Fast Identification of Possible Drug Treatment of Coronavirus Disease -19 (COVID-19) Through Computational Drug Repurposing Study

Junmei Wang*

Department of Pharmaceutical Sciences and Computational Chemical Genomics Screening Center, School of Pharmacy, University of Pittsburgh, Pittsburgh, PA 15261, USA.

Corresponding author: Junmei Wang, junmei.wang@pitt.edu

*To whom correspondence should be addressed

Abstract

The recent outbreak of novel coronavirus disease -19 (COVID-19) calls for and welcomes possible treatment strategies using drugs on the market. It is very efficient to apply computer-aided drug design techniques to quickly identify promising drug repurposing candidates, especially after the detailed 3D-structures of key virous proteins are resolved. Taking the advantage of a recently released crystal structure of COVID-19 protease in complex with a covalently-bonded inhibitor, N3,[1] I conducted virtual docking screening of approved drugs and drug candidates in clinical trials. For the top docking hits, I then performed molecular dynamics simulations followed by binding free energy calculations using an endpoint method called MM-PBSA-WSAS.[2-4] Several promising known drugs stand out as potential inhibitors of COVID-19 protease, including Carfilzomib, Eravacycline, Valrubicin, Lopinavir and Elbasvir. Carfilzomib, an approved anti-cancer drug acting as a proteasome inhibitor, has the best MM-PBSA-WSAS binding free energy, -13.82 kcal/mol. Streptomycin, an antibiotic and a charged molecule, also demonstrates some inhibitory effect, even though the predicted binding free energy of the charged form (-3.82 kcal/mol) is not nearly as low as that of the neutral form (-7.92 kcal/mol). One bioactive, PubChem 23727975, has a binding free energy of -12.86 kcal/mol. Detailed receptor-ligand interactions were analyzed and hot spots for the receptor-ligand binding were identified. I found that one hotspot residue HIS41, is a conserved residue across many viruses including COVID-19, SARS, MERS, and HCV. The findings of this study can facilitate rational drug design targeting the COVID-19 protease.

1. Introduction

A great application of drug repurposing is to identify drugs which were developed for treating other diseases to treat a new disease. Drug repurposing can be achieved by conducting systematic drug-drug target interaction (DTI) and drug-drug interaction (DDI) analyses. We have conducted a survey on DTIs collected by the DrugBank database[5] and found that on average each drug has 3 drug targets and each drug target has 4.7 drugs.[6] The analysis demonstrates that polypharmacology is a common phenomenon. It is important to identify potential DTIs for both approved drugs and drug candidates, which serves as the basis of repurposing drugs and selection of drug targets without DTIs that may cause side-effects. Polypharmacology opens novel avenues to rationally design next generation of more effective but less toxic therapeutic agents. Computer-aided drug design (CADD) has been playing essential roles in modern drug discovery and development. To balance the computational efficiency and accuracy, a hieratical strategy employing different types of scoring functions are applied in both the drug lead identification and optimization phases. A docking scoring function, such as the one employed by the Glide docking program,[7] is very efficient and thus can be utilized to screen a large library, but it is not very accurate. On the other hand, the molecular mechanical force field (MMFF)-based scoring functions, are physical and more accurate, but much less efficient. With the ever increasing computer power, MMFF-based free energy calculation methods, such as the endpoint MM-PB/GBSA (molecular mechanics-Poisson Boltzmann/ Generalized Born Surface Area) methods[2, 3, 8-21] and the alchemical thermodynamic integration (TI) and free energy perturbation (FEP) methods,[22, 23] have been extensively applied in structure-based drug discovery projects. Recently we've developed a hierarchical virtual screening (HVS)to balance the efficiency and accuracy and improve the success rate of rational drug design.[8, 24]The newly released crystal structure of COVID-19[1] provides a solid structural basis for identification of drugs that might interact with this protein target. In this work, I applied multiscale modeling techniques to identify drugs that may be repurposed to target COVID-19 protease Flexible docking and MM-PBSA-WSAS were applied as the 1st and 2nd filters, respectively, to improve the efficiency and accuracy of HVS to identify inhibitors of COVID-19. Compared to the experimental means, CADD-based approaches are more efficient in providing possible treatment solutions for epidemic disease outbreaks like COVID-19. The detailed ligand-residue interaction profile as well as the decomposition of binding free energy

into different components provide insight into rationally designing potent and selective inhibitors targeting COVID-19 protease.

2. Methodologies.

I conducted a hierarchical virtual screening (HVS) using the newly resolved crystal structure of COVID-19 protease (Resolution 2.16Å).[1] Two types of HVS filters were employed: Glide[7] flexible docking followed by MM-PBSA-WSAS.[2,4] Detailed computational methods are described below.

2.1 Docking Screening

The crystal structure was first treated using the protein structure preparation wizard provided by the Schrodinger software, followed by docking grid generation. Glide flexible docking was performed using the default settings except that the formation of intramolecular hydrogen bonds was rewarded and the enhancement of planarity of conjugated pi groups was turned on. The co-crystal ligand, N3, was covalently bonded to CYS145. I generated a new version of N3, N3' by breaking the covalent bond and filling in open valence. I then evaluated whether Glide flexible docking can reproduce the native binding pose. In addition, dataset of approved drugs was prepared using DrugBank,[5] and a set of PubChem compounds which are structurally similar to Lopinavir were enriched for docking screenings. Lopinavir, a potent inhibitor of HIV-1 protease,[25] was found effective in treating COVID-19 patients. Top hits from the docking screenings were advanced to the next HVS filter – MM-PBSA-WSAS.

2.2 System setup for molecular dynamics (MD) simulation and free energy calculation

MD simulations were first performed for a docking hit for two purposes: (1) studying the relative stability of the ligand residing in the binding pocket; (2) sampling a set of conformations for MM-PBSA-WSAS binding free energy calculations and MM-GBSA residue-ligand binding free energy decomposition analysis. A MD system consisted of one copy of COVID-19 protease, one copy of docked ligand, 17597 TIP3P[26] water molecules, about 50 Na^+ and Cl^- ions depending on the charge state of the ligand. The whole system was neutralized. For the force field parameters, the partial atomic charges of ligands were derived using the RESP[27] program to fit the HF/6-31G* electrostatic potentials generated using the GAUSSIAN 16 software package[28]. The other force

field parameters were derived from GAFF[29] and the AMBER FF14SB[30] force field to model proteins. The residue topologies for ligands were prepared using the Antechamber module.[31] For the covalently-bonded N3 ligand, I applied the residuegen program to generate non-standard amino acid residue topology.

2.3 MD Simulation Protocols.

For a protein-ligand complex, the MD system was first relaxed through a series of minimization procedures. The mainchain atoms of the receptor and the bound ligand were restrained using a harmonic potential and its force constant decreased from 20 to 10, 5, 1 and 0 kcal/mol/Å^2, progressively in five 10,000-step minimizations. Note that the last step applied no restraint at all as the force constant is 0. The system was further relaxed by a set of 100-picosecond atomistic MD simulations with the same restrain setting of minimizations.

There were three phases for a MD simulation: the relaxation phase, the equilibrium phase, and the sampling phase. In the relaxation phase, the simulation system was heated up progressively from 50 K to 250 K at steps of 50 K. At each temperature, a 1-nanosecond MD simulation was performed without any restraints or constraints. In the next equilibrium phase, the system was equilibrated at 298 K, 1 bar for 20 ns. Finally, a 100-nanosecond MD simulation was performed at 298 K, 1 bar to produce NTP (constant temperature and pressure) ensembles. In total, 10,000 snapshots were recorded from the last simulation. 200 snapshots were evenly selected for the MM-PBSA-WSAS binding free energy calculation and 5000 were selected for the MM-GBSA ligand-protein binding free energy decomposition analysis. Additional settings for constant pressure MD simulations performed in this work are listed as follows: temperature was regulated using Langevin dynamics[32] with a collision frequency of 5 ps^{-1}; pressure was regulated using the isotropic position scaling algorithm with the pressure relaxation time set to 1.0 ps; integration of the equations of motion was conducted at a time step of 1 fs for the two relaxation phases and 2 fs for the equilibrium and sampling phases. The Particle Mesh Ewald (PME) method[33] was used to calculate the full electrostatic energy of a unit cell in a macroscopic lattice of repeating images. All bonds were constrained using the SHAKE algorithm[34] in both the minimization and MD simulation stages. All MD simulations were performed using the pmemd program in AMBER 18.[35]

2.4 MM-PBSA-WSAS Binding Free Energy Calculation

200 MD snapshots were evenly selected for the binding free energy calculations. For each selected MD snapshot, the molecular mechanical (MM) energy (E_{MM}) and the MM-PBSA solvation free energy were calculated without further minimization.[8, 10, 11, 36-38] Key parameters controlling the MM-PBSA-WSAS analyses are listed as follows: external dielectric constant: 80; internal dielectric constant: 4; and the surface tension for estimating the nonpolar solvation energy by using solvent assessible surface area: 0.054. The Parse radii[39] were used in the MM-PBSA solvation calculation using the Delphi package (http://compbio.clemson.edu/delphi). The entropic term was estimated using a method coined WSAS (weighted solvent accessible surface area) described elsewhere.[4] It is noted that the entropic contribution cannot be neglected for this protein target as most ligands are large and have many rotatable bonds.

