A Multiple-step *In Silico* Screening Protocol to Identify Allosteric Inhibitors of Spike-hACE2 Binding

Jingchen Zhai, Xibing He, Viet Hoang Man, Yuchen Sun, Beihong Ji, Lianjin Cai, 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@pitt.edu

Jingchen Zhai is a PhD student in the School of Pharmacy, University of Pittsburgh. Her research interest is drug PK profile prediction and large-scale virtual screening.

Xibing He is a Research Scientist in the School of Pharmacy, University of Pittsburgh. His research interests include force-field development, free-energy calculations and computer-aided drug design.

Viet Hoang Man is a Research Instructor in the School of Pharmacy, University of Pittsburgh. His research focuses on amyloid aggregation, protein-ligand/protein interactions and force field development.

Yuchen Sun is a Master student in the School of Pharmacy, University of Pittsburgh. His research interest includes computer-aided drug design and free energy calculation improvement.

Beihong Ji is a PhD student in the School of Pharmacy, University of Pittsburgh. Her research interest is the development of computational tools for *in silico* drug discovery and design.

Lianjin Cai is a Master student in the School of Pharmacy, University of Pittsburgh. His research interests currently are in pharmacometrics and modeling application in translational research.

Junmei Wang is an associate professor in School of Pharmacy, University of Pittsburgh. His research interest includes computer modeling and simulation of protein–ligand interactions and other biological events, and pharmacometrics and computational systems pharmacology.

Abstract

While the COVID-19 pandemic keeps deteriorating, the effective medicines target the life cycle of the SARS-CoV-2 are still under development. With more highly infective and dangerous variants of the coronavirus emerged, the protective power of vaccines decreased or vanished. Thus, the development of drugs which are free of drug resistance is in dire need. The aim of this study is to identify allosteric binding modulators from a large compound library to inhibit the binding between the Spike protein of the SARS-CoV-2 virus and human angiotensin-converting enzyme 2 (hACE2). The Spike protein binding to hACE2 is the first step of the coronavirus to infect the host cells. We first built a compound library containing 77,448 antiviral compounds. Molecular docking was then conducted to preliminarily screen compounds which can potently bind to the Spike protein at two allosteric binding sites. Next, molecular dynamic simulations were performed to accurately calculate the binding affinity between the Spike protein and an identified compound from docking screening, and investigate whether the compound can interfere with the binding between Spike and hACE2. We successfully identified two possible drug binding sites on the Spike protein and discovered a series of antiviral compounds which can weaken the interaction between the Spike protein and hACE2 receptor through the conformational changes of the key Spike residues at the Spike-hACE2 binding interface induced by the ligand binding at the allosteric binding site. We also applied our screening protocol to another compound library which consists of 3,407 compounds for which the inhibitory activities of Spike/hACE2 binding were measured. Encouragingly, *in vitro* data supports that the identified compounds can inhibit the Spike-ACE2 binding. Thus, we developed a promising computational protocol to discover allosteric binding inhibitors for the binding of Spike protein of SARS-CoV-2 to hACE2 receptor, and several promising allosteric modulators were discovered.

Keywords: COVID-19; SARS-CoV-2; MD Simulation; Allosteric Binding Inhibitor; Virtual Screening; Drug Resistance; Binding Affinity Prediction.

1. Introduction

The outbreak of coronavirus disease 2019 (COVID-19) caused by the epidemic SARS-CoV-2 virus continues being uncontrolled around the US and many other countries all over the world. More seriously, new variants of the coronavirus emerged due to detrimental mutations on viral functional proteins and structural proteins especially the Spike protein.[1-3] The only countermeasure recommended by governments is vaccination.[4] However, the reinfection of the recovered patients is still a rising concern,[5-7] and the scientific research data is still not sufficient enough to confirm whether those vaccinated populations can acquire immunity for different variants of SARS-CoV-2 coronavirus or not. Moreover, the treatment of COVID-19 patients is still facing a great challenge, with many proposed drugs continuously being reported of their severe side effects. For example, Remdesivir may give rise to unwanted effects of hepatic disorder and cardiac adverse events,[8-10] and Hydroxychloroquine may raise the risk of cardiotoxicity.[11-13]

SARS-CoV-2 virus, a positive-sense single-stranded RNA virus,[14] shares a similar genomic organization to other beta-coronavirus.[15] It consists of four structural proteins and many nonstructural proteins (NSP). The structural proteins include Spike protein (S), envelope protein (E), membrane (M), and nucleocapsid (N). S, E, and M constitute the viral coat, and N packages the viral genome.[16] Specifically, the Spike protein mediates the virus entry into human angiotensin-converting enzyme 2 (hACE2) through the fusion of viral and cellular membranes.[17] Therefore, the interference of the interaction between the Spike protein and hACE2 should be a feasible way to inhibit the virus attachment to the host and thus decrease the infection.

Nowadays, many studies have been conducted to discover and develop therapeutics to inhibit the Spike-hACE2 protein-protein binding.[18-21] The development of drugs which can

weaken the Spike/hACE2 interaction through allostery could be a valid avenue to overcome the detrimental mutation effect. In this work, we attempted to discover potential inhibitors which bind to Spike protein at allosteric sites. With the development of computer-aided drug design (CADD) technology, *in silico* prediction accuracy and efficacy is continuously increasing. Combining different types of scoring functions greatly contributes to the successful identification of promising compounds through large-scale virtual screening. Since antiviral compounds are to date the most promising candidate category to treat COVID-19,[22, 23] this study conducted *in silico* screening among a library consisting of 77,448 antiviral compounds. Molecular docking screening, molecular dynamic (MD) simulations, and MM-PBSA-WSAS binding free energy calculation were conducted as a multiple-step screening pipeline to identify compounds which can inhibit the binding of the Spike protein to hACE2. Key residues for the protein-ligand interaction were also analyzed during MD simulation process. To strengthen our screening reliability, the same screening protocol was also performed using another compound library with *in vitro* experimental data to inhibit the Spike-ACE2 protein-protein interaction. Overall, our findings can provide insights to COVID-19 drug development.

2. Methods And Experimental Design

Allosteric binding site identification and validation. The crystal structure of the Spike-ACE2 complex was downloaded from Protein Data Bank (PDB ID: 6M17)[24]. The latest published crystal structure of the Spike protein is still lacking the information of the binding site for drug attachment. We first used the SiteID module of the Sybyl software to detect possible binding pockets of the Spike protein which was extracted from the Spike-ACE2 complex.[25] Each binding pocket is represented by key surrounding residues. The potential of an allosteric binding site was validated by the docking performance of a great number of structurally diverse compounds as described below.

Screening library construction. We compiled three compound libraries as detailed below. Library 1 consists of 123,192 druglike compounds collected by ZINC database.[26] Library 1 was applied to evaluate the potential of an allosteric binding site to serve as a drug binding pocket. To compile Library 2, we downloaded a ligand library from NIH OpenData Portal,[27] which possesses a series of compounds with SARS-CoV-2 *in vitro* screening data. The experimental data describes the inhibition extent of Spike-hACE2 interaction under the influence of each compound. The third library, Library 3 is a collection of antiviral compounds from various sources. [28-30]

Validation of screening protocol. In this part of our research, we aimed to design a protocol to gradually screen out a series of compounds through a large compound library with experimental data describing their activity of inhibiting the interaction between the Spike protein and hACE2 protein. After being applied to these reported compounds, our protocol can be validated by comparing the experimental data with our prediction results. Our general protocol is divided into three steps: 1) dock the compound library to the Spike protein to preliminary screen

out ligands which bind to the Spike protein with high affinities. 2) conduct MD simulations for top-ranking ligands in Step 1 with the Spike protein for more accurate binding affinity prediction. 3) perform MD simulations for promising ligands from Step 2 with the Spike-hACE2 complex and calculate binding free energy between the Spike protein and hACE2 protein under the interference of an allosteric ligand. We also conducted binding free energy calculation and energy decomposition analyzation based on MD simulation results, and the final screening results were compared to the experimentally reported data.

Glide Docking Screening. By using the Receptor Grid Generation module from Maestro software platform (Schrödinger Release 2017-3: Schrödinger, LLC, New York, NY, 2017.), binding pockets were located by the key residues identified by SiteID. Glide docking was performed with the major settings described as follows: (1) the scaling factor of van der Waals radii is 0.8 and the partial charge cut off is 0.15; (2) flexible ligand docking was performed under standard precision; (3) intramolecular hydrogen bonds were rewarded; (4) other settings were kept default. One pose per ligand was written out. Generally, a binding site should have a higher possibility of serving as a drug binding site when more drug-like compounds tend to have better binding affinities at the site, which is reflected by the lower docking scores.

