Accelerating the Discovery of the beyond Rule of Five Compounds that have High Affinities toward SARS-CoV-2 Spike RBD

Abd Al-Aziz A. Abu-Saleh,^{*,†} Arpita Yadav,[‡] and Raymond A. Poirier^{*,†}

[†]*Department of Chemistry, Memorial University, St. John's, NL A1B 3X7, Canada*

[‡]*Department of Chemistry, University Institute of Engineering & Technology, Chhatrapati Shahu Ji Maharaj University, Kanpur 208024, India*

E-mail: aabusaleh@mun.ca; rpoirier@mun.ca

Abstract

The battle against SARS-CoV-2 coronavirus is the focal point for the global pandemic that has affected millions of lives worldwide. The need for effective and selective therapeutics for the treatment of the disease caused by SARS-CoV-2 is critical. Herein, we performed computational *de novo* design incorporating molecular docking studies, molecular dynamics simulations, absolute binding energy calculations, and steered molecular dynamics simulations for the discovery of potential compounds with high affinity towards SARS-CoV-2 spike RBD. By leveraging ZINC15 database, a total of 1282 in-clinical and FDA approved drugs were filtered out from nearly 0.5 million protomers of relatively large compounds ($MW > 500$, and $LogP \leq 5$). Our results depict plausible mechanistic aspects related to the blockage of SARS-CoV-2 spike RBD by the top hits discovered. We found that the most promising candidates, namely, ZINC95628821, ZINC95617623, and ZINC261494658, strongly bind to the spike RBD

and interfere with the human ACE2 receptor. These findings accelerate the rational design of selective inhibitors targeting the spike RBD protein of SARS-CoV-2.

Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that emerged in late 2019 has caused serious illness and death all over the world. COVID-19, a disease caused by SARS-CoV-2, led the world health organization (WHO) to announce a global pandemic on March 11, 2020. To date, about 110 million infections and more than two million deaths have been reported by the WHO, and the numbers continue to rise.¹ In late 2020, among several vaccine candidates, two vaccines made by Moderna and Pfizer-BioNTech companies were approved by the U.S. food and drug administration (FDA) for an emergency use against COVID-19. Moreover, veklury (remdesivir), an antiviral drug, has also been approved by the U.S. FDA for the use as a treatment for COVID-19 in adults and pediatric patients (12 years of age and older) requiring hospitalization. Although studies reported that remdesivir failed to show clinical benefits for moderately severe COVID-19 patients,^{2,3} it is the only drug that has been approved so far for COVID-19 by the U.S. FDA. Therefore, accelerating the discovery of a safe and an effective COVID-19 medication is a must to control the pandemic.

The coronavirus genome encodes distinct structural and nonstructural proteins. The structural proteins (i.e., the membrane, the envelop, the spike, and the nucleocapsid) are responsible for key functions such as host infection,⁴ membrane fusion,⁵ self-assembly,⁶ release of virus-like particles,⁷ and other functions.⁸ The nonstructural proteins facilitate viral replication-transcription processes.⁹ The extensive amount of ongoing research is incredibly focused on both non-structural and structural proteins of SARS-CoV-2 as drug targets in order to develop effective therapeutics for COVID-19.^{10–13} Among structural proteins, the spike protein is the key machinery that empowers virus entry into the host cell.¹⁴ Structural characterization of the spike protein would give atomic-level information to guide structure-based

drug design. Recently, crystallographic and cryo-EM methods have been utilized to determine various structures and conformational states of the SARS-CoV-2 spike proteins.^{5,14,15} Amaro and coworkers have carried out massive molecular dynamics (MD) simulations to promote and extend the available structural data.¹⁶ These computational results added a detailed insight on the structural and dynamics of the full-length glycosylated SARS-CoV-2 spike protein at an atomic level.¹⁶

Detailed mechanisms of the cell entry of SARS-CoV-2 have been recently investigated.^{5,17,18} The spike protein has two subunits S1 and S2, which are critical for receptor recognition and membrane fusion, respectively. The S1 contains the receptor-binding domain (RBD) of the viral spike that specifically binds to the human angiotensin-converting enzyme 2 (hACE2). The RBD can be in either an up conformation, which is the receptor-accessible state (i.e., enables binding to the host receptor), or a down conformation, which is the receptor-inaccessible state.^{14,17} For membrane fusion, host cell proteases cleave SARS-CoV-2 spikes at the S1/S2 boundary followed by structural changes of the S2 that promote a host cell entry.^{17,19,20} Interestingly, experimental results have revealed that SARS-CoV-2 RBD has stronger hACE2 binding affinity than that possessed by the previous SARS-CoV RBD. However, the entire SARS-CoV-2 spike has hACE2 binding affinity similar to or lower than SARS-CoV-2 spike.^{15,17,21,22} Moreover, a computational study of the RBD -accessible and -inaccessible conformations and their binding strength to the hACE2 is in agreement with the experimental findings.²³ Therefore, the RBD is considered as a key target for designing and developing compounds that suppress the virus entry into the human cell.

