JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

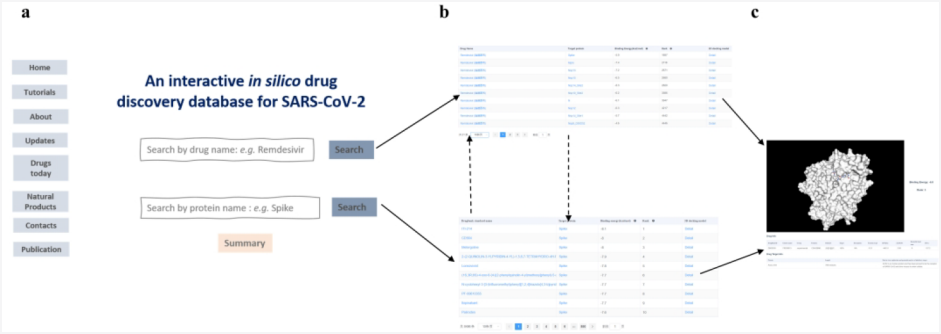
pubs.acs.org/jcim Article

Systemic *In Silico* Screening in Drug Discovery for Coronavirus Disease (COVID-19) with an Online Interactive Web Server

Chi Xu,[○] Zunhui Ke,[○] Chuandong Liu,[○] Zhihao Wang,[○] Denghui Liu, Lei Zhang, Jingning Wang, Wenjun He, Zhimeng Xu, Yanqing Li, Yanan Yang, Zhaowei Huang, Panjing Lv, Xin Wang, Dali Han,* Yan Li,* Nan Qiao,* and Bing Liu*

Cite This: https://dx.doi.org/10.1021/acs.jcim.0c00821 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: The emergence of the new coronavirus (nCoV-19) has impacted human health on a global scale, while the interaction between the virus and the host is the foundation of the disease. The viral genome codes a cluster of proteins, each with a unique function in the event of host invasion or viral development. Under the current adverse situation, we employ virtual screening tools in searching for drugs and natural products which have been already deposited in DrugBank in an attempt to accelerate the drug discovery process. This study provides an initial evaluation of current drug candidates from various reports using our systemic in silico drug screening based on structures of viral proteins and human ACE2 receptor. Additionally, we have built an interactive online platform (https://shennongproject.ai/) for browsing these results with the visual display of a small molecule docked on its potential target protein, without installing any specialized structural software. With continuous maintenance and incorporation of data from laboratory work, it may serve not only as the assessment tool for the new drug discovery but also an educational web site for the public.

INTRODUCTION

The notorious coronaviruses, belonging to the family *Coronaviridae* and subfamily *Coronavirinae*, are pathologically significant to many mammals, including humans. Just after the millennium, two betacoronaviruses from this group of viruses also named severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East respiratory syndrome coronavirus (MERS-CoV) swept part of the world and brought impacts on health and the economy in 2003 and 2012, respectively.[1] Recently, another member of the family—severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) became an unavoidable topic for almost everyone around the globe, and the disease it brings (COVID-19), declared a pandemic by the world health organization (WHO), has so far caused over 148 000 cases and 5400 fatalities in 149 countries and territories. The early cases of the diseases emerged in Wuhan—a Chinese metropolis with over 11 million people—in December of 2019; these cases were diagnosed as cryptogenic pneumonia in several hospitalized patients.[2] Since then, it has become a world-wide panic during which most countries have taken stringent measures to tighten border controls, the movement of people, *etc.*

The origin of the virus remains undefined, although homology comparison shows that its genome has a 96.3% sequence similarity compared with BatCoV RaTG13 (a coronavirus of bat origin) and 79% compared with SARS-CoV.[3,4] The most characteristic feature shared by coronaviruses is the single-strand, positive-sense RNA genomes which are 26−32 kilobases in length containing 6−12 open reading frames (ORFs).[5] The first ORF takes up to two-thirds of the whole genome of the coronavirus and contains genetic codes

Special Issue: COVID19 - Computational Chemists Meet the Moment

Received: July 21, 2020
Published: August 11, 2020

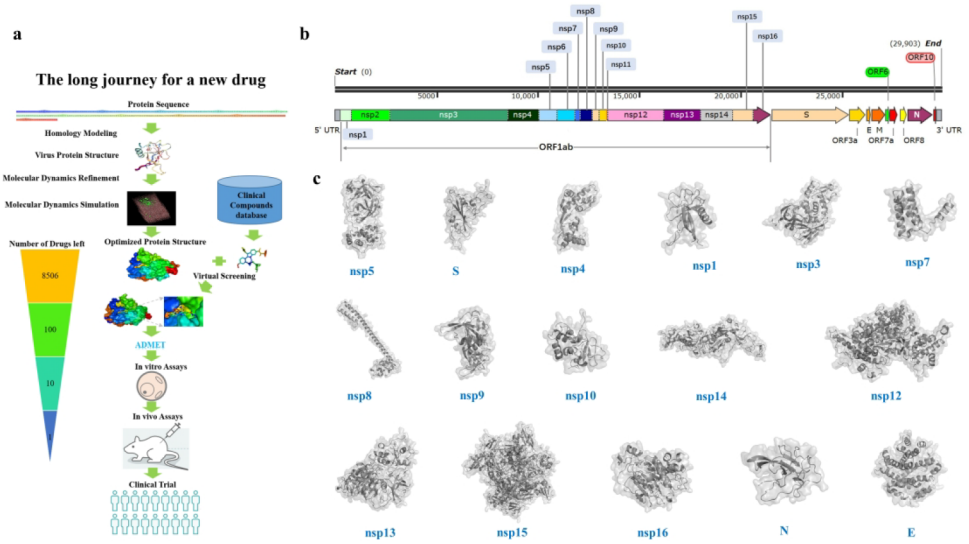


Figure 1. Structure-based *in silico* screening and homology modeling: (a) schematic description of the drug discovery process; (b) annotation of SARS-CoV-2 genome; (c) structures obtained from PDB (PDB ID for 6CS2 nsp5 and 6LU7 for S protein) and homology models built for SARS-CoV-2 using their SARS and mouse hepatitis virus A59 counterparts. PDB entries 2GDT, 6VXS, 3VCB, 6NUR, 6NUS, 1UW7, 2G9T, 6NUS, 6JYT, 5C8S, 2OZK, 3R24, 2GIB, and 1SSK were used as templates to model the structures for nsp1, nsp3, nsp4, nsp7, nsp8, nsp9, nsp10, nsp12, nsp13, nsp14, nsp15, nsp16, N, and E, respectively.

for two polyproteins named ppla and pplab, both of which are autoproteolytically cleaved into 15 or 16 nonstructural proteins: nsp1−nsp16 (nsp1 is absent in deltacoronavirus and gammacoronavirus). Meanwhile, the remaining ORFs encode some accessory proteins, including four indispensable structural proteins: spike glycoprotein, small envelope protein, matrix protein, and nucleocapsid protein.[6] These proteins play different roles at various stages during the viral invasion and viral development, many of which are vital for the survival of the coronaviruses.[7−9] The genome of SARS-CoV-2 is comprised of 29 891 nucleotides, which encode the 12 putative ORFs, coding for about 28 structural and nonstructural proteins (NCBI reference sequence: NC_045512.2).

There are four dispensable structural proteins coded by the viral genome—membrane (M), envelope (E), nucleocapsid (N), and spike (S). M protein is a small membrane protein with three transmembrane domains and the most abundant of the four, whose presence is required to form the shape of the virion.[10−12] The E protein is a small protein within the virion with functions like assembling and releasing of the virus.[13−15] The N protein only presents in the nucleocapsid and handles RNA structure and functions.[16,17] For a successful infection, the virus needs to recognize the host cell via the interaction between its S protein and the host cellular receptors—Angiotensin-Converting Enzyme 2 (ACE2) receptors in human. The S protein has two subunits—S1 which contains a so-called receptor-binding domain (RBD) allowing the virus to bind to the peptidase domain (PD) of ACE2 and the S2 subunit which helps the viral particle fuse with the host membrane.[18−21]

