Alp Tegin Şahin and Kemal Yelekçi 📵

Department of Bioinformatics and Genetics, Faculty of Engineering and Natural Sciences, Kadir Has University, Istanbul, Turkey

Communicated by Ramaswamy H. Sarma

ABSTRACT

In December 2019, COVID-19 epidemic was described in Wuhan, China, and the infection has spread widely affecting hundreds of thousands. Herein, an effort was made to identify commercially available drugs in order to repurpose them against coronavirus by the means of structure-based virtual screening. In addition, ZINC15 library was used to identify novel leads against main proteases. Human TMPRSS2 3D structure was first generated using homology modeling approach. Our molecular docking study showed four potential inhibitors against Mpro enzyme, two available drugs (Talampicillin and Lurasidone) and two novel drug-like compounds (ZINC000000702323 and ZINC000012481889). Moreover, four promising inhibitors were identified against TMPRSS2; Rubitecan and Loprazolam drugs, and compounds ZINC000015988935 and ZINC000103558522. ADMET profile showed that the hits from our study are safe and drug-like compounds. Furthermore, molecular dynamic (MD) simulation and binding free energy calculation using the MM-PBSA method was performed to calculate the interaction energy of the top-ranked drugs.

Introduction

Coronavirus disease (COVID-19) has become an important public issue across the globe since December 2019. As of 12th of April 2020, more than 1.79 million cases have been reported in 210 countries and territories (Worldometer, 2020). It affects people worldwide and there is no vaccine yet for this virus. Certain types of pneumonia are implicated to the new coronavirus, which is considered a big threat to global public health. There is an urgent need to develop potent anti-COVID-19 agents for the prevention of the outbreak and stop viral infections. Repurposing of known small molecules seems to be very efficient way in order to develop potent drugs to combat coronavirus in this short time period. Recently, a number of efforts were made to design novel inhibitors or employ drug repurposing approach to identify anti-COVID-19 drugs, which can act as promising inhibitors against coronavirus protease (Khan et al., 2020; Sarma et al., 2020). The known coronavirus 3-chymotrypsin-like protease (3CLpro), also known as Mpro, is the main protease, which is required for proteolytic maturation of the corona virus. This Mpro have an essential role in the immune regulation and cleaving the polyproteins pp1a and pp1ab, which making them attractive and important targets for anti-COVID-19 drugs (Zhou et al., 2019). Functional proteins such as RNA polymerase, endoribonuclease and exoribonuclease are generated by cleavage of pp1a and pp1ab polyproteins by Mpro (Khan et al., 2020). Therefore, targeting Mpro enzyme will inhibit the viral maturation and enhance the host innate immune response against COVID-19. 3D crystal structure of 3CL hydrolase from specific coronavirus (PDB ID: 6LU7) is recently reported in the public domain. This could be an excellent target to be used to screen small molecule libraries to inhibit the cleavage of the viral polyprotein and stop the spread of infection. In the Protein Data Bank (PDB) archive, there are also some proteins whose sequence identity at least 90% similar to the COVID-19 coronavirus 3CL hydrolase (Mpro). These proteins structures include bound inhibitors, which could lead to repurpose novel drug against coronavirus 3CL hydrolase. Recent study has shown that the cellular transmembrane protease serine 2 (TMPRSS2) is used by COVID-19 for the purpose of entering the host cells, and blocking TMPRSS2 might offer a promising treatment option and prevent the virus from entering the host cells (Hoffmann et al., 2020). Transmembrane serine proteases are connected to some viral respiratory infections, where they facilitate the entry of the virus into the lungs (Shulla et al., 2011). TMPRSS2 belongs to the serine protease transmembrane family type II, and it was recognized by its involvement in the cleavage of the influenza virus hemagglutinin protein in epithelial cells (Böttcher et al., 2006). In addition, studies have shown that TMPRSS2 can drive the cleavage spike protein, which is a coronavirus fusion glycoprotein. Spike proteins on the host cell surface are activated by TMPRSS2 to facilitate the virus cell-membrane fusion and the entry of the virus (Gierer et al., 2013; Matsuyama et al., 2010). Targeting TMPRSS2 in some animal studies decreased the pathological severity of influenza virus infection (Iwata-Yoshikawa et al., 2019). Hence, TMPRSS2 is an attractive target for designing and developing antiviral drugs (Laporte & Naesens, 2017). Herein, our study aims to help in coronavirus inhibition in one of two ways, preventing the virus from
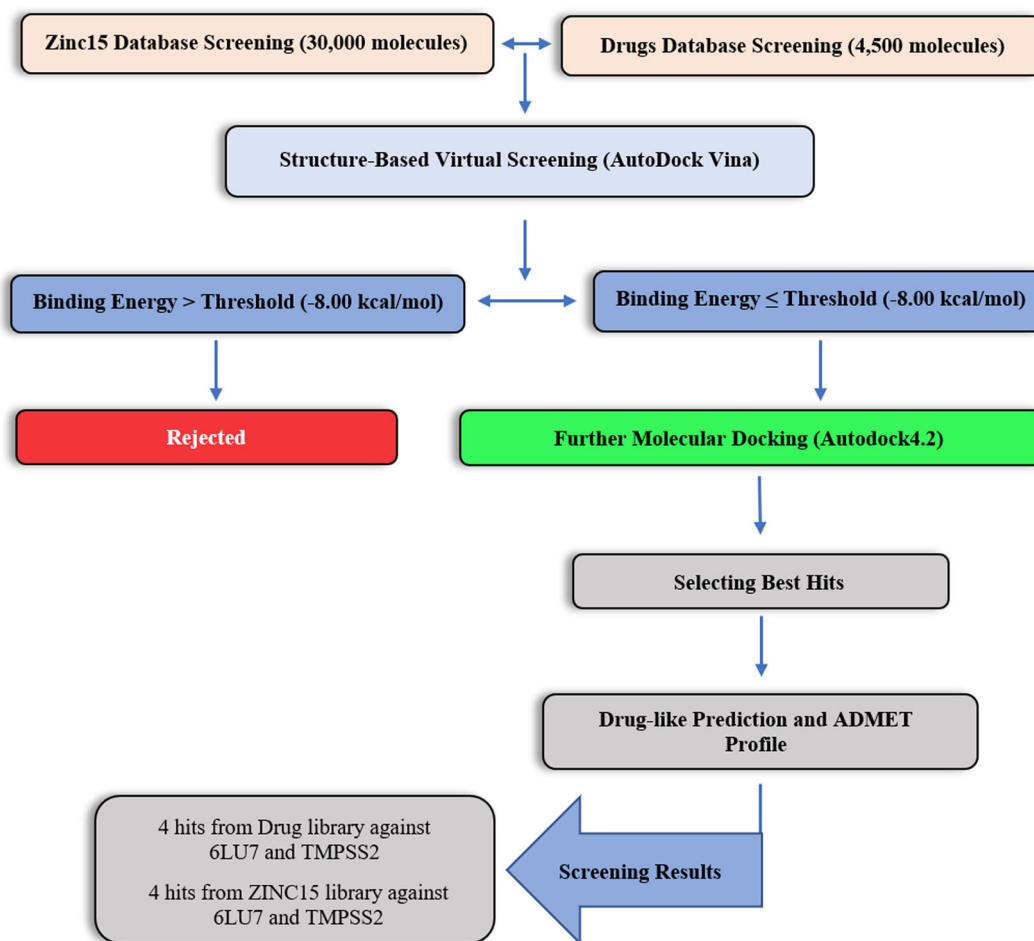
Figure 1. Workflow for identification of potential inhibitors against Mpro and TMPSS2 enzymes via structure-based virtual screening.

Table 1. Grid box coordinates and size parameters used in AutoDock Vina.

	Mpro	TMPRSS2
Center (Å)		
x	−12.149	−8.951
y	14.097	28.972
z	69.719	19.716
Box dimension (Å)		
x	18.75	18.75
y	18.75	18.75
z	18.75	18.75

Table 2. Grid box mapping parameters used in AutoDock 4.2.

	Mpro	TMPRSS2
Center (Å)		
x	−12.149	−8.951
y	14.097	28.972
z	69.719	19.716
Box dimension (Å)		
x	50	50
y	50	50
z	50	50

entering the host cells, and blocking the virus maturation inside the host cells. Computer-aided drug design (CADD) has been used for the identification of potent inhibitors against coronavirus (Berry et al., 2015; Oany et al., 2014). In this study, we have applied virtual screening approach, homology modeling of human TMPRSS2 enzyme, molecular docking study and ADMET profile analysis in order to identify novel and potential inhibitors against COVID-19 (Figure 1).

