Identifying the Hot Spot Residues of the SARS-CoV-2 Main Protease Using MM-PBSA and Multiple Force Fields

Jinyoung Byun

Kangwon National University

Juyong Lee (✉ drfaust23@gmail.com)

Kangwon National University

Research Article

Keywords: SARS-CoV-2 virus, hot spot, main protease, MM-PBSA, multiple force fields

Posted Date: June 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-598309/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Life on December 31st, 2021. See the published version at <https://doi.org/10.3390/life12010054>.

Identifying the hot spot residues of the SARS-CoV-2 main protease using MM-PBSA and multiple force fields

Jinyoung Byun¹ and Juyong Lee^{1,*}

¹Department of Chemistry, Division of Chemistry and Biochemistry, Kangwon National University, Chuncheon, 24341, Republic of Korea

*juyong.lee@kangwon.ac.kr

ABSTRACT

In this study, we investigated the binding affinities between the main protease of SARS-CoV-2 virus and its various ligand to identify the hot spot residues of the protease. To investigate the effect of various force fields, we performed MD simulations with three different force fields: GROMOS54a7, Amber99-SB, and CHARMM36. The total amount of MD simulation time was 1.1 μ s. To investigate how known ligands interact with M^{Pro} of SARS-CoV-2, the binding affinities were calculated by using the MMPBSA approach. It is identified that no single force field succeeded in predicting the relative rankings of experimental binding affinities. When compared between different force fields, Amber99-SB and GROMOS54a7 results are fairly correlated while CHARMM36 results show weak or almost no correlations with the others. Additionally, we identified specific residues of M^{Pro}, which contribute more importantly to the binding energies with ligands. It is identified that the residues of the S4 subsite of the binding site, N142, M165, and R188, contribute strongly to ligand binding. In addition, the terminal residues, D295, R298, and Q299 are identified to have attractive interactions with ligands via electrostatic and solvation energy. We believe that our findings will help facilitate develop novel inhibitors of SARS-CoV-2.

Introduction

The new type of coronavirus, SARS-CoV-2 virus (COVID-19), outbreak in China at the end of 2019. Since its first outbreak, it has led to more than 160 million infections and over 3.4 million deaths^{1,2}. Although many vaccines are approved and being developed, no effective small molecule drug has not been discovered³. Because it plays an important role in replication of the virus and transcription in processing the poly-proteins that are translated from the viral RNA in the replication of the virus, the main protease (M^{Pro}) of COVID-19 is considered as a promising drug target to cure COVID-19^{4,5}. In developing new drugs, computer-aided drug discovery (CADD) is one of the most efficient ways to reduce consuming time and cost. Thus it is critical to evaluate the binding free energy between a ligand and its receptor both accurately and efficiently. Because the shape of the

binding site of M^{pro} was well characterized⁴⁻⁷, docking studies between ligands and the M^{pro} receptor has been carried out to find out possible hits or leads^{8,9}. The docking has been used to perform a structure-based study to reveal potential covalent inhibitors from the AFCL library consisting of 1000 covalent molecules and 116 FDA-approved molecules⁸. Peele et al. screened 24 natural plant-derived molecules with antiviral property, 22 FDA-approved antiviral drugs, and 16 antimalarial drugs with the docking approach⁹. They found several molecules, lopinavir, amodiaquine, and theaflavin digallate that showed a good binding affinity with M^{pro} . To further refine docking results, more rigorous absolute binding free energy calculations are performed in general. There are many approaches to obtain the absolute binding free energies of protein-ligand complexes: attach-pulling of a ligand from its receptor, linear interaction energy (LIE), molecular mechanism/Poisson-Boltzmann surface area and generalized Born surface area (MM-PBSA and MM-GBSA), free energy perturbation (FEP), thermodynamic integration (TI), and nonequilibrium molecular dynamics simulation^{10,11}. All of these computational methods have a trade-off between computational cost and accuracy. Both FEP and TI approaches are considered most rigorous and accurate than the other methods. However, they demand much more computational resources to obtain well-converged binding free energies¹²⁻¹⁵.

The ligand attach-pulling approach pulls ligand from its binding site and calculates a binding free energy based on the Jarzynski or Crooks theorem^{16,16}. This non-equilibrium approach depends on a pulling pathway and it is still not clear how much computational resources are required to obtain a reliable binding free energy (ref). As an alternative approach to these perturbation methods, the MM-PB/GBSA approach, also known as an end-state method, is widely used to estimate absolute binding free energies of protein-ligand complexes due to its relatively lower computational burden. However, the approach suffers from rather large computational errors, lower accuracy, than the perturbation-based methods. Despite its lower accuracy, the MM-PB/GBSA method is becoming more popular to rank the binding free energies of different ligands with a single receptor quickly¹⁷⁻²⁰.

Here, we will briefly review previous computational studies that used free energy calculation methods to discover possible drug candidates that bind to M^{pro} . Ngo et al. used a total 3 methods, molecular docking, fast pulling of ligand (FPL), and FEP, to develop a potential SARS-CoV-2 M^{pro} from the existing drugs and natural compounds²¹. Throughout the three steps, they screened potential inhibitor candidates from the large library. Finally, they obtained the two natural compounds that showed high binding affinity in comparison with 13b ligand, which was from the recently reported experimental work². Moreover, they suggested the E166 residue of M^{pro} is critical in ligand binding through alanine mutation calculation. Sk et al. have evaluated the binding free energies of two molecules (α -ketoamide and Z31792168) with M^{pro} by using the MM-PB/GBSA approach²². They performed per-residue energy decomposition analysis and showed that the ΔE_{vdw} , ΔE_{elec} , and $\Delta E_{nonpolar}$ terms contributed favorably to ligand binding, and the ΔE_{vdw} contributed most. Based on the per-residue decomposition analysis, they identified hot-spot residues, which have higher interaction energy than -1.0 kcal/mol. For both ligands, H163, H164, and M165 were identified as hot spot residues. Nutho et al. selected lopinavir and ritonavir as inhibitors, which were found to be effective against HIV-protease²³. From the MM-PB/GBSA analysis, they found that the ΔE_{vdw} is the most favoured force to binding free energy more twice than ΔE_{elec} with both ligands. Moreover, they also took a residue which showed its decomposed binding

free energy lower than -1.0 kcal/mol into account; there were 4 residues, MET49, M165, P168, and Q189, in M^{Pro}-lopinavir complex and 9 residues, L27, H41, M49, F140, N142, G143, H164, M165, and E166, of the M^{Pro}-ritonavir complex. Wang et al. also used MM-PBSA-WSAS flexible docking for identifying potential inhibitor for SARS-CoV-2 M^{Pro} protease²⁴. They suggested that the five molecules, namely, carfilzomib, ervacycline, valrubicin, lopinavir, and elbasvir, were the potential inhibitor against M^{Pro}. Among them, carfilzomib was identified as the strongest binder.