After entering the cell, the virus hijacks the host translational machinery and starts to express its own proteins. A polyprotein is then translated via ORF1ab and subsequently cleaved into 16 different nsp proteins, some of which are better characterized than others.[22,23] For example, nsp1 suppresses the host gene expression by inducing template-dependent endonucleolytic cleavage of host mRNAs and preventing the accumulation of IFN-beta, which may provide a susceptive condition for viral infection and replication in cellular.[24−27] Papain-like protease (PL^{pro}), also named nsp3, is the largest multidomain proteins encoded by the virus. Among a dozen domains of nsp3, ubiquitin-like domain mediating multitudinous viral protein interactions with themselves or host proteins and papain-like domain responsible for releasing nsp1, nsp2, and nsp3 from the polyprotein become a potential target for antiviral drug exploitation.[28−37] Main protease (M^{pro}), being synonymous with 3C-like protease ($3CL^{pro}$) or nsp5, is able to cleave the polyprotein at 11 sites, generating at least 10 essential nonstructural proteins.[6,38−40] And its importance in viral development makes it one of the most popular drug targets. Nsp8, an RNA-dependent RNA polymerase (RdRp), is verified to be capable of *de novo* initiation of RNA which initiate the synthesis of complementary oligonucleotides of <6 residues in a reaction and has been proposed to operate as a primase with the cooperation of nsp7.[41−44] Similarly, nsp12 is the second RdRp of the virus which contains the canonical viral RdRp motifs in its C-terminal part and employs a primer-dependent RNA synthesis mechanism with the assistance of primase nsp8.[45,46] Nsp13 is the viral helicase which has both RNA and DNA duplex-unwinding activities considering natural nucleotides and deoxynucleotides as its substrates.[47−49] Nsp16, activated by cofactor nsp10, functioning as 2′-O methyltransferase, exerts pivotal roles in the capping process, similar to the C-terminal of nsp14 which acts as N7-methyltransferase (N7Tase).[50−52] Besides, in the presence of nsp10, the N-terminal of nsp14 serves as exoribonuclease and cooperates with the endoribonuclease (nsp15) to ensure the accurate cleavage of the coronavirus RNA genome in the host cells.[53−58] As the pandemic affects our health and lifestyles, there is still no vaccine for CVOID-19. The priority remains to find drugs for the treatment of infected patients. Considering the above-mentioned proteins and their importance alone or synergistically during virus infection and replication, finding drugs to interdict their functions and interactions would stop viral development and, thus, spread.

https://dx.doi.org/10.1021/acs.jcim.0c00821
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Table 1. Active Sites Used in Ligand Screening[a]

protein	target sites	expected biological effect	source
ACE2_1	H34	prevent ACE2−S protein interaction	PDB: 6cs2
ACE2_2	K353	prevent ACE2−S protein interaction	PDB: 6cs2
S	F456	prevent ACE2−S protein interaction	PDB: 6vyb
Mpro (nsp5)	L27, H41, H164	block main protease activity	PDB: 6lu7
nsp4	unspecified	automatic docking by VINA	homologue modeling and model refinement
nsp1	unspecified	automatic docking by VINA	homologue modeling and model refinement
nsp3	unspecified	automatic docking by VINA	homologue modeling and model refinement
nsp7	K7, H36, N37	prevent nsp7 forming complex with nsp12	PDB: 6m71
nsp8_1	C115	block interaction of nsp8 with nsp12	PDB: 7bv1
nsp8_2	M130	block interaction of nsp8 with nsp12	PDB: 7bv1
nsp9	unspecified	automatic docking by VINA	PDB: 6w4b
nsp10_1 nsp10_2 nsp10_3	Ala1, Asn3, Glu6, Phe16, Phe19, Val21, Asn40, Lys43, Leu45, Thr58, Ser72, Lys93, Tyr96, His80, Cys90	block interaction of nsp10 with nsp14 and nsp16	PDB: 6zct
nsp12	K545, R555	remdesivir binding site	PDB: 7bv2
nsp13	unspecified	automatic docking by VINA	homologue modeling and model refinement
nsp14_1	D90, E92, E191, D273, H268	block exonuclease activity	homologue modeling and model refinement
nsp14_2	C378, F367	block exonuclease activity	homologue modeling and model refinement
nsp15	K289, H234, H249, Y342	block exonuclease activity	PDB: 6w01
nsp16	L100, N101, D130, M131	Block SAM binding pocket	PDB: 6w4h
N	unspecified	automatic docking by VINA	homologue modeling and model refinement
E	unspecified	automatic docking by VINA	homologue modeling and model refinement

[a]The drug target sites, and the expected biological effects, are listed for each protein; maximized space search and automatic docking were performed if no active site was given.

Drug discovery is a very lengthy process, and virtual screening is regarded as the fastest and most accurate method in the early stage of drug design (Figure 1a). Many studies based on *in silico* tools have virtually screened small molecule databases and published a huge amount of information on new drug discoveries for the coronavirus disease (COVID-19).[59] However, these results are neither based on the approved drugs in the DrugBank nor very user-friendly to scientists outside its niche. Here, we carried out structure-based virtual screening using FDA approved drugs and drugs that are currently undergoing phase 3 clinical trials as the library and constructed an interactive online platform for quick browsing—Shennong (https://shennongproject.ai/). The advantages of the platform include the following: searching drug name or protein target name, 3D display of drugs docked on their potential target proteins, and a dedicated section for natural products and continuous maintenance. Shennong is a collaborative effort with more data to be incorporated in the pipeline and possibly the prototype of its kind.

RESULTS

SARS-CoV-2 Protein Sequence Variations Compared to SARS and Homology Modeling. Structure-based virtual screening requires the three-dimensional structure of its protein target and a function to estimate the likelihood of the ligand-binding affinity to the protein. To use the best available structures for screening, we listed all 28 putative viral proteins encoded in its genome (Figure 1b) and removed the small peptides (ORF6, ORF7, ORF10, and nsp11) which are less likely to be druggable. Then we further removed 10 more proteins from the list as there is no structure for either SARS-CoV-2 or SARS. Among the 16 proteins left, S protein, ansp5, nsp7, nsp8, nsp9, nsp10, nsp12, nsp15, and nsp16 of SARS-CoV-2 have structures deposited in the protein data bank (PDB) with PDB IDs 6VYB, 6LU7, 6M71, 7BV1, 6W4B, 6ZET, 7BV2, 6W01, and 6W4H, respectively. The remaining viral proteins share high sequence identities with their SARS counterparts, ranging from 76.60% in nsp3 to 99.84% in nsp13 (Table S1). The high sequence identities ensured the reliabilities of homologous structure prediction using SARS proteins as templates. Nsp4, whose template was using the homologous structure of mouse hepatitis virus A59 (61.36% sequence identity to nsp4 of SARS-CoV-2), has no other close homology. Using SWISS-MODEL[60] and structures of SARS proteins and nsp4 of mouse hepatitis virus A59 as templates, we built 16 structural models, followed by molecular dynamics refinement and simulation for optimized protein structures (Figure 1c).

https://dx.doi.org/10.1021/acs.jcim.0c00821
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

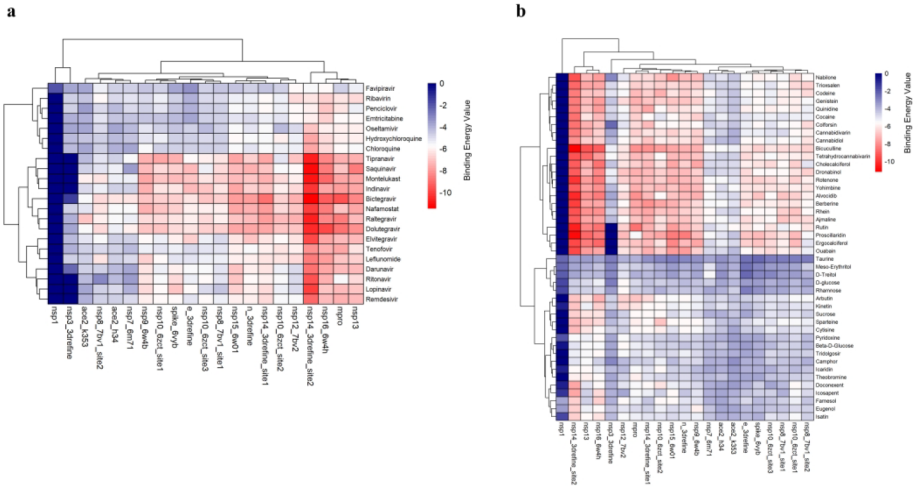


Figure 2. Overall result heatmap of binding energy for the predicted drugs. (a) The listed drugs have been reported to be in clinical trials. (b) Common natural compounds. The predicted energy rank from the most antagonistic pair to the most synergistic pair is colored from blue to red.