Computational methods

Protein and ligand setup

The recent resolved crystal structure of 3CL hydrolase enzyme was retrieved from the Protein Data Bank (PDB ID: 6LU7) with a resolution of 2.16 Å. The enzyme has one chain with a total of 306 amino acids. The enzyme was prepared at physiological pH

(7.4) using "Prepare Protein" protocol in BIOVIA Discovery Studio 4.5 (DS 4.5) and missing residues insertion was allowed if found. Water molecules and the inhibitor were removed from the structure. For the purpose of repurposing of known inhibitors, a total of 4,500 approved drugs were downloaded in SDF file format from ChEMBL (https://www.ebi.ac.uk/chembl/), DrugBank (https://www.drugbank.ca/) and Selleckchem (https://www.selleckchem.com/) databases. In addition, a total of 30,000 drug-like molecules were downloaded from ZINC15 database (https://zinc15.docking.org/). All small molecules were prepared and protonated at physiological pH (7.4) and their geometry was minimized using DS 4.5.

Homology modeling

Lack of information of crystal structures of particular proteins has delayed and obstructed the drug industry. Homology
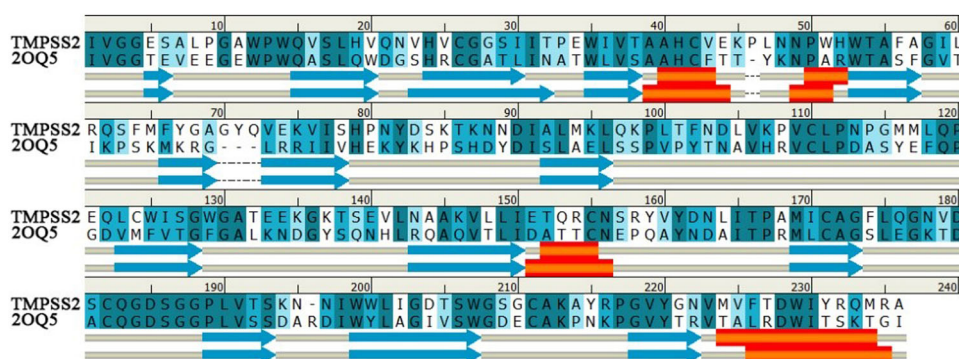
Figure 2. TMPRSS2 and human DESC1sequences alignment. Amino acids identity is shown in dark blue, and similarity is represented in light blue.
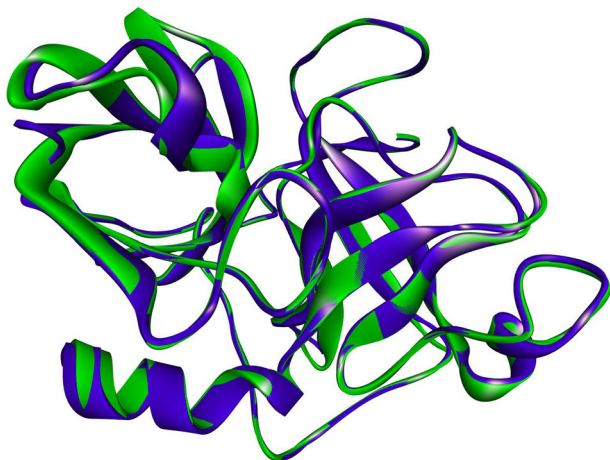

Figure 3. Structural alignment between the homology modeling of TMPRSS2 (target) and the crystal structure of DESC1 (template) showed an RMSD value of 0.40 Å.

modeling is widely applied to create a reliable structure of a protein using its own amino acid sequence (Vyas et al., 2012). To date, TMPRSS2 enzyme structure has not been resolved yet. Therefore, Basic Local Alignment Search Tool (BLAST) (https://blast.ncbi.nlm.nih.gov/Blast.cgi) search was carried out to identify the most suitable template for our target (TMPRSS2). According to the results, the crystal structure of human transmembrane protease serine (DESC1) showed high identity of 41%, 187 Bit-Score and 2.55333e-55 E-value with the query target. Consequently, the 3D structure of DESC1 was retrieved from PDB website (PDB ID: 2OQ5, 1.61 Å) (Kyrieleis et al., 2007). TMPRSS2 sequence in FASTA file format (UniProt ID: O15393) was downloaded from UniProt (http://www.uniprot.org/). The template and the target sequences were aligned using "Align Sequences" in DS 4.5. Twenty models were created using "Homology Modeling" protocol in DS 4.5 using MODELLER plug-in (Webb & Sali, 2016). MODELLER was then used to verify the 20 built models. The best model was found to be M0001 with a Normalized DOPE score of −1.106201. The created homology modeling was then minimized using the "Clean Geometry" tool provided by DS 4.5, which minimizes the energy under CHARMM forcefield. Additionally, the protein was prepared using "Prepare Protein" protocol and protonated at pH, 7.4. M0001 model was then refined using 3Drefine web server (http://sysbio.rnet.missouri.edu/3Drefine/) (Bhattacharya et al., 2016).

Further verification of our homology modeling was carried out using web-based tools ProSA (Wiederstein & Sippl, 2007), PROCHECK (Laskowski et al., 1993) and ERRAT (Colovos & Yeates, 1993) to assess the quality of the model structure.

Structure-based virtual screening

Structure-based virtual screening refers to in silico identification of potential chemical molecules out of large number of compound libraries, which have high affinity to proteins of known structure, based on the binding of the small molecule with the protein binding pocket (Sliwoski et al., 2014). To perform the virtual screening approach in our study, AutoDock Vina was used. AutoDock Vina offers an accurate and high performance docking tool, which relies on both empirical and knowledge-based scoring functions (Trott & Olson, 2009). The docking parameters for both Mpro and TMPRSS2 are given in Table 1. AutoDock Vina was used to assess the binding affinity of the native inhibitor (N3) obtained from the crystal structure of Mpro, as well as the binding affinity of selected known inhibitors of TMPRSS2 retrieved from the literature (Sielaff et al., 2011). The binding energy threshold was based on the average binding energy scores of docked known inhibitors against Mpro and TMPRSS2, which found to be −6.2 kcal/mol and −7.2 kcal/mol, respectively. In the present study, the threshold was set to −8 kcal/mol due to the fact that hit compounds were expected to show lower binding energy scores, and to narrow down the search space. Selection criterion was set to ensure that the virtual screening will end up only with those molecules with the highest binding affinities to their respective targets. A total of 3451 molecules satisfied our selection criterion ($\Delta G° \leq -8$ kcal/mol) and showed the highest binding affinity against Mpro and TMPRSS2 enzymes, respectively.