The main objectives of this study are 1) to identify the hot spots of M^{Pro} of COVID-19 via MM-PB/GBSA calculations using different force fields and 2) to assess the force field dependency of MM-PB/GBSA calculation results. In this study, we performed the MD simulations of 11 M^{Pro}-ligand complexes and calculated the binding affinities using MM-PB/GBSA calculations to identify the hot spot residues of M^{Pro}. Six ligands are from Diamond Light Source (diamond.ac.uk/covid-19/for-scientists/Main-protease-structure-and-XChem/Downloads.html), other recently reported ligands (11r, 13a, and 13b)²⁵, and three HIV protease inhibitors (darunavir, lopinavir, and indinavir), respectively. For all 11 M^{Pro}-ligand complexes, we performed three MD simulations using different force fields to investigate the force field dependency of binding free energy estimates. Three different force fields were used: GROMOS54a7 (g54a7)^{26,27}, AMBER99SB (ff99SB)²⁸, and CHARMM36 force field (c36)²⁹. In total, 100 ns of MD simulations were performed. The MM-PB/GBSA results obtained with three force fields are compared and the hot spot residues are identified via energy decomposition analyses.

Methods

Preparation of complex structures

The three-dimensional crystal structures of SARS-CoV-2 M^{Pro} complexes bound with various known ligands were downloaded from Diamond Light Source (diamond.ac.uk/covid-19). Among the complexes, we selected the complexes that featured the annotation-High Confidence (M^{Pro}-X0072, X0689, X0691, X0749, X1336). In addition to the complexes with known complex crystal structures, we modelled the initial conformations of the experimentally reported inhibitors whose complex structures are not known yet: (11r, 13a) and HIV inhibitors (darunavir, lopinavir, and indinavir). The structures of each ligand are in Supplementary information Figure S1. The initial complex structures were prepared by performing docking simulations using the docking software, GalaxyDock3³⁰.

Molecular docking simulations

When we used GalaxyDock3 to generate the initial M^{Pro}-ligand configurations, we selected the center of mass of the 13b (PDB ID: 6Y2F) as the center of a grid box with the SARS-CoV-2 M^{Pro} structure whose PDB ID: 6YB7 downloaded from diamond.ac.uk/covid-19. The atomic partial charges of each ligand were calculated with the Gasteiger approach^{31,32}. The lowest energy docking pose was selected as the initial configuration for MD simulations of 11r, 13a, darunavir, lopinavir, and indinavir.

Molecular dynamics simulations

We used two MD programs (GROMACS-2019.4³³ and AMBER18³⁴) and three force fields: g54a7^{26,27}, and ff99SB²⁸, c36²⁹.

MD simulations with GROMACS

The g54a7 was used to perform MD simulations with GROMACS-2019.4³³. The ligand topology files were obtained from the ATB (Automated Topology Builder) server³⁵, which combines a knowledge-based approach with quantum mechanical calculations to generate parameters compatible with the desired force field. We used the original geometry output of ATB. The systems were solvated with the SPC water model³⁶ with a 12 Å padding region. We added Na⁺ ions to neutralize the system. The time step of MD simulation was set to 2 fs with the LINCS algorithm to constraint hydrogen-connected covalent bonds³⁷. The long-range electrostatic interaction was calculated by the fast smooth particle-mesh Ewald electrostatics method^{38,39}. The solvated system was minimized by using the steepest descent method for 10,000 steps. After minimization, a series of equilibrium MD simulations were performed: the 200 ps of NVT, 1 ns of NPT, and 50 ns of NVT with the weak harmonic position restraints. At each step, the simulations were performed with the Noose-Hoover thermostat⁴⁰ and Parrinello-Rahman barostat⁴¹ at 300 K. Lastly, the production step was performed with NPT condition during 100 ns.

MD simulations with AMBER

We performed MD simulations with the two force fields, ff99SB²⁸ and c36²⁹ using AMBER18 MD software³⁴.

- Preparing systems with an AMBER99SB force field (ff99SB)

The M^{Pro} and ligand topology files were generated using AMBER Leap⁴² and Antechamber⁴³ programs implemented in AMBER18 with ff99SB and generalized AMBER force field (GAFF)⁴⁴. The systems were solvated with the TIP3P water model with a padding region of 12 Å. Energy minimization calculations of the solvated systems were performed by using the steepest descent method. After energy minimization, equilibrium MD simulations were performed as follows: 200 ps of NVT, 1 ns of NPT, and 50 ns of NVT with weak harmonic positional restraints. The time step for MD simulations was set to 2 fs. To constrain all bonds involving hydrogen atoms, the SHAKE algorithm was used⁴⁵. The particle mesh Ewald (PME) method also calculated the long-range electrostatic interactions⁴⁶. The temperature was controlled with Langevin thermostat⁴⁷. After equilibrium simulations are finished, we performed production simulations for 100 ns with NPT condition and the trajectory of each simulation was recorded every 2 ps.

- Preparing systems with CHARMM36 force field (c36)

The parameter and topology files of complexes were generated from CHARMM-GUI⁴⁸ with c36²⁹. The parameters for ligands were generated by CHARMM generalized force field (CGenFF)^{49,50}. The subsequent NVT 125 ps, NPT 1 ns, NVT 50 ns equilibration simulations were carried out at 300K, followed by production runs for 100 ns under NPT ensemble condition. The TIP3P water model was used to represent explicit water molecules. We added the sodium ions in order to neutralize the system. The time step for MD simulation was set to 2 fs with the SHAKE scheme. After the energy of the solvated system was minimized by using the steepest descent method, the subsequent steps were carried

out using the identical parameters with the AMBER setting described above.