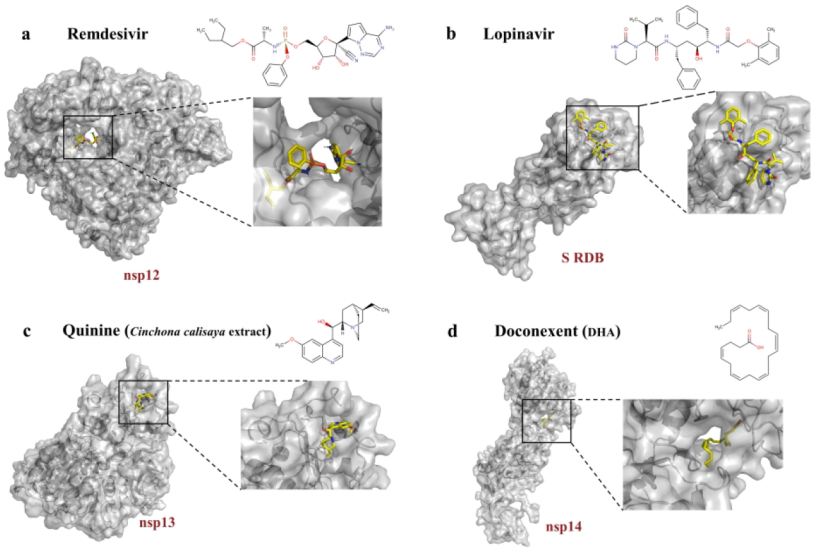


Figure 3. Low-energy binding conformations of ligand and protein complexes generated by AutoDock VINA: (a) antiviral drug remdesivir docked in the active pocket of SARS-CoV-2 nsp12 at its interface with RNA; (b) antiretroviral drug lopinavir docked in RDB of S protein; (c) quinine from Cinchona calisaya extract docked on nsp13; (d) deconexent from fish oil docked on nsp14 at its interface with nsp10.

Screening Library and Targets. Virtual screening is a technique largely based on its libraries of small molecules and the target sites. DrugBank has a collection of 9591 drug entries, including 2037 FDA-approved small molecule drugs, 241 FDA-approved polypeptide drugs, 96 nutraceuticals, and over 6000 experimental drugs.[61] As repurposing current drugs is the fastest way to meet the urgency of COVID-19, we built our library by selecting only FDA-approved drugs and drugs currently in clinical trials in DrugBank. Then we selected a list of active sites from structures of the 16 viral proteins and ACE2 protein (PDB ID: 6CS2) to use as the ligand targets for screening (Table 1). An individual protein has a biological role, and a successful drug should be able to specifically block its function by directly acting on the active site or indirectly via conformational change of the structure. For example, drugs screened based on human ACE2 protein and viral S protein were designed to block the interaction between the human cell and the virus while those for nsp5 were ought to have an effect on preventing its protease activity.

Docking Results Overview. To avoid overinterpretation of the results by ourselves, we uploaded the data to our web server for individual assessment. The complete set of the docking results (178, 626 in total) are available at our interactive server—https://shennongproject.ai/. In addition, we built two heatmaps for drugs with the lowest binding energies and natural compounds (some of which do not require a doctor's prescription), respectively (Figure 2a and b).

In general, the binding energies are relatively high for the dockings at active sites we chose for nsp1, nsp3, and nsp7. No specific active sites for nsp1 and nsp3 were given during screening due to the lack of characterization while the key residues (K7, H36, and N37) of nsp7 at its interaction interface with nsp12 were selected for screening. It is likely that these sites, either automatic generated or specified, were not suitable as drug targets, at least not for the candidates in our library. The absences of hydrophobic residues at these sites are the likely explanation for this phenomenon. Meanwhile, the binding energies for nsp5 (Mpro), nsp16, nsp14, and nsp13 are

https://dx.doi.org/10.1021/acs.jcim.0c00821
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

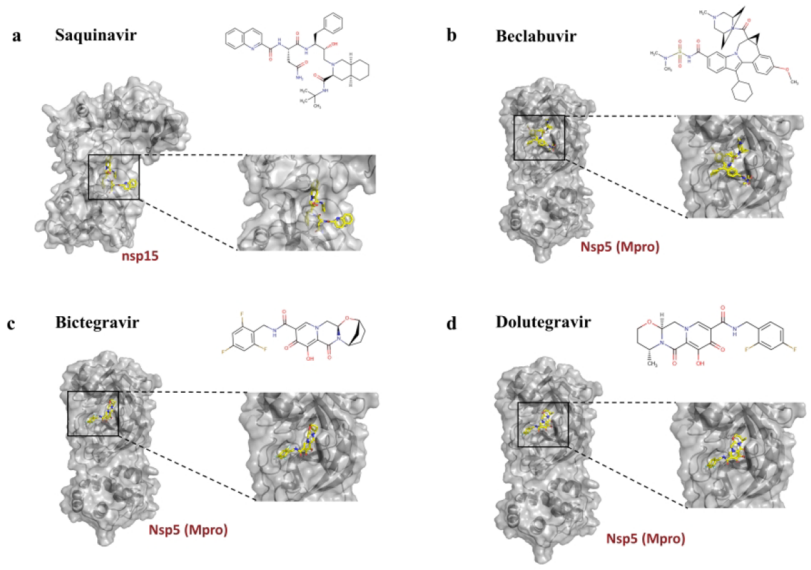


Figure 4. Best performing drugs in our docking but not currently in the clinical trial to our knowledge: (a) antiviral drug saquinavir docked on the RNase site of nsp15; (b) anti-HIV drug beclabuvir docked at the protease site of nsp5; (c) anti-HIV drug bictegravir docked at the protease site of nsp5; (d) antiretroviral drug dolutegravir docked on nsp14 at the protease site of nsp5.

generally low as the surface geometry and hydrophobicity of the active sites make them more druggable (discussed in detail later). Antiviral drugs like saquinavir, lopinarvir, darunavir nafamostat, raltegravir dolutegravir, bictegravir, tipranavir, indinarvir, and montelukast are among the highest scoring drugs in our screening (Figure 2a). In the other hand, natural products have higher binding energy in general although they still have a similar preference for nsp14 (Figure 2b). Proscillaridin extracted plants of the genus *Scilla* and in *Drimia maritima*, which is used for treating congestive heart failure and cardiac arrhythmia, achieved comparable reading as the above antiviral drugs.

A group of chemotherapeutic drugs, including tivantinib, lifirafenib, entrectinib, nilotinib, and radotinib, should not be neglected either. These tyrosine kinases (or tyrosine kinase receptor) inhibitors are either approved or investigational to be used in the therapy of certain hematopathy and metastatic cancers like acute myeloid leukemia (AML), acute lymphocytic leukemia (ALL), and lung cancers. In our docking results, these drugs are ranked among the top with main protease and exonuclease of SARS-CoV-2, as well as other nonstructural and structural proteins, indicating that they are worthy for further investigations in treatment for coronaviruses.

Drugs under Clinical Trials. Our results coincide with much of the current research in drug development. Our web site offers detailed docking results for most of them. For example, remdesivir is a nucleotide analog used for antiviral purposes. Although it was designed as a treatment for Ebola virus disease, it has also been found to show antiviral activity against other single-stranded RNA viruses and used in the treatment of COVID-19.[62−65] In our screening, remdesivir is predicted to interact with nsp12 by forming hydrogen bonding with K521, D623, R553, and extensively with R555 and additional hydrophobic interactions between the corresponding residues in the binding pocket (Figure 3a). By comparison, the triphosphate form of remdesivir is bound to the published nsp12−nsp7−nsp8 complex via the side chains of K545 and R555[66] and occupies the same binding pocket. Since the NTP entry channel is formed by the hydrophilic residues such as K545, R553, and R555,[67] the occupation of this binding pocket by remdesivir is proposed to inhibit the activity of the complex.

Lopinavir, an anti-HIV drug in the category of protease inhibitor, is another popular drug that has been reported to have strong positive results in a few trials.[68−73] In our docking, lopinavir binds to the receptor binding domain of S protein with strong binding affinity (−7.1 kcal/mol) (Figure 3b). The $\pi-\pi$ stacking between lopinavir and the side chain of F456 help to stabilize the interaction, and the hydrogen bonding with T470 and the backbone of F456 and R467 also contributes to the high binding affinity. Thus, lopinavir may be a potent spike inhibitor based on our results.