MD simulation. During the MD simulation for Step 2, each ligand-Spike protein system includes a ligand, a Spike protein, approximately 20,500 TIP3P water molecules,[24] 60 Cl^- to keep the salt concentration at 0.15 M and necessary Na^+ to neutralize the system. For Step 3 MD simulation, each system includes a ligand, a Spike-hACE2 complex, approximately 43,800 TIP3P water molecules, 120 Cl^- ions, and necessary Na^+ ions to neutralize the system. The RESP partial charges were obtained by fitting the HF/6-31G* electrostatic potentials generated using the Gaussian 16 package.[31, 32] The force field parameters for ligands were from the General Amber

Force Field (GAFF),[33] and the parameters of proteins came from AMBER FF14SB [34]. The topologies were all prepared utilizing the Antechamber module from AMBER software package.[35, 36] For each MD simulation for a ligand-protein complex, a five 10,000-step energy minimization was first performed to remove possible steric crashes in the system. The restrained harmonic force constant was reduced from 20 to 10, 5, 1, and 0 kcal/mol/Å^2, step by step. The followed MD simulations consist of three phases: relaxation, equilibrium, and sampling. During the relaxation phase, the system was heated up from 50 K to 298 K. For the equilibrium phase and the sampling phase, a 10-ns and a 100-ns MD simulation were carried out at 298 K, respectively. For each system, the equilibrium and sampling phases were repeated multiple times as individual runs. Additional parameter setting during the MD simulations includes: the temperature was controlled utilizing Langevin dynamics with a collision frequency of 5 ps^{-1};[37] pressure was regulated using the isotropic scaling algorithm with the relaxation time of 1.0 ps; The time step was 1 fs for the relaxation phase and 2 fs for the other two phases. The SHAKE algorithm was utilized to constrain the bonds involving Hydrogen atoms. [38] All minimization, relaxation and the equilibration were performed using the Central Processing Unit (CPU) version of the PMEMD module in the AMBER 18 package,[33] and the sampling phase was conducted using the Graphic Processing Unit (GPU) version of PMEMD.[39]

MM-PB/GBSA-WSAS Binding Free Energy Calculation and Energy Decomposition. The solvation free energy for each MD system was calculated using the molecular mechanics/Poisson Boltzmann surface area (MM/PBSA) method or molecular mechanics/ Generalized Born surface area (MM/GBSA) method, and the conformational entropy was estimated with the WSAS method.[40-43] For a molecule in the solvent, calculation equation of the free energy is as follows:

$$\Delta \boldsymbol{G}_{\boldsymbol{MM\text{-}PB}/\mathrm{GBSA}} = \Delta \mathrm{H} - \boldsymbol{T}\Delta \boldsymbol{S} = \Delta \boldsymbol{E}_{\boldsymbol{vdw}} + \Delta \boldsymbol{E}_{\boldsymbol{ele}} + \Delta \boldsymbol{E}_{\mathrm{inter}} + \Delta G_p^{sol} + \Delta G_{np}^{sol} - \boldsymbol{T}\Delta \boldsymbol{S}$$

$\Delta \boldsymbol{E}_{\boldsymbol{vdw}}$ is the van der Waals energy, $\Delta \boldsymbol{E}_{\boldsymbol{ele}}$ is the gas phase electrostatic energy, $\Delta \boldsymbol{E}_{\mathrm{inter}}$ is the internal energy, ΔG_p^{sol} and ΔG_{np}^{sol} are the polar and nonpolar solvation energy, T is the absolute temperature, and $\Delta \boldsymbol{S}$ is the change of conformational entropy. 362 snapshots were evenly collected for MD simulation systems containing a ligand and a Spike protein (Step 2), and 456 snapshots were evenly collected for MD simulation systems containing the Spike-hACE2 complex with ligands (Step 3), in order to perform MM-PBSA-WSAS calculation. For energy decomposition, the solvent effect was taken into account using a Generalized Born model.[44] The external and internal dielectric constant is 80 and 1, respectively. The calculation was conducted utilizing the Sander program in AMBER 18.[33] Residues are considered as hot spot residues when the MM-GBSA ligand-residue interaction is stronger than -1.0 kcal/mol. We conducted free energy decomposition analyses for both Spike-ligand (Step 2) and Spike-hACE2-ligand (Step 3) systems, utilizing 1000 and 5000 snapshots, respectively.

3. Results

3.1 Identification and Validation of Allosteric Binding Sites

Seven binding sites were detected by SiteID module of the Sybyl software package,[25] with the surrounding residues listed in **Table 1** and shown in **Figure 2.** Docking screenings were performed at each binding site using Library 1. According to the docking results of this druglike compound library to each defined binding site (**Table S1**), Cluster1 and Cluster4 have 33 and 21 compounds with docking scores better than -7.0 kcal/mol, respectively, and thus were identified as highly possible drug binding sites on the Spike protein. Other clusters all have none or no more

than three ligands with docking scores lower than -7.0 kcal/mol, indicating their poor binding affinities between the compounds and Spike protein at these binding sites.

3.2 Validation of Screening Protocol Using Library 2

Docking screening. Library 2 contains 3,407 compounds with *in vitro* data to cripple the Spike-ACE2 interaction. The docking results of this compound library to Cluster1 and Cluster4 have been summarized in **Table S2 and S3**. 30 top ligands for Cluster1 and 20 top ligands in Cluster4 have been chosen for compound-Spike protein MD simulation.

MD simulations for a ligand/Spike protein complex. All the top compounds from docking screening are further filtered by this first-round MD simulation. For all simulation systems, the starting conformations of the ligands came from the docking results. For Cluster1, 12 compounds were eliminated by this MD simulations filter because those ligands drifted away from the binding pocket during the simulation process. 18 compounds can bind to the Spike protein throughout the simulation process and were performed two more rounds of MD simulations using time-dependent random seed for the Langevin dynamics. 8 of these compounds possess promising predicted binding free energy according to the calculation and were performed another two rounds of MD simulation for each compound. The predicted binding free energy for 8 compounds during multiple rounds of simulation are listed in **Table 2** and detailed information for 18 compounds has been summarized in **Table S4.** Lower binding free energy means better binding affinity of the ligand targeting to the Spike protein. On the other hand, although Cluster4 has 14 compounds which can steadily bind to the Spike protein throughout the first-round simulation, none was identified with low binding free energy after 3 rounds of MD simulations. Thus, Cluster4 does not have a candidate compound survived for the next-step research. The detailed information of predicted binding free energy for 14 compounds is shown in **Table S5.**

Confirmation of an allosteric ligand to compromise Spike/hACE2 binding. According to the predicted binding affinity for the 8 compounds for Cluster1 in the previous step, 5 representative conformations were generated during multiple simulations for each compound. Considering the calculated binding free energy for each simulation, 7 Spike-hACE2-ligand systems, involving 4 different compounds which have promising binding affinity to the Spike protein, were constructed for further study. The selection has been labeled in **Table 2** (N1 to N7)**.** To prepare the initial structure of a Spike-hACE2-ligand complex, the Spike protein in a representative MD conformation of Spike-ligand complex was aligned to the Spike protein in the Spike-hACE2 crystal structure resulting in a translation-rotation matrix. Then ligand structure in the Spike-ligand complex was transformed using the matrix and then combined with the Spike-hACE2 complex.

From another research publication by our lab, the predicted binding free energy between the wild-type Spike protein and hACE2 is -16.04 kcal/mol using the same MD simulation setting as described in the Method section.[45] During this research, the higher predicted binding free energy between the Spike protein and the hACE2 protein under the influence of a compound represents the stronger inhibitory effect the allosteric ligand has. Every Spike-hACE2-ligand system underwent 6 separate simulations and the calculated Spike-hACE2 binding free energies were listed in **Table 3** and the decomposition of different energy terms was summarized in **Table S6.** The docking score and *in vitro* experimental value for each compound were also summarized in **Table 3** for comparison**.** Encouragingly, the prediction results are consistent with the experimental values. N1 system (Spike-hACE2 complex with compound NCGC00389662-01) has the highest predicted binding free energy between the Spike protein and the hACE2 protein under the influence of the compound based on the 6 simulations, indicating the interaction between the

Spike protein and the hACE2 has been compromised. And the *in vitro* experiment result for N1 also shows the efficacy of -24.30 to inhibit the interaction between the Spike protein and hACE2 (**Table 3**). For other systems (N2 to N7), the ligands either enhanced or slightly decreased the binding affinities between Spike and hACE2, thus the inhibitory effect was not obvious. This result is consistent with the experimental findings that those ligands (NCGC00521952-01, NCGC00167505-03, NCGC00379053-02) cannot decrease the binding between Spike and hACE2 proteins, with their efficacy of 0 in *in vitro* experiment.

Furthermore, to facilitate the development of more potent allosteric inhibitors, we conducted free energy decomposition analysis to identify hotspot residues which make significant contributions to the protein-ligand binding utilizing all six MD trajectories for the N1 system. A heatmap exhibiting the identified hotspot residues was shown in **Figure 3.** As shown in **Figure 4C**, most hotspot residues are around the protein-protein binding interface, and the allosteric effect upon ligand binding can be effectively transferred to those residues. The identified hotspot residues here, Ile32, Cyx40, Glu744 and Cyx748, were also recognized as key residues in our previous study to interact with the hACE2 receptor at the protein-protein binding interface.[45] The inhibitory effect is the result of conformational changes on those hotspots located at the Spike-hACE2 binding interface.