MM-PBSA and MM-GBSA Analysis

In this study, the interactions between the M^{pro} receptor and each ligand were calculated with the MM-PB/GBSA analysis. We utilized the two MM-PB/GBSA tools; one is `g_mmpbsa`⁵¹ for the GROMACS MD results and the other is `MMPBSA.py`⁵² from the AMBER18 package respectively. Note that the `g_mmpbsa` supports only MM-PBSA calculation but not MM-GBSA. Thus, we performed MM-PBSA analysis for the trajectories obtained with GROMACS.

- `g_mmpbsa` setting

Each topology and parameter files required for MM-PB/GBSA calculations are obtained from the preparation section. For a trajectory of 100 ns, 10,000 frames were selected for MM-PBSA calculations as input. For non-polar solvation energy calculations, we selected the SASA model^{53–55} which is one of the most used non-polar models. For the other parameters, the default values of GROMACS were used.

- `MMPBSA.py` setting

Each parameter and system topology files for a solvated complex, non-solvated complex, receptor, and ligand were used from the preparation steps above. Additionally, we performed the solvated M^{pro} receptor simulation and solvated ligand simulation to estimate the receptor and ligand contribution in implementing MM-PB/GBSA analysis. Similar to the `g_mmpbsa` setting, 10,000 frames also were used for MM-PB/GBSA analyses with `MMPBSA.py` from trajectories of 100 ns. For PB calculations, the SASA (solvent accessible surface area) for the non-polar solvation energy used the LCPO model⁵⁶ and the ionic strength of solution was set to 0.1 M and the `mbondi`⁵⁷ radius set was used. For GB calculations, we used a modified GB model developed by A. Onufriev, and co-works^{58,59} represented by `igb=5` and the ionic strength was set to 0.1 M (`saltcon=0.100`). The rest of the options were set to the default values of `MMPBSA.py` in Amber18 package.

Results and Discussion

The structure of M^{pro} consists of three domains: domain I (residues 10-99), domain II (residues 100-184), domain III (residues 201-303), and a long loop region (residues 185-200) (Fig. 1-(a)). Domain III is connected to domain II via the long loop structure. The substrate-binding site is located in a cleft between domain I, II with the loop region involved. It is composed of a series of subsites (S1, S1', S2 and S4) (Fig. 1-(b) and -(c)). Each subsite is spatially located and composed of consecutive amino acids in the substrate-binding site. The S1 subsite consists of F140, N142, C145, H163, E166, and H172 residues. (Fig. 1-(b) red region). The S1' subsite is composed of T26 and L27 (Fig. 1-(b) lime region). The side chains of H41, M49, Y54, M164, E165, D187 are involved in the formation of the S2 subsite (Fig. 1-(b) light blue). The E165, L167, F185, Q189, Q192 consists of S4 subsite (the orange region in Fig. 1-(b) yellow).

Figure 1. The structure of a single M^{pro} receptor. (a) The three domains are colored and the pink colored region represents a long loop region. Domain I (green), II (sea green), III (cornflower blue), and the pinked long loop which is located between domains II and III. The binding site is located within the dashed square box, composed of Domain I, Domain II, and the loop region. (b) The surface was represented by the catalytic binding center. (c) The surface removed view of the binding site. The residues involved in each subsite are represented.

MD results of complexes

Most MD simulations of the M^{pro} complexes showed stable bound-state conformations, except the M^{pro} -X0072 and -X0691 complexes. In the simulations of the M^{pro} -X0072 complex with both ff99SB and c36, the X0072 ligand was detached from M^{pro} . The X0691 ligand with c36 also was detached from M^{pro} during simulations. The trajectory showed that both X0072 and X0691 ligands occupied a region between E166 and Q189 in S4 subsite for a while and finally were detached from M^{pro} . These results suggest that both residues may play the important roles of gateways of ligand-binding.

The spatial occupancy patterns of ligands sampled during simulations are visually investigated in order to identify which residues participate in ligand binding of M^{pro} (Fig. 3). To perform this investigation, the *grid* command of the CPPTRAJ⁶⁰ program of the AMBER18 was used to bin the distribution of a ligand into a 3D grid. Before performing the analyses, periodic boundary conditions were removed and centered the system to put a ligand at the origin of each trajectory.

The MD result of darunavir showed a significantly localized distribution than those of lopinavir and indinavir, which are the inhibitors of HIV protease. The sulfonate group of darunavir stayed at the S4 subsite and the benzene ring also was located in

Figure 2. 11r, 13a, and 13b ligands were at the upper row, The branches of 11r, 13a, and 13b ligands were represented in colored-circle. Darunavir, lopinavir, and indinavir are in the lower row.

Figure 3. The examples of a 3D histogram, comparing results from ff99SB and c36. The colored region is a 3D histogram on the ligand with ff99SB and the representative conformations from both ff99SB and c36 are imposed. (a) 11r (b) 13a (c) 13b (d) darunavir ligand.

S1 subsite and did not fluctuate much. However, in the case of lopinavir, the functional groups of lopinavir fluctuated much, except that the dimethyl substituted benzene group stayed at the S4 subsite. The branches of indinavir showed that the rings did not stay stably, and did not occupy the S4 subsite similar to other HIV protease inhibitors

The other ligands, 11r and 13a showed different binding behaviors from 13b. The ligands, 11r and 13a showed more

concentrated spatial distribution consistently than the 13b ligand in simulations performed with all force fields. On the contrary, the trajectory of 13b showed more widely distributed 3D histograms because of the fluctuation of its P1, P3 branches (Fig. 2 and 3). The P1 branch of both 11r and 13a remained stably at the S1 subsite during simulations. The P3 branch of 11r and the P1' part of 13a with ff99SB, and the cyclohexyl ring of 13a with c36 also remained steadily at the S4 subsite, where most strong binders are occupied significantly during simulations. However, the P1 branch of 13b was located at a different region compared with 11r and 13a and fluctuated more than them, and 13b did not occupy the S4 subsite between E166 and Q189 during the simulations.