Molecular docking study

The hits compound from the virtual screening study were further docked into the studied targets using AutoDock 4.2 (Morris et al., 2009). Grid box mapping parameters are shown in Table 2. Grid box parameters were set to cover the whole binding pocket and its adjacent residues. Twenty million energy evaluations were performed for each compound with
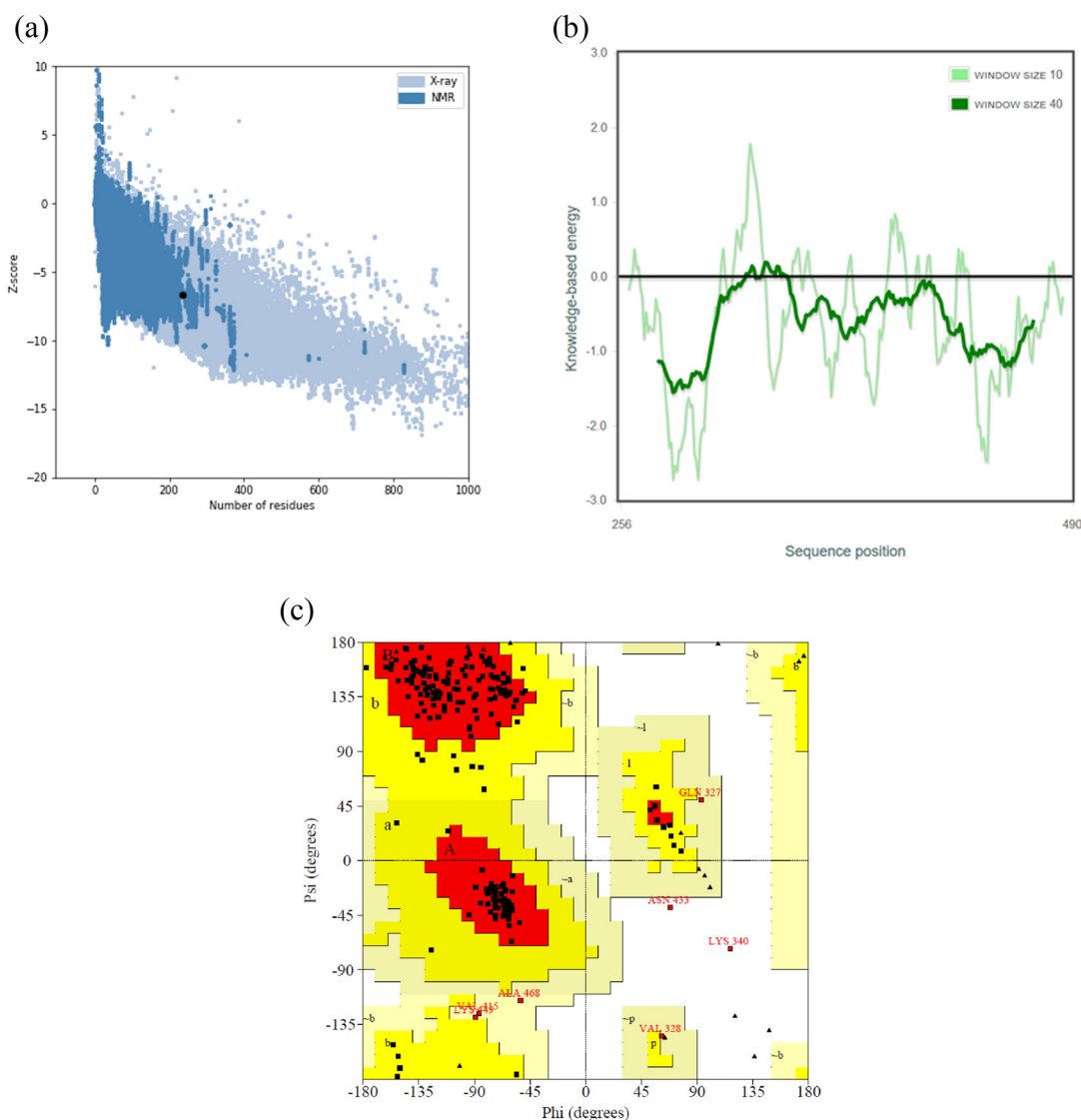
(a)

(b)

(c)

Figure 4. Homology modeling (M0001) validation. (a) z-score plot shows model M0001 (black dot) within the range of native structure. (b) Local energy plot. (c) Ramachandran plot.
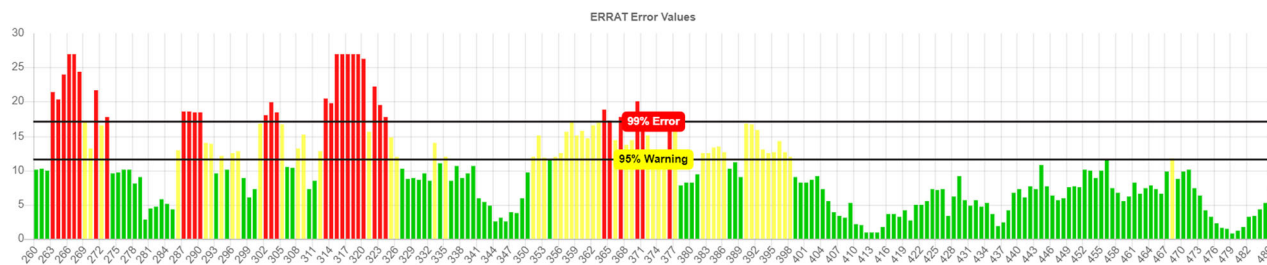

Figure 5. ERRAT scheme plot for the TMPRSS2 model. Red bars refer to the misfolded region at the 99% confidence level. Regions that exhibit misfolding at 95% confidence level are in yellow bars. Green bars specify the regions with the proper protein folding.

a total of twenty runs using AutoDock 4.2. Lamarckian Genetic Algorithm was used and Gasteiger partial charges were added onto all studied proteins.

Electrostatic potential calculation

Electrostatic potential was predicted using DelPhi program included in DS 4.5, which employs the Poisson-Boltzmann formula on a cubic lattice using the finite-difference

technique to calculate the electrostatic charge distribution of TMPRSS2 and Mpro enzymes (Nicholls & Honig, 1991).

ADMET profile and drug-like prediction

Absorption, distribution, metabolism, excretion and toxicity (ADMET) profile has a great importance in drug industry and is widely used in CADD to reduce undesired effects. In our study, web-based tools admetSAR (Yang et al., 2019) and

Table 3. Computed binding energy (ΔG°) and calculated constant inhibition (Ki) values of the top-ranked ligands from ZINC15 and approved drugs libraries.

	Mpro	
	Binding Energy ΔG° (kcal/mol)	Inhibition Constant Ki (nM)
Talampicillin	−11.17	6.49
Lurasidone	−11.17	6.52
ZINC000015988935	−12.39	0.826
ZINC000103558522	−12.36	0.866
	TMPRSS2	
	Binding Energy ΔG° (kcal/mol)	Inhibition Constant Ki (nM)
Rubitecan	−9.47	115
Loprazolam	−9.31	149
ZINC000000702323	−12.18	1.18
ZINC000012481889	−12.14	1.26

SwissADME (Daina et al., 2017) were used to predict the drug-likeness and ADMET profile.

Molecular dynamic (MD) simulation

Molecular dynamic approach is widely used to assess atoms behavior, structural stability and to study the conformational changes on atomic level. Herein, top-ranked compounds, homology model of TMPSS2 as well as the crystal structure of Mpro enzyme were subjected to MD simulation by nanoscale molecular dynamics (NAMD) software (Phillips et al., 2005), and the temperature was set to 303 K. Configuration files for MD simulations were generated by CHARMM-GUI website (http://www.charmm-gui.org/) (Jo et al., 2008). Ligand parameterization was performed using CHARMM General Force Field (CGenFF) web-based tool (https://cgenff.umaryland.edu/). All systems were solvated using transferable intermolecular potential water molecules (TIP3P) model (Boonstra et al., 2016). MD simulations were performed in two steps; equilibration under NVT ensemble for 5 ns timescale, where the systems' energy was minimized for 10,000 steps, and production for 50 ns timescale under the NPT ensemble.

Binding free energy calculation

The binding free energy (ΔG$_{bin}$) between protein and ligand is controlled by three factors; the gas-phase free energy (ΔG$_{MM}$), the solvation free energy (ΔG$_{sol}$) and the change in the system entropy (−TΔS), which can be calculated by molecular mechanics Poisson-Boltzmann Surface Area (MM-PBSA) module, as expressed in Equation (1) (Kollman et al., 2000):

$$\Delta G_{bin} = \Delta G_{MM} + \Delta G_{sol} - T\Delta S \tag{1}$$

In the present study, MM-PBSA calculations were performed using Calculation of Free Energy (CaFE) tool to calculate the interaction energy between the ligand and the residues of the protein active site (Liu & Hou, 2016). In the MM-PBSA calculations, the external dielectric constant was set to 80.0, the internal dielectric constant was 1.0 and the reciprocal of grid spacing of 1.0 Å was employed.

Results and discussion

Homology modeling

Out of the 20 models generated by MODELLER, we selected model M0001, which had the lowest Normalized DOPE score (−1.106201). Sequence alignment for M0001 and the human DESC1were performed unsing DS 4.5, and the sequence identity was 39.4%, while sequence similarity was 61% (Figure 2). In fact, if the target and template sequences share 30% and more of amino acid identity, the homology modeling is considered to be reliable and successful (Xiang, 2006). Structural alignment of our homology modeling structure (target) with the 3 D structure of DESC1 (template) was performed using DS 4.5, and the calculated RMSD was 0.40 Å, which suggests the high similarity between the target and the template (Figure 3). Our model (M0001) was verified using the Web-based version of Protein Structure Analysis tool, ProSA. ProSA is used to calculate the z-score of a specific model and correlate this score to those calculated scores from all publicly available structures on PDB website. Accordingly, the z-score of our model M0001 was -6.69, which is in the range of all PDB conformations (Figure 4a). Moreover, the energy plot describes amino acid position depending on its local energy, where negative values relate to the accuracy of the structure (Figure 4b). In addition, PROCHECK web-based tool was used to generate Ramachandran plot to assess the energetically allowed regions of our model (Figure 4c). 83.8% of the residues are in the most favored regions, 15.2% are in additional allowed regions and 1.0% of the residue are in disallowed regions. ERRAT (https://servicesn.mbi.ucla.edu/ERRAT/) was used to calculate the overall quality factor (OQF) for non-bonded atomic interactions, and generally any structure with OQF of 50% and more is considered as a high quality model. For our homology modeling, the overall quality factor calculated by the ERRAT server was 62% (Figure 5).