3.3 Identification of Promising Ligands from Library 3.

Glide docking screening. 77,448 compounds in Library 3 were screened and ranked according to their docking scores, which were summarized in **Table S7** for Cluster1 pocket and **Table S8** for the Cluster 4 pocket. For those antiviral compounds, the best docking score is -9.76 kcal/mol for Cluster1 and -9.19 kcal/mol for Cluster4. All compounds with docking scores lower than -8.0 kcal/mol were selected for the next step MD simulation. In total, 46 compounds for

Cluster1 and 43 compounds for Cluster4 were selected. In addition, 15 top-ranked approved drugs were also selected to undergo the MD simulation screening for both binding pockets.

MD simulations for a ligand/Spike protein complex. Again, the docking structure of a compound served as the starting conformation for the subsequent MD simulations. For Cluster1, only 25 ligands, including approved drugs and antiviral compounds, completed the whole procedure of MD simulations and the calculated binding free energies after two rounds of MD simulations were summarized in **Table S9**. Three more MD runs were performed for five antiviral compounds which had the best MM-PBSA-WSAS binding free energies calculated using the snapshots collected in two MD runs. As shown in **Table 4**, two compounds, 1426855-10-8 and 14350-38-0, outperformed others in terms of potency and reliability of prediction. As to Cluster4, only 16 ligands completed the whole procedure of MD simulations and the MM-PBSA-WSAS binding free energies were listed in **Table S10**. For six antiviral compounds, three more MD runs were performed and the binding free energies were summarized in **Table 5**. Three promising compounds, 1053055-44-9, 149297-82-5 and 2001001-59-6 were recognized as promising inhibitors.

However, strong binding may not necessarily be a potent allosteric inhibitor as the perturbation upon ligand binding must be propagated to the protein-protein binding interface and triggered conformational changes on the hotspot residues. We conducted energy decomposition to those systems which exhibit good binding affinities to the Spike protein. The heatmap for Cluster1 and Cluster4 were depicted in **Figure 5A** and **5B**, respectively. Spike protein residues with strong ligand-residue interactions are marked in darker colors. As Cluster 1 is around the Spike-hACE2 binding interface, naturally most hotspot residues are located at the protein-protein binding interface. Those hotspot residues located around the protein-protein binding interface are colored

in yellow (**Figure 6).** Ile732, Tyr733, Cys740, Glu744, and Tyr749, which are identified as key residues in this study, were also recognized as hotspot residues at the protein-protein interface to facilitate the binding of the Spike protein to hACE2 in our previous research.[45] The following compounds were selected for further study: 1426855-06-2, 1426855-06-8, and 20777-72-4. Those compounds are most potent binding to the Spike protein (**Table 4**) and have protein-protein binding interface hotspot residues (**Figure 5A**).

As to Cluster4, the binding pocket is far away from the protein-protein binding interface (**Figure 2**), thus the hotspot residues marked in darker green are mostly around the binding site but not near the interface (**Figure 7**). To investigate whether the ligand-residue interaction can trigger the conformational changes on the residues located at the Spike-hACE2 protein-protein binding interface, we performed residue correlation analysis and the results were shown in **Figure 7**. The binding site surrounding residues which tend to interact with the selected compounds and at the same time have a high correlation with interface residues were marked in orange, and the correlated interface residues were colored in yellow. The calculated residue-residue correlations were normalized and listed in **Table 6**. Ser609, Tyr611, Ala612, and Arg726, the hotspot residues with the ligand-residue interaction energies lower than -1.0 kcal/mol, have strong correlation with 8 interface residues applying 0.3 as the threshold of the correlation value. Among these 8 interface residues, Tyr713 and Tyr755 were also key interface residues according to our previous research result.[45] This means the interaction between the ligand and the binding site hotspot residues may contribute to a cascade of the conformational change around the Spike-hACE2 binding interface which weakens the Spike-hACE2 binding. Finally, 1002334-80-6 and 2001001-59-6 in total having three promising binding modes (**Table 5** and **Figure 5B)** entered the next stage.

Investigation on the ability of an allosteric ligand compromising Spike/hACE2 binding. We applied a similar protocol to prepare the starting conformations of the Spike/hACE2-ligand complex as detailed above. 8 independent MD simulations were performed for most systems. The predicted binding free energies of the protein-protein interaction were summarized in **Table 7**. Comparing with the wild type binding free energy of -16.04, B1_C1 (compound 1426855-06-2 for Cluster1) shows very promising outcomes in reducing the binding affinity between the protein complex with their influence (**Part A of Table 7**). Only 6 MD runs were performed for B2_C1 as the compounds had little inhibitory effect on the Spike-hACE2 binding. As all the 8 rounds of simulation for the rest complexes contributed to large standard deviation (SD), we eliminated three calculated binding affinity values which were considered as outliers for each system. The results were updated and exhibited in **Part B** of **Table 7**. It is shown that B2_C4 and B3_C4 from Cluster4 have the best performance in disrupting the protein-protein interaction, and the rest systems, two from Cluster1 and one from Cluster4 can also decrease the binding affinity of the Spike protein to the hACE2 receptor. All the SD structural information of the 4 compounds is shown in **Figure 8.** This result is very promising as most selected systems exhibited potentiality to inhibit Spike-hACE2 protein-protein interaction, suggesting the screening protocol is effective.

4. Discussion

The importance of this research. Computational prediction possesses the advantage of lower cost and higher efficiency compared to *in vitro* and *in vivo* research during drug discovery and development. By utilizing different *in silico* screening methods, algorithms with high efficiency and with high accuracy cooperated as a multiple-step screening and prediction method. Under the current pandemic situation, the rapid spread and variants of coronavirus as well as the lack of clinical treatment of this disease is urgently calling for the discovery of promising antiviral compounds for further research and clinical validation. Unlike most research in the field, we focused the development of allosteric inhibitors targeting to compromise Spike-hACE2 binding. An effective and potent allosteric inhibitor has better chance to escape from drug resistance, hence it can be used to combat different variants of the coronavirus. We have developed and evaluated a screening pipeline to screen allosteric inhibitors utilizing multiscale molecular modeling techniques, including binding pocket prediction, docking simulations, molecular dynamics simulations, binding free energy calculations, free energy decomposition, and correlation analysis. Moreover, we have also come out with guidelines on how to select compounds to enter the next stages of screening. The research protocol can be applied to develop allosteric inhibitors of the Spike protein of Sars-CoV2 through large-scale screenings.

The hierarchical screening pipeline. According to the research, compounds collected from different libraries were subjected to a series of hierarchical virtual screenings. The selection of compounds was under cautious consideration, especially for Library 3, which includes antiviral compounds from different sources and approved drugs. Considering the docking scores are affected by the ligand size, we selected top-ranking ligands not purely based on docking scores, but also the sources. For example, we selected a number of drugs to enter the next stage of

screening pipeline even though their docking scores are not very satisfactory. When we selected potent inhibitors to evaluate their inhibitory effect on weaken Spike-hACE2 binding, not only the binding affinity of ligands to the Spike protein, but also the allosteric effect were considered. The allostery occurs when the hotspot residues themselves are binding surface residues or the hotspots residues have strong correlation with binding surface residues. For the two allosteric binding pockets, Cluster1 is located around the protein-protein binding interface, while Cluster4 is far away from the protein-protein binding interface. We are more interested in identifying allosteric inhibitors binding in the Cluster4 binding pocket, as the allosteric inhibitors have better ability to escape drug resistance effect. For the last stage of the screening, although 1053055-44-9 and 149297-82-5 have promising binding affinity to the Spike protein at the Cluster4 binding pocket (**Table 5**), the residue-ligand interaction analysis did not show strong correlations between their hotspot residues and protein-protein binding interface residues. Thus, they were not selected to enter the last stage of the screening pipeline, i.e., to study the inhibitory ability of a compound weakening Spike-hACE2 binding.

In our screening pipeline of identifying allosteric inhibitors, two types of MD simulations were performed for Library 2 and Library 3. The MD simulations of a Spike-Ligand complex and a Spike-hACE-ligand serves different purpose. For the first type of MD simulations, we calculated binding free energy between a ligand and the Spike protein. The lower the binding free energy, the better chance it is a strong allosteric inhibitor. Many selected docking hits were dropped off at this stage, as they drifted away from the binding pocket during the 100 ns MD simulations. Thus, MD simulation is an efficient filter in screening allosteric inhibitors, especially for a not well-defined binding pocket. For the second type of MD simulations, we calculated the binding free energy of

Spike and hACE2 binding to confirm the allosteric effect. The higher the binding free energy, the better the inhibitory effect.