We assumed that this trend is consistent with the stronger binding free energy of 13b than 11r and 13a, which will be discussed in the next section. The fact that the S4 subsite was occupied by many strong binders was employed in a previous study that developed a new drug candidate from a weak hit molecule named perampanel⁶¹. The authors performed a substrate scope analysis, which considers the interaction and environments between M^{Pro} and hit molecules, and rationally changed the branches of perampanel to fill the void of the S4 subsite. Most modified ligands interacted strongly with the E166, Q189, and Q192 residues. Especially, most modified ligands whose the chloride atom was substituted for various alkoxy groups such as propoxy, butoxy group or more bulky benzyloxy group to occupy the S4 subsite region, became more than 10 times stronger than the hit molecule. This tendency is consistent with the MM-PBSA results, which will be discussed in the next section and may be originated from hydrogen bonds between ligands and M^{Pro}. The complete results of all the 3D histogram results are illustrated in Supplementary information (Fig. S15-S17)

Comparison of MD simulation results of complexes using different force fields

In this section, we compared the MM-PBSA and MM-GBSA calculation results with three different force fields to investigate the effect of force field parameters on binding free energy estimates (Figure 4 and Fig. S2-S11).

A comparison of AMBER and CHARMM force field results

The MM-PB/GBSA results obtained with ffSB99 and c36 showed weak correlations between them, $\rho = 0.004$ and $= 0.098$, respectively (Fig. 4-(a)). Overall, the MM-PBSA binding free energy estimates obtained with ffSB99 were more favorable than those of c36 simulations. Most binding energy results obtained with MM-PBSA calculations using ffSB99 had lower bind energy by more than about 10 kcal/mol compared to the c36 results. The strongest binders of each force field are 11r (-54.990 kcal/mol) with ffSB99 and 13a (-28.523 kcal/mol) with c36. 13a also showed that it was the top3 strong binder with ffSB99 (-37.971 kcal/mol). However, 11r was not identified as a strong binder with c36. From the experimental data^{25,62}, the binding free energies of 11r, 13a, 13b, darunavir ligands were -9.23, -7.70, -8.45, and -6.14 kcal/mol, respectively. In terms of consistency with experimental data, MM-PBSA with ffSB99 correctly identified 11r as the strongest binder. However, c36 results predicted that 13a would be a stronger binder than 11r, which is inconsistent with the experiment. Both ffSB99 (-29.535 kcal/mol) and c36 (5.579 kcal/mol) estimated that the 13b ligand is the weakest binder among the 11r, 13a, and 13b ligands despite 13b is the second strong binder based on the experiment.

We assumed the fact that 13b did not occupy the S4 subsite unlike 11r and 13a is one of the reasons why the MM-PBSA

Figure 4. A comparison of the ΔE_{MM} obtained with MM-PB/GBSA calculations using ffSB99, c36 and g54a7. The solid diagonal line corresponds to the perfect agreement between the two results. The dotted lines correspond to a range of an absolute error of 10 kcal/mol. The error bars are presented with points. Because the X0072 and X0691 ligand were detached from M^{Pro} , they were omitted.

results were not consistent with the experiment among 11r, 13a, and 13b ligands. Moreover, although darunavir is the weakest binder from the experimental data, the two force field failed to estimate the fact that darunavir is the weakest binder.

The MM-GBSA results obtained with the ffSB99 and c36 also showed weak correlations, $\rho = 0.098$. Based on the MM-GBSA calculations, except the two cases, the 11r and indinavir ligands, the rest of the data points are located upper side of the diagonal line, meaning that the estimated binding energies with MM-GBSA using ff99SB were more favorable than those of MM-GBSA with c36 (Fig. 4-(b)). This trend is similar to the MM-PBSA results. Similar to MM-PBSA results, the MM-GBSA calculations had the same binding energy trend among 11r, 13a, 13b, and darunavir ligands. The strongest binders of each force field with MM-GBSA are 11r (-50.107 kcal/mol) and 13a (-26.602 kcal/mol), respectively. Moreover, MM-GBSA with both ff99SB and c36 estimated the 13b ligand had the lowest binding energy compared with 11r and 13a. Although MM-PBSA and MM-GBSA analysis did not estimate the ranking order of binding energies with 11r, 13a, 13b, and darunavir ligands, the two approaches had a similar trend.

A comparison of CHARMM and GROMOS force field results

Next, we compared the MM-PBSA results obtained with c36 and g54a7 (Fig. 4-(c)). It showed a moderate correlation with a correlation coefficient, $\rho = 0.335$. Overall, the binding energies obtained with g54a7 showed more favorable values than those of c36. MM-PBSA results with g54a7 estimated that darunavir would be the strongest binder with a binding energy of -46.640 kcal/mol. The top contributing residues for the binding of darunavir with g54a7 were M49, E166, M165, and Q189 with decomposed energies (-3.625, -2.740, -1.972, and -1.491 kcal/mol, respectively). Because the M165, E166, and Q189 residues are in the S4 subsite, g54a7 also suggested that the S4 subsite was important in M^{Pro}-ligand binding. Although c36 showed less favorable energies than g54a7, it is valuable to compare the binding energy of darunavir, which is the strongest binding with g54a7, between g54a7 and c36. With c36 both E166 and Q189 had the unfavorable decomposed energy (4.748 and 1.314 kcal/mol, respectively). These energies were opposite that E166 and Q189 had decomposed energy with g54a7 (-2.740 and -1.491, respectively). We found that these energies were from both residues were calculated with highly unfavorable E_{elec} (19.815, 4.431 kcal/mol, respectively), although H-N of E166 has frequently formed hydrogen bonds with darunavir.

A comparison of GROMOS and AMBER force field results

The MM-PBSA binding energies obtained with ff99SB were relatively well-correlated from those of g54a7 with a correlation coefficient, $\rho = 0.618$ (Fig. 4-(d)). g54a7 also estimated that 13b is the weakest binder among 11r, 13a, 13b, and darunavir. These results were consistent with all three force fields. The MM-PBSA results with ff99SB had shown a similar trend with the MM-PBSA results with g54a7 except for 11r and lopinavir cases. If the three outliers are excluded, a correlation coefficient was $\rho = 0.933$ between g54a7 and ff99SB. This result suggested that g54a7 was able to present similar binding free energies with the all-atom model from ff99SB, although g54a7 used united-atom model parameters. The strongest binder without the outlier ligand was darunavir: -50.288 kcal/mol with ff99SB and -46.640 kcal/mol with g54a7. Darunavir was also the second strongest binder with outlier ligands with ff99SB.