Natural Products in the Screening. We picked two natural products of our interest (quinine and doconexent) from the 924 docking results from our screening (https://shennongproject.ai/#/naturalProducts). Quinine is a famous antimalarial drug which was recently repurposed quinine as an antiviral against dengue virus infection. It has a binding energy of −7.5 kcal/mol against nsp13, which is comparable to some of the drugs under clinical trials (Figure 3c). Its interaction with nsp13 includes $\pi-\pi$ stacking with F499 and hydrophobic interaction with the hydrophobic side chains in the binding pocket, thus making it a potential inhibitor for nsp13. Meanwhile doconexent is a mixture of fish oil and primrose oil and used as a high-docosahexaenoic acid (DHA) with minor anti-inflammatory effects. It is ranked at the bottom half against all active sites, likely due to the lack of $\pi-\pi$ stacking and limited hydrogen bonding to A353, L366, and Y368 of nsp14 and hydrophobic interactions due to unfavorable distances. However, it has a low binding energy with nsp14 at −7.4 kcal/mol (Figure 3d). Although it is undoubtedly a less preferred ligand in our screening, the ability to purchase DHA or fish oil without a prescription makes it a potential mild viral inhibitor for self-protection.

Drugs Perform Well in Our Screening but Not under Clinical Trial. A few drugs, including saquinavir, beclabuvir, bictegravir, and dolutegravir are not currently under investigation for the treatment of COVID-19 to our knowledge. However, the antiviral mechanisms of these drugs, together

https://dx.doi.org/10.1021/acs.jcim.0c00821
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

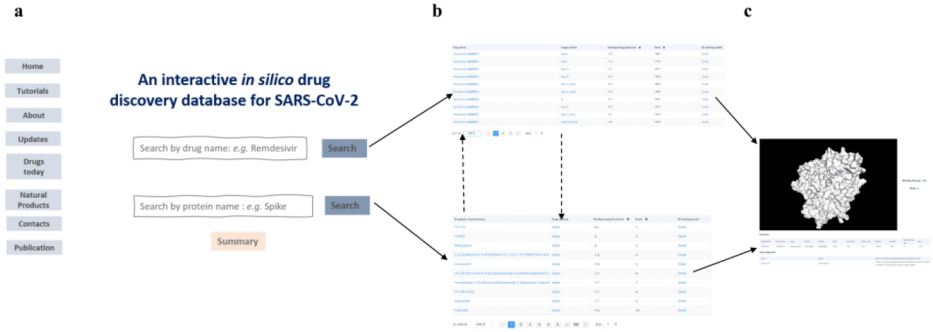


Figure 5. Shennong web server. (a) The home page for the Shennong server provides two search engines that support enquiries by drug name or by target protein name. (b) The results for the enquiries by drug name or by target protein name are ranked. (c) The ranked results contain detailed docking models including information on the drug, the protein, and graphic interfaces.

with their performance in our screening, make them the recommendable drugs for COVID-19.

Saquinavir is an antiretroviral drug used in a cocktail for treating HIV patients[74] and has a binding energy of −7.2 kcal/mol to nsp15 in our screening that arose from the strong hydrogen bonding with K89, N199, D272, and Y278, $\pi-\pi$ stacking with Y278 and hydrophobic interaction with hydrophobic side chains in the binding pocket (Figure 4a). Among them, beclabuvir is the only antiviral drug with the purpose for the treatment of HCV infection,[75,76] while the rest are drugs for HIV infection. With a low binding energy of −10.4 kcal/mol to nsp5, beclabuvir is one of the drugs that performed the best in the docking. With strong hydrogen bonding with Y54 and N142, hydrophobic interaction with the hydrophobic side chains in the binding pocket, and $\pi-\pi$ stacking with H41, it is likely a stronger inhibitor for the exonuclease activity inhibitor of nsp15 (Figure 4b). It is possibly the best nsp15 inhibitor at least in our screening. Bictegravir and dolutegravir are integrase inhibitors used in combination with other drugs for the treatment of HIV infection. They are structurally related, as the former is a derivation from the latter.[77] And their binding energies to nsp5 are also very similar (−9.5 kcal/mol for bictegravir and −8.9 kcal/mol for dolutegravir), with bictegravir forming hydrogen bonding with S144, C145, E166, and Q189, and dolutegravir forming hydrogen bonding with H41, G143 and Q189 (Figure 4c and d). Interestingly, all three drugs are in the category of protease inhibitors and have low binding energies against nsp5—the main protease of the virus. These underlying similarities make them worthy of repurposing for potential COVID-19 treatment.

Shennong Web Server and Results Reporting. To give users the familiar search engine style experience, we adopted a user-friendly homepage and a graphic interface for viewing the docking results (Figure 5). The web server supports searches by either drug name or protein target name, with additional features like updates for drugs under clinical trials and a tab dedicated for natural compounds. For example, the user wishing to look for docking results of dexamethasome could type the name in the first search bar, and the results would be shown in a new page and ranked according to their binding energy to the respective proteins. The binding of dexamethasone isonicotinate with Nsp16 is the ranked top with −9 kcal/mol binding energy. It is ranked 13th among all the drugs docked with nsp16, and the user could click on the Nsp16 in the target protein column to view all the docking results for nsp16.

The three-dimensional docking model of dexamethasone with nsp16 can be viewed by clicking detail in the 3D docking model column. The new page includes the interactive three-dimensional visualization and the annotation information on the drug and the function of nsp16.

Alternatively, the user can type the name of protein of interest in the second search bar to search the results of drugs docked with the protein. The new page would show the ranked the results and contains links to each drug for viewing the docking results of other proteins. This online platform may not only assist fast and cost-efficient drug discovery but also serve as an educational web site for the general public.

DISCUSSION

To provide a fast track solution, we performed virtual screening using drugs from the DrugBank, targeting some of the viral proteins and human ACE2 receptors. Our results coincide with some of the most popular drugs currently under clinical trials and provide some potential new candidates. The drugs on the top of our list are related anti-HIV drugs, anti-HCV drugs, influenza virus antagonists, chemotherapeutic drugs, and asthma drugs.

Anti-HIV drugs are popularized across our docking list and can be divided into two groups: enzyme inhibitors which are generally located on the top of our list and dideoxynucleoside (or nucleoside) analogs, generally at the bottom of our list. Nucleoside reverse transcriptase inhibitors (NRTIs), including emtricitabine and tenofovir, may not work well in coronaviruses; this can be attributed to the fact that coronavirus is a positive-sense single-stranded RNA virus which lacks nucleoside reverse transcriptase, which is also reflected in our docking as most of the NRTIs ranked at the bottom with low binding affinity. Among the enzyme inhibitors of HIV in our docking results, dolutegravir and raltegravir exhibit strong binding affinity with multiple target sites, especially at the catalytic sites of main protease and exonuclease, suggesting the great potential of clinical drugs in therapies for COID-19. For example, saquinavir, acting on HIV protease cleavage site, is a highly specific inhibitor of HIV-1 and HIV-2 proteases. Interestingly, it shows a strong affinity with the main protease of SARS-CoV-2, which is coincident to the recent results of other researchers. It is also worth noting that S protein, RdRp (nsp12 and nsp8), exonuclease (nsp14), 2′-*O* methyltransferase (nsp16), helicase (nsp13), and nsp10 of SARS-CoV-2 are potential targets of saquinavir. The binding energies of nsp13, nsp14, and nsp16 with saquinavir even surpass that of nsp5, suggesting that saquinavir might be a multitarget inhibitor of

https://dx.doi.org/10.1021/acs.jcim.0c00821
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

SARS-CoV-2. Not surprisingly, other enzyme inhibitors of HIV such as ritonavir, tipranavir, elvitegravir, nelfinavir, darunavir, and fosamprenavir have a relatively high binding affinity with the chosen targets in our docking. Six anti-HCV drugs including five RdRp (NS5B of HCV) inhibitors, including bictegravir, filibuvir, ribavirin-monophosphate, sofosbuvir, and one protease (NS3/4B) inhibitor—bictegravir—are also our best-performing drugs. It is worth noting that bictegravir has an impressively strong affinity with Mpro (binding energy −10.4 kcal/mol), nsp13 (binding energy −9.8 kcal/mol), nsp14 (binding energy −8.8 kcal/mol), and nsp15 (binding energy −8.3 kcal/mol), making it one of best-performing drugs in our docking. The comprehensive score of filibuvir does not fall far behind that of bictegravir and even exceeds it in some docking sites. Therefore, anti-HCV drugs should be tested for battling with SARS-CoV-2. Last but not least, tivantinib, lifirafenib, entrectinib, nilotinib, and radotinib, the chemotherapeutic drugs also for cancer treatments, and montelukast and zafirlukast which are used in the therapy of asthma are also on the top of our list.