Structure-based virtual screening and molecular docking results

Calculated binding energies and estimated inhibition constants of the top-ranked ligands retrieved in our study are given in Table 3.

Mpro potential inhibitors

Our molecular docking study revealed two of the approved drugs Talampicillin and Lurasidone with the highest binding affinities to Mpro enzyme. Talampicillin was found to have a binding energy score of −11.17 kcal/mol and an inhibition constant of 6.49 nM. Talampicillin is a prodrug of ampicillin, an antibacterial drug with a high potency against gram-negative bacteria. Talampicillin formed conventional and carbon hydrogen bonds with residues Gly143, Glu166, Gln189 and Gln192, π-σ bond with His41, two alkyl interactions with the residues Met165 and Met49, two π-alkyl bonds with the residues Cys145 and Pro168, a π-stacked interaction with Leu167 and many van der Waals interactions were found as well (Figures 6a, 7a). Lurasidone was the second-best candidate inhibitor against Mpro that showed and estimated inhibition constant value of 6.52 nM with a predicted binding energy score of −11.17 kcal/mol. Lurasidone was approved by
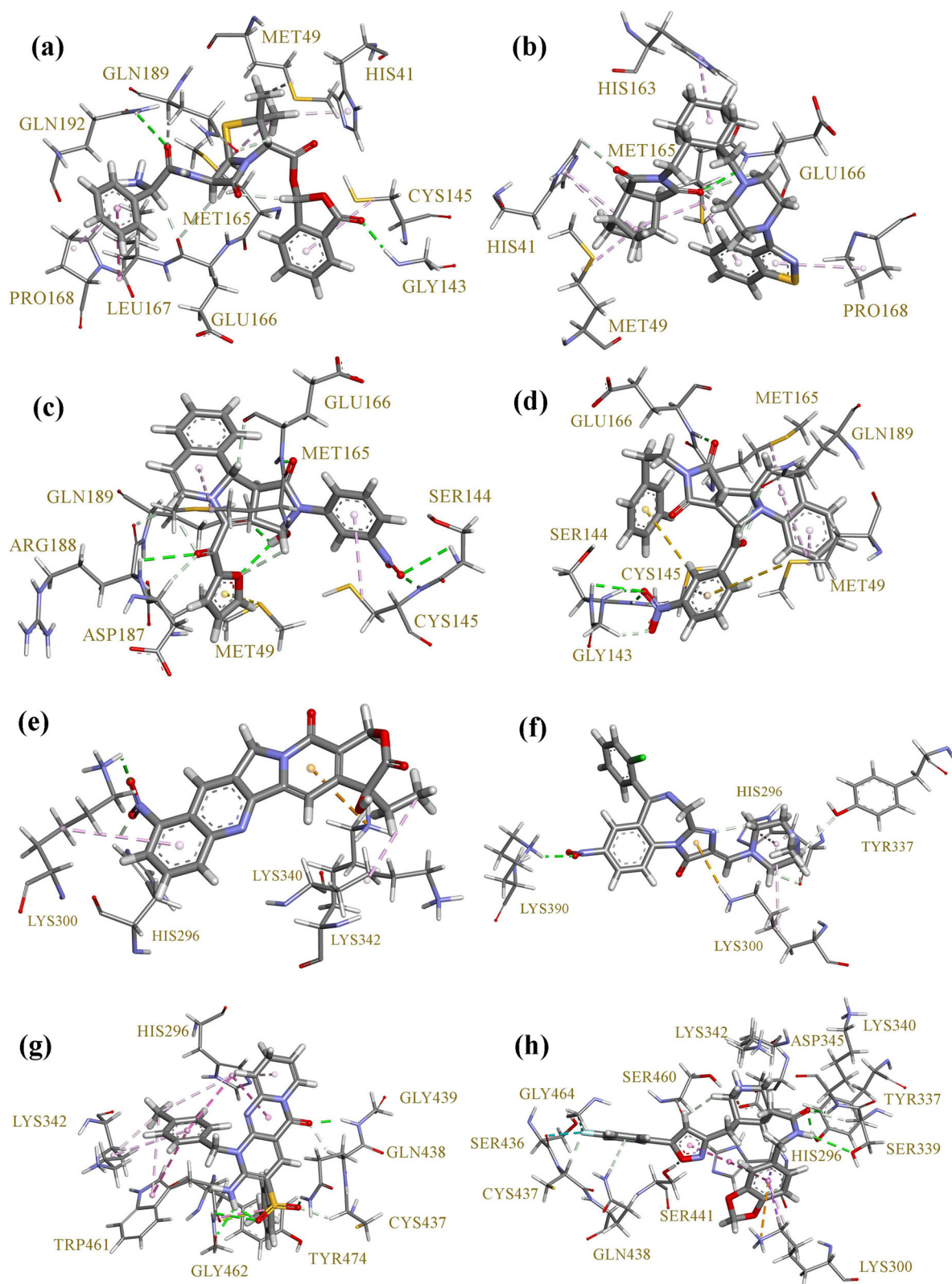
Figure 6. 3D-interaction diagram for the top-ranked compounds. Mpro complexes: (a) Talampicillin, (b) Lurasidone, (c) ZINC000015988935, (d) ZINC000103558522. TMPRSS2 complexes: (e) Rubitecan, (f) Loprazolam, (g) ZINC000000702323, (h) ZINC000012481889.

the FDA in 2010 as an antipsychotic drug for treating schizophrenia patients (Corponi et al., 2019). Different ligand-protein interactions formed after the molecular docking study with Lurasidone, which includes hydrogen bonds with the residues His41 and Glu166, alkyl interactions with Met165 and Met49, π-alkyl interaction with Pro168, Met165 and His41 and multiple van der