Summary of MM-PB/GBSA results

In summary, the MM-PBSA estimates obtained with c36 is showing a different trend with those obtained with ff99SB and g54a7 (Table 1). The absolute values of binding energies obtained with c36 are smaller than those with ff99SB and g54a7. Unfortunately, all three force fields failed to predict the rankings of known experimental binding affinity data. All three force fields estimated that 13b ligand was the weakest binder among 11r, 13a, and 13b, while 13b was the second strong binder between them. Only ff99SB correctly identified the 11r ligand as the strongest binder. However, more experimental data are necessary to evaluate the accuracy of force fields more rigorously in obtaining correct absolute binding free energies of M^{Pro} complexes. These results suggest that, based on the MM-PBSA calculations, no force field is accurate enough to obtain the absolute binding affinities of protein-ligand complexes yet and the calculation results may show large variations according to the tested force field. Thus, our results also show that a care should be taken to interpret the absolute binding free energy estimates obtained with MM-PBSA calculations. The complete results of all energies are Supplementary information Table S1

and S2.

Ligand	MM-PBSA ^a	MM-GBSA ^a	MM-PBSA ^b	MM-GBSA ^b	MM-PBSA ^c	NMA ΔS	QH ΔS
X0072	-34.857	-14.958	-17.302	-12.022	-13.659	-16.455	-12.441
X0689	-23.223	-18.827	3.943	-20.393	-22.241	-16.899	-12.723
X0691	-9.075	-20.779	12.182	-14.610	-20.549	-18.119	-12.761
X0749	-18.167	-35.717	-13.242	-16.781	-22.040	-17.327	-12.677
X1336	-23.727	-24.571	-8.932	-12.958	-22.734	-19.280	-12.791
11r	-54.990	-50.107	-10.177	-24.645	-37.905	-31.088	-13.369
13a	-37.971	-34.261	-28.523	-26.602	-43.613	-27.803	-13.389
13b	-29.535	-25.480	5.579	-20.019	-31.927	-25.499	-13.404
darunavir	-50.288	-47.902	-1.556	-27.556	-46.640	-29.247	-13.330
indinavir	-29.238	-37.580	-21.343	-19.073	-36.142	-30.431	-13.430
lopinavir	-17.064	-41.404	-13.166	-24.922	-40.744	-27.693	-13.451

Table 1. Overall MM-PBSA difference energy results. All values have a kcal/mol energy unit. The ^a means that the energies are from the AMBER18 MD engine with ff99SB, while the ^b is from the AMBER18 MD engine with c36. Also, the ^c is that the energy values are from GROMACS-2019.4 MD engine with g54a7. NMA ΔS is that the entropy is calculated with normal mode analysis and QH ΔS means that the entropy is estimated by quasi-harmonic approximation with only c36 system because MMPBSA.py package only supports Normal mode analysis with ff99SB, not c36.

Hot spot residues for ligand binding

To identify hot spot residues, which residues play important roles in ligand binding, we performed a per-residue energy decomposition analysis and calculated the average binding energy contribution of all ligands. The top 10 residues that contribute most to ligand-binding of M^{Pro} based on the average per-residue interaction energy is illustrated in Fig. 5. All the top 10 contributing residues have ΔE^{MM} lower than -1.0 kcal/mol on average.

Figure 5. Averaged contribution energy from MM-PBSA and corresponding hot spot residues (left). Hot spot residues are highlighted on the surface of M^{Pro} (right). Because the X0072 and X0691 ligands were detached from M^{Pro}, they were omitted when generating the figure. Because the terminal residues are in the backside of the figure, their orange-colored regions were hidden.

The M165 residue is identified as the top contributing residue consistently across all tested force fields. M165 showed $\Delta E_{\text{avg}}^{\text{contrib}} = -5.0$ kcal/mol which is 2~3 fold stronger than other residues' ΔE^{MM} . This result is consistent with the fact that most strongly binding ligands stayed stably around the S4 subsite. The decomposition analysis showed that M165 interacts

favorably with ligands via both electrostatic and van der Waals interactions. With the strongest binder 11r with ff99SB and 13a with c36, M165 was the most favorable residue. For the 11r case, E_{vdw} , E_{elec} , and $E_{\text{polarsolv}}$ were -2.508, -1.977, and 1.462 kcal/mol with ff99SB and -1.564, -1.614, and 1.714 kcal/mol with c36, respectively. For the 13a case, E_{vdw} , E_{elec} , and $E_{\text{polarsolv}}$ were -1.304, 0.842, and 1.003 kcal/mol with ff99SB and -1.721, -0.910, and 1.343 kcal/mol with c36, respectively.

Based on the fact that 11r with ff99SB and 13a with c36 are the strongest binders, we performed a residue-wise binding free energy decomposition analysis to identify the hot-spot residues of M^{Pro} . The top contribution residues of 11r with ff99SB were D295, D187, D176, and H163 with binding energies of -5.538, -4.983, -4.350 and -3.624 kcal/mol, respectively. Except for D176 (4.185 kcal/mol), their E_{elec} were lower than -5.0 kcal/mol. The top contribution residues of 13a with c36 were D295, Y161, R298, L58, L50 with binding energies of -3.457, -3.308, -3.195, and -3.025 kcal/mol, respectively. They also had low E_{elec} except for Y161 (1.435 kcal/mol).

Interestingly, the terminal charged residues, D295, R298, and Q299, which are far from the binding site interacted strongly with ligands. Although they were far from the binding site over about 30 Å during MD simulation, the energy decomposition analysis showed that these residues have strong electrostatic or polar solvation attractive interaction for ligand's binding. Recently T. Sztain et al. reported that the C-terminal region was examined as a potential allosteric site⁶³, which is consistent with our per-residue decomposition analysis result.

It is observed that the residues consisting of the S1' subsite contributed less to ligand binding than the other subsites, which is consistent with the 3D histogram analysis. Except for the residues that are far from the known binding site, the N142, M165, D176 and, R188 residue are in S1 or S4 subsites: S1: N142, D176 and S4: M165, R188. This result suggests that the residues of the S4 and S1 subsites may interact stronger than the residues of the S1' and S2 subsites with ligands.