2.5 MM-GBSA Ligand-Residue Free Energy Decomposition Analysis

I conducted ligand-residue free energy decomposition analysis for 5000 snapshots evenly selected from the sampled snapshots. Besides the electrostatic and van der Waals interactions, the solvation effect was taken into account using a Generalized GB model developed by Onufriev et al.[40] The ligand-residue MM-GBSA interaction energies were calculated using the Sander program in AMBER18.[35] Data analysis was performed using an internal program developed by us. A hotspot residue is recognized when its ligand-residue MM-GBSA interaction is stronger than -1.0 kcal/mol.

3. Results

In this work, I have performed two-step hierarchical virtual screenings to identify repurposing drugs targeting COVID-19 protease.

3.1 Docking Screenings.

After downloading the crystal structure of COVID-19, I performed Glide docking screenings for a set of datasets (approved drugs, investigational drugs, and experimental drugs) downloaded from DrugBank.[5] I first evaluated the docking power of Glide for the co-crystal ligand of COVID-19, N3. The ligand RMSD of the best docking pose based on docking score (-9.398), 3.32 Å, was acceptable for a big ligand of about 100 atoms in flexible docking. I then applied the docking setting to conduct docking screenings. All the drug molecules that had docking scores better than

-8.5 kcal/mol, which accounted for about 1% of total screening compounds, were selected as hits and advanced to the next filter – MM-PBSA-WSAS.

3.2 MD Simulations

For the promising docking hits, I conducted molecular dynamics (MD) simulations using the AMBER software package.[35] In total 39 ligands including the co-crystal N3 ligand, were studied in the second phase of HVS. The Top 5 approved neutral drugs that have excellent MM-PBSA-WSAS binding free energies (ΔG_{bind} <=-5.0 kcal/mol) are shown in Figure 1. The 2D structures of charged drugs with at least one form achieved ΔG_{bind} <=-5.0 kcal/mol are shown in Figure 2. I also found two bio-actives (Figure 3), which are structurally similar to Lopinavir, have excellent binding free energies (Section 3.3). It is noted that Lopinavir was observed to be effective in treating COVID-19.

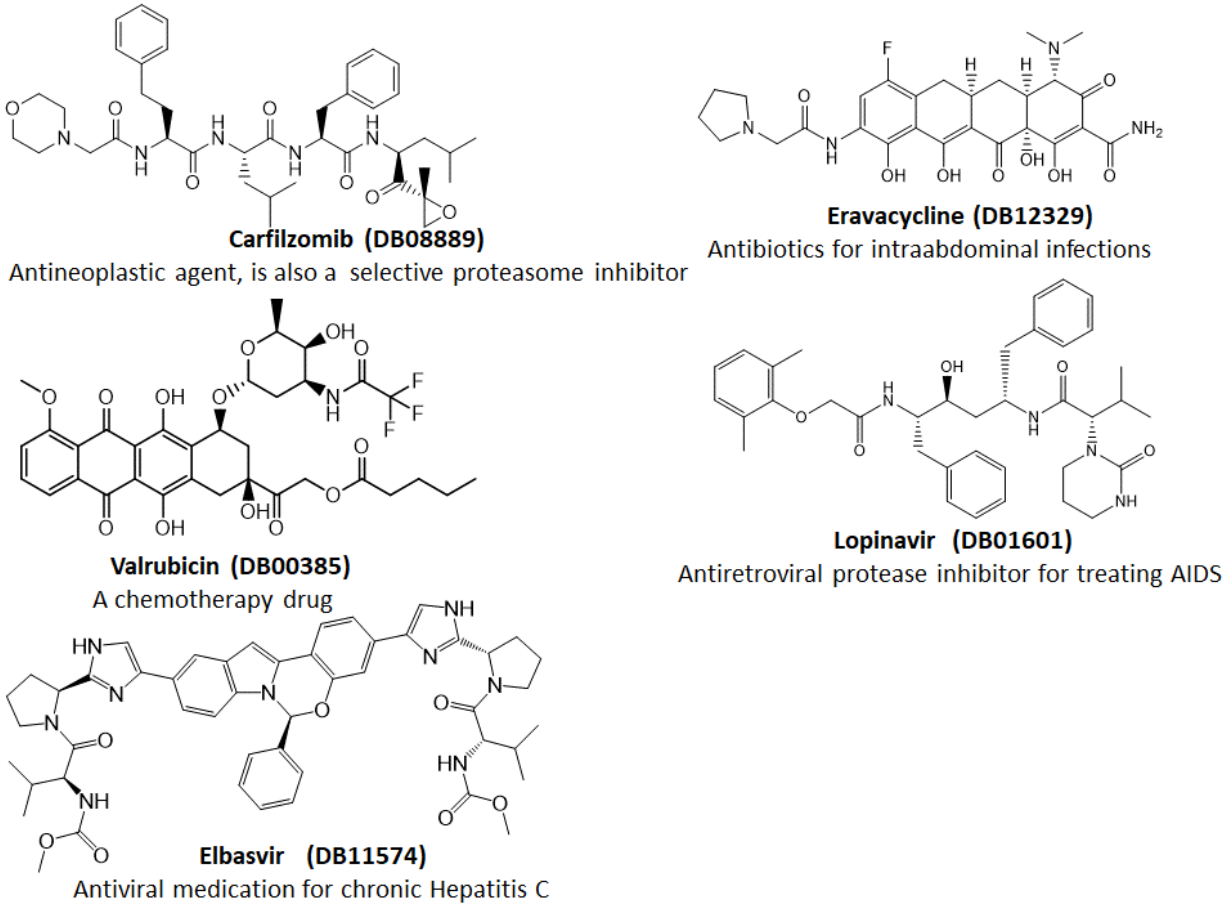
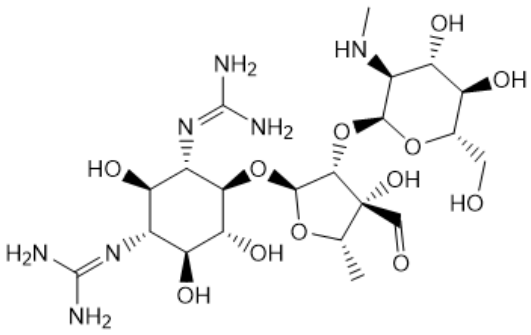
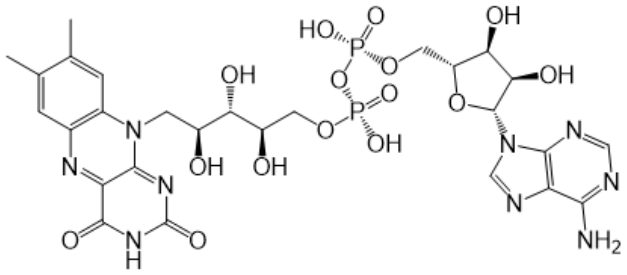


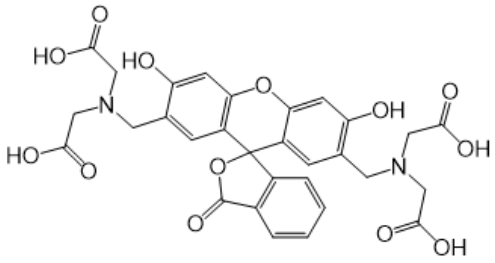
Figure 1. 2D-Structures of promising repurpose drugs. All five approved drugs are in neural form under physiological conditions.