The validation of our protocol using compounds with *in vitro* experimental data. To the best of our knowledge, there is no report on the allosteric inhibitors of the Spike protein of Sars-CoV2. We indirectly evaluated our screening protocol using a compound library consisting of compounds with known *in vitro* experimental data of inhibiting Spike protein and the hACE2 protein binding. Encouragingly, our screening results are consistent with the measured data, suggesting the screening protocol is valid. However, we want to emphasize that the experiment did not provide information on how and where an inhibitor binds to the Spike or Spike-hACE2 proteins.

The prediction stability of MD simulation result. According to the calculated binding affinity for the same compound in separate simulations, the binding free energy sometimes variates a lot during multiple simulations as illustrated in **Table 7**. One reason lies that the allosteric binding pockets are not well-defined thus the docking poses may not have high quality. Another reason for this phenomenon should be related to the compound size. Most compounds in this study are very bulky and flexible, leading to high fluctuations during the MD simulations. This high uncertainty problem can be mitigated by performing multiple independent MD simulations. For example, we conducted up to eight MD runs at the last stage of the screening pipeline and the prediction results became more reliable after removing outliers.

5. Conclusion

In this research, we successfully developed a screening pipeline to develop allosteric inhibitors to weaken the interaction between the Spike protein and hACE2 protein. The

compromised Spike binding can reduce the infectability of the coronavirus. Two allosteric binding sites were identified and evaluated through virtual screenings of three compound libraries. Active ligands in Cluster1 tend to directly interfere with the binding of the Spike protein to the hACE2 receptor, while active ligands in Cluster4 can serve as allosteric binding modulators to inhibit the interaction between the Spike protein and the hACE2 receptor. The screening protocol is at least partially validated by a set of compounds in Library 2 for which their inhibitory activities of Spike-hACE2 binding were measured. Applying the multiple-step screening pipeline, four out of 77,448 compounds were identified as allosteric inhibitors which can decrease the binding affinity of the Spike-hACE2 complex as indicated by more positive MM-PBSA-WSAS binding free energies. Compared to *in vitro* and *in vivo* experiments, our computational method is not only highly efficient and cost-effective, but also elucidates the molecular mechanism of the allosteric effect to guide rational drug development.

Key Points

- We proposed to develop allosteric inhibitors for Spike-hACE2 of Sars-CoV2 which have potential to escape from drug resistance.
- We provided a hierarchical *in silico* screening protocol to identify allosteric inhibitors targeting the Spike-hACE2 protein-protein binding. This proposed method possesses higher screening efficiency and less cost compared to *in vitro* and *in vivo* experiments.
- This multistep screening protocol has been validated with a compound library with *in vitro* inhibitory activities of the Spike-hACE2 binding.
- Four allosteric compounds were identified from an antiviral compound library. One of them can lower the binding affinities of Spike-hACE2 binding more than 2.0 kcal/mol.

Supplementary Data

Table S1 summarizes the docking scores of top 200 drug-like compounds (library 1) in seven different clusters predicted by Sybyl. **Table S2** and **Table S3** list the top 200 docking score of compounds with *in vitro* data (library 2) in cluster1 and cluster4, respectively. **Table S4** describes detailed MM-PBSA binding free energies of the 18 compounds from library 2 during multiple simulations in cluster1 and **Table S5** describes MM-PBSA binding free energies of the 14 compounds from library 2 in cluster4 during multiple simulations. **Table S6** shows detailed MM-PBSA binding free energies for N1 to N7 during multiple simulations. **Table S7 - S8** lists the top 200 docking scores of the antiviral compounds (library3) in cluster1 and cluster4, respectively. **Table S9 – S10** includes the detailed MM-PBSA binding free energies of the 25 ligands in cluster1 and 16 ligands in cluster4, respectively, from library 3 during multiple simulations.

Acknowledgments

This work was supported by National Institutes of Health (Grant No: R01GM079383 to JW) and the National Science Foundation (Grant No: 1955260 to JW). The authors also thank the computing resources provided by the Center for Research Computing (CRC) at University of Pittsburgh.

Conflict of Interest

There are no conflicts to declare.

References

1. Tegally, H., et al., *Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa.* medRxiv, 2020: p. 2020.12.21.20248640.
2. Kemp, S.A., et al., *Recurrent emergence and transmission of a SARS-CoV-2 spike deletion H69/V70.* bioRxiv, 2021: p. 2020.12.14.422555.
3. Wang, L., L. Wang, and H. Zhuang, *Profiling and characterization of SARS-CoV-2 mutants' infectivity and antigenicity.* Signal Transduction and Targeted Therapy, 2020. **5**(1): p. 1-2.
4. Li, R., et al., *Differential efficiencies to neutralize the novel mutants B. 1.1. 7 and 501Y. V2 by collected sera from convalescent COVID-19 patients and RBD nanoparticle-vaccinated rhesus macaques.* Cellular & molecular immunology, 2021. **18**(4): p. 1058-1060.
5. Ota, M., *Will we see protection or reinfection in COVID-19?* Nature Reviews Immunology, 2020. **20**(6): p. 351-351.
6. Stokel-Walker, C., *What we know about covid-19 reinfection so far.* bmj, 2021. **372**.
7. Costa, A.O.C., et al., *COVID-19: Is reinfection possible?* EXCLI journal, 2021. **20**: p. 522.
8. Montastruc, F., S. Thuriot, and G. Durrieu, *Hepatic Disorders With the Use of Remdesivir for Coronavirus 2019.* Clinical Gastroenterology and Hepatology, 2020. **18**(12): p. 2835-2836.
9. Gupta, A.K., et al., *Cardiac adverse events with remdesivir in COVID-19 infection.* Cureus, 2020. **12**(10).
10. Beigel, J.H., et al., *Remdesivir for the treatment of Covid-19.* New England Journal of Medicine, 2020. **383**(19): p. 1813-1826.
11. Alanagreh, L.a., F. Alzoughool, and M. Atoum, *Risk of using hydroxychloroquine as a treatment of COVID-19.* International Journal of Risk & Safety in Medicine, 2020. **31**(3): p. 111-116.
12. Stevenson, A., et al., *Hydroxychloroquine use in COVID-19: is the risk of cardiovascular toxicity justified?* Open heart, 2020. **7**(2): p. e001362.
13. Liu, J., et al., *Hydroxychloroquine, a less toxic derivative of chloroquine, is effective in inhibiting SARS-CoV-2 infection in vitro.* Cell discovery, 2020. **6**(1): p. 1-4.
14. Chowdhury, S.M., et al., *Antiviral Peptides as Promising Therapeutics against SARS-CoV-2.* The Journal of Physical Chemistry B, 2020. **124**(44): p. 9785-9792.
15. Zhu, N., et al., *A novel coronavirus from patients with pneumonia in China, 2019.* New England journal of medicine, 2020.
16. Wu, C., et al., *Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods.* Acta Pharmaceutica Sinica B, 2020. **10**(5): p. 766-788.
17. Li, F., *Structure, function, and evolution of coronavirus spike proteins.* Annual review of virology, 2016. **3**: p. 237-261.
18. Kumar, V., H. Liu, and C. Wu, *Drug repurposing against SARS-CoV-2 receptor binding domain using ensemble-based virtual screening and molecular dynamics simulations.* Computers in Biology and Medicine, 2021. **135**: p. 104634.
19. Lehrer, S. and P.H. Rheinstein, *Ivermectin docks to the SARS-CoV-2 spike receptor-binding domain attached to ACE2.* in vivo, 2020. **34**(5): p. 3023-3026.
20. Huo, J., et al., *Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2.* Nature structural & molecular biology, 2020. **27**(9): p. 846-854.
21. Esparza, T.J., et al., *High affinity nanobodies block SARS-CoV-2 spike receptor binding domain interaction with human angiotensin converting enzyme.* Scientific reports, 2020. **10**(1): p. 1-13.
22. Jean, S.-S., P.-I. Lee, and P.-R. Hsueh, *Treatment options for COVID-19: The reality and challenges.* Journal of Microbiology, Immunology and Infection, 2020. **53**(3): p. 436-443.