Waals interactions with the receptor (Figures 6b and 7b). Talampicillin and Lurasidone showed a reliable binding mode in Mpro and blocking the active site of the enzyme. The two top-ranked compounds from the ZINC15 library were ZINC000015988935 and ZINC000103558522. ZINC000015988935 was found to have a binding energy of −12.39 kcal/mol and an
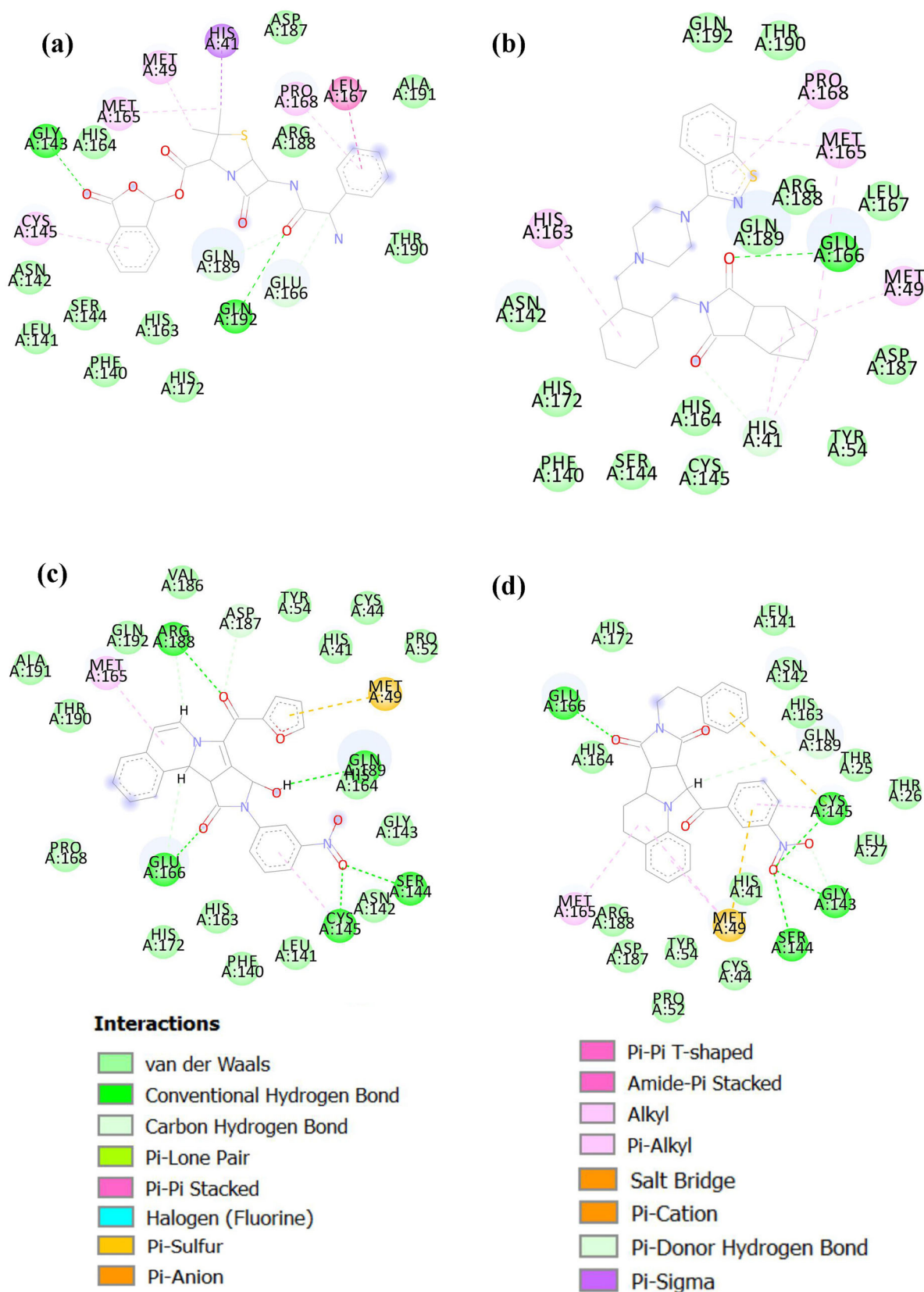
Figure 7. 2D-interaction diagram for the top-ranked Mpro-complexes: (a) Talampicillin, (b) Lurasidone, (c) ZINC000015988935, (d) ZINC000103558522.

inhibition constant of 826 pM, while ZINC000103558522 had a binding energy of -12.36 kcal/mol and an inhibition constant of 866 pM. ZINC000015988935 showed a π-sulfur interaction with Met49, seven H-bonds with the residues Arg188, Asp187, Gln189,

Ser144, Cys145 and Glu166, alkyl interaction with Met165 and a π-alkyl interaction with Cys145 (Figures 6c and 7c). ZINC000103558522 shared two π-sulfur interactions with residues Met49 and Cys145, hydrogen bonds with Glu166, Gln189,
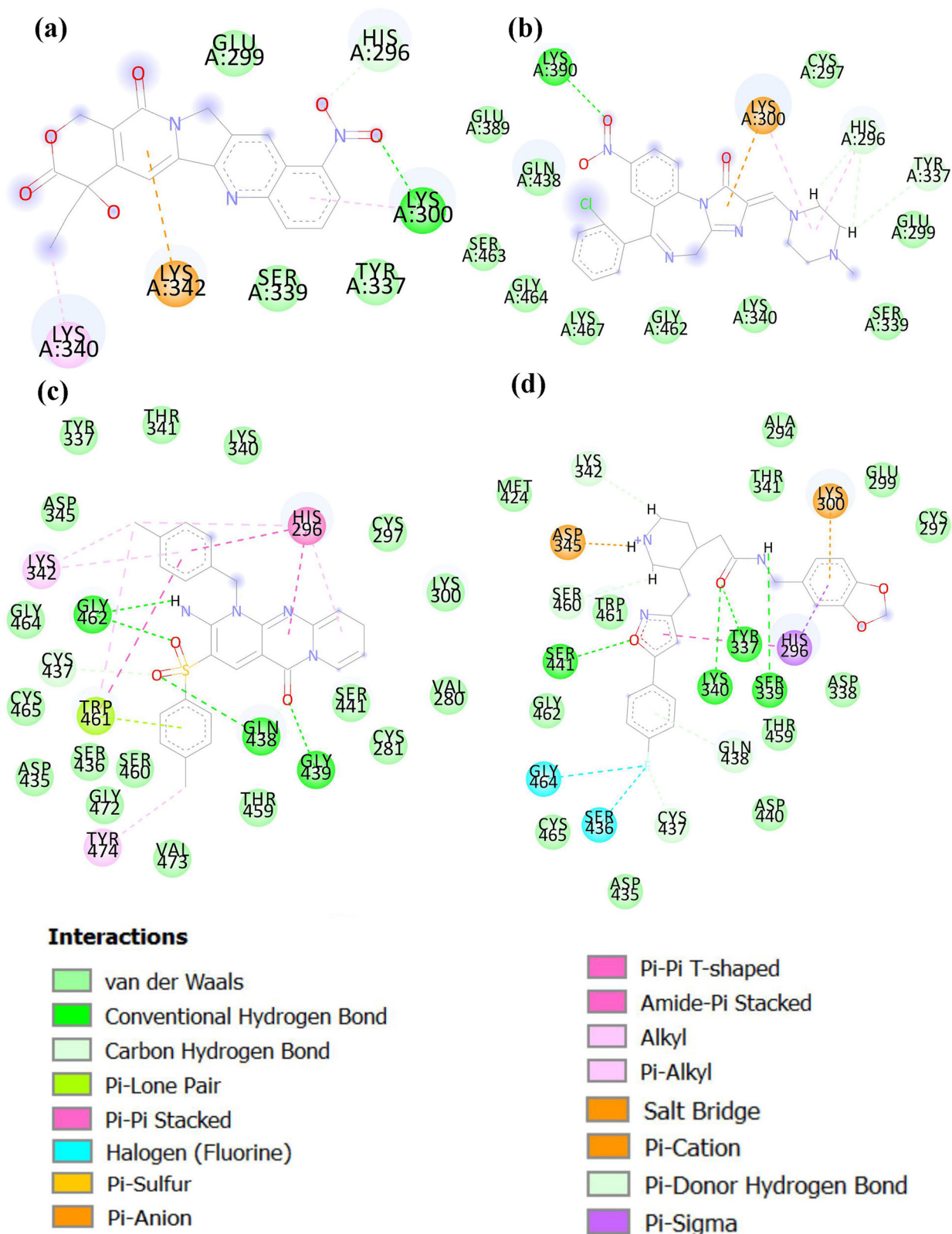
Figure 8. 2D-interaction diagram for the top-ranked TMPRSS2-complexes: (a) Rubitecan, (b) Loprazolam, (c) ZINC000000702323, (d) ZINC000012481889.