Streptomycin (DB01082)
An aminoglycoside antibiotic

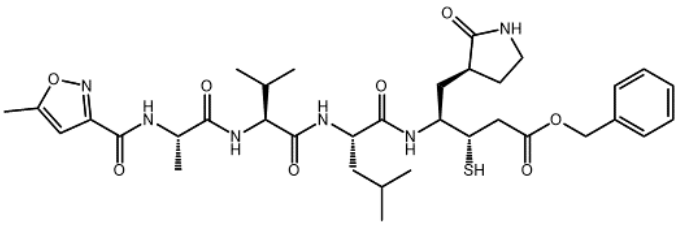


Flavin adenine dinucleotide (DB03147)
Ophthalmic treatment for vitamin B2 deficiency

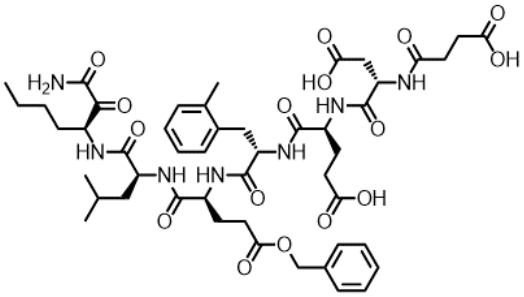


Oftasceine (DB11184)
A staining agent

Figure 2. 2D-Structures of promising repurposing drugs. All three approved drugs are in charged form under physiological conditions.



PubChem: 23727975



PubChem: 88143175

Figure 3. 2D-Structures of promising bio-actives which are structurally similar to Lopinavir. PubChem 88143175, although studied in neutral form, bears -3 charges under physiological conditions.

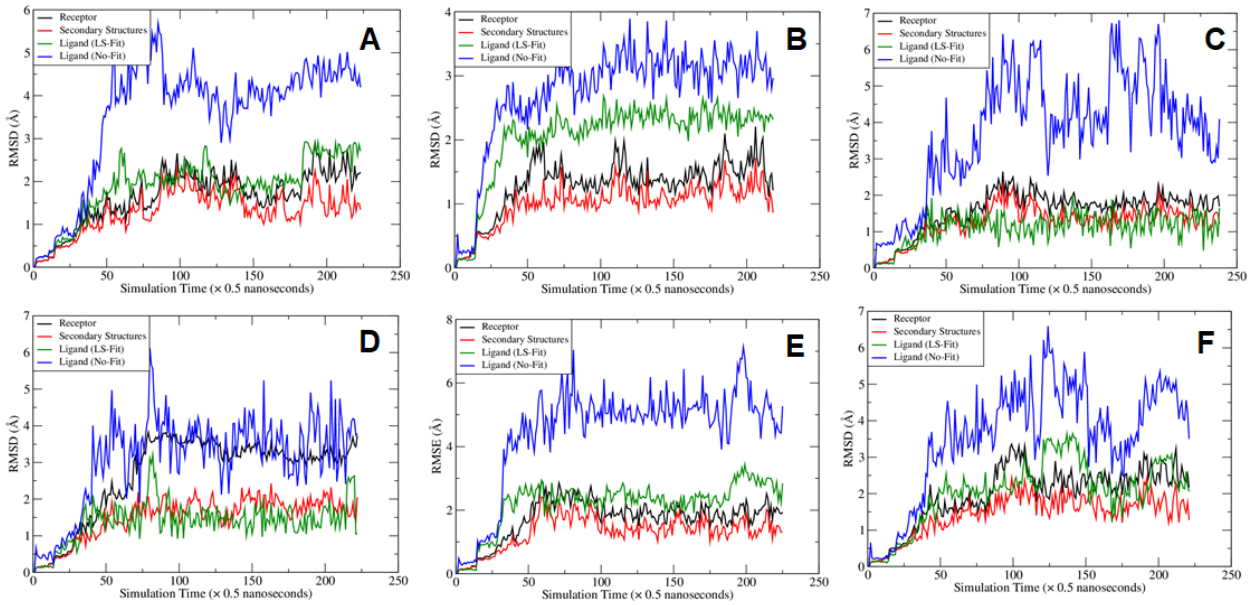


Figure 4. Plots of Root-mean-square deviations of receptor main chain atoms and ligand heavy atoms along the MD simulation time for (A) co-crystal ligand N3, (B) DB08889, (C) DB12329, (D) DB00385, (E) DB01601 and (F) DB11574.

I explored the MD stability of each MD system. Figure 4 showed the RMSD fluctuations along the MD simulation time. It is shown that the mainchain atoms of the receptor (black curves) and the secondary structures (red curves) reached equilibrium after 20 nanoseconds. The least-square (LS) fitted RMSDs of the ligands (green curves) are around 2 Å, which is reasonable for large ligands like COVID-19 protease inhibitors. A ligand's No-Fit RMSD was calculated by first performing LS-fitting for the main chain atoms of the receptor and the resulting translation-rotation matrix was applied to the ligand, and then the RMSD was calculated directly. Evidently, the ligand No-Fit RMSDs measures not only the conformational changes, but also its translational and rotational movements inside the binding pocket. The ligand No-Fit RMSDs (blue curves) are larger than the LS-Fit values, however, those values around 3-4 Å are still acceptable for large ligands. In summary, the RMSD fluctuation analysis suggests that the MD trajectories are overall stable during the sampling phase for all the studied MD systems.

I first generated the average structure of the collected snapshots, and then the MD snapshot that had the smallest main chain atom RMSD against the average structure was chosen as the representative conformation. The comparisons between the crystal structure and the representative MD conformations are shown in Figure 5 for the native ligand N3, and Figures 6-8 for the other

ligands. As shown in Figure 5, the benzene motif located in the dashed red cycle was inserted between two hotspot residues (HIS41 and MET49) for the MD structure, which is quite distinct from the crystal structure, which shows that the benzene motif has no direct interactions with HIS4 and MET49. I believe that under physiological conditions, the benzene motif becomes less solvent exposed and has more favorable interactions with HIS4 and MET49 by inserting itself between the side chains of the two residues. As shown below, both HIS4 and MET49 were hotspot residues in our MM-GBSA free energy decomposition analysis.

3.3 MM-PBSA-WSAS Binding Free Energy Calculations

I measured the ligand binding affinity using the endpoint MM-PBSA method. Considering the ligands of COVID-19 protease are flexible molecules with large sizes, the contribution from conformational entropy cannot be neglected. Instead of applying normal mode analysis to estimate the entropic effect, I applied an efficient method called WSAS[4] to calculate this energy term. This scoring function is therefore called MM-PBSA-WSAS. The calculated binding free energies and the Glide docking scores are summarized in Table 1. The calculated entropic term, $T\Delta S$, is quite different for different ligands as shown in Table 1, suggesting the necessity of including this term in binding free energy calculations. The structures of the promising drug repurposing candidates, which have both excellent docking scores and MM-PBSA-WSAS binding affinities are shown in Figures 1-3. All the known drugs shown in Figure 1 are neutral and have a better MM-PBSA-WSAS affinity than -5.0 kcal/mol. It should be noted that the cocrystal ligand, N3 is covalently bonded to the receptor, therefore its binding free energy is not directly comparable to those non-covalent ligands. The individual terms of MM-PBSA-WSAS binding free energies of other less potent ligands were summarized in Table S1.

The values of each energy term, van der Waals (ΔE_{VDW}), electrostatics ($\Delta E_{VDW} + \Delta G_{PB}$), nonpolar solvation term (ΔG_{SA}), and entropy ($T\Delta S$), vary significantly from one system to another (Table 1 and Table S1), suggesting there is no single energy term that dominates the protein-ligand interaction.

For the charged drug molecules, caution should be taken in result interpretation. For example, the neural form of Streptomycin (DB01082) has a MM-PBSA-WSAS binding free energy of -7.92 kcal/mol, much better than the charged form (-3.82 kcal/mol). However, the latter is dominant

under physiological conditions. We therefore should use the result of the charged form or take the penalty of protonation into consideration when using the result of the neutral form.

Table 1. List of Glide docking scores and MM-PBSA-WSAS binding free energies for potential inhibitors binding to COVID-19 protease (in kcal/mol). PubChem IDs are listed for the two bio-actives. The entropic contribution was estimated using T of 298.15 K.