23. YAVUZ, S. and S. Ünal, *Antiviral treatment of COVID-19.* Turkish journal of medical sciences, 2020. **50**(SI-1): p. 611-619.
24. Yan, R., et al., *Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2.* Science, 2020. **367**(6485): p. 1444-1448.
25. Gelius-Dietrich, G., et al., *sybil – Efficient constraint-based modelling in R.* BMC Systems Biology, 2013. **7**(1): p. 125.
26. Wang, J., Y. Ge, and X.-Q. Xie, *Development and testing of druglike screening libraries.* Journal of chemical information and modeling, 2018. **59**(1): p. 53-65.
27. Duan, Y., et al., *A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations.* J Comput Chem, 2003. **24**(16): p. 1999-2012.
28. Morton, W.A. and G.G. Stockton, *Methylphenidate Abuse and Psychiatric Side Effects.* Prim Care Companion J Clin Psychiatry, 2000. **2**(5): p. 159-164.
29. Zhuang, X. and C. Lu, *PBPK modeling and simulation in drug research and development.* Acta Pharm Sin B, 2016. **6**(5): p. 430-440.
30. Faraone, S.V., et al., *The worldwide prevalence of ADHD: is it an American condition?* World Psychiatry, 2003. **2**(2): p. 104-13.
31. Bayly, C.I., et al., *A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model.* The Journal of Physical Chemistry, 1993. **97**(40): p. 10269-10280.
32. Frisch, M., et al., *Gaussian 16*. 2016, Gaussian, Inc. Wallingford, CT.
33. Wang, J., et al., *Development and testing of a general amber force field.* Journal of computational chemistry, 2004. **25**(9): p. 1157-1174.
34. Maier, J.A., et al., *ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB.* Journal of chemical theory and computation, 2015. **11**(8): p. 3696-3713.
35. Wang, J., et al., *Automatic atom type and bond type perception in molecular mechanical calculations.* Journal of molecular graphics and modelling, 2006. **25**(2): p. 247-260.
36. Case, D.A., et al., *Amber 2020.* 2020.
37. Larini, L., R. Mannella, and D. Leporini, *Langevin stabilization of molecular-dynamics simulations of polymers by means of quasisymplectic algorithms.* The Journal of Chemical Physics, 2007. **126**(10): p. 104101.
38. Miyamoto, S. and P.A. Kollman, *Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models.* Journal of Computational Chemistry, 1992. **13**(8): p. 952-962.
39. Le Grand, S., A.W. Götz, and R.C. Walker, *SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations.* Computer Physics Communications, 2013. **184**(2): p. 374-380.
40. Wang, E., et al., *Assessing the performance of the MM/PBSA and MM/GBSA methods. 10. Impacts of enhanced sampling and variable dielectric model on protein–protein interactions.* Physical Chemistry Chemical Physics, 2019. **21**(35): p. 18958-18969.
41. Wang, J., et al., *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.* Journal of the American Chemical Society, 2001. **123**(22): p. 5221-5230.
42. Hou, T., et al., *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): p. 69-82.
43. Page, C.S. and P.A. Bates, *Can MM-PBSA calculations predict the specificities of protein kinase inhibitors?* Journal of Computational Chemistry, 2006. **27**(16): p. 1990-2007.

44. Hawkins, G.D., C.J. Cramer, and D.G. Truhlar, *Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium.* The Journal of Physical Chemistry, 1996. **100**(51): p. 19824-19839.
45. Zhang, Y., et al., *In silico binding profile characterization of SARS-CoV-2 spike protein and its mutants bound to human ACE2 receptor.* Briefings in bioinformatics, 2021: p. bbab188.

Tables

Table 1. Identified key residues for 7 predicted binding sites.

Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
ARG714	ASN682	PHE660	ALA657	VAL655	VAL667	LEU773
GLU731	TYR683	ILE670	SER659	PHE775	ILE670	PHE775
ILE732	LYS684	ILE678	PHE607	ILE618	ALA671	TYR625
TYR733	THR690	ASN682	VAL771	SER619	VAL693	LEU628
TYR749	PHE724	TYR683	ASN614	TYR625	ALA695	PHE637
PRO751	GLU725	VAL610	LYS616	CYS651	VAL770	LEU647
	VAL772	TRP613	VAL601		VAL772	
	SER774					
	TRP613					
	ARG615					

Table 2. Cluster1 for Library 2: calculated binding free energies (kcal/mol) between allosteric ligand and Spike protein using the MM-PBSA-WSAS method. N1-N7 were further evaluated their ability to compromise Spike/hACE2 binding.

	Round1	Round2	Round3	Round4	Round5
NCGC00386028-03	-3.95±0.11	-0.59±0.88	-5.7±0.33	-4.24±1.05	4.76±0.85
NCGC00485882-01	2.01±0.73	-6.58±0.42	0.51±0.35	-5.37±0.12	7.65±0.58
NCGC00389662-01	-0.37±0.29	-3.67±0.33	-7.5±0.90	3.13±0.16	-13.34±0.68 (N1)
NCGC00521952-01	-7.94±0.44	-0.83±0.44	-10.09±0.29 (N2)	-10.29±1.09 (N3)	-5.02±0.38
NCGC00167505-03	-9.16±0.89	-15.42±1.5 (N4)	0.17±0.35	-11.51±0.35 (N5)	-10.04±0.70 (N6)
NCGC00179313-05	-6.52±0.29	4.49±0.27	-1.7±1.02	-0.25±1.51	-0.13±0.51
NCGC00379053-02	-0.67±0.64	3.65±0.38	-13.33±0.18 (N7)	-2.18±0.73	-1.92±0.11
NCGC00482516-02	-0.82±0.32	-2.61±1.08	-5.68±1.21	-7.66±1.00	-3.22±0.54

Table 3. Cluster1 for library 2: Calculated complex energies and binding free energies (kcal/mol) using the MM-PBSA-WSAS method of Spike-hACE2 complex systems under the influence of compounds with docking score and experiment data**.**

System	N1	N2	N3	N4	N5	N6	N7
Round1	-14.86±0.53	-12.55±0.14	-25.07±0.14	-18.37±0.31	-17.46±0.43	-17.87±0.23	-7.54±0.08
Round2	-6.39±0.16	-16.75±1.01	-19.95±0.77	-10.96±0.45	-16.43±0.12	-28.47±0.67	-19.75±0.09
Round3	-15.33±0.46	-17.69±0.16	-18.06±0.41	-15.77±0.32	-11.15±0.22	-14.87±0.49	-19.76±0.44
Round4	-9.05±0.38	-14.80±0.17	-14.16±0.28	-9.83±0.34	-12.46±0.30	-22.17±0.37	-10.87±0.57
Round5	-14.37±0.33	-17.79±0.45	-19.42±0.15	-21.22±0.21	-16.28±0.38	-7.70±0.49	-8.53±0.45
Round6	-10.30±0.10	-17.82±0.29	-14.96±0.51	-10.20±0.22	-13.48±0.30	-12.33±0.05	-16.93±0.60
SD	3.35	1.95	3.60	4.37	2.31	6.73	5.10
AVE	-11.72	-16.23	-18.60	-14.39	-14.54	-17.24	-13.90

Docking score	-9.37	-6.93	-7.06	-8.10
Inhibition efficacy*	-24.30	0.00	0.00	0.00

* Disrupting of the Spike protein-ACE2 interaction may cripple the ability of SARS-CoV-2 virions to infect host cells. The AlphaLISA assay determines the ability of a compound to disrupt the important protein-protein interaction. The measured inhibition efficiency data in AlphaLISA assay is normalized to a high-signal vehicle control (ACE2 + RBD + Beads + DMSO) and a low-signal control (ACE2 + Beads + DMSO). A high signal approaching zero indicates no deviation from the ACE2 and RBD protein-protein interaction; on the contrary, a low signal approaching -100 indicates the interaction between ACE2 and RBD does not occur. More details on the experimental protocol can be found from https://opendata.ncats.nih.gov/covid19.

Table 4. Cluster1 for Library 3: Calculated binding free energies (kcal/mol) between ligands and the Spike protein using the MM-PBSA-WSAS method of compound-Spike protein systems.

	Round1	Round2	Round3	Round4	Round5
1246392-09-5	-4.29±0.26	-7.85±1.21	1.17±0.58	-4.79±1.83	2.16±0.60
1426855-06-2	6.52±0.84	-7.14±0.76	-10.22±0.49 (B1_C1)	2.07±0.79	-3.60±1.53
1426855-10-8	-16.10±0.95	-10.18±0.85	-10.59±0.45	-16.31±0.68 (B3_C1)	-7.27±1.02
14350-38-0	-6.35±0.86	-7.61±1.06	-7.61±1.06	-23.08±0.29 (B2_C1)	-5.59±0.84
20777-72-4	0.12±0.83	-9.58±0.29	-2.82±0.14	-9.58±0.29	-9.04±0.13

Table 5. Cluster4 for Library 3: Calculated binding free energies (kcal/mol) between ligands and the Spike protein using the MM-PBSA-WSAS method of compound-Spike protein systems.

	Round1	Round2	Round3	Round4	Round5
1002334-80-6	0.31±2.16	-10.93±0.97 (B1_C4)	-4.28±1.10	4.91±0.83	-6.60±1.16
1023756-44-6	-3.41±0.27	2.75±1.05	4.10±0.60	0.27±0.19	4.10±1.44
1053055-44-9	-3.09±0.71	-10.70±0.19	-7.47±1.10	-12.31±0.80	-6.27±0.90
149297-82-5	-5.09±0.58	-7.33±0.78	-11.46±0.65	-5.63±0.61	-12.39±1.46
1807796-52-6	1.77±0.57	-5.05±0.59	-7.93±0.60	3.91±1.10	1.45±1.12
2001001-59-6	-14.70±0.69 (B2_C4)	-10.88±0.93 (B3_C4)	-6.57±1.01	-7.92±0.41	-11.01±0.10

Table 6. Residue correlation value of the Spike Protein.