Cys145, Ser144 and Gly143, alkyl interactions with Met165 and Met49 and lastly π-alkyl interactions with Cys145 and Met49 and various van der Waals interactions with the enzyme (Figures 6d and 7d). ZINC000015988935 and ZINC000103558522 depicted more hydrogen bond numbers shared with the enzyme in comparison to Talampicillin and Lurasidone, and this explains the higher binding affinity estimated for the Mpro enzyme (Figures 6 and 7).
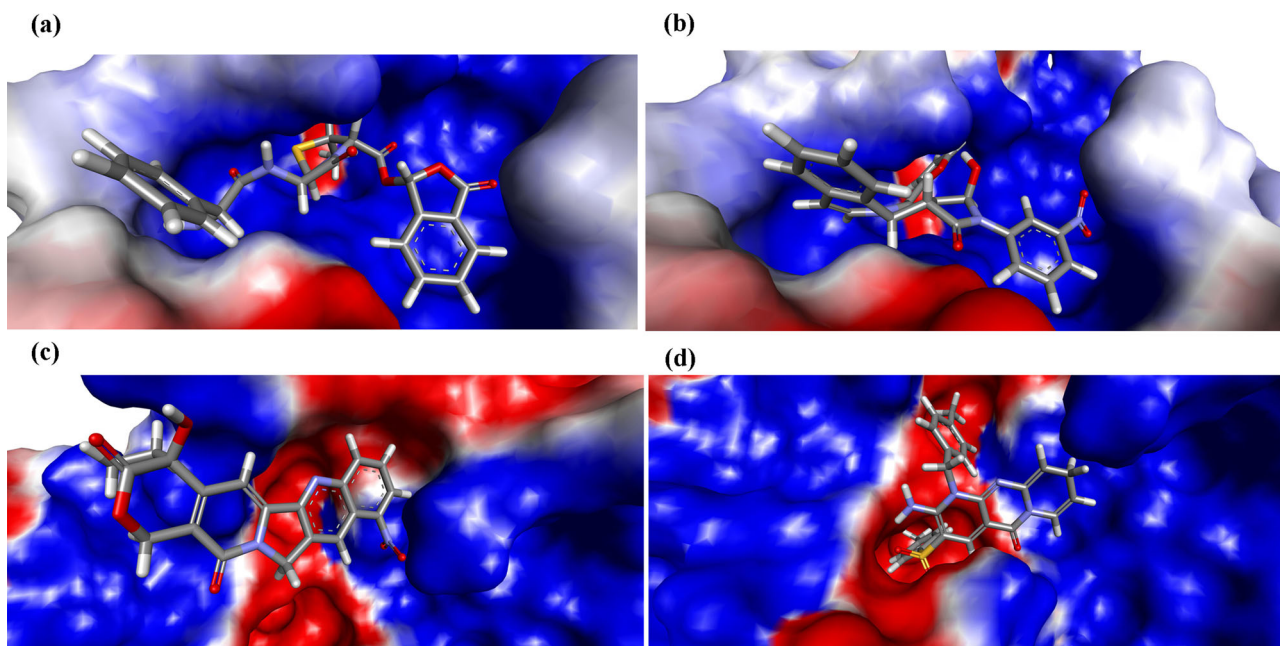
(a) **(b)**

(c) **(d)**

Figure 9. Electrostatic surface potential of top-ranked ligands. The electrostatic potential was predicted onto the surface of the protein of (a) Mpro_Talampicillin, (b) Mpro_ZINC000015988935 (c) TMPRSS2_Rubitecan and (d) TMPRSS2_ZINC000000702323. Positive and negative electrostatic charges are colored in blue and red, respectively.

Table 4. Physicochemical properties of potential drug candidates.

ChEMBL ID	MW <500	HA <10	HD <5	log P <5	TPSA ≤140	Caco-2 (cm/s)	log S > −5
ZINC000000702323	470.54	5	1	3.9	105.67	0.6357	−3.594
ZINC000012481889	451.49	7	2	2.63	85.62	0.8268	−3.471
ZINC000015988935	455.42	6	0	1.29	116.65	0.6492	−3.664
ZINC000103558522	493.51	5	0	2.83	103.51	0.6635	−3.582

TMPRSS2 potential inhibitors

Rubitecan drug bound to our model with a binding energy value of −9.47 kcal/mol and an inhibition constant of 115.06 nM. Rubitecan is a potent antitumor drug and has a high potency against solid tumors and pancreatic cancer (Clark, 2006). The nitroso group (NO−) of Rubitecan formed hydrogen bonds with His296 and Lys300 (Figure 6e, 8a). Rubitecan also formed a π-cation bond with residue Lys342, an alkyl interaction with Lys340 and a π-alkyl interaction with Lys300. According to our molecular docking study, Loprazolam was found to have a binding energy score of −9.31 kcal/mol and an inhibition constant of 149 nM against our homology modeling of TMPRSS2. Loprazolam is described for treating insomnia, hypnotic and muscle spasm (Clark et al., 1986). Loprazolam showed several interactions including π-cation with Lys300, H-bonds with Lys390, His296 and Tyr337, π-alkyl interaction with His296 and an alkyl interaction with residue Lys300 (Figure 6f, 8b). Our molecular docking study showed compound ZINC000000702323 form ZINC15 with the highest binding affinity against TMPRSS2. ZINC000000702323 was found to have a binding energy of −12.18 kcal/mol and a Ki value of 1.18 nM. ZINC000000702323 was found to have hydrogen bonds with the residues Gly462, Cys437, Gln438 and Gly439, π-lone pair interaction with Trp461, π–π interactions with residues His296 and Trp461, amide-π interaction with Trp461 and π-alkyl interactions with His296, Lys342 and Tyr474 and an alkyl interaction with the Lys342 (Figure 6g and 8c).

ZINC000012481889 compound was the second hit in our study with a binding energy of −12.14 kcal/mol and a Ki of 1.26 nM. ZINC000012481889 formed π–σ interactions with Lys300 and His296, a salt bridge with Asp345, a π-cation with Lys300, halogen bonds with Gly464 and Ser436, multiple H-bonds with the residues Lys342, Ser460, Ser441, Cys437, Gln438, Ser339, Lys340 and Tyr337, a π–π T-shaped bond with His296, a π-alkyl interaction with Lys300 as well as many van der Waals interactions (Figures 6h and 8d).

Electrostatic potential distribution

Electrostatic potential helps understanding the structural aspects of the protein and the ligands and on those interactions between them. Electrostatic charges were mapped on our homology modeling, TMPRSS2, and the coronavirus protease Mpro for better understanding the surface nature of both enzymes. Intensity of positive potential was found to cover the binding pocket of the protease Mpro (Figure 9a, b). The inner cavity of the binding pocket of the TMPRSS2 model was found to have a negative potential and the rest of the active site is covered by positive charge (Figure 9c, d).

ADMET profile and drug-likeness

Lipinski's rule of 5 was used to predict the drug-likeness of our hits from the ZINC15 library. Lipinski's rule of 5 states that for
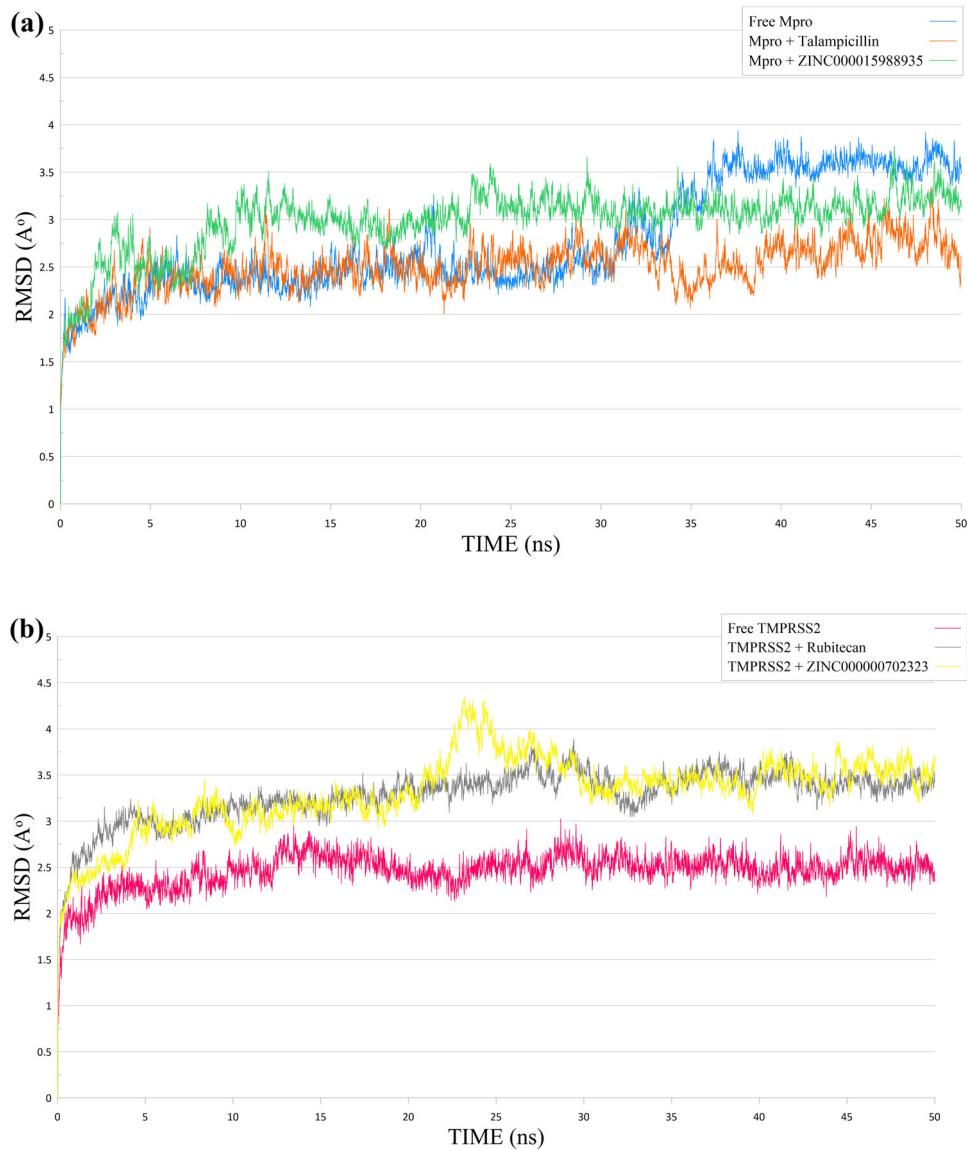
Figure 10. 50 ns-MD simulation RMSD plots of unbounded and bound enzymes. (a) Free Mpro in blue line, Mpro_Talampicillin in orange colored line and Mpro_ZINC000015988935 complex in green line. (b) Free TMPRSS2 in pink line, TMPRSS2_Rubitecan complex in gray colored line and TMPRSS2_ ZINC000000702323 in yellow line.