Compound Name	Docking Score	ΔE_{EEL}	ΔE_{VDW}	ΔG_{PB}	ΔG_{SA}	$T\Delta S$	ΔG_{bind}
Co-crystal ligand covalently bonds to sulfur of CYS145	-	-75.13 ± 0.27	-81.97 ± 0.39	84.01 ± 0.23	-6.47 ± 0.02	-29.75 ± 0.14	-38.79 ± 0.64
Co-crystal ligand (no covalent bond formed)	-9.40	-52.28 ± 0.27	-26.90 ± 0.30	56.32 ± 0.15	-4.82 ± 0.04	-24.14 ± 0.05	-3.55 ± 0.36
Neutral Approved Drugs							
DB08889	-8.56	-75.66 ± 0.41	-40.93 ± 0.09	**78.31 ± 0.42**	-5.97 ± 0.02	-30.43 ± 0.07	-13.82 ± 0.20
DB12329	-8.75	-45.70 ± 0.39	-25.46 ± 0.55	45.00 ± 0.28	-3.40 ± 0.01	-21.82 ± 0.04	-7.73 ± 0.52
DB00385	-9.19	-59.84 ± 0.23	-21.49 ± 0.21	52.85 ± 0.41	-4.55 ± 0.01	-25.86 ± 0.07	-7.16 ± 0.13
DB01601 (Lopinavir)	-9.77	-52.46 ± 0.33	-20.09 ± 0.63	46.58 ± 0.56	-4.59 ± 0.02	-23.93 ± 0.01	-6.63 ± 0.28
DB11574	-9.89	-70.57 ± 0.36	-21.78 ± 0.41	65.64 ± 0.64	-6.38 ± 0.01	-26.57 ± 0.17	-6.53 ± 0.31
Charged Approved Drugs							
DB01082 (NC=0)	-8.61	-45.99 ± 0.35	-71.69 ± 0.95	89.35 ± 0.88	-3.78 ± 0.02	-24.19 ± 0.04	-7.92 ± 0.41
DB01082 (NC=2)	-6.88	-33.35 ± 0.56	-279.99 ± 1.55	291.70 ± 1.28	-3.41 ± 0.03	-21.24 ± 0.09	-3.82 ± 0.52
DB03147 (NC=0)	-10.22	-67.82 ± 0.26	-67.58 ± 0.92	106.61 ± 0.62	-5.00 ± 0.01	-26.31 ± 0.06	-7.48 ± 0.53
DB03147 (NC=-2)	-8.30	-52.86 ± 0.05	123.04 ± 0.26	-74.27 ± 0.51	-4.49 ± 0.01	-23.22 ± 0.06	14.65 ± 0.29
DB11184 (NC=0)	-8.64	-57.38 ± 0.82	-52.20 ± 0.36	82.23 ± 0.29	-4.86 ± 0.01	-25.38 ± 0.06	-6.82 ± 0.26
DB11184 (NC=-4)	-7.44	-50.10 ± 0.35	175.38 ± 1.59	-148.74 ± 0.82	-4.45 ± 0.02	-22.90 ± 0.06	-5.00 ± 0.52
Bio-actives Structurally Similar to Lopinavir							
23727975	-8.84	-63.78 ± 0.10	-50.11 ± 0.66	77.91 ± 0.10	-5.16 ± 0.01	-28.28 ± 0.09	-12.86 ± 0.51
88143175 (NC = 0)	-10.05	-73.93 ± 0.24	-104.49 ± 1.29	148.93 ± 0.86	-6.64 ± 0.02	-30.48 ± 0.04	-5.65 ± 0.62

3.4 MM-GBSA Free Energy Decomposition.

I performed MM-GBSA binding free energy decomposition to identify the hotspot residues which make substantial contributions to the protein-ligand binding. The identified hotspots could enable us to rationally design potent and selective inhibitors of this drug target. To obtain statistically meaningful results, I studied 5000 MD snapshots for each system, and both the average ligand-residue interaction energies ($\Delta G_{lig\text{-}res}$) and their RMSD values were calculated.

A hotspot residue is defined as a residue with $\Delta G_{lig\text{-}res}$ equal to or smaller than -1.0 kcal/mol. The identified hotspots of each ligand are summarized in Table 2. The most significant hotspot residues ($\Delta G_{lig\text{-}res}$<-3.0) are illustrated in Figures 5-8. The common significant hotspot residues for

most ligands (in bold in Table 2) are as follows: HIS41, MET49, ASN142, HIS164, MET165, GLU166, and GLN189.

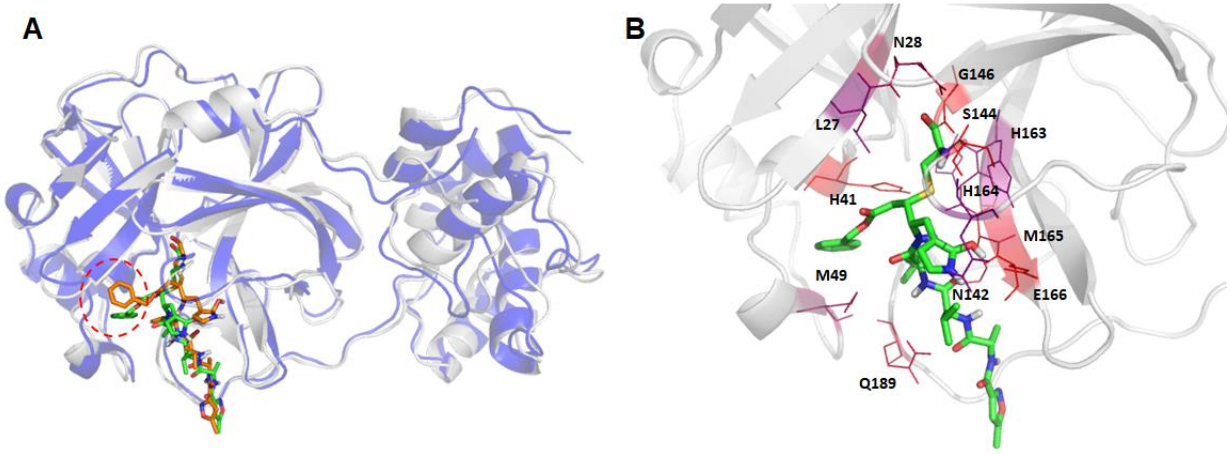


Figure 5. Structural comparison between the crystal structure and a representative MD structureof COVID-19 protease bound to the known ligand, N3 . The crystal structure is shown as blue cartoon with the co-crystal ligand shown as brown sticks, while the representative MD structure is shown in grey cartoon and the ligand as green sticks (Panel **A**). The hotspot residues ($\Delta G_{Lig\text{-}Res}$ < -3.0 kcal/mol) revealed by MM-GBSA analysis are shown in Panel B; the more bluish a residue is colored, the stronger the interaction between the residue and the ligand.

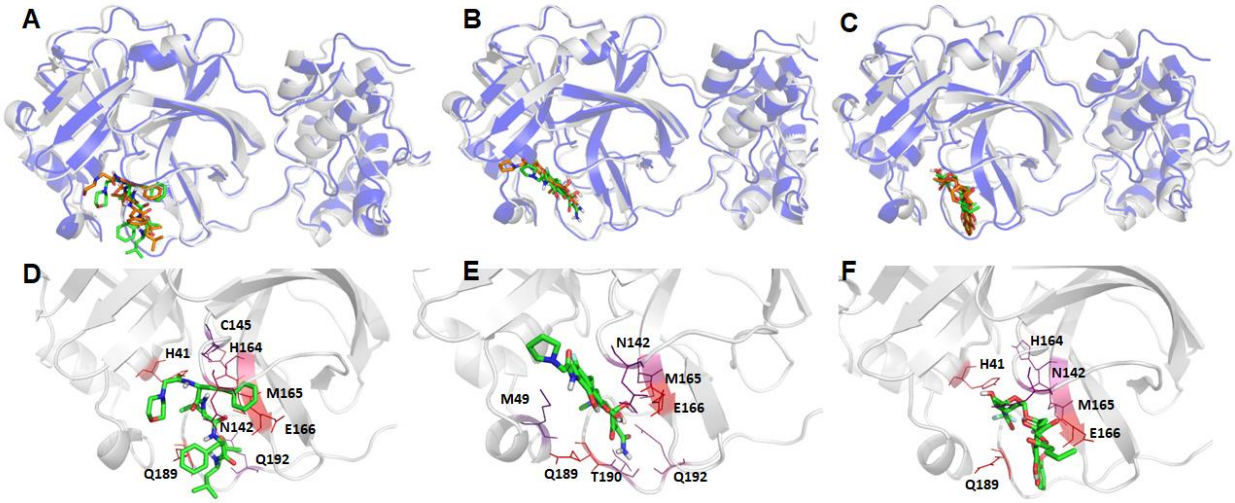


Figure 6. Structural comparison between the crystal structure and a representative MD structure of COVID-19 protease bound to three neutral ligands DB08889, DB12329, and DB00385. The crystal structure is shown as blue cartoon with the docked ligand shown as brown sticks, while the representative MD structure is shown in grey cartoon and the ligand as green sticks. A: DB08889,

B: DB12329, and C: DB00385. The detailed ligand-receptor interactions are shown in the bottom panel (D-F). All the hotspot residues ($\Delta G_{Lig\text{-}Res} < -3.0$) revealed by MM-GBSA analyses are labeled and colored by a blue to red spectrum, the more bluish a residue is colored, the stronger the interaction between the residue and the ligand. D: DB08889, E: DB12329, and F: DB00385.