At the beginning of the COVID-19 pandemic, two drugs used for influenza virus, oseltamivir and arbidol, were widely used in treatments. However, there is no further evidence, so far, to show that oseltamivir has an obvious clinical effect. Both arbidol and oseltamivir are thought be interacting with mainly binds to surface hemagglutinin (HA) of the H2 strain of influenza viruses to block infections.[78] However, no proteins having such functions have been found in SARS-CoV-2 so far. Coincidentally, our docking results also display the low binding energies of oseltamivir with different targeted proteins of SARS-CoV-2.

Another interesting finding in our results is the performance of natural compounds. Although most of them are at the bottom of the league and one should not overinterpret the results, the fact that many of them could be found in large quantity without prescriptions make them potentially the best household compounds, especially when half of the world is in self-isolation.

There are still limitations to our study. For example, remdesivir in the previous studies acting as RdRp inhibitors had a promising efficiency in interdicting the infection of MERS-CoV.[65,79,80] Whereas, the binding affinity of remdesivir with RdRp (binding energy −6.3 kcal/mol) is lower than that with endonuclease (binding energy −8.3 kcal/mol), due likely to the differences and the absence of metal ions to stabilize the drug in the binding pocket.

Overall, our web server—Shennong—offers a new way to browse drug−protein docking results. It supports searches by either drug name or protein target name, with additional features like updates for drugs under clinical trials and a tab dedicated for natural compounds. This online platform may not only assist fast and cost-efficient drug discovery but also serves as an educational web site for the general public.

METHODS

Compound Libraries. We prepared a large-scale library consisting of 8506 small molecular compounds from DrugBank. It covers all FDA-approved drugs and compounds in the midst of clinical trials and molecules under experimental investigations. The SDF files were downloaded for each compound from DrugBank, whereas the SMILES files were downloaded for compounds without 3D SDF files, for example saquinavir, lopinavir, ritonavir, and carfilzomlib. We converted the SMILES files to 3D SDF files for the four drugs using python rdkit library. We also listed FDA-approved covalent inhibitors and known covalent small-molecule kinase inhibitors that filtered by identifier mapping with other public sources[81,82](Table S2).

SARS-CoV-2 Genome Annotation. The reference genome of SARS-CoV-2 was downloaded from NCBI with accession number: NC_045512.2. But due to the lack of genome annotation, the protein sequence of SARS-CoV-2 cannot be obtained directly. Considering the high similarity between SARS-CoV and SARS-CoV-2, we aligned the protein sequence of SARS-CoV to SARS-CoV-2 genome and selected the best match region as the corresponding protein sequence for SARS-CoV-2. Using this method, we obtained all the 28-protein sequence of SARS-CoV-2, including 16 nonstructural proteins (nsp1−16), 4 structural proteins, spike (S), membrane (M), nucleocapsid (N), and envelope (E), and 8 putative accessory proteins.

Homology Modeling of SARS-CoV-2 Proteins. Homology modeling is performed by SWISS-MODEL (https://swissmodel.expasy.org/). SWISS-MODEL takes the protein sequence and template protein structure as inputs. Protein sequence is obtained as described previously. An optimal template protein is selected for homologous modeling based on the following criteria: (1) The identity between the target and template proteins in the sequence should be over 30%. The template protein with the highest identity is selected preferentially. (2) The SARS-CoV template protein is preferred for homologous modeling. (3) The template protein constructed with the high-precision X-ray method is preferred. If X-ray is unavailable, check the protein structure resolution in the PDB database and choose the structure with a higher resolution. (4) If Oligo-State has two values, homo and hetero, select both of them.

After selecting the optimal template protein, SWISS-MODEL builds protein structure with default parameters. After the modeling is completed, the PDB files of the template and target proteins can be downloaded. Ions and waters are deleted before downstream analysis.

PDB entries 2GDT, 6VXS, 3VCB, 6JYT, 5C8S, and 1SSK were used as templates to model the structures for nsp1, nsp3, nsp4, nsp13, nsp14, and E, respectively. Structures of ACE2 protein, S protein, nsp5, nsp7, nsp8, nsp9, nsp10, nsp12, nsp15, and nsp16 were extracted from PDB entries 6CS2, 6VYB, 6LU7, 6M71, 7BV1, 6W4B, 6ZCT, 7BV2, 6W01, and 6W4H respectively. The 3D-refine server (http://sysbio.rnet.missouri.edu/3Drefine/) was used for the refinement of the protein structures to make structure models closer to native states.[83] It first optimizes the hydrogen bond network of the structure models, then performs atomic-level energy minimization on the models using a composite physics and knowledge-based force fields and outputs five optimized models. The $3D^{refine}$ score is the potential energy of the refined model according to the $3D^{refine}$ force field and the lower score indicates a better quality model. The Top-1 $3D^{refine}$ score ranked structure model for each protein was selected for further analysis.

Virtual Docking. *Preparation of Proteins and Ligands.* The structures of proteins to be used in docking were first examined, and any ligand, metal ion, or other substances presenting in the structure is removed. Then, Gasteiger charges were added, bonds of hydrogens were repaired, and nonpolar hydrogens were removed. Besides, structures of proteins were

https://dx.doi.org/10.1021/acs.jcim.0c00821
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

already refined by the 3D-refine server. The PDB format was then converted to a PDBQT format to meet the requirement of AutoDock Vina[84] To prepare a ligand file for docking, chemical files of FDA approved and investigational drugs were downloaded from DrugBank and then converted into PDBQT format file by OpenBabel or AutoDock Tools.

Docking Parameters. Following our selection criteria (Table 1), amino acids of interests were highlighted, and the corresponding coordinates and size of binding box were obtained using AutoDock Tools.

Large-Scale Docking between Protein Receptors and Chemicals. Protein receptors and chemical ligands were docked using over 10 thousand of CPU nodes in parallel. The values of binding energy of the first model in docking PDBQT output files were used to represent and compare the binding strength for each receptor−chemical pair.

Drug-likeness Analysis. We calculated five drug-likeness indexes for each compound−the ratio of sp^3 hybridized carbons over the total carbon count of the molecule (Fraction Csp3) for saturation, the molecular weight for size, TPSA for polarity, XLOGP for lipophilicity, and the number of rotatable bonds for flexibility using python rdkit library. We set corresponding thresholds for each drug-likeness index to evaluate whether a compound could be drug-like, Fraction Csp3 $\geq$ 0.25, 150 $\leq$ MW $\leq$ 500, 20 $\leq$ TPSA $\leq$ 130, 0.7 $\leq$ XLOGP3 $\leq$ 6, rotatable bond num. $\leq$ 9.[60]

Shengnong Web Server. The Vue.js framework (https://cn.vuejs.org/index.html) was used to construct Shennong server. Spring Boot (https://spring.io/projects/spring-boot) was used for data query and search. The nglview plugin (https://github.com/arose/nglview) was used for 3D docking visualization.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.0c00821.

> Structure availability of SARS-CoV-2 proteins and the sequence identity compared with SARS proteins (Table S1) and the list of FDA-approved covalent inhibitors and covalent small molecule kinase inhibitors (Table S2) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Bing Liu − *BioBank, The First Affiliated Hospital of Xi'an Jiaotong University, Shaanxi 710061, China; MRC Centre for Molecular Bacteriology and Infection, Imperial College London, London SW7 2AZ, U.K.; Instrument Analysis Centre, of Xi'an Jiaotong University, Shaanxi 710049, China;* orcid.org/0000-0003-1670-6115; Email: bliu2018@xjtu.edu.cn

Nan Qiao − *Laboratory of Health Intelligence, Huawei Technologies Co., Ltd, Shenzhen 518100, China;* Email: qiaonan3@huawei.com

Yan Li − *Department of Pathogen Biology, School of Basic Medicine, Tongji Medical College and Department of Pediatrics, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China;* Email: yanli@hust.edu.cn

Dali Han − *Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China; College of Future Technology, Sino-Danish College, University of Chinese Academy of Sciences, Beijing 100049, China; Institute for Stem Cell and Regeneration, Chinese Academy of Sciences, Beijing 100101, China; China National Center for Bioinformation, Beijing 100101, China;* Email: handl@big.ac.cn

Authors

Chi Xu − *Laboratory of Health Intelligence, Huawei Technologies Co., Ltd, Shenzhen 518100, China*

Zunhui Ke − *Wuhan Children's Hospital, Tongji Medical College, Huazhong University of Science & Technology, Wuhan 430000, China*

Chuandong Liu − *Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China; College of Future Technology, Sino-Danish College, University of Chinese Academy of Sciences, Beijing 100049, China*