Different approaches have been conducted that are based on interfering with the RBD binding to the hACE2 including neutralizing antibodies,²⁴⁻²⁶ miniproteins,²⁷ peptides,²⁸⁻³⁰ and small molecules.³¹⁻³⁵ As far as we are aware, no high-throughput virtual screening has been conducted in relatively large molecules in the “beyond rule of 5” chemical space that have high RBD binding affinity. Herein, we performed a comprehensive computational protocol, which incorporates molecular docking, MD simulations, absolute binding energy cal-

culations, and steered MD simulations for the discovery of relatively large compounds that would bind to SARS-CoV-2 spike RBD very tightly, therefore, blocking viral attachment to the host cell.

Computational Methods

Model System Preparation

The initial structure of SARS-CoV-2 RBD bound to the hACE2 was retrieved from the RCSB protein data bank [PDB entry 6LZG].⁵ The protonation states of the titratable residues were determined using the H++ Web server under the physiological pH.³⁶ Moreover, salinity, internal and external dielectric constants were set to be 0.15, 10.0, and 80.0 respectively.

Molecular Docking

3-D chemical structures ($MW > 500$, and $LogP \leq 5$) were retrieved from ZINC15 database,³⁷ resulting in about 0.5 million protomers. Among these protomers, only in-clinical trials and FDA approved drugs (a total of 1283 protomers) were selected for docking. OpenBabel software³⁸ was used to minimize selected compounds. These compounds were docked into a preprocessed SARS-CoV-2 RBD protein pocket by utilizing the AutoDock Vina.³⁹ For improving the reliability of docking, these compounds were also docked using the BindScope web application.⁴⁰ The xyz grid cell origin of -37.0 Å, 30.0 Å, and 3.5 Å with dimensions of 26.0, 45.5, and 22.0 in the x, y, and z directions, respectively, was built for docking calculations. The grid cell was constructed at key residues located within receptor binding motif of the RBD that are directly involved in interacting with the hACE2. The spike RBD–ligand binding affinities were ranked based on a consensus docking results from two scoring functions Vina³⁹ and 3D-convolutional neural networks method (Kdeep).⁴¹ Consensus docking scores were chosen as the measure of protein–ligand binding affinities to rank the poses of the studied compounds. For the next high-throughput virtual screening, we selected the top

13 compounds that have a consensus docking score of ≤ -38 kJ mol⁻¹ (-9 kcal mol⁻¹). Moreover, we selected one compound from each scoring functions Vina³⁹ and Kdeep algorithm⁴¹ that has the highest protein–ligand binding affinity. All molecular docking results are provided in the supporting information (SI).

Molecular Dynamics (MD) Simulations

All-atom MD simulations were done using the NAMD 2.13 package.⁴² We performed MD simulations on the model systems with top docking results (total of 15 RBD–ligand complexes). Model systems of the RBD–ligand complexes were prepared in an explicit solvent using the TIP3P water model⁴³ in a simulation cell with dimensions of $86 \times 71 \times 96$ Å³. All crystallographically resolved water molecules were removed from model systems. The parameters for the ligands and the spike RBD protein model structures were all set using the CHARMM general force field (CGenFF),⁴⁴ and the CHARMM36 force field,⁴⁵ respectively. Model systems were neutralized using chloride ions. The MD protocols involve minimization, annealing, equilibration, and production. The protein backbone atoms were restrained in the minimization and annealing simulations. The C α atoms of the protein backbone were restrained in the 1 ns equilibration simulations. For the 100 ns MD production simulations, no atoms were restrained. The isothermal-isobaric (NPT) ensemble and a 2 fs time-step was chosen for all MD simulations. The pressure was set at 1 atm using the Nosé-Hoover Langevin piston barostat^{46,47} with the Langevin piston decay of 0.2 ps and a period of 0.4 ps. The temperature was set at 298.15 K using the Langevin thermostat⁴⁸ and a damping frequency of 1 ps⁻¹. A cutoff distance of 10.0 Å was applied for Lennard-Jones interactions with a pair list distance of 12 Å, interactions were smoothly truncated at 8.0 Å. The particle-mesh Ewald (PME) method was utilized to treat the long-range electrostatic interactions,⁴⁹ where a grid spacing of 1.0 Å was used for all simulation cells. Covalent bonds involving hydrogen atoms were constrained using the SHAKE algorithm.⁵⁰ See the SI for more information about the MD simulations protocol.