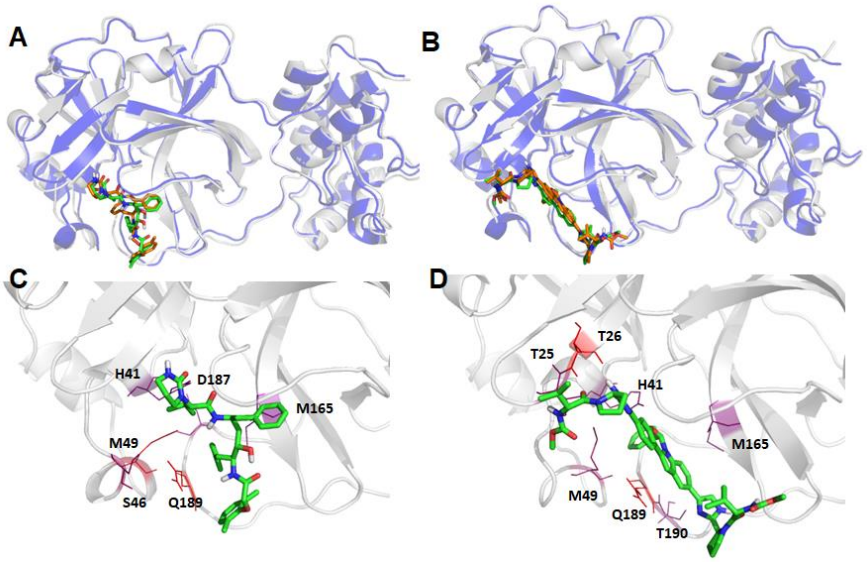


Figure 7. Structural comparison between the crystal and a representative MD structure of COVID-19 protease bound to two neutral ligands DB01601 and DB11574. The crystal structure is shown as blue cartoon with the docked ligand shown as brown sticks, while the representative MD structure is shown in grey cartoon and the ligand as green sticks. A: DB01601 and B: DB11574. The detailed ligand-receptor interactions are shown in the bottom panel (C-D). All the hotspot residues ($\Delta G_{Lig\text{-}Res} < -3.0$) revealed by MM-GBSA analyses are labeled and colored by a blue to red spectrum, the more bluish a residue is colored, the stronger interaction between the residue and the ligand. C: DB01601 and D: DB11574.

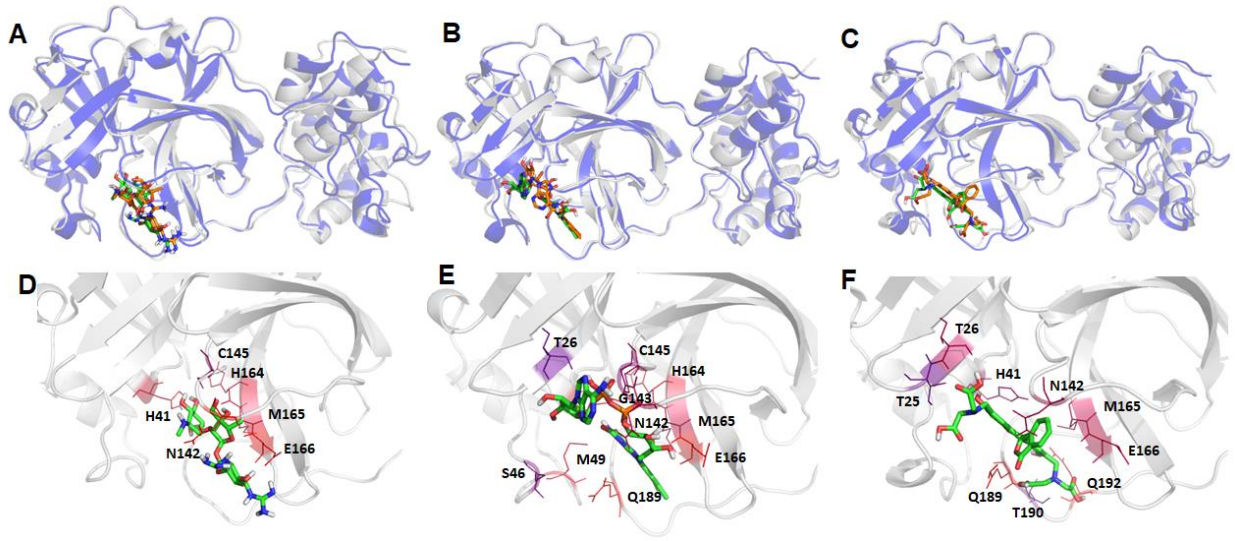


Figure 8. Structural comparison between the crystal structure and a representative MD structure of COVID-19 protease bound to three charged ligands DB01082, DB03147, and DB11184. The crystal structure is shown as blue cartoon with the docked ligand shown as brown sticks, while the representative MD structure is shown in grey cartoon and the ligand as green sticks. A: DB01082, B: DB03147, and C: DB11184. The detailed ligand-receptor interactions are shown in the bottom panel (D-F). All the hotspot residues ($\Delta G_{Lig\text{-}Res} < -3.0$) revealed by MM-GBSA analyses are labeled and colored by a blue to red spectrum, the more bluish a residue is colored, the stronger the interaction between the residue and the ligand. D: DB01082, E: DB03147, and F: DB11184.

Table 2: Ligand-residue MM-GBSA interaction energies (kcal/mol). PubChem IDs are listed for the two bio-actives. The common hotspots for all the ligands are: HIS41, MET49, ASN142, HIS164, MET165, GLU166, and GLN189.

Residue ID	Residue Type	Co-crystal Ligand	Neutral Approved Drug					Charged Approved Drug			Bioactive	
			DB08889	DB12329	DB00385	DB01601	DB11574	DB01082	DB03147	DB11184	23727975	88143175
24	THR	-0.09	-0.15	-0.04	-0.04	-0.06	-0.93	-0.01	-0.10	-0.28	-0.08	-1.40
25	THR	-1.51	-1.50	-0.15	-0.55	-1.56	-4.04	-0.10	-1.60	-3.18	-0.64	-3.68
26	THR	-0.61	-0.90	-0.17	-0.21	-0.27	-6.19	-0.07	-3.08	-4.51	-0.17	-4.33
27	LEU	-3.74	-1.24	-0.52	-1.03	-1.26	-2.20	-0.71	-2.60	-2.48	-0.25	-1.64
28	ASN	-4.78	-0.09	-0.03	-0.06	-0.05	-0.19	-0.10	-0.22	-0.07	-0.08	-0.14
39	PRO	-0.84	-0.19	-0.09	-0.13	-0.24	-0.20	-0.31	-0.32	-0.11	-0.13	-0.17
41	**HIS**	**-6.74**	**-6.56**	**-2.93**	**-5.41**	**-3.58**	**-3.79**	**-5.38**	**-7.78**	**-4.18**	**-3.02**	**-2.99**
44	CYS	-1.02	-0.13	-0.09	-0.07	-1.37	-0.67	-0.05	-2.50	-0.37	-0.09	-0.43
45	THR	-0.60	-0.32	-0.24	-0.06	-1.10	-0.59	-0.02	-1.76	-0.26	-0.17	-0.47