Zhihao Wang − *BioBank, The First Affiliated Hospital of Xi'an Jiaotong University, Shaanxi 710061, China; MRC Centre for Molecular Bacteriology and Infection, Imperial College London, London SW7 2AZ, U.K.*

Denghui Liu − *Laboratory of Health Intelligence, Huawei Technologies Co., Ltd, Shenzhen 518100, China*

Lei Zhang − *Laboratory of Health Intelligence, Huawei Technologies Co., Ltd, Shenzhen 518100, China*

Jingning Wang − *Department of Pathogen Biology, School of Basic Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China*

Wenjun He − *Laboratory of Health Intelligence, Huawei Technologies Co., Ltd, Shenzhen 518100, China*

Zhimeng Xu − *Laboratory of Health Intelligence, Huawei Technologies Co., Ltd, Shenzhen 518100, China*

Yanqing Li − *BioBank, The First Affiliated Hospital of Xi'an Jiaotong University, Shaanxi 710061, China*

Yanan Yang − *BioBank, The First Affiliated Hospital of Xi'an Jiaotong University, Shaanxi 710061, China*

Zhaowei Huang − *Laboratory of Health Intelligence, Huawei Technologies Co., Ltd, Shenzhen 518100, China*

Panjing Lv − *Department of Pathogen Biology, School of Basic Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China*

Xin Wang − *BioBank, The First Affiliated Hospital of Xi'an Jiaotong University, Shaanxi 710061, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.0c00821

Author Contributions

○These authors contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is supported by National Natural Science Foundation of China (82041004), HUST COVID-19 Rapid Response Call (2020kfyXGYJ036), Zhejiang University special scientific research fund for COVID-19 prevention and control (2020XGZX089), Shandong University special funding for COVID-19 (2020XGB02), and the Key Research Program of Frontier Sciences, Chinese Academy of Sciences (ZDBS-LY-SM013 to D.H.).

REFERENCES

(1) Hui, D. S. Epidemic and Emerging Coronaviruses (Severe Acute Respiratory Syndrome and Middle East Respiratory Syndrome). *Clin Chest Med.* **2017**, *38*, 71−86.

(2) Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; Niu, P.; Zhan, F.; Ma, X.; Wang, D.; Xu, W.; Wu, G.; Gao, G. F.; Tan, W. China Novel Coronavirus, I.; Research, T.; A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **2020**, *382*, 727−733.

(3) Paraskevis, D.; Kostaki, E. G.; Magiorkinis, G.; Panayiotakopoulos, G.; Sourvinos, G.; Tsiodras, S. Full-genome evolutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event. *Infect., Genet. Evol.* **2020**, *79*, 104212.

(4) Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; Bi, Y.; Ma, X.; Zhan, F.; Wang, L.; Hu, T.; Zhou, H.; Hu, Z.; Zhou, W.; Zhao, L.; Chen, J.; Meng, Y.; Wang, J.; Lin, Y.; Yuan, J.; Xie, Z.; Ma, J.; Liu, W. J.; Wang, D.; Xu, W.; Holmes, E. C.; Gao, G. F.; Wu, G.; Chen, W.; Shi, W.; Tan, W. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **2020**, *395*, 565−574.

(5) Luk, H. K. H.; Li, X.; Fung, J.; Lau, S. K. P.; Woo, P. C. Y. Molecular epidemiology, evolution and phylogeny of SARS coronavirus. *Infect., Genet. Evol.* **2019**, *71*, 21−30.

(6) Ziebuhr, J. Molecular biology of severe acute respiratory syndrome coronavirus. *Curr. Opin. Microbiol.* **2004**, *7*, 412−9.

(7) Weiss, S. R.; Leibowitz, J. L. Coronavirus pathogenesis. *Adv. Virus Res.* **2011**, *81*, 85−164.

(8) Brian, D. A.; Baric, R. S. Coronavirus genome structure and replication. *Curr. Top. Microbiol. Immunol.* **2005**, *287*, 1−30.

(9) Narayanan, K.; Huang, C.; Makino, S. SARS coronavirus accessory proteins. *Virus Res.* **2008**, *133*, 113−21.

(10) Arndt, A. L.; Larson, B. J.; Hogue, B. G. A conserved domain in the coronavirus membrane protein tail is important for virus assembly. *J. Virol.* **2010**, *84*, 11418−28.

(11) Neuman, B. W.; Kiss, G.; Kunding, A. H.; Bhella, D.; Baksh, M. F.; Connelly, S.; Droese, B.; Klaus, J. P.; Makino, S.; Sawicki, S. G.; Siddell, S. G.; Stamou, D. G.; Wilson, I. A.; Kuhn, P.; Buchmeier, M. J. A structural analysis of M protein in coronavirus assembly and morphology. *J. Struct. Biol.* **2011**, *174*, 11−22.

(12) Siu, K. L.; Chan, C. P.; Kok, K. H.; Chiu-Yat Woo, P.; Jin, D. Y. Suppression of innate antiviral response by severe acute respiratory syndrome coronavirus M protein is mediated through the first transmembrane domain. *Cell. Mol. Immunol.* **2014**, *11*, 141−9.

(13) Schoeman, D.; Fielding, B. C. Coronavirus envelope protein: current knowledge. *Virol. J.* **2019**, *16*, 69.

(14) Liu, D. X.; Yuan, Q.; Liao, Y. Coronavirus envelope protein: a small membrane protein with multiple functions. *Cell. Mol. Life Sci.* **2007**, *64*, 2043−8.

(15) Ruch, T. R.; Machamer, C. E. The coronavirus E protein: assembly and beyond. *Viruses* **2012**, *4*, 363−82.

(16) McBride, R.; van Zyl, M.; Fielding, B. C. The coronavirus nucleocapsid is a multifunctional protein. *Viruses* **2014**, *6*, 2991−3018.

(17) Chang, C. K.; Hou, M. H.; Chang, C. F.; Hsiao, C. D.; Huang, T. H. The SARS coronavirus nucleocapsid protein–forms and functions. *Antiviral Res.* **2014**, *103*, 39−50.

(18) Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020**, *367*, 1444−1448.

(19) Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181*, 281−292.

(20) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Kruger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; Muller, M. A.; Drosten, C.; Pohlmann, S. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181*, 271−280.

(21) Dimitrov, D. S. The Secret Life of ACE2 as a Receptor for the SARS Virus. *Cell* **2003**, *115*, 652−653.

(22) Subissi, L.; Imbert, I.; Ferron, F.; Collet, A.; Coutard, B.; Decroly, E.; Canard, B. SARS-CoV ORF1b-encoded nonstructural proteins 12−16: replicative enzymes as antiviral targets. *Antiviral Res.* **2014**, *101*, 122−30.

(23) Graham, R. L.; Sparks, J. S.; Eckerle, L. D.; Sims, A. C.; Denison, M. R. SARS coronavirus replicase proteins in pathogenesis. *Virus Res.* **2008**, *133*, 88−100.

(24) Kamitani, W.; Narayanan, K.; Huang, C.; Lokugamage, K.; Ikegami, T.; Ito, N.; Kubo, H.; Makino, S. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 12885−90.

(25) Narayanan, K.; Huang, C.; Lokugamage, K.; Kamitani, W.; Ikegami, T.; Tseng, C. T.; Makino, S. Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. *J. Virol.* **2008**, *82*, 4471−9.

(26) Tohya, Y.; Narayanan, K.; Kamitani, W.; Huang, C.; Lokugamage, K.; Makino, S. Suppression of host gene expression by nsp1 proteins of group 2 bat coronaviruses. *J. Virol.* **2009**, *83*, 5282−8.

(27) Huang, C.; Lokugamage, K. G.; Rozovics, J. M.; Narayanan, K.; Semler, B. L.; Makino, S. SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to nsp1-induced RNA cleavage. *PLoS Pathog.* **2011**, *7*, No. e1002433.

(28) Lei, J.; Kusov, Y.; Hilgenfeld, R. Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein. *Antiviral Res.* **2018**, *149*, 58−74.

(29) Chen, X.; Yang, X.; Zheng, Y.; Yang, Y.; Xing, Y.; Chen, Z. SARS coronavirus papain-like protease inhibits the type I interferon signaling pathway through interaction with the STING-TRAF3-TBK1 complex. *Protein Cell* **2014**, *5*, 369−81.

(30) Baez-Santos, Y. M.; Mielech, A. M.; Deng, X.; Baker, S.; Mesecar, A. D. Catalytic function and substrate specificity of the papain-like protease domain of nsp3 from the Middle East respiratory syndrome coronavirus. *J. Virol.* **2014**, *88*, 12511−27.