Hydrogen Bonds Analysis

We carried out a hydrogen bond analysis and investigated its fraction to find which residues are forming hydrogen bonds strongly with the ligands. The complete hydrogen bond analysis results are in Supplementary information Fig. S12-S14. It is identified that if a specific hydrogen bond lasts long during a simulation, it contributed strongly to binding free energy. For the cases of 11r and 13a, which were the strongest binders with ff99SB and c36, respectively, the hydrogen bonds that existed for more than 60% of the entire trajectory were identified to be the most strongly contributing residues H142, S144, 163, and E166. Especially, the 11r ligand with ff99SB formed two stable hydrogen bonds, which existed more than 70% of the entire trajectory (Fig. 6).

For the simulation of 11r with ff99SB, HE2-NE2@H163 and H-N@E166 atoms formed hydrogen bonds with 11r for 79% and 73% of the trajectory, respectively. Moreover, for the simulation of darunavir with ff99SB, O@H164 and H-N@E166 atoms also formed the hydrogen bonds with darunavir with fractions of 0.84 and 0.66, respectively. For the case of 13a with c36, which is the strongest binder with c36, the H-N@E166 also existed for a fraction of 0.61 while other hydrogen bonds existed lower than a fraction of 0.1. Except for the ligands discussed above (11r, 13a and darunavir), the other ligands showed that the M^{Pro} -ligand complexes formed hydrogen bonds under a fraction of 0.3 with both ff99SB and c36 (Table 2). For example,

Figure 6. The hydrogen bond analysis of 11r, which was strongest binder with ff99SB at left. The residues that were frequently captured in hydrogen bond analysis at right

X0749, which is the weakest binder with ff99SB, showed that all hydrogen bonds lasted lower than a fraction of 0.1.

Threshold	0.1		0.2		0.3		0.5		0.7	
Ligand	ff99SB	c36	ff99SB	c36	ff99SB	c36	ff99SB	c36	ff99SB	c36
X0072	0	0	0	0	0	0	0	0	0	0
X0689	2	2	1	2	1	1	0	1	0	1
X0691	1	0	0	0	0	0	0	0	0	0
X0749	0	0	0	0	0	0	0	0	0	0
X1336	3	1	1	0	0	0	0	0	0	0
11r	7	2	4	1	4	1	2	0	2	0
13a	6	1	6	1	2	1	0	1	0	0
13b	2	3	0	1	0	1	0	0	0	0
darunavir	4	4	3	3	3	2	3	1	1	0
lopinavir	5	2	1	1	1	0	0	0	0	0
indinavir	3	1	2	1	1	0	1	0	0	0

Table 2. The number of hydrogen bonds of each ligand which satisfies the thresholds with both ff99SB and c36. The total 10,000 number of frames was used to perform hydrogen bond analysis.

The residues, consisting the subsites formed diverse hydrogen bonds during MD simulation with all of ligands (Table 3). N142, E166, and Q189 formed highly stable hydrogen bonds with the ligands. The atoms that were frequently involved in

Residue Force field	Backbone		Side chain	
	ff99SB	c36	ff99SB	c36
T26	H-N(5), O(2)	H-N(7)	HG1-OG1(2), OG1(1)	HG1-OG1(3)
H41	O(1)	O(1)	HE2-NE2(7), ND1(2), NE2(1)	HD1-ND1(5), NE2(2)
S46	O(3), H-N(5)	O(1), H-N(5)	HG-OG(10), OG(7)	HG-OG(7), OG(4)
N142	H-N(1), O(1)	H-N(1).	HD21-ND2(10), HD22-ND2(8), OD1(12), ND2(3)	HD22-ND2(8), HD21-ND2(7), OD1(6), ND2(3)
E166	H-N(10), O(4)	H-N(7)	OE1(6), OE2(7)	OE1(6), OE2(4)
Q189	O(3), H-N(2)	O(3), H-N(1)	HE22-NE2(10), HE21-NE2(10), OE1(1), NE2(1)	HE21-NE2(11), HE22-NE2(11), OE1(15), NE2(1)

Table 3. The atoms on residues that frequently formed hydrogen bonds with ff99SB and c36 during simulations. The atoms were based on the hydrogen bonds in the top5 existence fraction of each ligand MD simulation by hydrogen bond analysis. The number in parentheses means that how many atoms formed hydrogen bonds in all ligand cases.

forming hydrogen bonds were similar both with ff99SB and c36. Based on these results, the stable hydrogen bonds appeared to be independent of the force field. Most stable hydrogen bonds were formed by the atoms in side chain. Except for the H-N

atom in E166, all the frequency of each backbone atom involved in hydrogen bonds were less than 10 times.

Conclusion

In this study, we utilized the MM-PBSA and MM-GBSA approaches to investigate the hot spot residues of SARS-CoV-2 M^{pro} with three different force fields: ff99SB, c36, and g54a7. We performed MD simulations of 11 M^{pro}-ligand complexes with three different force fields, and performed both MM-PBSA and MM-GBSA analysis, while other studies only performed MM-GBSA^{8,24} or investigated only two ligands²³. The estimated binding energies from MM-PBSA were not consistent between three different force fields. By comparing the MM-PBSA results from the tested force fields, we found that there is a correlation between the ff99SB and g54a7 results with the interested ligands. However, the results obtained with c36 showed weakly or almost no correlation with the g54a7 and ff99SB results. It is also identified that the binding energies from c36 were consistently higher than those of both ff99SB and g54a7. It is also identified that no single force field is able to predict the correct relative ranking of experimental binding affinities. All MM-PBSA calculations with three force field estimated that 13b was the weakest binder among 11r, 13a, and 13b ligand. However, based on the experiment, 13b is the second strong binder among them. This result was from that 13b did not occupy the S4 subsite during simulation.