46	SER	-0.91	-1.79	-1.54	-0.62	-4.26	-1.45	-0.07	-3.53	-1.28	-1.79	-2.70
49	**MET**	**-4.39**	**-2.37**	**-3.14**	**-2.14**	**-5.57**	**-4.18**	**-1.07**	**-5.46**	**-2.23**	**-2.17**	**-3.93**
52	PRO	-0.76	-0.06	-0.08	-0.08	-0.03	-0.19	-0.01	-0.07	-0.02	-0.06	-0.04
54	TYR	-0.65	-0.12	-0.45	-0.22	-0.14	-0.54	-0.15	-0.17	-0.11	-0.09	-0.05
119	ASN	-0.16	-0.02	-0.07	-0.09	-0.02	-1.41	-0.02	-0.44	-0.05	-0.02	-0.05
140	PHE	-0.19	-0.98	-0.04	-0.09	-0.02	-0.06	-0.80	-0.08	-0.17	-0.35	-1.81
141	LEU	-0.28	-1.16	-0.10	-0.23	-0.03	-0.09	-1.00	-0.05	-0.27	0.49	-2.64
142	**ASN**	**-3.50**	**-4.39**	**-3.46**	**-3.66**	**-0.59**	**-2.89**	**-6.70**	**-4.56**	**-4.58**	**-5.50**	**-6.82**
143	GLY	-2.23	-0.75	-0.96	-0.36	-0.37	-2.34	-1.87	-3.77	-1.50	-2.48	-2.11
144	SER	-13.05	-1.42	-0.12	-0.17	-0.10	-0.31	-2.09	-1.19	-0.60	-5.29	-2.33
145	CYS	-66.34	-3.22	-0.71	-1.53	-1.59	-2.09	-3.95	-4.27	-1.50	-2.96	-3.83
146	GLY	-10.20	-0.02	-0.01	-0.03	-0.03	-0.02	-0.06	-0.04	-0.03	-0.06	-0.06
147	SER	-0.64	-0.05	-0.01	-0.02	-0.01	-0.02	-0.06	-0.05	-0.02	-0.02	-0.10
163	HIS	-3.88	-1.65	-0.13	-0.38	-0.14	-0.17	-2.82	-0.40	-0.63	-1.22	-1.18
164	**HIS**	**-3.92**	**-4.89**	**-1.41**	**-3.65**	**-1.80**	**-1.56**	**-5.26**	**-5.44**	**-1.30**	**-2.51**	**-2.53**
165	**MET**	**-5.49**	**-7.83**	**-4.16**	**-4.63**	**-3.46**	**-3.65**	**-5.27**	**-4.91**	**-4.74**	**-6.50**	**-4.43**
166	**GLU**	**-7.31**	**-6.67**	**-6.05**	**-6.76**	**-0.82**	**-1.77**	**-17.06**	**-6.44**	**-4.91**	**-14.38**	**-23.80**
167	LEU	-1.67	-1.36	-1.26	-2.11	-0.63	-1.28	-1.92	-0.78	-1.31	-2.06	-1.16
168	PRO	-2.73	-1.17	-1.11	-2.50	-1.26	-2.66	-2.55	-0.35	-1.18	-2.15	-0.91
170	GLY	-0.05	-0.08	-0.01	-0.03	-0.02	-0.07	-0.54	-0.01	-0.03	-0.11	-0.17
172	HIS	-0.17	-1.15	-0.09	-0.18	-0.07	-0.10	-0.70	-0.09	-0.33	-1.07	-1.07
186	VAL	-0.31	-0.61	-0.28	-0.18	-0.33	-0.37	-0.09	-0.60	-0.27	-0.28	-0.29
187	ASP	-1.49	-1.29	-2.24	-1.36	-3.74	-1.90	-2.85	-2.05	-1.64	-1.34	-1.17
188	ARG	-1.60	-2.04	-1.94	-1.74	-1.68	-2.34	-0.57	-1.75	-1.85	-1.82	-1.17
189	**GLN**	**-5.25**	**-13.79**	**-5.90**	**-11.52**	**-8.10**	**-7.69**	**-2.27**	**-6.11**	**-8.78**	**-10.07**	**-6.93**
190	THR	-1.42	-1.56	-3.94	-2.43	-1.55	-3.84	-0.05	-1.08	-3.22	-0.99	-0.60
191	ALA	-1.65	-2.48	-0.23	-0.68	-1.52	-1.56	-0.03	-0.16	-0.81	-0.46	-0.13
192	GLN	-2.24	-3.48	-3.96	-2.28	-1.13	-1.80	-0.16	-1.23	-6.28	-1.72	-0.75

4. Discussion

The outbreak of highly infectious diseases such as COVID-19 demands to work out multiple treatment plans as soon as possible. Computational drug repurposing study can provide treatment options in a short period of time. For this study, amounts of computational time used for individual tasks are as follows. Docking screenings of all the 2201 approved drugs with a single CPU core (Intel Xeon CPU E5-2683) took 11 hours. For each docking hit, we need to perform *ab initio* calculations to derive point charges. The *ab initio* calculations using wB97XD/6-31G*//HF/6-31G* consumed about 1 day using four CPU cores; then it took us about 1.2 days to sample 120 nanoseconds using one GTX-1080 ti GPU; the following MM-PBSA-WSAS calculation consumed one day. Therefore, equipped with sufficient numbers of CPUs and GPUs and the

current hardware, we can finish the drug repurposing screenings within four to five days using a reliable HVS strategy. Given that the inhibitors of COVID-19 protease have relatively large sizes, the screening time can be even shorter for other drug targets with smaller ligands.

Another consideration is the availability of high-quality drug target structures. Luckily, a high-resolution crystal structure of COVID-19 protease in complex with a ligand was resolved timely, allowing us to conduct this drug repurpose screening. If no high-quality structure is available, one can rely on homology modeling technique, probably with a reduced success rate of identifying repurposing drugs. Take COVID-19 protease as an example, I performed structural alignments using an internal program which takes a multiple-sequence-alignment (MSA) as an input. The MSA was generated by using the Promals3D web server.[41] The structure of COVID-19 protease is found to be most similar to those of SARS protease (PDB Code 3TNT[42]) and less similar to MERS protease (PDB Code 5WKK[43]) (Figure 9A). In comparison, the structure of HCV NS3/4A (PDB Code 3M5L[44]) is quite different: the RMSD of 2.26 Å between HCV and COVID-19 is much larger and with only 108 residues participating the least-square fitting (Figure 9B). I also compared the sequences of the four proteases around the seven hotspot residues, which are colored in red in Table 3. It is shown that COVID-19 and SARS share all the seven hotspot residues. MERS and COVID-19 have four of the seven common hotspot residues, while HCV NS3/4A and COVID-19 have only one common hotspot residue (H41). Even though the sequence identity is low between COVID-19 and HCV NS3/4A, as shown in Figure 9B, the co-crystal ligands, N3 (green sticks) for COVID-19 and ITMN-191 (brown sticks) for HCV NS3/4A, largely overlap. This suggests that homology models can be constructed using Modeller[45] with SARS, MERS and even HCV NS3/4A as templates.

Table 3. Sequence comparison around hotspot residues for proteases of four types of viruses.

Virus Protease	PDB Code	Residue ID of Hot Spots (colored in red)				
		41	49	142	164-166	189
COVID-19	6LU7	...CPRHVIC...	SEDMLNP...	FLNGSC...	CYMHHMELP...	VDRQTAQ...
SARS	3TNT	...CPRHVIC...	AEDMLNP...	FLNGSC...	CYMHHMELP...	VDRQTAQ...
MERS	5WKK	...CPRHVMC...	ADQLSDP...	FLCGSC...	CYMHQMELA...	MDKQVHQ...
HCV NS3/4A	3M5L	...TVYHGAG...	-------...	YLKGSA...	VGIFRAAVS...	-------...

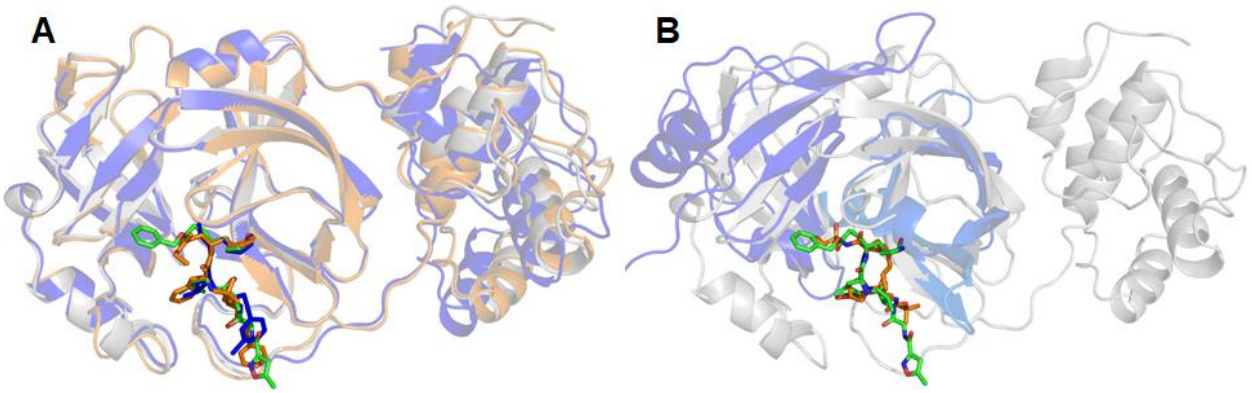


Figure 9. Structural comparison of proteases among three coronavirus viruses (COVID-19, SARS and MERS) (**A**), and between COVID-19 and hepatitis C NS3/4A proteases (B). The COVID-19 protease is colored in grey and its ligands are shown as green sticks. The following are the color codes for the other proteases: SARS protease and its co-crystal ligand – brown, MERS protease and its co-crystal ligand – blue, HCV NS3/4A – blue, co-crystal ligand of HCV NS3/4A – brown. Backbone RMSD between SARS and COVID-19 is 0.4711 Å, with 284 residues participating in the least-square fitting and 22 omitted, and the backbone RMSD between MERS and COVID-19 is 0.41 Å, but with 195 residues participating in the least-square fitting and 104 omitted. In contrast, the backbone RMSD between COVID-19 protease and HCV NS3/4A is 2.2632 Å, with 108 residues participating in the least-square fitting and 43 omitted.