Figure 11. RMSF profiles of free enzymes and complexes. (a) Free Mpro in blue line, Mpro_Talampicillin in orange colored line and Mpro_ZINC000015988935 complex in green line. (b) Free TMPRSS2 in pink line, TMPRSS2_Rubitecan complex in gray colored line and TMPRSS2_ ZINC000000702323 in yellow line.
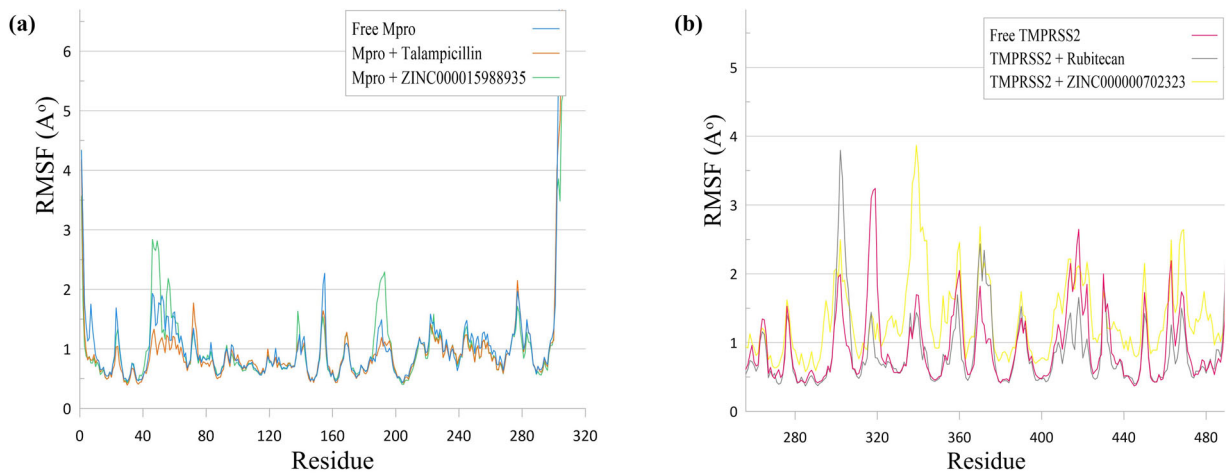
Table 5. The computed (MM-PBSA) binding free energies (ΔG_{bind}) of the identified ligands against Mpro and TMPRSS2 enzymes.

Complex	MM-PBSA (kcal/mol)
Mpro_Talampicillin	−2.80
Mpro_ZINC000015988935	−4.62
TMPRSS2_Rubitecan	−10.41
TMPRSS2_ZINC000000702323	−2.26

any ligand to be considered as a drug-like, a molecule should obey these criteria: molecular weight <500 Dalton, number of H-bond donors <5, number of H-bond acceptors <10 and LogP <5 (Lipinski, 2004). Accordingly, all of our hits obeyed this rule and thus considered as drug-like compounds. Furthermore, our compounds were found to be within the reference range considering their water solubility (LogS), Caco-2 permeability and topological polar surface area (TPSA) (Table 4).

Molecular dynamic simulation analysis

In present study, to evaluate the stability of the homology model and the drug candidates, these structures were subjected to MD simulation. Root mean square deviation (RMSD) was calculated considering the proteins backbone with respect to the initial conformations. As depicted in Figure 10a, RMSD of free Mpro enzyme remained stable between 5 ns and 34 ns timescale at 2.4 Å, then slightly increased and persisted at 3.6 Å from 35 ns till the end of the run. The RMSD of Mpro_ZINC000015988935 reached its steady state at 3.1 Å at 10 ns. The RMSD of Mpro_Talampicillin showed that the system was balanced after 5 ns and slightly fluctuated between 2.3 and 3 Å and displayed the greatest effect on Mpro's stability. As illustrated in Figure 10b, The RMSD curves of the free TMPRSS2 had reached the equilibrium state after 3 ns. The RMSD of compounds TMPRSS2_Rubitecan and TMPRSS2_ZINC000000702323 revealed that the two complexes had shown almost the same behavior during the MD simulation and achieved structural stability the last 15 ns. A thorough study of the root mean square fluctuation (RMSF) curves of the free Mpro and its complexes showed all amino acids located in the active site of the enzyme had RMSF fluctuations between 0.8 Å and 1.7 Å, which indicate that the studied compounds kept close contact with their binding pockets during the MD simulations. However, Mpro_ZINC000015988935 showed higher fluctuations for residues 43–45 and 187–190 and this observation is consistent with the RMSD profiles of Mpro_ZINC000015988935 (Figure 11a). The higher fluctuations depicted in the RMSF profiles of TMPRSS2 and its complexes are located within the loop regions of the protein, and this is expected considering the high flexibility nature of loops (Figure 11b).

Binding free energy calculation

The binding free energy (MM-PBSA) was computed after the MD simulation the last 10 ns for all the complexes and the results are given in Table 5. Compound ZINC000015988935 depicted the lowest binding free energy with Mpro enzyme, while Talampicillin exhibited relatively higher binding energy.

On the other hand, Rubitecan showed the lowest binding free energy with the homology modeling TMPRSS2 enzyme in comparison to compound ZINC000000702323.

Conclusion

Mpro and TMPRSS2 are shown to be important and highly potent targets for the inhibition of COVID-19 infection. In this study, the structure-based virtual screening identified four commercially available drugs (Talampicillin, Lurasidone, Rubitecan and Loprazolam) with a potential inhibition of Mpro and TMPRSS2 enzymes. These drugs might be repurposed against COVID-19. In addition, four novel compounds were identified from ZINC15 library showed promising high affinities against Mpro and TMPRSS2 enzymes. These hits were described as drug-like compounds and showed harmless ADMET properties and may aid in developing and optimizing more efficient and potent COVID-19 inhibitors. Trajectories analysis showed that the studied complexes have displayed structural stability during the MD runs.

Disclosure statement

No potential conflict of interest was reported by the authors.

Author contributions statement

All authors conceived, conducted and analyzed the results of experiment(s) equally. A.D.E. and K. Y. drafted the manuscript; K.Y. proofread and reviewed the manuscript.

ORCID

Ammar D. Elmezayen http://orcid.org/0000-0002-2926-7070
Anas Al-Obaidi http://orcid.org/0000-0002-6089-8188
Kemal Yelekçi http://orcid.org/0000-0002-0052-4926

References

Berry, M., Fielding, B. C., & Gamieldien, J. (2015). Potential broad spectrum inhibitors of the Coronavirus 3CLpro: A virtual screening and structure-based drug design study. *Viruses*, 7(12), 6642–6660. https://doi.org/10.3390/v7122963

Bhattacharya, D., Nowotny, J., Cao, R., & Cheng, J. (2016). 3Drefine: An interactive web server for efficient protein structure refinement. *Nucleic Acids Research*, 44(W1), W406–W409. https://doi.org/10.1093/nar/gkw336

Boonstra, S., Onck, P. R., & Van Der Giessen, E. (2016). CHARMM TIP3P Water Model Suppresses Peptide Folding by Solvating the Unfolded State. *The Journal of Physical Chemistry B*, 120(15), 3692–3698. https://doi.org/10.1021/acs.jpcb.6b01316

Böttcher, E., Matrosovich, T., Beyerle, M., Klenk, H.-D., Garten, W., & Matrosovich, M. (2006). Proteolytic activation of influenza viruses by serine proteases TMPRSS2 and HAT from human airway epithelium. *Journal of Virology*, 80(19), 9896–9898. https://doi.org/10.1128/JVI.01118-06