Absolute Binding Energy Calculations

One of the main goal in computer-aided drug discovery is the accurate and reliable prediction of the absolute binding energy of receptor-ligand systems. Here, we calculated the absolute binding energies of the stable RBD–ligand complexes that resulted from the MD simulations. Starting from the equilibrated RBD–ligand structures, we conducted the absolute binding energy calculations using the potential of mean force (PMF) approach as described by Chipot and coworkers.^{51,52} It has been reported that the PMF approach accurately predicts protein–ligand binding affinities.^{51,53} A set of geometrical restraints on collective variables⁵⁴ are exploited to accurately determine the translational, rotational, and conformational entropies that accompany the binding process and, thus, computing protein–ligand absolute binding energies. These restraints include the root mean-square deviation (RMSD) of the ligand in the bound (site) and unbound (bulk) states compared to an equilibrated reference structure of the RBD–ligand complex. Additional restraints include three Euler (Θ, Φ, Ψ) and two spherical (θ, φ) angles that describe the relative orientation and position of the ligand with respect to the RBD protein, respectively, see Figure S3 in the SI. The equilibrium binding constant (K_{eq}) can be determined as,

$$K_{eq} = \exp \left[-\beta \left(\Delta G_c^{site} + \Delta G_c^{bulk} + \Delta G_o^{site} + \Delta G_a^{site} + \Delta G_o^{bulk} - \frac{1}{\beta} \ln (S^* I^* C^o) \right) \right] \quad (1)$$

where the ΔG_c^{site} and ΔG_c^{bulk} denote the Gibbs energy changes of the conformational ligand and RMSD in the site and bulk states, respectively, the ΔG_o^{site} and ΔG_a^{site} represent the Gibbs energy changes associated with the orientational and the positional of the ligand in the binding site, respectively, and the ΔG_o^{bulk} represents the Gibbs energy change associated with the ligand orientation in the bulk. The last term in Eq. (1) represents the separation of the ligand from the binding site into the bulk. It should be mentioned that the ΔG_o^{bulk} is calculated analytically as reported by Chipot and coworkers.⁵¹ More details on theoretical background of the binding energy calculations using the PMF route can be obtained

elsewhere.⁵¹ Eventually, the standard binding energy is

$$\Delta G_{bind}^{\circ} = k_B T \ln K_{eq} C^{\circ} \quad (2)$$

where the k_B and T are the Boltzmann constant and simulation temperature, respectively, and $C^{\circ} = 1/1661 \text{ \AA}^3$ is the standard one molar concentration. For each model system, the RBD–ligand separation, RMSD (site), RMSD (bulk), and angular course variables were run for 50 ns, 50 ns, 100 ns, and 50 ns, respectively. By leveraging GPU-accelerated NAMD engine, the simulations were performed in triplicate, yielding an aggregate total of 750 ns simulation time for each model system. Absolute binding energies reported in this study are the averages of the triplicate runs. See the SI for more details about the absolute binding energy protocol.

Steered Molecular Dynamics (SMD) Simulations

To probe the selectivity and the potency of the top hits binding to the spike RBD protein, SMD simulations of the RBD–ligand–hACE2 model systems were conducted. The RBD–ligand–hACE2 model systems were solvated using a TIP3P water model with a 12 $\text{\AA}$ buffer around the system and then extended 30.0 $\text{\AA}$ in the SMD pulling axis, resulting in a simulation cell with dimensions of $121 \times 121 \times 167 \text{ \AA}^3$. The SMD simulations were employed by harmonically restraining the position of the spike RBD protein center-of-mass, and moving the second restraint center-of-mass of the hACE2 protein, with a constant pulling speed of $1.5 \text{ \AA ns}^{-1}$ in the z-axis direction. More details about the SMD protocol are provided in the SI.

Results and discussion

We performed an integrated computational strategy including molecular docking, MD simulations, absolute binding energy calculations, and SMD simulations to accelerate the identi-

fication of promising drug candidates for COVID-19. 1283 compounds (in-clinical trials and FDA approved drugs) were retrieved from the ZINC15 database³⁷ and advanced to flexible molecular docking using both AutoDock Vina³⁹ and Kdeep⁴¹ scoring functions. Consensus docking scores were chosen as an initial measure of the RBD-ligand binding affinities to rank the poses of the screened compounds. Docking results of the top 15 compounds are listed in Table 1. The two-dimensional structures of these compounds are also provided in the SI (Figure S1). Docking results of all screened compounds are provided in the SI.