RES_1	RES_2	CORR
SER609	TYR711	0.351
SER609	TYR713	0.333
TYR611	LEU712	0.331
TYR611	TYR713	0.325
TYR611	AGR714	0.332
TYR611	LEU752	0.345
ALA612	LEU712	0.349
TYR711	TYR755	0.331

TYR711	GLY756	0.328
AGR726	AGR717	0.397

Table 7. Calculated binding free energies (kcal/mol) using the MM-PBSA-WSAS method of the Spike-hACE2 complex with the influence of each compound

A: predicted binding free energy and system energy for each simulation process						
	B1_C1	B2_C1	B3_C1	B1_C4	B2_C4	B3_C4
Round1	-11.70	-22.26	-28.24	-13.19	-19.13	-15.57
Round2	-17.72	-15.87	-14.09	-17.39	-17.37	-21.62
Round3	-16.44	-20.20	-18.10	-19.06	-12.56	-13.61
Round4	-18.42	-21.82	-3.40	-26.99	-15.31	-25.71
Round5	-15.19	-20.16	-12.27	-12.52	-12.33	-9.44
Round6	-14.20	-16.01	-12.27	-12.65	-20.32	-25.86
Round7	-18.40		-25.43	-21.10	-16.88	-10.92
Round8	-5.10		-15.04	-18.69	-10.87	-7.14
AVE	-14.65	-19.39	-16.10	-17.70	-15.60	-16.23
SD	4.20	2.56	7.37	4.66	3.20	6.85

B: eliminate several simulation results to keep the SD of predicted binding free energy within 3.0										
B1_C1		B3_C1		B1_C4		B2_C4		B3_C4	
Round1	-11.70	Round2	-14.09	Round1	-13.19	Round3	-12.56	Round1	-15.57
Round2	-17.72	Round3	-18.10	Round2	-17.39	Round4	-15.31	Round3	-13.61
Round3	-16.44	Round5	-12.27	Round5	-12.52	Round5	-12.33	Round5	-9.44
Round5	-15.19	Round6	-12.27	Round6	-12.65	Round7	-16.88	Round7	-10.92
Round6	-14.20	Round8	-15.04	Round8	-18.69	Round8	-10.87	Round8	-7.14
AVE	-15.05		-14.35		-14.89		-13.59		-11.34
SD	2.05		2.16		2.61		2.18		2.98

Figures

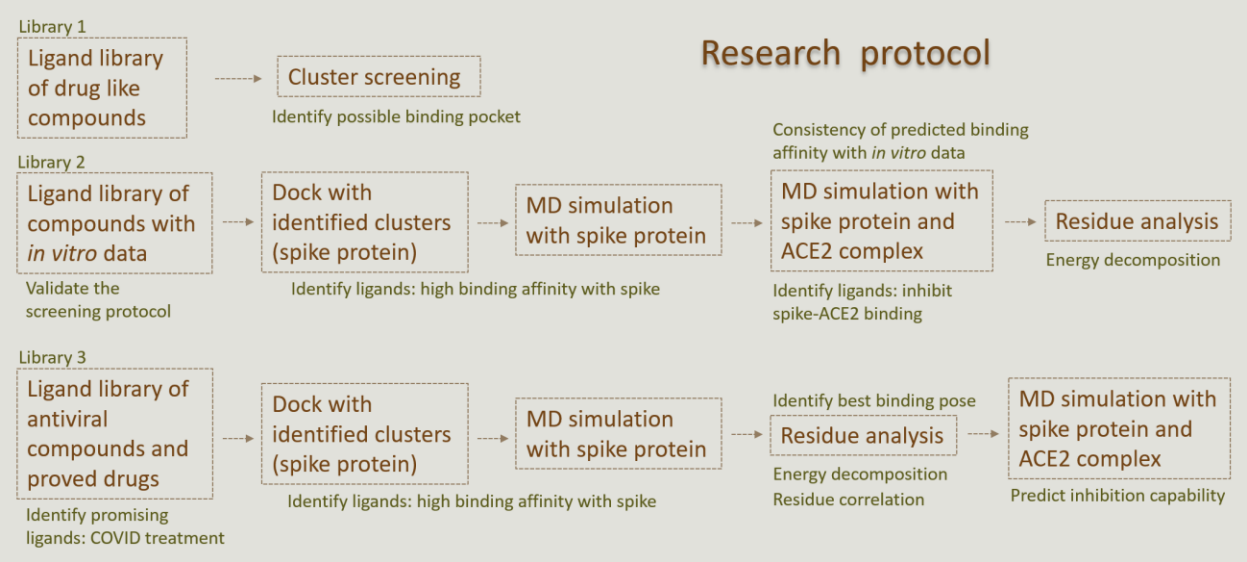


Figure 1. The workflow of the research protocol.

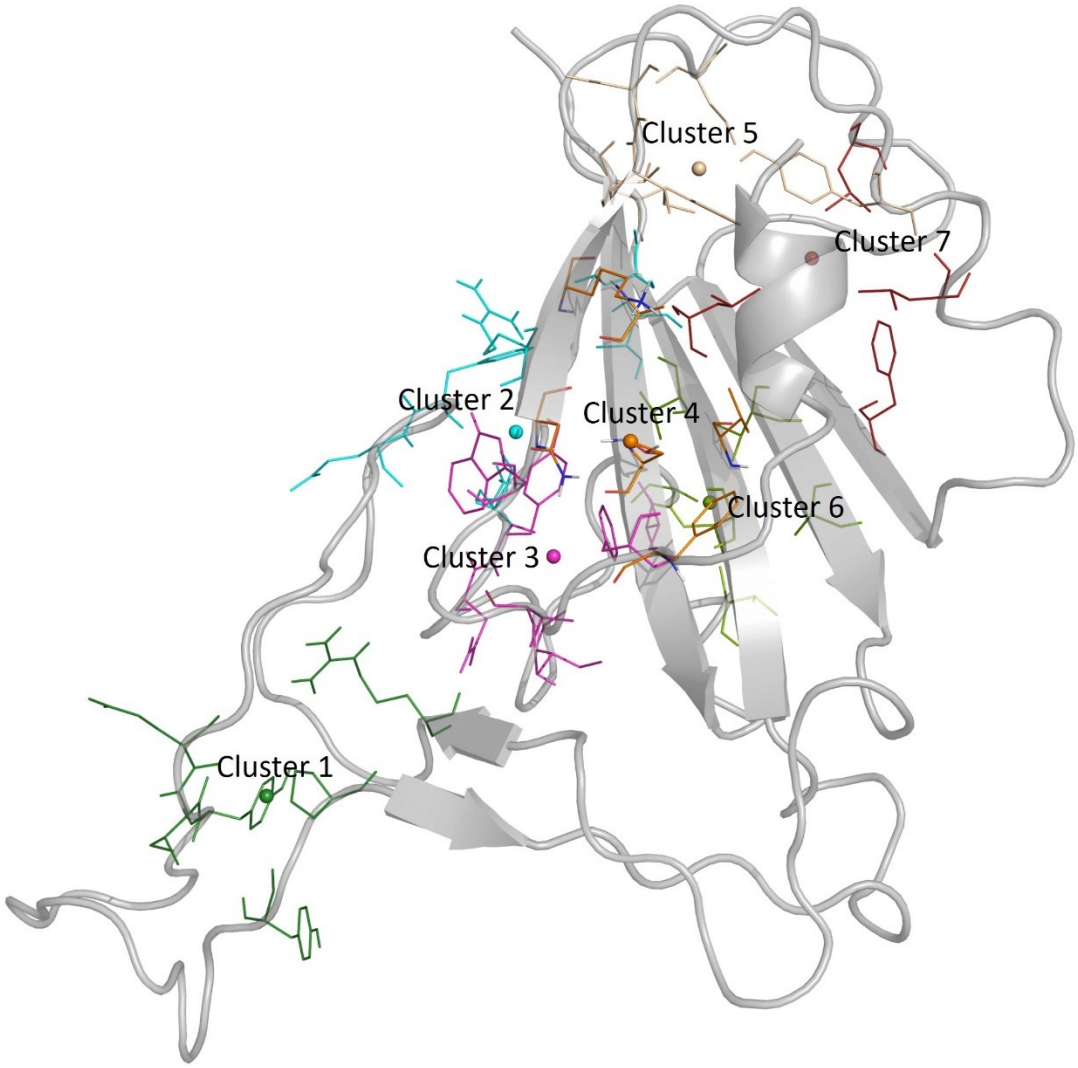


Figure 2. Seven binding sites predicted by Sybyl. The red dots show the center of each binding site.