Based on the residue-wise decomposed energies, the terminal residues (D295, R298, and Q299) were identified to be hot spot residues. Because they were far from the known binding site, they were not considered as a residue that contributes the binding energy at first glance. However, the terminal residues had averaged decomposed energy in our ligands: -4.065, -2.582, and -1.383 kcal/mol, respectively, which has been suggested by the other study that the terminal residues were important in ligand binding.⁶³ The decomposed residue energies showed that the residues of the S1 or S4 subsites contributed more strongly to binding energy than the residues in S1' or S2 subsites. The branches of ligand in S1' subsite fluctuated more freely than those residues in the S1 or S4 subsites. We believe that the obtained results could suggest information on the further design and development of more potency inhibitor against SARS-CoV-2 M^{pro}.

References

1. The Johns Hopkins Coronavirus Resource Center. <https://coronavirus.jhu.edu/>.
2. Ioannidis, J. P. Global perspective of covid-19 epidemiology for a full-cycle pandemic. *Eur. J. Clin. Investig.* **50**, e13423 (2020).
3. Altay, O. *et al.* Current Status of COVID-19 Therapies and Drug Repositioning Applications. *iScience* **23**, 101303, DOI: <https://doi.org/10.1016/j.isci.2020.101303> (2020).
4. Jin, Z. *et al.* Structure of m pro from sars-cov-2 and discovery of its inhibitors. *Nature* **582**, 289–293 (2020).
5. Douangamath, A. *et al.* Crystallographic and electrophilic fragment screening of the sars-cov-2 main protease. *Nat. Commun.* **11**, 1–11 (2020).

6. Jin, Z. *et al.* Structural basis for the inhibition of sars-cov-2 main protease by antineoplastic drug carmofur. *Nat. Struct. & Mol. Biol.* **27**, 529–532 (2020).
7. Bhardwaj, V. K., Singh, R., Das, P. & Purohit, R. Evaluation of acridinedione analogs as potential sars-cov-2 main protease inhibitors and their comparison with repurposed anti-viral drugs. *Comput. Biol. Medicine* **128**, 104117 (2021).
8. Alamri, M. A. *et al.* Pharmacoinformatics and molecular dynamics simulation studies reveal potential covalent and fda-approved inhibitors of sars-cov-2 main protease 3clpro. *J. Biomol. Struct. Dyn.* 1–13 (2020).
9. Peele, K. A. *et al.* Molecular docking and dynamic simulations for antiviral compounds against sars-cov-2: A computational study. *Informatics Medicine Unlocked* **19**, 100345 (2020).
10. Jarzynski, C. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Phys. Rev. E* **56**, 5018 (1997).
11. Macchiagodena, M. *et al.* Virtual double-system single-box: A nonequilibrium alchemical technique for absolute binding free energy calculations: Application to ligands of the sars-cov-2 main protease. *J. Chem. Theory Comput.* **16**, 7160–7172 (2020).
12. Zwanzig, R. W. High-temperature equation of state by a perturbation method. i. nonpolar gases. *The J. Chem. Phys.* **22**, 1420–1426 (1954).
13. Beveridge, D. L. & DiCapua, F. Free energy via molecular simulation: applications to chemical and biophmolecular systems. *Annu. Rev. Biophys. Biophys. Chem.* **18**, 431–492 (1989).
14. Kollman, P. Free energy calculations: applications to chemical and biochemical phenomena. *Chem. Rev.* **93**, 2395–2417 (1993).
15. Heinzlmann, G., Chen, P.-C. & Kuyucak, S. Computation of standard binding free energies of polar and charged ligands to the glutamate receptor glua2. *The J. Phys. Chem. B* **118**, 1813–1824 (2014).
16. Collin, D. *et al.* Verification of the crooks fluctuation theorem and recovery of rna folding free energies. *Nature* **437**, 231–234 (2005).
17. Srivastava, H. K. & Sastry, G. N. Molecular dynamics investigation on a series of hiv protease inhibitors: assessing the performance of mm-pbsa and mm-gbsa approaches. *J. Chem. Inf. Model.* **52**, 3088–3098 (2012).
18. Williams, A. H. & Zhan, C.-G. Fast prediction of binding affinities of the sars-cov-2 spike protein mutant n501y (uk variant) with ace2 and miniprotein drug candidates. *The J. Phys. Chem. B* (2021).
19. King, E., Qi, R., Li, H., Luo, R. & Aitchison, E. Estimating the roles of protonation and electronic polarization in absolute binding affinity simulations. *J. Chem. Theory Comput.* **17**, 2541–2555 (2021).
20. Srivastava, H. K. & Sastry, G. N. Molecular dynamics investigation on a series of hiv protease inhibitors: assessing the performance of mm-pbsa and mm-gbsa approaches. *J. Chem. Inf. Model.* **52**, 3088–3098 (2012).

21. Ngo, S. T., Quynh Anh Pham, N., Thi Le, L., Pham, D.-H. & Vu, V. V. Computational determination of potential inhibitors of sars-cov-2 main protease. *J. Chem. Inf. Model.* **60**, 5771–5780, DOI: [10.1021/acs.jcim.0c00491](https://doi.org/10.1021/acs.jcim.0c00491) (2020).
22. Sk, M. F., Roy, R., Jonniya, N. A., Poddar, S. & Kar, P. Elucidating biophysical basis of binding of inhibitors to sars-cov-2 main protease by using molecular dynamics simulations and free energy calculations. *J. Biomol. Struct. Dyn.* 1–13 (2020).
23. Nutho, B. *et al.* Why are lopinavir and ritonavir effective against the newly emerged coronavirus 2019? atomistic insights into the inhibitory mechanisms. *Biochemistry* **59**, 1769–1779 (2020).
24. Wang, J. Fast identification of possible drug treatment of coronavirus disease-19 (covid-19) through computational drug repurposing study. *J. Chem. Inf. Model.* **60**, 3277–3286 (2020).
25. Zhang, L. *et al.* Crystal structure of sars-cov-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science* **368**, 409–412 (2020).
26. Poger, D., Van Gunsteren, W. F. & Mark, A. E. A new force field for simulating phosphatidylcholine bilayers. *J. Comput. Chem.* **31**, 1117–1125, DOI: <https://doi.org/10.1002/jcc.21396> (2010).
27. Schmid, N. *et al.* Definition and testing of the gromos force-field versions 54a7 and 54b7. *Eur. Biophys. J.* **40**, 843–856 (2011).
28. Hornak, V. *et al.* Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Struct. Funct. Bioinforma.* **65**, 712–725 (2006).
29. Huang, J. & MacKerell Jr, A. D. Charmm36 all-atom additive protein force field: Validation based on comparison to nmr data. *J. Comput. Chem.* **34**, 2135–2145 (2013).
30. Yang, J., Baek, M. & Seok, C. Galaxydock3: Protein–ligand docking that considers the full ligand conformational flexibility. *J. Comput. Chem.* **40**, 2739–2748 (2019).
31. Gasteiger, J. & Marsili, M. A new model for calculating atomic charges in molecules. *Tetrahedron Lett.* **19**, 3181–3184 (1978).
32. Gasteiger, J. & Marsili, M. Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron* **36**, 3219–3228 (1980).
33. Kutzner, C. *et al.* More bang for your buck: Improved use of gpu nodes for gromacs 2018. *J. Comput. Chem.* **40**, 2418–2431 (2019).
34. Lee, T. S. *et al.* GPU-Accelerated Molecular Dynamics and Free Energy Methods in Amber18: Performance Enhancements and New Features. *J. Chem. Inf. Model.* **58**, 2043–2050, DOI: [10.1021/acs.jcim.8b00462](https://doi.org/10.1021/acs.jcim.8b00462) (2018).
35. Malde, A. K. *et al.* An automated force field topology builder (atb) and repository: version 1.0. *J. Chem. Theory Comput.* **7**, 4026–4037 (2011).