Table 1: Summary of top consensus docking results in kJ mol^{-1} for SARS-CoV-2 spike RBD target

Compound ID	Vina score	Kdeep score	Consensus score
ZINC6717782	-38	-40	-39
ZINC95628821	-36	-41	-39
ZINC253638647	-35	-43	-38
ZINC3979524	-36	-41	-38
ZINC253387884	-35	-41	-38
ZINC255977094	-36	-40	-38
ZINC150339052	-37	-39	-38
ZINC261494659	-35	-41	-38
ZINC936069565	-34	-42	-38
ZINC51951669	-37	-38	-38
ZINC49637509	-32	-43	-38
ZINC261494658	-33	-42	-38
ZINC95617623	-32	-43	-38
ZINC261494590 ^a	-25	-47	-36
ZINC3978005 ^b	-40	-23	-32

^a The compound was selected based on the Kdeep scoring function; ^b The compound was selected based on the Vina scoring function.

The best 15 RBD-ligand complexes from molecular docking were used as starting structures for the conventional MD simulations. Ensemble MD simulations consolidate the effect of solvation, flexibility of both ligand and protein target, and better estimate of protein-ligand nonbonded interactions. The structure stability of the RBD-ligand model systems were evaluated by the RMSD with respect to the equilibrated structures (see Figure S2 in the SI). The time-evolution of the RMSD show that most ligands with invalid binding modes reveal significant signs of binding instability in the RBD binding pocket. The use of abso-

lute binding energy calculations is crucial to enhance the hit identification. Consequently, rigorous absolute binding energy calculations were only conducted on the stable RBD-ligand systems resulting from the MD simulations. In absolute binding energy calculations, the resulting contributions from the aforementioned restraints are computed sequentially using the extended adaptive biasing force (eABF) method⁵⁵ with the unbiased corrected z-averaged restraint (CZAR) estimator⁵⁶ implemented in the Colvars module⁵⁴ of NAMD. The contributions from the course variables and the final standard binding energy (ΔG_{bind}°) of ligands to the SARS-CoV-2 spike RBD protein are provided in Table 2.

Table 2: The calculated contributions accompanying the RBD–ligand binding process and the predicted ΔG_{bind}° in kJ mol^{-1}

Protein-ligand complex	ΔG_c^{site}	$\Delta G_{angular}^{site}$	$-\frac{1}{\beta} \ln (S^* I^* C^{\circ})$	ΔG_c^{bulk}	ΔG_{bind}°
RBD–ZINC95628821	-73 ± 6	-5 ± 0	-44 ± 6	72 ± 3	-23 ± 3
RBD–ZINC3979524	-22 ± 4	-5 ± 0	-38 ± 2	22 ± 2	-12 ± 3
RBD–ZINC150339052	-41 ± 5	-9 ± 3	-12 ± 3	38 ± 3	$+3 \pm 2$
RBD–ZINC261494658	-33 ± 2	-7 ± 0	-24 ± 3	29 ± 1	-8 ± 2
RBD–ZINC95617623	-38 ± 4	-9 ± 1	-26 ± 4	27 ± 1	-19 ± 1
RBD–ZINC3978005	-21 ± 5	-9 ± 1	-12 ± 1	19 ± 3	$+4 \pm 4$

Unambiguously, the four ligands, namely, ZINC95628821, ZINC95617623, ZINC3979524, and ZINC261494658 bind spontaneously to the SARS-CoV-2 spike RBD with ΔG_{bind}° of -23 , -19 , -12 , and -8 , respectively. In contrast, the two ligands, namely, ZINC150339052 and ZINC3978005 have non-spontaneous binding behavior toward the RBD with $\Delta G_{bind}^{\circ} = 3$ and 4 kJ mol^{-1} , respectively. These values represent the non-covalent interaction energies between the ligands and the spike RBD and give a good estimate of their binding affinities.

To validate the potency of the RBD–ligand binding affinities, we need to investigate if the top hits are able to interfere with the RBD–hACE2 binding event. For this purpose, SMD simulations were employed for distinguishing active from decoy ligands. Both strong and weak ligands listed in Table 2 were considered for SMD simulations as a validation strategy. Figure 1 shows the final results of 20 ns SMD simulations for the RBD–hACE2 bounded and unbounded to ligands.

Figure 1: (a) The structure of the apo RBD-hACE2 (left) along with its time-averaged SMD results (right); (b) the structures of the RBD-ligand-hACE2 (top) along with their time-averaged SMD results (bottom). All structures were depicted at the 20 ns SMD simulations. The RBD, the hACE2, and the ligands are shown as red cartoon, cyan cartoon, spheres, respectively.