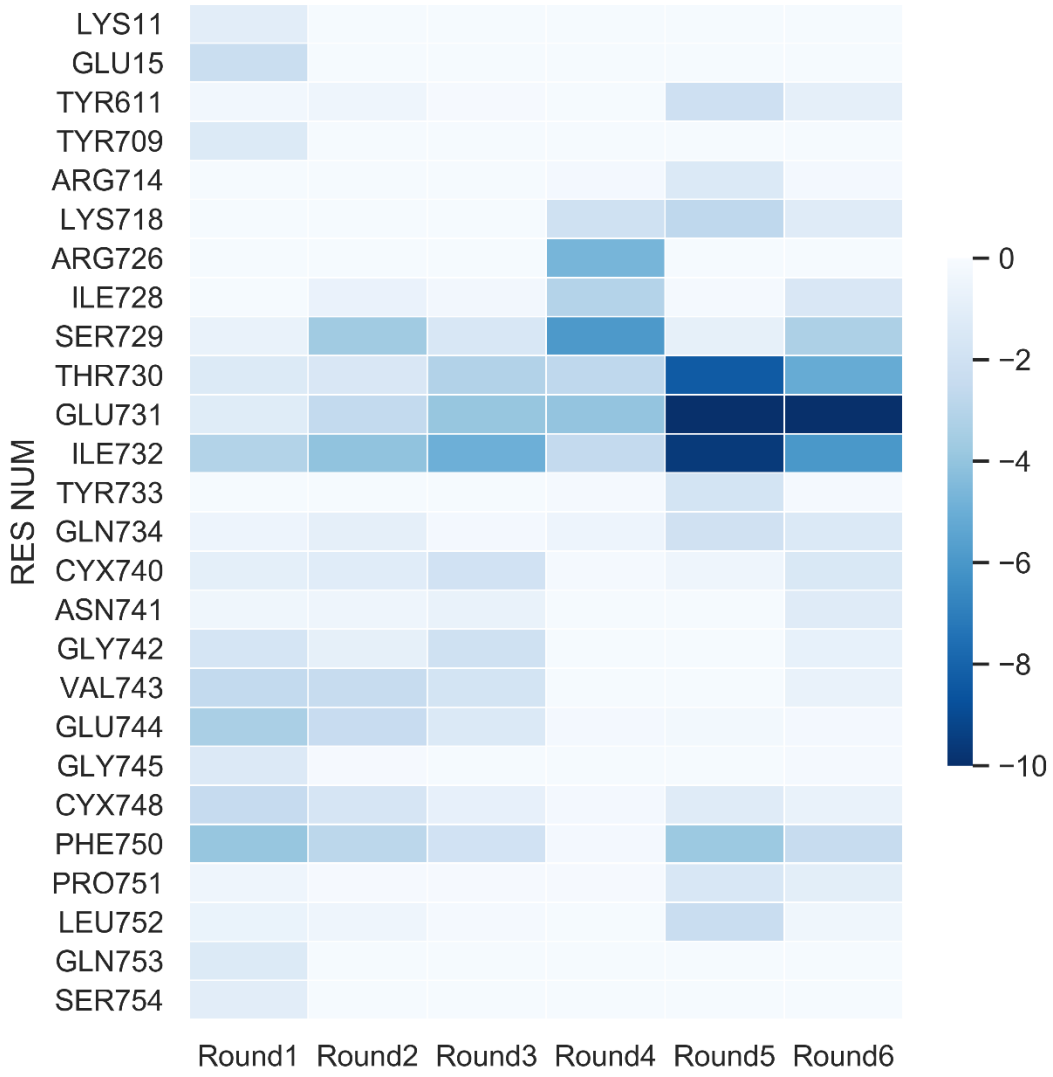


Figure 3. The heatmap of residues from the Spike protein. The residues were selected if their energy contribution exceeded -1 kcal/mol during compound NCGC00389662-01 binding.

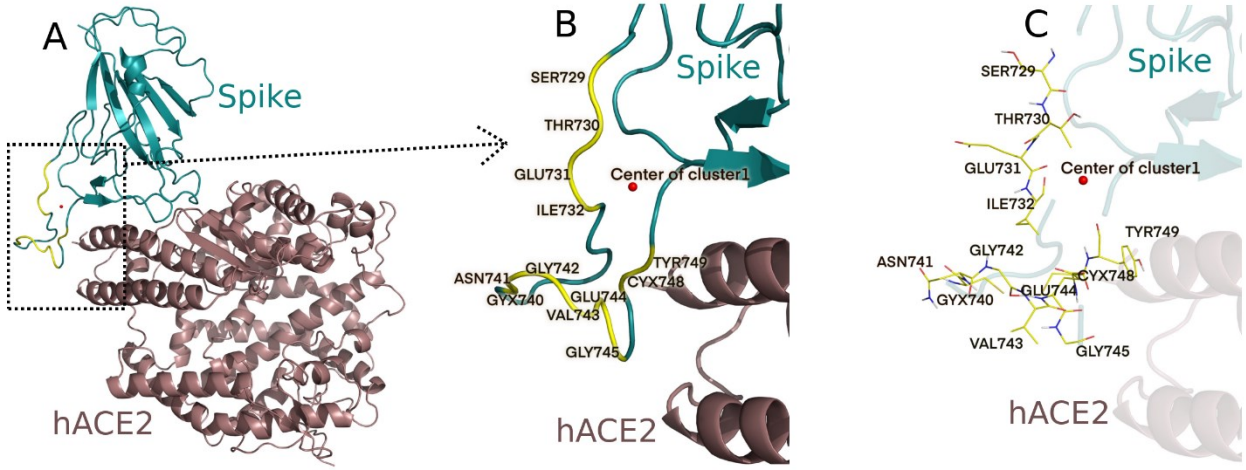


Figure 4. The residues from the Spike protein which interact frequently with compound NCGC00389662-01 are around the Spike-hACE2 binding interface. (A): The overview position of the residues shown in yellow cartoon. (B): A close-up view of the hot residues labeled with residue names. (C): Hotspot residues are shown in yellow lines and labeled with residue names. In all panels, the center of cluster1 is marked as a red dot.