5. Conclusion

In this study, I took advantage of the recently released crystal structure of COVID-19 protease and conducted multiscale drug repurposing screenings. Five neutral drugs, namely, Carfilzomib, Eravacycline, Valrubicin, Lopinavir and Elbasvir, are identified to have inhibitory activities against COVID-19 protease. Streptomycin, a charged molecule may also be an inhibitor of this COVID-19 protease. Our study suggests that computational drug repurposing screening is very efficient and it can provide potential repurposing drug candidates in less than five days. A set of hotspot residues which make substantial contributions to the protein-ligand binding are also identified, which can facilitate us to rationally design novel selective inhibitors targeting COVID-19.

ASSOCIATED CONTENT

Supporting Information.

The MM-PBSA-WSAS binding free energies for other less promising docking hits ($\Delta G_{lig\text{-}res}$ worse than -5.0 kcal/mol) are listed in Table S1.

AUTHOR INFORMATION

Corresponding Author

* Phone: (412) 383-3268. Fax: (412) 383-7436. E-mail: juw79@pitt.edu

ORCID

Junmei Wang: 0000-0002-9607-8229

Notes

The author declares no competing financial interest.

ACKNOWLEDGMENT

The author gratefully acknowledges the funding support from the National Institutes of Health (NIH) to J.W. (R01GM079383 and R21GM097617). The author also thanks for the computing resources provided by the Center for Research Computing (CRC) at University of Pittsburgh. The author is grateful to Dr. Xiang Liu at Nankai University who kindly provided the PDB file of 6LU7 to me prior its release at www.rcsb.org.

SUPPORTING INFORMATION

Table S1. List of Glide docking scores and MM-PBSA-WSAS binding free energies for potential inhibitors binding to COVID-19 protease (in kcal/mol). The net charge (NC) of charged molecules is shown in the parentheses.

Compound Name	Docking Score	ΔE_{EEL}	ΔE_{VDW}	ΔG_{PB}	ΔG_{SA}	ΔTS	ΔG_{bind}
DB01601 Docking Pose 2	-9.41	-51.59 ± 0.19	-18.20 ± 0.30	47.37 ± 0.25	-4.75 ± 0.02	-23.40 ± 0.06	-3.76 ± 0.24
10011418	-8.88	-46.92 ± 0.35	-69.66 ± 0.68	95.64 ± 0.69	-4.76 ± 0.02	-22.70 ± 0.13	-3.01 ± 0.30
6102717	-9.95	-47.25 ± 0.18	-22.44 ± 0.48	51.55 ± 0.60	-4.34 ± 0.02	-22.84 ± 0.06	0.36 ± 0.31
DB00183	-10.39	-43.26 ± 0.19	-27.36 ± 0.64	54.68 ± 1.23	-4.36 ± 0.02	-21.57 ± 0.11	1.27 ± 0.60
DB01072	-9.08	-46.84 ± 0.17	-16.77 ± 0.51	47.33 ± 0.46	-4.74 ± 0.01	-22.81 ± 0.15	1.80 ± 0.62

DB01204	-10.10	-51.07 ± 0.20	-31.24 ± 0.17	58.54 ± 0.58	-4.07 ± 0.03	-23.35 ± 0.03	-4.49 ± 0.43
DB01232	-9.05	-44.17 ± 0.40	-8.55 ± 0.45	34.73 ± 0.35	-4.43 ± 0.03	-20.98 ± 0.13	-1.42 ± 0.31
DB01264	-9.86	-37.23 ± 0.11	-31.21 ± 0.57	47.83 ± 0.64	-2.91 ± 0.01	-20.33 ± 0.07	-3.19 ± 0.20
DB01282	-8.68	-49.14 ± 0.52	-66.51 ± 0.39	91.76 ± 0.20	-4.39 ± 0.04	-23.79 ± 0.04	-4.50 ± 0.33
DB01328 (NC = 0)	-8.52	-37.19 ± 0.14	-23.92 ± 0.67	49.46 ± 0.56	-3.20 ± 0.02	-19.82 ± 0.06	4.96 ± 0.07
DB01328 (NC = -2)	-6.12	-44.85 ± 0.33	114.98 ± 0.81	-80.85 ± 0.74	-3.72 ± 0.01	-21.93 ± 0.06	7.49 ± 0.84
DB01698	-8.55	-48.39 ± 0.15	-46.01 ± 0.39	73.67 ± 0.44	-3.91 ± 0.02	-22.73 ± 0.08	-1.91 ± 0.77
DB03141 (NC = 0)	-9.85	-53.00 ± 0.27	-27.43 ± 0.42	62.85 ± 0.31	-4.63 ± 0.01	-22.45 ± 0.03	0.24 ± 0.24
DB03247 (NC = -2)	-7.38	-47.63 ± 0.48	73.97 ± 1.86	-39.54 ± 1.98	-3.64 ± 0.01	-21.75 ± 0.14	4.91 ± 0.33
DB03310 (NC = 0)	-8.67	-54.38 ± 0.40	-76.02 ± 0.92	107.76 ± 1.01	-4.24 ± 0.03	-26.13 ± 0.04	-0.75 ± 0.34
DB03310 (NC = -2)	-6.93	-39.48 ± 0.22	45.88 ± 2.57	-20.92 ± 2.79	-3.26 ± 0.03	-21.76 ± 0.05	3.99 ± 0.26
DB06441 (NC = -4)	-6.93	-52.54 ± 0.39	194.37 ± 2.84	-141.28 ± 2.09	-4.34 ± 0.01	-24.44 ± 0.04	20.66 ± 0.85
DB06717 (NC = -2)	-8.54	-34.85 ± 0.68	177.39 ± 4.93	-136.79 ± 2.22	-3.43 ± 0.01	-21.06 ± 0.08	23.38 ± 2.09
DB08818 (NC = 0)	-9.06	-44.07 ± 0.31	-80.21 ± 0.27	105.43 ± 0.52	-4.63 ± 0.02	-23.85 ± 0.08	0.38 ± 0.44
DB08818 (NC = -2)	-9.33	-44.91 ± 0.13	75.93 ± 1.54	-49.16 ± 1.31	-4.26 ± 0.01	-22.78 ± 0.04	0.37 ± 0.61
DB09059	-9.02	-61.28 ± 0.16	-33.85 ± 0.68	75.13 ± 0.81	-5.00 ± 0.03	-25.76 ± 0.06	0.77 ± 0.63
DB09065	-9.01	-58.19 ± 0.12	-14.29 ± 0.11	48.25 ± 0.39	-5.29 ± 0.02	-25.38 ± 0.08	-4.14 ± 0.29
DB11602	-8.88	-52.58 ± 0.45	-33.10 ± 0.43	73.03 ± 0.19	-5.07 ± 0.03	-24.74 ± 0.11	7.02 ± 0.19
DB13074	-9.22	-40.00 ± 0.31	-22.68 ± 0.80	45.24 ± 0.44	-3.27 ± 0.04	-19.97 ± 0.12	-0.74 ± 0.08

REFERENCES

1. Liu, X.; Zhang, B.; Jin, Z.; Yang, H.; Rao, Z., The crystal structure of COVID-19 main protease in complex with an inhibitor N3. **2020**.
2. Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T., End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem Rev* **2019,** *119* (16), 9478-9508.
3. Wang, J. M.; Hou, T. J.; Xu, X. J., Recent Advances in Free Energy Calculations with a Combination of Molecular Mechanics and Continuum Models. *Curr Comput-Aid Drug* **2006,** *2* (3), 287-306.
4. Wang, J.; Hou, T., Develop and test a solvent accessible surface area-based model in conformational entropy calculations. *Journal of chemical information and modeling* **2012,** *52* (5), 1199-212.
5. Law, V.; Knox, C.; Djoumbou, Y.; Jewison, T.; Guo, A. C.; Liu, Y.; Maciejewski, A.; Arndt, D.; Wilson, M.; Neveu, V.; Tang, A.; Gabriel, G.; Ly, C.; Adamjee, S.; Dame, Z. T.; Han, B.; Zhou, Y.; Wishart, D. S., DrugBank 4.0: shedding new light on drug metabolism. *Nucleic acids research* **2014,** *42* (Database issue), D1091-7.
6. Wang, J.; Ge, Y.; Xie, X. Q., Development and Testing of Druglike Screening Libraries. *Journal of chemical information and modeling* **2019,** *59* (1), 53-65.
7. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S., Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry* **2004,** *47* (7), 1739-49.
8. Wang, J.; Morin, P.; Wang, W.; Kollman, P. A., Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. *J Am Chem Soc* **2001,** *123* (22), 5221-5230.
9. Swanson, J. M.; Henchman, R. H.; McCammon, J. A., Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy. *Biophys J* **2004,** *86* (1 Pt 1), 67-74.