The time-evolution force profile of the unbinding event of the RBD–ACE2 model system has a maximum rupture force of 704 pN, in a good agreement with the computed maximum rupture force of 751 pN reported recently.⁵⁷ It is well established that the rupture force is directly proportional with the receptor-ligand binding affinity.⁵⁸ The moving average (black line) of the time-evolution force profiles in Figure 1 indicates that ligands (particularly ZINC95628821, ZINC3979524, and ZINC95617623) are disrupting the RBD-hACE2 binding by lowering the rupture force. Additionally, the structures in Figure 1b show that ligands, namely, ZINC95628821, ZINC261494658, and ZINC95617623 are strong binders to the RBD protein. This is anticipated due to the fact that these ligands have relatively better binding affinity to the RBD than other ligands, namely, ZINC150339052, and ZINC3978005. On the contrary, during the SMD simulations, ZINC3979524 and ZINC3978005 left the RBD and attached to the hACE2 receptor, revealing that these ligands have more favourable interactions to the hACE2 than the RBD. SMD simulations also revealed that ZINC150339052 has non promising binding affinity toward neither the RBD nor the hACE2 proteins (see Figure 1b).

Detailed analyses of the ligand interaction diagram of the equilibrated RBD–ligand–hACE2 complexes are shown in Figure 2. Ligands that exhibit robust RBD binding stability during the SMD simulations, namely, ZINC95628821, ZINC261494658, and ZINC95617623, indeed form more non-covalent interactions (i.e. hydrogen bonding, pi-cation, pi-pi stacking) with key amino acid residues of the RBD and enjoy more exposure to the RBD binding pose than other ligands. Notably, these robust binders have direct polar contact with ARG403 of the RBD. Moreover, these ligands also interfere with key residues of the hACE2 protein (i.e. ASP38, TYR41, and LYS353). Interestingly, ARG403 of the RBD and residues ASP38, TYR41, and LYS353 of the hACE2 were reported as true hot spots that contribute to the stability of the RBD–hACE2 interface.⁵⁹ Wang et al reported that hydrophobic interactions also play a critical role in anchoring the RBD to the hACE2 receptor.⁶⁰ All ligands that possess spontaneous standard binding energies have at least four direct contacts with the RBD

Figure 2: Representative poses (left) detailed ligand interaction diagrams (right) of the RBD–ligand–ACE2 complexes. (a) ZINC95628821, (b) ZINC3979524, (c) ZINC150339052, (d) ZINC261494658, (e) ZINC95617623, (f) ZINC3978005. In the ligand interaction diagram, the hACE2 and the RBD residues were labeled A and B, respectively.

hydrophobic residues (see Figure 2). The polar and non-polar interactions of the RBD-ligand complexes contribute to their total standard binding energy. Based on rigorous binding energy calculations and SMD simulations, we suggest that the top binders are ZINC95628821 followed by ZINC95617623 and ZINC261494658. These top hits certify critical interactions associated with the RBD-ligand-hACE2 binding, which in turn disrupt the RBD-hACE2 binding event.

Conclusions

In this work, we performed high-throughput virtual screening that comprise of molecular docking, MD simulations, non-covalent absolute binding energy calculations, and SMD simulations for the discovery of SARS-CoV-2 spike RBD-based inhibitors. Vina³⁹ and Kdeep⁴¹ scoring functions were leveraged for structure-based screening and ranking of in-clinical trials and FDA approved drugs. Absolute binding energy calculations in combinations with the SMD simulations provided more effective ranking procedures of the top hits. Based on integrated *in silico* approach, our findings suggest that ZINC95628821, ZINC95617623, and ZINC261494658 have potential antiviral effects by blocking the RBD-hACE2 recognition, and thus, need to be considered for further *in vitro* and *in vivo* investigation. This study provides a rational drug design of RBD-based inhibitors for accelerating the research and development of effective COVID-19 medications.

Acknowledgement

The authors thank Compute Canada for computer time. R.A.P. is grateful to the Natural Sciences and Engineering Council of Canada (NSERC) for financial support. A.A.A thanks the School of Graduate Studies of Memorial University for funding. A.Y. thanks Science and Engineering Research Board, New Delhi for research grant number EMR/2016/000769. A.Y is also grateful to Shastri Indo Canadian Institute, New Delhi for Shastri Scholar Travel

Subsidy Grant in the year 2016.

Supporting Information Available

Protocols and parameters used for MD simulations, absolute binding energy calculations, and SMD simulations.

Notes

The authors declare no competing financial interest.