Clark, B. G., Jue, S. G., Dawson, G. W., & Ward, A. (1986). Loprazolam: A preliminary review of its pharmacodynamic and pharmacokinetic

properties and therapeutic efficacy in insomnia. *Drugs*, *31*(6), 500–516. doi.org/

Clark, J. W. (2006). Rubitecan. *Expert Opinion on Investigational Drugs*, *15*(1), 71–79. https://doi.org/10.1517/13543784.15.1.71

Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science*, *2*(9), 1511–1519. https://doi.org/10.1002/pro.5560020916

Worldometer. (2020). Coronavirus Update (Live): 1,790,824 cases and 109,662 Deaths from COVID-19 virus pandemic. https://www.worldometers.info/coronavirus/

Corponi, F., Fabbri, C., Bitter, I., Montgomery, S., Vieta, E., Kasper, S., Pallanti, S., & Serretti, A. (2019). Novel antipsychotics specificity profile: A clinically oriented review of lurasidone, brexpiprazole, cariprazine and lumateperone. *European Neuropsychopharmacology*, *29*(9), 971–985. https://doi.org/10.1016/j.euroneuro.2019.06.008

Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, *7*(1), 1–13. (October 2016), https://doi.org/10.1038/srep42717

Gierer, S., Bertram, S., Kaup, F., Wrensch, F., Heurich, A., Kramer-Kuhl, A., Welsch, K., Winkler, M., Meyer, B., Drosten, C., Dittmer, U., von Hahn, T., Simmons, G., Hofmann, H., & Pohlmann, S. (2013). The spike protein of the emerging betacoronavirus EMC uses a novel coronavirus receptor for entry, can be activated by TMPRSS2, and is targeted by neutralizing antibodies. *Journal of Virology*, *87*(10), 5502–5511. https://doi.org/10.1128/JVI.00128-13

Hoffmann, M., Kleine-Weber, H., Krüger, N., Müller, M., Drosten, C., & Pöhlmann, S. (2020). The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells. *Cell, 181*, 271–280. https://doi.org/10.1101/2020.01.31.929042

Iwata-Yoshikawa, N., Okamura, T., Shimizu, Y., Hasegawa, H., Takeda, M., & Nagata, N. (2019). TMPRSS2 contributes to virus spread and immunopathology in the airways of murine models after coronavirus infection. *Journal of Virology*, *93*(6), e01815–18. https://doi.org/10.1128/JVI.01815-18

Jo, S., Kim, T., Iyer, V. G., & Im, W. (2008). CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry*, *29*(11), 1859–1865. https://doi.org/10.1002/jcc.20945

Khan, R. J., Jha, R., Amera, G. M., Jain, M., Singh, E., Pathak, A., & Singh, A. (2020). Targeting novel coronavirus 2019: A systematic drug repurposing approach to identify promising inhibitors against 3C-like proteinase and 2′-O-ribose methyltransferase. *Journal of Biomolecular Structure and Dynamics*, (2), 1–40. https://doi.org/10.26434/chemrxiv.11888730

Khan, S. A., Zia, K., Ashraf, S., Uddin, R., & Ul-Haq, Z. (2020). Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 via integrated computational approach. *Journal of Biomolecular Structure and Dynamics*, 1–13. https://doi.org/10.1080/07391102.2020.1751298

Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., & Cheatham, T. E. (2000). Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Accounts of Chemical Research*, *33*(12), 889–897. https://doi.org/10.1021/ar000033j

Kyrieleis, O. J. P., Huber, R., Ong, E., Oehler, R., Hunter, M., Madison, E. L., & Jacob, U. (2007). Crystal structure of the catalytic domain of DESC1, a new member of the type II transmembrane serine proteinase family. *FEBS Journal*, *274*(8), 2148–2160. https://doi.org/10.1111/j.1742-4658.2007.05756.x

Laporte, M., & Naesens, L. (2017). Airway proteases: An emerging drug target for influenza and other respiratory virus infections. *Current Opinion in Virology*, *24*, 16–24. https://doi.org/10.1016/j.coviro.2017.03.018

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, *26*(2), 283–291. https://doi.org/10.1107/S0021889892009944

Lipinski, C. A. (2004). Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discovery Today: Technologies*, *1*(4), 337–341. https://doi.org/10.1016/j.ddtec.2004.11.007

Liu, H., & Hou, T. (2016). CaFE: A tool for binding affinity prediction using end-point free energy methods. *Bioinformatics*, *32*(14), 2216–2218. https://doi.org/10.1093/bioinformatics/btw215

Matsuyama, S., Nagata, N., Shirato, K., Kawase, M., Takeda, M., & Taguchi, F. (2010). Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. *Journal of Virology*, *84*(24), 12658–12664. https://doi.org/10.1128/JVI.01542-10

Morris, G. M., Ruth, H., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, *30*(16), 2785–2791. https://doi.org/10.1002/jcc.21256

Nicholls, A., & Honig, B. (1991). A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson–Boltzmann equation. *Journal of Computational Chemistry*, *12*(4), 435–445. https://doi.org/10.1002/jcc.540120405

Oany, A. R., Emran, A., Al., & Jyoti, T. P. (2014). Design of an epitope-based peptide vaccine against spike protein of human coronavirus: An in silico approach. *Drug Design, Development and Therapy*, *8*, 1139–1149. https://doi.org/10.2147/DDDT.S67861

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., & Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, *26*(16), 1781–1802. https://doi.org/10.1002/jcc.20289

Sarma, P., Sekhar, N., Prajapat, M., Avti, P., Kaur, H., Kumar, S., & Medhi, B. (2020). In-silico homology assisted identification of inhibitor of RNA binding against 2019-nCoV N-protein (N terminal domain). *Journal of Biomolecular Structure & Dynamics*, 1–11. https://doi.org/10.1080/07391102.2020.1753580

Shulla, A., Heald-Sargent, T., Subramanya, G., Zhao, J., Perlman, S., & Gallagher, T. (2011). A Transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. *Journal of Virology*, *85*(2), 873–882. https://doi.org/10.1128/JVI.02062-10

Sielaff, F., Böttcher-Friebertshäuser, E., Meyer, D., Saupe, S. M., Volk, I. M., Garten, W., & Steinmetzer, T. (2011). Development of substrate analogue inhibitors for the human airway trypsin-like protease HAT. *Bioorganic & Medicinal Chemistry Letters*, *21*(16), 4860–4864. https://doi.org/10.1016/j.bmcl.2011.06.033

Sliwoski, G., Kothiwale, S., Meiler, J., & Lowe, E. W. (2014). Computational methods in drug discovery. *Pharmacological Reviews*, *66*(1), 334–395. https://doi.org/10.1124/pr.112.007336

Trott, O., & Olson, A. J. (2009). Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, *31*(2), NA–461. https://doi.org/10.1002/jcc.21334

Vyas, V. K., Ukawala, R. D., Ghate, M., & Chintha, C. (2012). Homology modeling a fast tool for drug discovery: Current perspectives. *Indian Journal of Pharmaceutical Sciences*, *74*(1), 1–17. https://doi.org/10.4103/0250-474X.102537

Webb, B., & Sali, A. (2016). Comparative protein structure modeling using MODELLER. *Current Protocols in Bioinformatics*, *54*(1), 5.6.1–5.6.37. https://doi.org/10.1002/cpbi.3

Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, *35*(Web Server), W407–410. https://doi.org/10.1093/nar/gkm290

Xiang, Z. (2006). Advances in homology protein structure modeling. *Current Protein & Peptide Science*, *7*(3), 217–227. https://doi.org/10.2174/138920306777452312

Yang, H., Lou, C., Sun, L., Li, J., Cai, Y., Wang, Z., Li, W., Liu, G., & Tang, Y. (2019). AdmetSAR 2.0: Web-service for prediction and optimization of chemical ADMET properties. *Bioinformatics*, *35*(6), 1067–1069. https://doi.org/10.1093/bioinformatics/bty707

Zhou, J., Fang, L., Yang, Z., Xu, S., Lv, M., Sun, Z., Chen, J., Wang, D., Gao, J., & Xiao, S. (2019). Identification of novel proteolytically inactive mutations in coronavirus 3C-like protease using a combined approach. *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, *33*(12), 14575–14587. https://doi.org/10.1096/fj.201901624RR