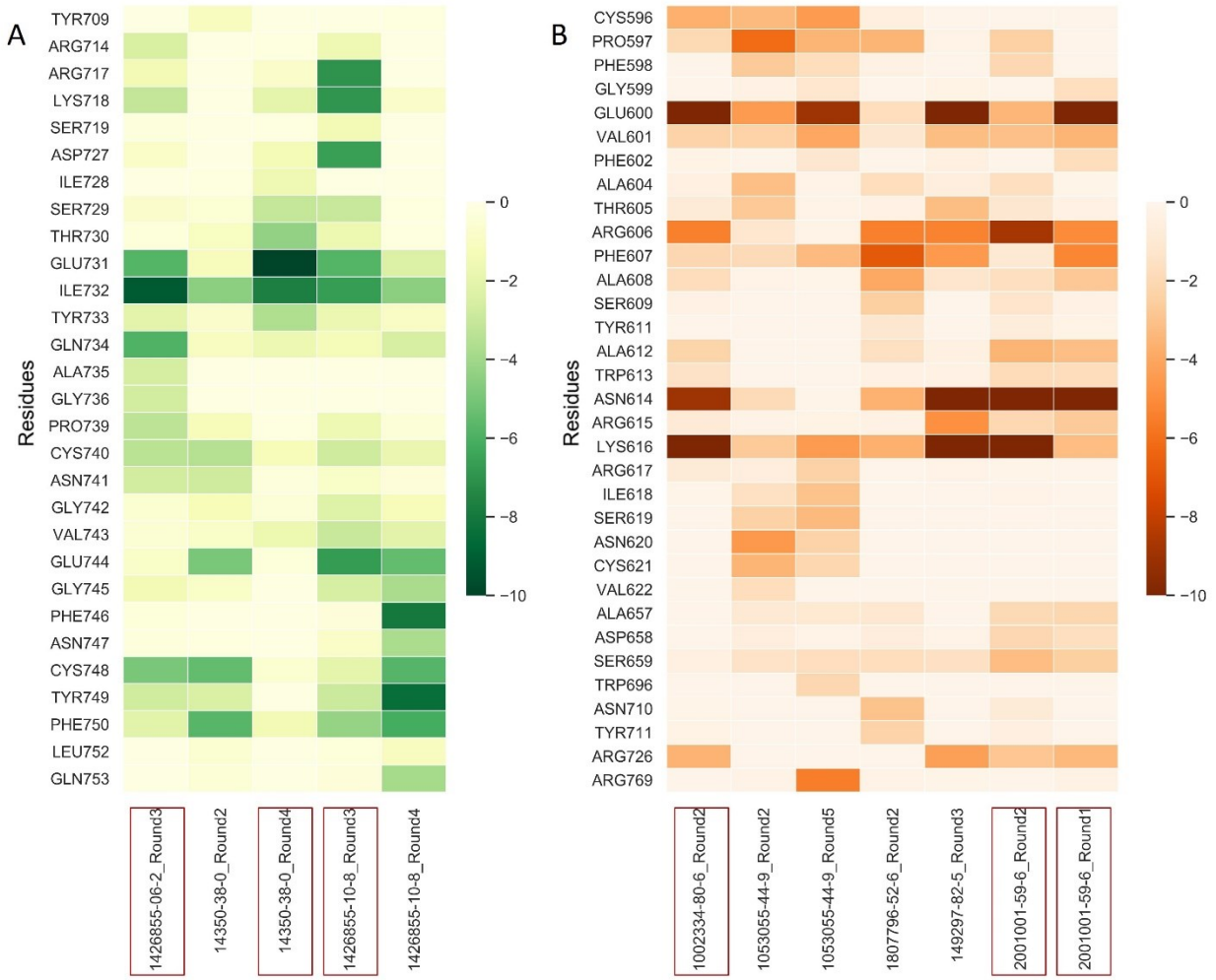


Figure 5. The heatmap of residues from the Spike protein. The residues were selected if their energy contribution exceeded -1 kcal/mol during the binding of selected compounds. Finally selected conformations for compounds are marked in red boxes for next step complex MD simulation with the compounds. (A): residue energy decomposition results for compounds selected in cluster1. (B): residue energy decomposition results for compounds selected in cluster4.

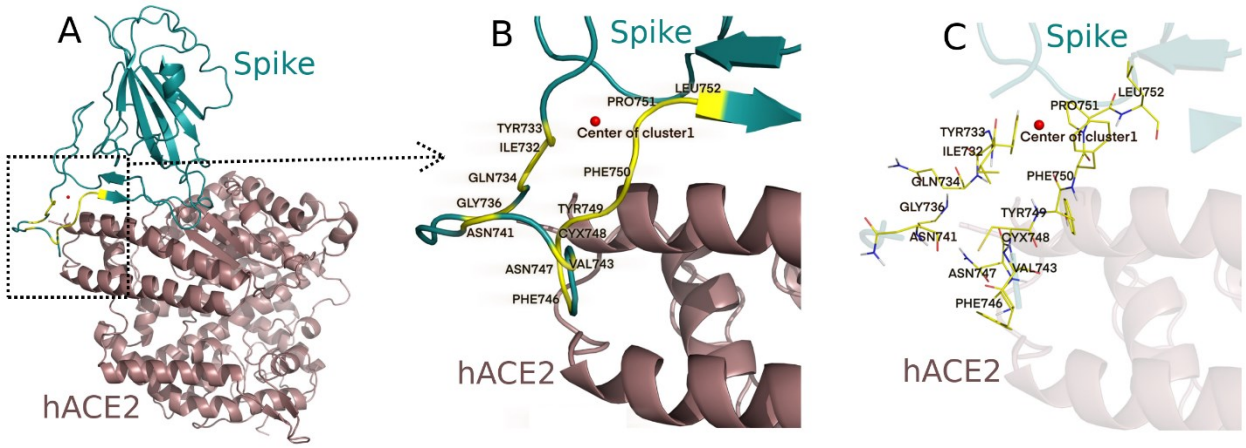


Figure 6. The residues from the Spike protein which interact frequently with selected compounds for Cluster1 and are around the Spike-hACE2 binding interface. (A): The overview position of the residues shown in yellow cartoon. (B): A close-up view of the hot residues labeled with residue names. (C): Hotspot residues are shown in yellow lines and labeled with residue names. In all panels, the center of cluster1 is marked as a red dot.

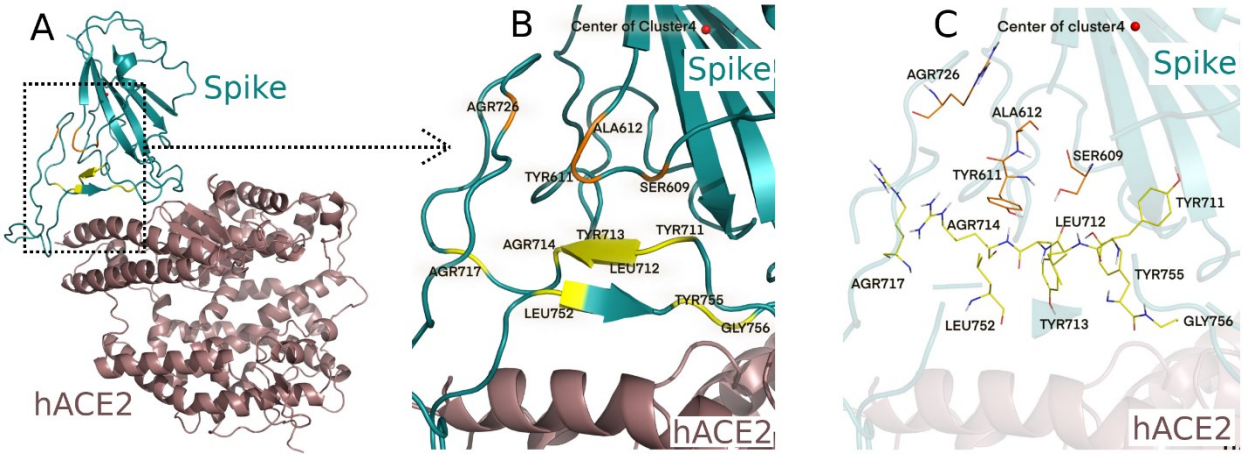


Figure 7. The residues from the Spike protein which interact frequently with selected compounds for Cluster4 and are around the Spike-hACE2 binding interface. (A): The overview position of the residues shown in yellow cartoon. (B): A close-up view of the hot residues labeled with residue names. (C): Hotspot residues are shown in yellow lines and labeled with residue names. In all panels, the center of cluster4 is marked as a red dot.

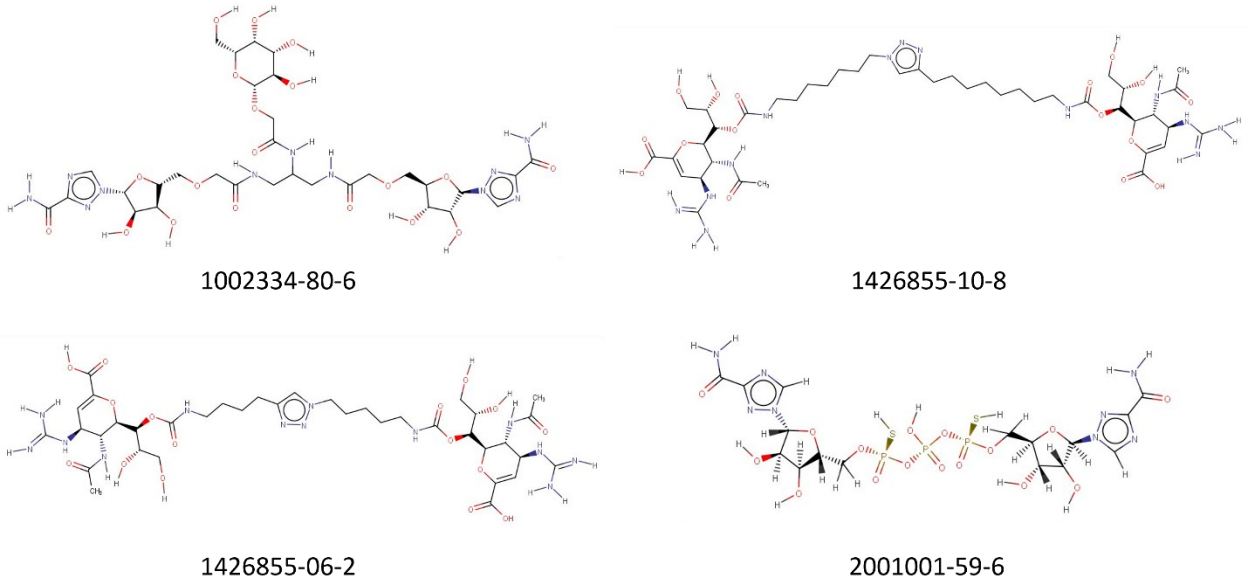
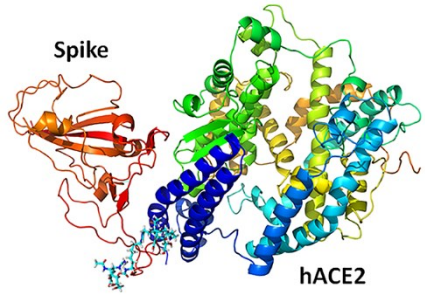


Figure 8. The chemical structures of the 4 identified compounds.



TOC (size 5.6 cm × 4 cm)

Text: An allosteric inhibitor shown as sticks on bottom-left corner could interfere with Spike and hACE2 binding.