References

- (1) *COVID-19 weekly epidemiological update*; World Health Organization, 23 February 2021.
- (2) Wang, Y.; Zhang, D.; Du, G.; Du, R.; Zhao, J.; Jin, Y.; Fu, S.; Gao, L.; Cheng, Z.; Lu, Q., et al. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. *The Lancet* **2020**, *395*, 1569–1578.
- (3) Young, B.; Tan, T. T.; Leo, Y. S. The place for remdesivir in COVID-19 treatment. *The Lancet. Infectious Diseases* **2021**, *21*, 20.
- (4) Hossain, M. G.; Akter, S.; Saha, S. SARS-CoV-2 host diversity: An update of natural infections and experimental evidences. *Journal of Microbiology, Immunology and Infection* **2020**,
- (5) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K.-Y., et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell* **2020**, *181*, 894–904.

- (6) Ye, Q.; West, A. M.; Silletti, S.; Corbett, K. D. Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. *Protein Science* **2020**, *29*, 1890–1901.
- (7) Siu, Y.; Teoh, K.; Lo, J.; Chan, C.; Kien, F.; Escriou, N.; Tsao, S.; Nicholls, J.; Altmeyer, R.; Peiris, J., et al. The M, E, and N structural proteins of the severe acute respiratory syndrome coronavirus are required for efficient assembly, trafficking, and release of virus-like particles. *Journal of virology* **2008**, *82*, 11318–11330.
- (8) Mittal, A.; Manjunath, K.; Ranjan, R. K.; Kaushik, S.; Kumar, S.; Verma, V. COVID-19 pandemic: Insights into structure, function, and hACE2 receptor recognition by SARS-CoV-2. *PLoS pathogens* **2020**, *16*, e1008762.
- (9) Haque, S. M.; Ashwaq, O.; Sarief, A.; Azad John Mohamed, A. K. A comprehensive review about SARS-CoV-2. *Future virology* **2020**, *15*, 625–648.
- (10) Izda, V.; Jeffries, M. A.; Sawalha, A. H. COVID-19: A review of therapeutic strategies and vaccine candidates. *Clinical Immunology* **2020**, 108634.
- (11) Liu, X.-H.; Zhang, X.; Lu, Z.-H.; Zhu, Y. S.; Wang, T. Potential molecular targets of nonstructural proteins for the development of antiviral drugs against SARS-CoV-2 infection. *Biomedicine & Pharmacotherapy* **2020**, 111035.
- (12) Mulholland, A. J.; Amaro, R. E. COVID19-Computational chemists meet the moment. **2020**, *60*, 5724–5726.
- (13) Abu-Saleh, A. A.-A. A.; Awad, I. E.; Yadav, A.; Poirier, R. A. Discovery of potent inhibitors for SARS-CoV-2's main protease by ligand-based/structure-based virtual screening, MD simulations, and binding energy calculations. *Physical Chemistry Chemical Physics* **2020**, *22*, 23099–23106.
- (14) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C.-L.; Abiona, O.;

- Graham, B. S.; McLellan, J. S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367*, 1260–1263.
- (15) Walls, A. C.; Park, Y.-J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veerler, D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **2020**, *181*, 281–292.
- (16) Casalino, L.; Gaieb, Z.; Goldsmith, J. A.; Hjorth, C. K.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; McLellan, J. S., et al. Beyond shielding: the roles of glycans in the SARS-CoV-2 spike protein. *ACS Central Science* **2020**, *6*, 1722–1734.
- (17) Shang, J.; Wan, Y.; Luo, C.; Ye, G.; Geng, Q.; Auerbach, A.; Li, F. Cell entry mechanisms of SARS-CoV-2. *Proceedings of the National Academy of Sciences* **2020**, *117*, 11727–11734.
- (18) Tang, T.; Jaimes, J. A.; Bidon, M. K.; Straus, M. R.; Daniel, S.; Whittaker, G. R. Proteolytic activation of SARS-CoV-2 spike at the S1/S2 boundary: potential role of proteases beyond furin. *ACS Infectious Diseases* **2021**, *7*, 264–272.
- (19) Letko, M.; Marzi, A.; Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nature microbiology* **2020**, *5*, 562–569.
- (20) Borkotoky, S.; Dey, D.; Banerjee, M. Computational insight into the mechanism of SARS-CoV-2 membrane fusion. *Journal of Chemical Information and Modeling* **2021**, *61*, 423–431.
- (21) Ou, X.; Liu, Y.; Lei, X.; Li, P.; Mi, D.; Ren, L.; Guo, L.; Guo, R.; Chen, T.; Hu, J., et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nature communications* **2020**, *11*, 1–12.