36. Berendsen, H. J., Postma, J. P., van Gunsteren, W. F. & Hermans, J. Interaction models for water in relation to protein hydration. In *Intermolecular forces*, 331–342 (Springer, 1981).
37. Hess, B., Bekker, H., Berendsen, H. J. & Fraaije, J. G. Lincs: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
38. Darden, T., York, D. & Pedersen, L. Particle mesh ewald: An $n \log(n)$ method for ewald sums in large systems. *The J. Chem. Phys.* **98**, 10089–10092 (1993).
39. Essmann, U. *et al.* A smooth particle mesh ewald method. *The J. Chem. Phys.* **103**, 8577–8593 (1995).
40. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. review A* **31**, 1695 (1985).
41. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
42. Salomon-Ferrer, R., Case, D. A. & Walker, R. C. An overview of the amber biomolecular simulation package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **3**, 198–210 (2013).
43. Wang, J., Wang, W., Kollman, P. A. & Case, D. A. Antechamber: an accessory software package for molecular mechanical calculations. *J. Am. Chem. Soc.* **222**, U403 (2001).
44. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
45. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Physics* **23**, 327–341 (1977).
46. 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* **7**, R55–R60 (1999).
47. Uberuaga, B. P., Anghel, M. & Voter, A. F. Synchronization of trajectories in canonical molecular-dynamics simulations: Observation, explanation, and exploitation. *The J. Chem. Phys.* **120**, 6363–6374 (2004).
48. Jo, S., Kim, T., Iyer, V. G. & Im, W. Charmm-gui: a web-based graphical user interface for charmm. *J. Comput. Chem.* **29**, 1859–1865 (2008).
49. Vanommeslaeghe, K. & MacKerell Jr, A. D. Automation of the charmm general force field (cgenff) i: bond perception and atom typing. *J. Chem. Inf. Model.* **52**, 3144–3154 (2012).
50. Vanommeslaeghe, K., Raman, E. P. & MacKerell Jr, A. D. Automation of the charmm general force field (cgenff) ii: assignment of bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* **52**, 3155–3168 (2012).
51. Kumari, R., Kumar, R., Consortium, O. S. D. D. & Lynn, A. g_mmpbsa-a gromacs tool for high-throughput mm-pbsa calculations. *J. Chem. Inf. Model.* **54**, 1951–1962 (2014).

52. Miller III, B. R. *et al.* Mmpbsa. py: an efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **8**, 3314–3321 (2012).
53. Sitkoff, D., Sharp, K. A. & Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *The J. Phys. Chem.* **98**, 1978–1988 (1994).
54. Still, W. C., Tempczyk, A., Hawley, R. C. & Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **112**, 6127–6129 (1990).
55. Kuhn, B. & Kollman, P. A. Binding of a diverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. *J. Medicinal Chem.* **43**, 3786–3791 (2000).
56. Weiser, J., Shenkin, P. S. & Still, W. C. Approximate atomic surfaces from linear combinations of pairwise overlaps (lcpo). *J. Comput. Chem.* **20**, 217–230 (1999).
57. Tsui, V. & Case, D. A. Theory and applications of the generalized born solvation model in macromolecular simulations. *Biopolym. Orig. Res. on Biomol.* **56**, 275–291 (2000).
58. Onufriev, A., Bashford, D. & Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Struct. Funct. Bioinforma.* **55**, 383–394 (2004).
59. Onufriev, A., Bashford, D. & Case, D. A. Modification of the generalized born model suitable for macromolecules. *The J. Phys. Chem. B* **104**, 3712–3720 (2000).
60. Roe, D. R. & Cheatham III, T. E. Ptraj and cpptraj: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).
61. Zhang, C.-H. *et al.* Potent noncovalent inhibitors of the main protease of sars-cov-2 from molecular sculpting of the drug perampanel guided by free energy perturbation calculations. *ACS Cent. Sci.* **7**, 467–475 (2021).
62. Mahdi, M. *et al.* Analysis of the efficacy of hiv protease inhibitors against sars-cov-2' s main protease. *Viol. J.* **17**, 1–8 (2020).
63. Sztain, T., Amaro, R. & McCammon, J. A. Elucidation of cryptic and allosteric pockets within the sars-cov-2 protease. *Biorxiv* (2020).

Acknowledgement

This work is supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No.2018R1C1B600543513). This work was also supported by the BK21 FOUR funded by the Ministry of Education(MOE, Korea) and National Research Foundation of Korea(NRF). MD simulations were performed with KISTI supercomputer resources(Neuron)

Author contributions statement

Jinyoung Byun and Juyong Lee conceived, designed, performed and analyzed the experiments and wrote the manuscript.

The corresponding author is responsible for submitting a [competing interests statement](#) on behalf of all authors of the paper.

This statement must be included in the submitted article file.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryIdentifyinghotspotsofSARSCOV2mainprotease.pdf](#)