(31) Mielech, A. M.; Deng, X.; Chen, Y.; Kindler, E.; Wheeler, D. L.; Mesecar, A. D.; Thiel, V.; Perlman, S.; Baker, S. C. Murine coronavirus ubiquitin-like domain is important for papain-like protease stability and viral pathogenesis. *J. Virol.* **2015**, *89*, 4907−17.

(32) Zhang, W.; Bailey-Elkin, B. A.; Knaap, R. C. M.; Khare, B.; Dalebout, T. J.; Johnson, G. G.; van Kasteren, P. B.; McLeish, N. J.; Gu, J.; He, W.; Kikkert, M.; Mark, B. L.; Sidhu, S. S. Potent and selective inhibition of pathogenic viruses by engineered ubiquitin variants. *PLoS Pathog.* **2017**, *13*, No. e1006372.

(33) Baez-Santos, Y. M.; Barraza, S. J.; Wilson, M. W.; Agius, M. P.; Mielech, A. M.; Davis, N. M.; Baker, S. C.; Larsen, S. D.; Mesecar, A. D. X-ray structural and biological evaluation of a series of potent and highly selective inhibitors of human coronavirus papain-like proteases. *J. Med. Chem.* **2014**, *57*, 2393−412.

(34) Chou, C. Y.; Chien, C. H.; Han, Y. S.; Prebanda, M. T.; Hsieh, H. P.; Turk, B.; Chang, G. G.; Chen, X. Thiopurine analogues inhibit papain-like protease of severe acute respiratory syndrome coronavirus. *Biochem. Pharmacol.* **2008**, *75*, 1601−9.

(35) Ghosh, A. K.; Takayama, J.; Aubin, Y.; Ratia, K.; Chaudhuri, R.; Baez, Y.; Sleeman, K.; Coughlin, M.; Nichols, D. B.; Mulhearn, D. C.; Prabhakar, B. S.; Baker, S. C.; Johnson, M. E.; Mesecar, A. D. Structure-based design, synthesis, and biological evaluation of a series of novel and reversible inhibitors for the severe acute respiratory syndrome-coronavirus papain-like protease. *J. Med. Chem.* **2009**, *52*, 5228−40.

(36) Ratia, K.; Pegan, S.; Takayama, J.; Sleeman, K.; Coughlin, M.; Baliji, S.; Chaudhuri, R.; Fu, W.; Prabhakar, B. S.; Johnson, M. E.; Baker, S. C.; Ghosh, A. K.; Mesecar, A. D. A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 16119−24.

https://dx.doi.org/10.1021/acs.jcim.0c00821
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(37) Zheng, D.; Chen, G.; Guo, B.; Cheng, G.; Tang, H. PLP2, a potent deubiquitinase from murine hepatitis virus, strongly inhibits cellular type I interferon production. *Cell Res.* **2008**, *18*, 1105−13.
(38) Yin, J.; Niu, C.; Cherney, M. M.; Zhang, J.; Huitema, C.; Eltis, L. D.; Vederas, J. C.; James, M. N. A mechanistic view of enzyme inhibition and peptide hydrolysis in the active site of the SARS-CoV 3C-like peptidase. *J. Mol. Biol.* **2007**, *371*, 1060−74.
(39) Chou, K. C.; Wei, D. Q.; Zhong, W. Z. Binding mechanism of coronavirus main proteinase with ligands and its implication to drug design against SARS. *Biochem. Biophys. Res. Commun.* **2003**, *308*, 148−51.
(40) Muramatsu, T.; Kim, Y. T.; Nishii, W.; Terada, T.; Shirouzu, M.; Yokoyama, S. Autoprocessing mechanism of severe acute respiratory syndrome coronavirus 3C-like protease (SARS-CoV 3CLpro) from its polyproteins. *FEBS J.* **2013**, *280*, 2002−13.
(41) te Velthuis, A. J.; van den Worm, S. H.; Snijder, E. J. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Res.* **2012**, *40*, 1737−47.
(42) Xiao, Y.; Ma, Q.; Restle, T.; Shang, W.; Svergun, D. I.; Ponnusamy, R.; Sczakiel, G.; Hilgenfeld, R. Nonstructural proteins 7 and 8 of feline coronavirus form a 2:1 heterotrimer that exhibits primer-independent RNA polymerase activity. *J. Virol.* **2012**, *86*, 4444−54.
(43) Zhai, Y.; Sun, F.; Li, X.; Pang, H.; Xu, X.; Bartlam, M.; Rao, Z. Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. *Nat. Struct. Mol. Biol.* **2005**, *12*, 980−6.
(44) Imbert, I.; Guillemot, J. C.; Bourhis, J. M.; Bussetta, C.; Coutard, B.; Egloff, M. P.; Ferron, F.; Gorbalenya, A. E.; Canard, B. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* **2006**, *25*, 4933−42.
(45) te Velthuis, A. J.; Arnold, J. J.; Cameron, C. E.; van den Worm, S. H.; Snijder, E. J. The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Res.* **2010**, *38*, 203−14.
(46) Kirchdoerfer, R. N.; Ward, A. B. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat. Commun.* **2019**, *10*, 2342.
(47) Seybert, A.; Hegyi, A.; Siddell, S. G.; Ziebuhr, J. The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5′-to-3′ polarity. *RNA* **2000**, *6*, 1056−68.
(48) Ivanov, K. A.; Thiel, V.; Dobbe, J. C.; van der Meer, Y.; Snijder, E. J.; Ziebuhr, J. Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. *J. Virol.* **2004**, *78*, 5619−32.
(49) Lee, N. R.; Kwon, H. M.; Park, K.; Oh, S.; Jeong, Y. J.; Kim, D. E. Cooperative translocation enhances the unwinding of duplex DNA by SARS coronavirus helicase nsP13. *Nucleic Acids Res.* **2010**, *38*, 7626−36.
(50) Decroly, E.; Imbert, I.; Coutard, B.; Bouvet, M.; Selisko, B.; Alvarez, K.; Gorbalenya, A. E.; Snijder, E. J.; Canard, B. Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2′O)-methyltransferase activity. *J. Virol.* **2008**, *82*, 8071−84.
(51) Chen, Y.; Cai, H.; Pan, J.; Xiang, N.; Tien, P.; Ahola, T.; Guo, D. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 3484−9.
(52) Wang, Y.; Sun, Y.; Wu, A.; Xu, S.; Pan, R.; Zeng, C.; Jin, X.; Ge, X.; Shi, Z.; Ahola, T.; Chen, Y.; Guo, D. Coronavirus nsp10/nsp16 Methyltransferase Can Be Targeted by nsp10-Derived Peptide In Vitro and In Vivo To Reduce Replication and Pathogenesis. *J. Virol.* **2015**, *89*, 8416−27.
(53) Ogando, N. S.; Ferron, F.; Decroly, E.; Canard, B.; Posthuma, C. C.; Snijder, E. J. The Curious Case of the Nidovirus Exoribonuclease: Its Role in RNA Synthesis and Replication Fidelity. *Front. Microbiol.* **2019**, *10*, 1813.
(54) Bouvet, M.; Lugari, A.; Posthuma, C. C.; Zevenhoven, J. C.; Bernard, S.; Betzi, S.; Imbert, I.; Canard, B.; Guillemot, J. C.; Lecine, P.; Pfefferle, S.; Drosten, C.; Snijder, E. J.; Decroly, E.; Morelli, X. Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes. *J. Biol. Chem.* **2014**, *289*, 25783−96.
(55) Minskaia, E.; Hertzig, T.; Gorbalenya, A. E.; Campanacci, V.; Cambillau, C.; Canard, B.; Ziebuhr, J. Discovery of an RNA virus 3′->5′ exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 5108−13.
(56) Kang, H.; Bhardwaj, K.; Li, Y.; Palaninathan, S.; Sacchettini, J.; Guarino, L.; Leibowitz, J. L.; Kao, C. C. Biochemical and genetic analyses of murine hepatitis virus Nsp15 endoribonuclease. *J. Virol.* **2007**, *81*, 13587−97.
(57) Bhardwaj, K.; Sun, J.; Holzenburg, A.; Guarino, L. A.; Kao, C. C. RNA recognition and cleavage by the SARS coronavirus endoribonuclease. *J. Mol. Biol.* **2006**, *361*, 243−56.
(58) Zhang, L.; Li, L.; Yan, L.; Ming, Z.; Jia, Z.; Lou, Z.; Rao, Z. Structural and Biochemical Characterization of Endoribonuclease Nsp15 Encoded by Middle East Respiratory Syndrome Coronavirus. *J. Virol.* **2018**, *92*, 92.
(59) Ton, A. T.; Gentile, F.; Hsing, M.; Ban, F.; Cherkasov, A. Rapid Identification of Potential Inhibitors of SARS-CoV-2 Main Protease by Deep Docking of 1.3 Billion Compounds. *Mol. Inf.* **2020**, *39*, No. e2000028.
(60) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296−W303.
(61) Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; Assempour, N.; Iynkkaran, I.; Liu, Y.; Maciejewski, A.; Gale, N.; Wilson, A.; Chin, L.; Cummings, R.; Le, D.; Pon, A.; Knox, C.; Wilson, M. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **2018**, *46*, D1074−D1082.
(62) Tchesnokov, E. P.; Feng, J. Y.; Porter, D. P.; Gotte, M. Mechanism of Inhibition of Ebola Virus RNA-Dependent RNA Polymerase by Remdesivir. *Viruses* **2019**, *11*, 11.
(63) Al-Tawfiq, J. A.; Al-Homoud, A. H.; Memish, Z. A. Remdesivir as a possible therapeutic option for the COVID-19. *Travel Med. Infect Dis* **2020**, *34*, 101615.
(64) Agostini, M. L.; Andres, E. L.; Sims, A. C.; Graham, R. L.; Sheahan, T. P.; Lu, X.; Smith, E. C.; Case, J. B.; Feng, J. Y.; Jordan, R.; Ray, A. S.; Cihlar, T.; Siegel, D.; Mackman, R. L.; Clarke, M. O.; Baric, R. S.; Denison, M. R. Coronavirus Susceptibility to the Antiviral Remdesivir (GS-5734) Is Mediated by the Viral Polymerase and the Proofreading Exoribonuclease. *mBio* **2018**, *9*, 9.
(65) de Wit, E.; Feldmann, F.; Cronin, J.; Jordan, R.; Okumura, A.; Thomas, T.; Scott, D.; Cihlar, T.; Feldmann, H. Prophylactic and therapeutic remdesivir (GS-5734) treatment in the rhesus macaque model of MERS-CoV infection. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 6771−6776.
(66) Gao, Y.; Yan, L.; Huang, Y.; Liu, F.; Zhao, Y.; Cao, L.; Wang, T.; Sun, Q.; Ming, Z.; Zhang, L.; Ge, J.; Zheng, L.; Zhang, Y.; Wang, H.; Zhu, Y.; Zhu, C.; Hu, T.; Hua, T.; Zhang, B.; Yang, X.; Li, J.; Yang, H.; Liu, Z.; Xu, W.; Guddat, L. W.; Wang, Q.; Lou, Z.; Rao, Z. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* **2020**, *368*, 779−782.
(67) Yin, W.; Mao, C.; Luan, X.; Shen, D. D.; Shen, Q.; Su, H.; Wang, X.; Zhou, F.; Zhao, W.; Gao, M.; Chang, S.; Xie, Y. C.; Tian, G.; Jiang, H. W.; Tao, S. C.; Shen, J.; Jiang, Y.; Jiang, H.; Xu, Y.; Zhang, S.; Zhang, Y.; Xu, H. E. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* **2020**, *368*, 1499−1504.
(68) Kim, J. Y. Letter to the Editor: Case of the Index Patient Who Caused Tertiary Transmission of Coronavirus Disease 2019 in Korea: the Application of Lopinavir/Ritonavir for the Treatment of COVID-19 Pneumonia Monitored by Quantitative RT-PCR. *J. Korean Med. Sci.* **2020**, *35*, No. e88.