- (22) Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020**, *581*, 221–224.
- (23) Peng, C.; Zhu, Z.; Shi, Y.; Wang, X.; Mu, K.; Yang, Y.; Zhang, X.; Xu, Z.; Zhu, W. Computational Insights into the Conformational Accessibility and Binding Strength of SARS-CoV-2 Spike Protein to Human Angiotensin-Converting Enzyme 2. *The journal of physical chemistry letters* **2020**, *11*, 10482–10488.
- (24) Piccoli, L.; Park, Y.-J.; Tortorici, M. A.; Czudnochowski, N.; Walls, A. C.; Beltramello, M.; Silacci-Fregni, C.; Pinto, D.; Rosen, L. E.; Bowen, J. E., et al. Mapping neutralizing and immunodominant sites on the SARS-CoV-2 spike receptor-binding domain by structure-guided high-resolution serology. *Cell* **2020**, *183*, 1024–1042.
- (25) Huo, J.; Zhao, Y.; Ren, J.; Zhou, D.; Duyvesteyn, H. M.; Ginn, H. M.; Carrique, L.; Malinauskas, T.; Ruza, R. R.; Shah, P. N., et al. Neutralization of SARS-CoV-2 by destruction of the prefusion spike. *Cell host & microbe* **2020**, *28*, 445–454.
- (26) Wu, N. C.; Yuan, M.; Liu, H.; Lee, C.-C. D.; Zhu, X.; Bangaru, S.; Torres, J. L.; Caniels, T. G.; Brouwer, P. J.; Van Gils, M. J., et al. An alternative binding mode of IGHV3-53 antibodies to the SARS-CoV-2 receptor binding domain. *Cell reports* **2020**, *33*, 108274.
- (27) Cao, L.; Goresnik, I.; Coventry, B.; Case, J. B.; Miller, L.; Kozodoy, L.; Chen, R. E.; Carter, L.; Walls, A. C.; Park, Y.-J., et al. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **2020**, *370*, 426–431.
- (28) Schütz, D.; Ruiz-Blanco, Y. B.; Münch, J.; Kirchhoff, F.; Sanchez-Garcia, E.; Müller, J. A. Peptide and peptide-based inhibitors of SARS-CoV-2 entry. *Advanced drug delivery reviews* **2020**, *167*, 47–65.
- (29) Pomplun, S.; Jbara, M.; Quartararo, A. J.; Zhang, G.; Brown, J. S.; Lee, Y.-C.; Ye, X.;

- Hanna, S.; Pentelute, B. L. De novo discovery of high-affinity peptide binders for the SARS-CoV-2 spike protein. *ACS Central Science* **2021**, *7*, 156–163.
- (30) Chowdhury, S. M.; Talukder, S. A.; Khan, A. M.; Afrin, N.; Ali, M. A.; Islam, R.; Parves, R.; Al Mamun, A.; Sufian, M. A.; Hossain, M. N., et al. Antiviral peptides as promising therapeutics against SARS-CoV-2. *The Journal of Physical Chemistry B* **2020**, *124*, 9785–9792.
- (31) Toelzer, C.; Gupta, K.; Yadav, S. K.; Borucu, U.; Davidson, A. D.; Williamson, M. K.; Shoemark, D. K.; Garzoni, F.; Stauffer, O.; Milligan, R., et al. Free fatty acid binding pocket in the locked structure of SARS-CoV-2 spike protein. *Science* **2020**, *370*, 725–730.
- (32) Wang, X.; Cao, R.; Zhang, H.; Liu, J.; Xu, M.; Hu, H.; Li, Y.; Zhao, L.; Li, W.; Sun, X., et al. The anti-influenza virus drug, arbidol is an efficient inhibitor of SARS-CoV-2 in vitro. *Cell discovery* **2020**, *6*, 1–5.
- (33) Awad, I. E.; Abu-Saleh, A. A.-A. A.; Sharma, S.; Yadav, A.; Poirier, R. A. High-throughput virtual screening of drug databanks for potential inhibitors of SARS-CoV-2 spike glycoprotein. *Journal of Biomolecular Structure and Dynamics* **2020**, 1–14.
- (34) Pomplun, S. Targeting the SARS-CoV-2-spike protein: from antibodies to miniproteins and peptides. *RSC Medicinal Chemistry* **2021**,
- (35) Trezza, A.; Iovinelli, D.; Santucci, A.; Prischi, F.; Spiga, O. An integrated drug repurposing strategy for the rapid identification of potential SARS-CoV-2 viral inhibitors. *Scientific reports* **2020**, *10*, 1–8.
- (36) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: automating p K prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic acids research* **2012**, *40*, W537–W541.

- (37) Sterling, T.; Irwin, J. J. ZINC 15–ligand discovery for everyone. *Journal of chemical information and modeling* **2015**, *55*, 2324–2337.
- (38) O’Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *Journal of cheminformatics* **2011**, *3*, 33.
- (39) Trott, O.; Olson, A. J. Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem* **2010**, *31*, 455–461.
- (40) Skalic, M.; Martínez-Rosell, G.; Jiménez, J.; De Fabritiis, G. PlayMolecule BindScope: large scale CNN-based virtual screening on the web. *Bioinformatics* **2019**, *35*, 1237–1238.
- (41) Jiménez, J.; Skalic, M.; Martinez-Rosell, G.; De Fabritiis, G. K deep: protein–ligand absolute binding affinity prediction via 3d-convolutional neural networks. *Journal of chemical information and modeling* **2018**, *58*, 287–296.
- (42) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *Journal of computational chemistry* **2005**, *26*, 1781–1802.
- (43) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics* **1983**, *79*, 926–935.
- (44) Yu, W.; He, X.; Vanommeslaeghe, K.; MacKerell Jr, A. D. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *Journal of computational chemistry* **2012**, *33*, 2451–2468.

- (45) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; MacKerell Jr, A. D. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ_1 and χ_2 dihedral angles. *Journal of chemical theory and computation* **2012**, *8*, 3257–3273.
- (46) Nosé, S.; Klein, M. Constant pressure molecular dynamics for molecular systems. *Molecular Physics* **1983**, *50*, 1055–1076.
- (47) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Molecular physics* **1984**, *52*, 255–268.
- (48) Grest, G. S.; Kremer, K. Molecular dynamics simulation for polymers in the presence of a heat bath. *Physical Review A* **1986**, *33*, 3628.
- (49) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *The Journal of chemical physics* **1995**, *103*, 8577–8593.
- (50) 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. *Journal of computational physics* **1977**, *23*, 327–341.
- (51) Fu, H.; Cai, W.; Hénin, J.; Roux, B.; Chipot, C. New coarse variables for the accurate determination of standard binding free energies. *Journal of chemical theory and computation* **2017**, *13*, 5173–5178.
- (52) Fu, H.; Gumbart, J. C.; Chen, H.; Shao, X.; Cai, W.; Chipot, C. BFEE: A user-friendly graphical interface facilitating absolute binding free-energy calculations. *Journal of chemical information and modeling* **2018**, *58*, 556–560.
- (53) Gumbart, J. C.; Roux, B.; Chipot, C. Standard binding free energies from computer

- simulations: What is the best strategy? *Journal of chemical theory and computation* **2013**, *9*, 794–802.
- (54) Fiorin, G.; Klein, M. L.; Hénin, J. Using collective variables to drive molecular dynamics simulations. *Molecular Physics* **2013**, *111*, 3345–3362.
- (55) Fu, H.; Shao, X.; Chipot, C.; Cai, W. Extended adaptive biasing force algorithm. An on-the-fly implementation for accurate free-energy calculations. *Journal of chemical theory and computation* **2016**, *12*, 3506–3513.
- (56) Lesage, A.; Lelievre, T.; Stoltz, G.; Hénin, J. Smoothed biasing forces yield unbiased free energies with the extended-system adaptive biasing force method. *The Journal of Physical Chemistry B* **2017**, *121*, 3676–3685.
- (57) Nguyen, H. L.; Lan, P. D.; Thai, N. Q.; Nissley, D. A.; O’Brien, E. P.; Li, M. S. Does SARS-CoV-2 bind to human ACE2 more strongly than does SARS-CoV? *The Journal of Physical Chemistry B* **2020**, *124*, 7336–7347.
- (58) Vuong, Q. V.; Nguyen, T. T.; Li, M. S. A new method for navigating optimal direction for pulling ligand from binding pocket: application to ranking binding affinity by steered molecular dynamics. *Journal of chemical information and modeling* **2015**, *55*, 2731–2738.
- (59) Laurini, E.; Marson, D.; Aulic, S.; Fermeglia, M.; Prisl, S. Computational alanine scanning and structural analysis of the SARS-CoV-2 spike protein/angiotensin-converting enzyme 2 complex. *ACS nano* **2020**, *14*, 11821–11830.
- (60) Wang, Y.; Liu, M.; Gao, J. Enhanced receptor binding of SARS-CoV-2 through networks of hydrogen-bonding and hydrophobic interactions. *Proceedings of the National Academy of Sciences* **2020**, *117*, 13967–13974.

Graphical TOC Entry

The representative pose of the top drug candidate, ZINC95628821, interfering with SARS-CoV-2 spike RBD-hACE2 recognition. The spike RBD and the hACE2 receptor are shown as red and cyan cartoons, respectively.