10. Kuhn, B.; Gerber, P.; Schulz-Gasch, T.; Stahl, M., Validation and use of the MM-PBSA approach for drug discovery. *Journal of medicinal chemistry* **2005,** *48* (12), 4040-4048.
11. Hou, T.; Wang, J.; Li, Y.; Wang, W., Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of chemical information and modeling* **2011,** *51* (1), 69-82.
12. Xu, L.; Sun, H.; Li, Y.; Wang, J.; Hou, T., Assessing the performance of MM/PBSA and MM/GBSA methods. 3. The impact of force fields and ligand charge models. *J Phys Chem B* **2013,** *117* (28), 8408-21.
13. Sun, H.; Li, Y.; Shen, M.; Tian, S.; Xu, L.; Pan, P.; Guan, Y.; Hou, T., Assessing the performance of MM/PBSA and MM/GBSA methods. 5. Improved docking performance using high solute dielectric constant MM/GBSA and MM/PBSA rescoring. *Phys Chem Chem Phys* **2014,** *16* (40), 22035-45.
14. Sun, H.; Li, Y.; Tian, S.; Xu, L.; Hou, T., Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set. *Phys Chem Chem Phys* **2014,** *16* (31), 16719-29.
15. Chen, F.; Liu, H.; Sun, H.; Pan, P.; Li, Y.; Li, D.; Hou, T., Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein-protein binding free energies and re-rank binding poses generated by protein-protein docking. *Phys Chem Chem Phys* **2016,** *18* (32), 22129-39.
16. Karami, M.; Jalali, C.; Mirzaie, S., Combined virtual screening, MMPBSA, molecular docking and dynamics studies against deadly anthrax: An in silico effort to inhibit Bacillus anthracis nucleoside hydrolase. *J Theor Biol* **2017,** *420*, 180-189.
17. Chen, F.; Sun, H.; Wang, J.; Zhu, F.; Liu, H.; Wang, Z.; Lei, T.; Li, Y.; Hou, T., Assessing the performance of MM/PBSA and MM/GBSA methods. 8. Predicting binding free energies and poses of protein-RNA complexes. *RNA* **2018,** *24* (9), 1183-1194.
18. Mishra, S. K.; Koca, J., Assessing the Performance of MM/PBSA, MM/GBSA, and QM-MM/GBSA Approaches on Protein/Carbohydrate Complexes: Effect of Implicit Solvent Models, QM Methods, and Entropic Contributions. *J Phys Chem B* **2018,** *122* (34), 8113-8121.
19. Sun, H.; Duan, L.; Chen, F.; Liu, H.; Wang, Z.; Pan, P.; Zhu, F.; Zhang, J. Z. H.; Hou, T., Assessing the performance of MM/PBSA and MM/GBSA methods. 7. Entropy effects on the performance of end-point binding free energy calculation approaches. *Phys Chem Chem Phys* **2018,** *20* (21), 14450-14460.
20. Wang, E.; Weng, G.; Sun, H.; Du, H.; Zhu, F.; Chen, F.; Wang, Z.; Hou, T., Assessing the performance of the MM/PBSA and MM/GBSA methods. 10. Impacts of enhanced sampling and variable dielectric model on protein-protein Interactions. *Phys Chem Chem Phys* **2019,** *21* (35), 18958-18969.
21. Weng, G.; Wang, E.; Chen, F.; Sun, H.; Wang, Z.; Hou, T., Assessing the performance of MM/PBSA and MM/GBSA methods. 9. Prediction reliability of binding affinities and binding poses for protein-peptide complexes. *Phys Chem Chem Phys* **2019,** *21* (19), 10135-10145.
22. Lee, T. S.; Hu, Y.; Sherborne, B.; Guo, Z.; York, D. M., Toward Fast and Accurate Binding Affinity Prediction with pmemdGTI: An Efficient Implementation of GPU-Accelerated Thermodynamic Integration. *J Chem Theory Comput* **2017,** *13* (7), 3077-3084.
23. Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R., Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *J Am Chem Soc* **2015,** *137* (7), 2695-2703.

24. Wang, J.; Kang, X.; Kuntz, I. D.; Kollman, P. A., Hierarchical database screenings for HIV-1 reverse transcriptase using a pharmacophore model, rigid docking, solvation docking, and MM-PB/SA. *Journal of medicinal chemistry* **2005,** *48* (7), 2432-44.
25. Capparelli, E. V.; Holland, D.; Okamoto, C.; Gragg, B.; Durelle, J.; Marquie-Beck, J.; van den Brande, G.; Ellis, R.; Letendre, S.; Group, H., Lopinavir concentrations in cerebrospinal fluid exceed the 50% inhibitory concentration for HIV. *AIDS* **2005,** *19* (9), 949-52.
26. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983,** *79* (2), 926.
27. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A., A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges- the RESP model. *J Phys Chem* **1993,** *97* (40), 10269-10280.
28. Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J., Gaussian 16, Revision B.01. *Gaussian, Inc., Wallingford CT* **2016**.
29. Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *Journal of Computational Chemistry* **2004,** *25* (9), 1157-1174.
30. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **2015,** *11* (8), 3696-713.
31. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* **2006,** *25* (2), 247-260.
32. Larini, L.; Mannella, R.; Leporini, D., Langevin stabilization of molecular-dynamics simulations of polymers by means of quasisymplectic algorithms. *J Chem Phys* **2007,** *126* (10), 104101.
33. Darden, T.; Perera, L.; Li, L.; Pedersen, L., New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure* **1999,** *7* (3), R55-60.
34. Miyamoto, S.; Kollman, P. A., Settle - an analytical version of the Shake and Rattle algorithm for rigid water models. *Journal of Computational Chemistry* **1992,** *13* (8), 952-962.
35. Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, I., T.E. ; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; D.M., Y.; Kollman, P. A., AMBER 2018. *University of California, San Francisco* **2018**.
36. Page, C. S.; Bates, P. A., Can MM-PBSA calculations predict the specificities of protein kinase inhibitors? *Journal of Computational Chemistry* **2006,** *27* (16), 1990-2007.
37. Kongsted, J.; Ryde, U., An improved method to predict the entropy term with the MM/PBSA approach. *J Comput Aided Mol Des* **2009,** *23* (2), 63-71.

38. Hou, T.; Wang, J.; Li, Y.; Wang, W., Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking. *J Comput Chem* **2011,** *32* (5), 866-77.
39. Sitkoff, D.; Sharp, K. A.; Honig, B., Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *J Phys Chem-Us* **1994,** *98* (7), 1978-1988.
40. Onufriev, A.; Bashford, D.; Case, D. A., Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* **2004,** *55* (2), 383-94.
41. Pei, J.; Grishin, N. V., PROMALS: towards accurate multiple sequence alignments of distantly related proteins. *Bioinformatics* **2007,** *23* (7), 802-8.
42. Hilgenfeld, R., From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. *FEBS J* **2014,** *281* (18), 4085-96.
43. Galasiti Kankanamalage, A. C.; Kim, Y.; Damalanka, V. C.; Rathnayake, A. D.; Fehr, A. R.; Mehzabeen, N.; Battaile, K. P.; Lovell, S.; Lushington, G. H.; Perlman, S.; Chang, K. O.; Groutas, W. C., Structure-guided design of potent and permeable inhibitors of MERS coronavirus 3CL protease that utilize a piperidine moiety as a novel design element. *Eur J Med Chem* **2018,** *150*, 334-346.
44. Romano, K. P.; Ali, A.; Royer, W. E.; Schiffer, C. A., Drug resistance against HCV NS3/4A inhibitors is defined by the balance of substrate recognition versus inhibitor binding. *Proc Natl Acad Sci U S A* **2010,** *107* (49), 20986-91.
45. Webb, B.; Sali, A., Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Protein Sci* **2016,** *86*, 5.6.1-5.6.37.