https://dx.doi.org/10.1021/acs.jcim.0c00821
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(69) Khot, W. Y.; Nadkar, M. Y. The 2019 Novel Coronavirus Outbreak - A Global Threat. *J. Assoc Physicians India* **2020**, *68*, 67−71.
(70) Zeng, Y. M.; Xu, X. L.; He, X. Q.; Tang, S. Q.; Li, Y.; Huang, Y. Q.; Harypursat, V.; Chen, Y. K. Comparative effectiveness and safety of ribavirin plus interferon-alpha, lopinavir/ritonavir plus interferon-alpha, and ribavirin plus lopinavir/ritonavir plus interferon-alpha in patients with mild to moderate novel coronavirus disease 2019: study protocol. *Chin Med. J. (Engl)* **2020**, *133*, 1132−1134.
(71) Wang, Z.; Chen, X.; Lu, Y.; Chen, F.; Zhang, W. Clinical characteristics and therapeutic procedure for four cases with 2019 novel coronavirus pneumonia receiving combined Chinese and Western medicine treatment. *BioSci. Trends* **2020**, *14*, 64−68.
(72) Martinez, M. A. Compounds with Therapeutic Potential against Novel Respiratory 2019 Coronavirus. *Antimicrob. Agents Chemother.* **2020**, *64*, 64.
(73) Rebello, K. M.; Andrade-Neto, V. V.; Zuma, A. A.; Motta, M. C. M.; Gomes, C. R. B.; de Souza, M. V. N.; Atella, G. C.; Branquinha, M. H.; Santos, A. L. S.; Torres-Santos, E. C.; d'Avila-Levy, C. M. Lopinavir, an HIV-1 peptidase inhibitor, induces alteration on the lipid metabolism of Leishmania amazonensis promastigotes. *Parasitology* **2018**, *145*, 1304−1310.
(74) Kitchen, V. S.; Skinner, C.; Ariyoshi, K.; Lane, E. A.; Duncan, I. B.; Burckhardt, J.; Burger, H. U.; Bragman, K.; Pinching, A. J.; Weber, J. N. Safety and activity of saquinavir in HIV infection. *Lancet* **1995**, *345*, 952−5.
(75) Tatum, H.; Thuluvath, P. J.; Lawitz, E.; Martorell, C.; DeMicco, M.; Cohen, S.; Rustgi, V.; Ravendhran, N.; Ghalib, R.; Hanson, J.; Zamparo, J.; Zhao, J.; Cooney, E.; Treitel, M.; Hughes, E. A randomized, placebo-controlled study of the NS5B inhibitor beclabuvir with peginterferon/ribavirin for HCV genotype 1. *J. Viral Hepat* **2015**, *22*, 658−64.
(76) Gentile, I.; Zappulo, E.; Buonomo, A. R.; Maraolo, A. E.; Borgia, G. Beclabuvir for the treatment of hepatitis C. *Expert Opin. Invest. Drugs* **2015**, *24*, 1111−21.
(77) Cook, N. J.; Li, W.; Berta, D.; Badaoui, M.; Ballandras-Colas, A.; Nans, A.; Kotecha, A.; Rosta, E.; Engelman, A. N.; Cherepanov, P. Structural basis of second-generation HIV integrase inhibitor action and viral resistance. *Science* **2020**, *367*, 806−810.
(78) Kadam, R. U.; Wilson, I. A. Structural basis of influenza virus fusion inhibition by the antiviral drug Arbidol. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 206−214.
(79) Gordon, C. J.; Tchesnokov, E. P.; Feng, J. Y.; Porter, D. P.; Gotte, M. The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J. Biol. Chem.* **2020**, *295*, 4773−4779.
(80) Sheahan, T. P.; Sims, A. C.; Leist, S. R.; Schafer, A.; Won, J.; Brown, A. J.; Montgomery, S. A.; Hogg, A.; Babusis, D.; Clarke, M. O.; Spahn, J. E.; Bauer, L.; Sellers, S.; Porter, D.; Feng, J. Y.; Cihlar, T.; Jordan, R.; Denison, M. R.; Baric, R. S. Comparative therapeutic efficacy of remdesivir and combination lopinavir, ritonavir, and interferon beta against MERS-CoV. *Nat. Commun.* **2020**, *11*, 222.
(81) Kumalo, H. M.; Bhakat, S.; Soliman, M. E. Theory and applications of covalent docking in drug discovery: merits and pitfalls. *Molecules* **2015**, *20*, 1984−2000.
(82) Zhao, Z.; Bourne, P. E. Progress with covalent small-molecule kinase inhibitors. *Drug Discovery Today* **2018**, *23*, 727−735.
(83) Bhattacharya, D.; Nowotny, J.; Cao, R.; Cheng, J. 3Drefine: an interactive web server for efficient protein structure refinement. *Nucleic Acids Res.* **2016**, *44*, W406−9.
(84) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2009**, *31*, 455−61.

https://dx.doi.org/10.1021/acs.jcim.0c00821
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX