ARTICLE

Discovery of potential small molecular SARS-CoV-2 entry blockers targeting the spike protein

Lin Wang[1,2], Yan Wu[3], Sheng Yao[4,5], Huan Ge[6], Ya Zhu[7], Kun Chen[7], Wen-zhang Chen[1], Yi Zhang[1], Wei Zhu[1], Hong-yang Wang[2], Yu Guo[8], Pei-xiang Ma[1], Peng-xuan Ren[1,2], Xiang-lei Zhang[1,2], Hui-qiong Li[1,2], Mohammad A. Ali[9], Wen-qing Xu[2], Hua-liang Jiang[1,7], Lei-ke Zhang[3], Li-li Zhu[6,2], Yang Ye[4,5], Wei-juan Shang[3] and Fang Bai[1,2]

An epidemic of pneumonia caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is spreading worldwide. SARS-CoV-2 relies on its spike protein to invade host cells by interacting with the human receptor protein Angiotensin-Converting Enzymes 2 (ACE2). Therefore, designing an antibody or small-molecular entry blockers is of great significance for virus prevention and treatment. This study identified five potential small molecular anti-virus blockers via targeting SARS-CoV-2 spike protein by combining in silico technologies with in vitro experimental methods. The five molecules were natural products that binding to the RBD domain of SARS-CoV-2 was qualitatively and quantitively validated by both native Mass Spectrometry (MS) and Surface Plasmon Resonance (SPR). Anti-viral activity assays showed that the optimal molecule, H69C2, had a strong binding affinity (dissociation constant K_D) of 0.0947 μM and anti-virus IC_{50} of 85.75 μM.

Keywords: SARS-CoV-2; natural products; virtual screening; spike protein; protein-protein interaction modulators

INTRODUCTION

A novel coronavirus connected to a cluster of acute respiratory illnesses, named COVID-19 by the World Health Organization (WHO), has spread worldwide and resulted in a pandemic. Genetic analysis has revealed that the virus is closely related to SARS-CoV with 79.6% sequence identity [1, 2]. Hence, it was named SARS-CoV-2 by the International Committee on the Taxonomy of Viruses. Up to February 27th of 2021, 112,902,746 patients have been confirmed worldwide, and 2,508,679 resulting deaths (https://covid19.who.int/). Unfortunately, the number of patients is increasing, but there is no effective drug treatment so far.

SARS-CoV-2 and SARS-CoV are beta coronaviruses, which are enveloped, single, and positive-stranded RNA viruses [3]. Their genome RNA encodes a part-structural replicase polyprotein and structural proteins, which include Spike (S), Envelope (E), Membrane (M), and Nucleocapsid (N) proteins [4]. Recent studies have systematically constructed the protein-protein interaction (PPI) map between SARS-CoV-2 and human cells, which have provided several potential PPI interfering strategies to inhibit the virus [5].

S protein is a major structural protein of SARS-CoV-2, that is essential for the interaction of the virions with host cell receptors, i.e., ACE2, and subsequent fusion of the viral envelope with the host cell membrane to allow infection [1, 6]. It consists of two subunits: S1 and S2. The S1 subunit contains a receptor-binding domain (RBD) that triggers the infection by its binding to a host cell receptor [7, 8] and another N-terminal domain. The S2 subunit is responsible for mediating fusion between the viral host cell membranes [9, 10]. Delineating virus entry mechanisms [11], designing molecular agents, that bind to the ACE2 binding site of the S protein to interfere with the protein-protein interaction (PPI) of ACE2 and S protein, may be a promising strategy to prevent viral infection. In addition, a determination of the cryo-EM structure of the S protein has provided the structural basis to develop molecular binders [6].

ACE2 is an angiotensin-converting enzyme, whose *Apo* state and *Holo* complex with small molecular inhibitor MLN-4760, have been elucidated [12]. Recently, many studies have reported that some FDA-approved drugs exhibit potential activity to inhibit the entry of SARS-CoV-2 into host cells [13, 14]. Some of these inhibitors can directly block the binding of ACE2 to S protein, such as clemastine [15] and heparan sulfate [16], whereas others act by inhibiting Transmembrane Serine Protease 2 (TMPRSS2) [13], which is responsible for S protein priming. The latter appear to

[1]Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, Shanghai 201210, China; [2]School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China; [3]State Key Laboratory of Virology, Wuhan Institute of Virology, Center for Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan 430064, China; [4]Department of Natural Products Chemistry, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; [5]State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; [6]State Key Laboratory of Bioreactor Engineering, Shanghai Key Laboratory of New Drug Design, School of Pharmacy, East China University of Science and Technology, Shanghai 200237, China; [7]CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; [8]College of Pharmacy and State Key Laboratory of Medicinal Chemical Biology, Nankai University, Tianjin 300071, China and [9]Department of Botany and Microbiology, College of Science, King Saud University, Riyadh, Saudi Arabia
Correspondence: Lei-ke Zhang (zhangleike@wh.iov.cn) or Li-li Zhu (zhulfl@ecust.edu.cn) or Yang Ye (yye@mail.shcnc.ac.cn) or Wei-juan Shang (Shangweijuan@wh.iov.cn) or Fang Bai (baifang@shanghaitech.edu.cn)
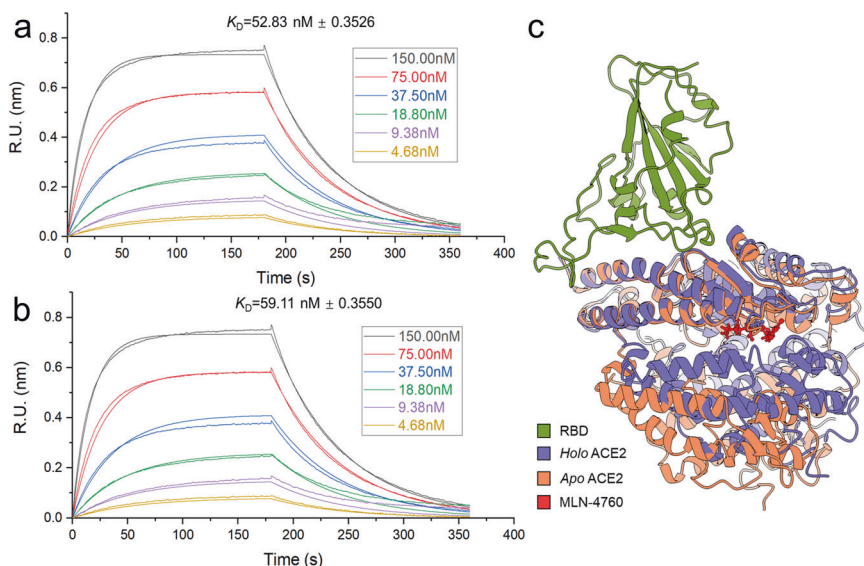These authors contributed equally: Lin Wang, Yan Wu, Sheng Yao, Huan Ge.

Fig. 1 Comparison of RBD binding affinities with *Apo* ACE2 and ACE2 treated with MLN-4760 by BLI. Association and dissociation curves of RBD with *Apo* ACE2 (**a**) or ACE2 treated with MLN-4760 (**b**) in a concentration range between 4.68 nM and 150 nM. The number that follows the ±sign is the standard deviation (SD). (**c**) Comparison of binding conformations between the binding complex of RBD with *Apo* ACE2 (PDB code: 1R42) and *Holo* ACE2 (PDB code: 1R4L). The conformational alignment was performed on the first 100 residues at the N-terminus of ACE2.

have attracted much attention in drug development to target SARS-CoV-2, whereas the former, i.e., drugs that interfere with the interface of ACE2 and RBD [17], have received little attention. A large number of antibodies, that potently neutralize the ACE2-RBD binding interface, have been reported [18–21]. However, small molecular blockers, which have lower costs and more flexible clinical use, have seldom been reported.

Interestingly, the inhibitor-binding ACE2 structure exhibits a large hinge-bending motion in the *Apo* state of ACE2 [12], which suggests that the binding between ACE2 and S protein is affected allosterically. Hence, it was expected that ACE2 inhibitors could be used to treat SARS-CoV-2. To further verify this hypothesis, we used a Bio-Layer Interferometry (BLI) Octet® RED 96 platform (Sarotirus, Fremont, California, USA) to characterize the binding interactions between RBD and ACE2 with or without ACE2 inhibitor MLN-4760 (Fig. 1a, b). MLN-4760 is a very potent ACE2 inhibitor with an IC_{50} of 0.44 nM [22]. However, our results indicated that the ACE2 inhibitor appeared to have a negligible effect on the binding affinity between RBD and ACE2. The dissociation constant K_D between the two proteins is 52.83 nM versus 59.11 nM treated with MLN-4760. However, although ACE2 exhibits some conformational distortion while bound to MLN-4760, there were no obvious changes in its binding interface to RBD (Fig. 1c). Overall, designing small molecular blockers to directly disrupt the binding of ACE2-RBD may be the most feasible approach to prohibit the infection.

Designing small molecular modulators to interfere with PPIs, is a challenging task in drug design and development, which is hampered by the large, flat, and featureless areas of interfaces [23, 24]. This also occurs when developing inhibitors to prevent binding of the RBD domain of the S protein to ACE2 [25]. Hotspot residues are groups of essential residues that dominantly contribute to binding affinities for PPIs [26, 27]. Detecting and targeting such important residues can greatly improve the success rate of identifying lead compounds. Thus, our hotspot identification module in Fd-DCA [28] was used to search over the cryo-EM structure of the RBD of S protein (PBD code:6M17) [7]. As shown in Fig. 2, multiple residues that may contribute to binding affinities between the RBD of the spike protein to ACE2 were identified and
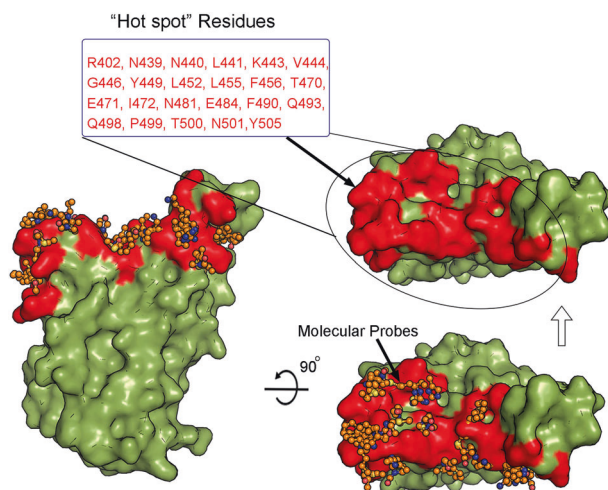

Fig. 2 Identification of hotspot residues (highlighted in red and listed in the text box) on the binding sites of the RBD to S protein by Fd-DCA. The orange stick-ball model represents the predicted low-energetic bound conformations of fragment-sized molecular probes introduced in Fd-DCA. The dark green surface model is the RBD domain of S protein.

focused on for our molecular docking simulations to search for candidate RBD binders.

On the basis of the previously reported structural information of binding complexes of ACE2–RBD [7, 8], some critical interactions between the two proteins can be found, i.e., Q498, K417, and Y453 of RBD, which form hydrogen bonds with Q24, Q42, K31, D30, and H34 of ACE2, and Y489 and F486 of RBD, which interact with F28, Y83, M82, and L79 of ACE2 via a cluster of hydrophobic contacts. Some residues were also considered as potential hotspot residues on the RBD for protein–protein or protein-ligand binding (Fig. 2). Therefore, they were regarded as the druggable binding pocket to search for candidate small molecular binders of RBD.

MATERIALS AND METHODS

Protein expression, and purification, and Bio-layer interferometry assay

The codon-optimized wild-type cDNA of SARS-CoV-2 receptor-binding domain (RBD) (residues 333–530) was synthesized by GENEWIZ. The SARS-CoV-2 RBD was expressed using the Bac-to-Bac baculovirus system. The supernatant of cell culture containing the secreted removal of glycosylated RBD was harvested 72 h after infection and concentrated and RBD was captured by Ni-NTA resin (GE Healthcare). The resin was washed 5–6 times with 30 mL of wash buffer (25 mM Tris, 150 mM NaCl, 40 mM imidazole, pH 7.5), the target protein was eluted with elution buffer containing 25 mM Tris, 150 mM NaCl, 500 mM imidazole, pH 7.5. The protein was further purified on a Superdex S75 (GE Healthcare) column equilibrated with 25 mM Tris, 150 mM NaCl, pH 7.5. The RBD protein was biotinylated by using Biotin-Protein Ligase kit (GeneCopoeia™), further purified on the Superdex S75 column, and concentrated to 15 mg/mL.

Bio-Layer Interferometry assay was conducted at 30 °C in PBS, 0.02% Tween 20, 0.1% BSA, pH 7.4, using an Octet Red 96 instrument (Sarotirus, Fremont, California, USA). Sensors were loaded with 10 µg/mL ligand (Fc-RBD). The dissociation wells were used only once to ensure the potency of the buffer. To characterize whether MLN4760 could inhibit ACE2 binding to immobilized Fc-RBD, the proA sensors, which were coated with Fc-RBD, were exposed to 150 nM, 75 nM, 37.5 nM, 18.75 nM, 9.375 nM, 4.68 nM ACE2 with 1.5 µM MLN4760. The proA sensors, which were coated with Fc, were exposed to 150 nM, 75 nM, 37.5 nM, 18.75 nM, 9.375 nM, 4.68 nM ACE2 with 1.5 µM MLN4760. The correction of any systematic baseline drift was accomplished by subtracting the shift recorded for sensors loaded with ligand but no analyte.

More details are shown in S1 File.

Ligand structure preparation and virtual screening

Natural products in the database were first prepared for docking using the LigPrep ligand preparation module (Schrödinger, Inc), which generates multiple minimized conformations and protonation/tautomerization states with default settings. Then, virtual screening was performed using a molecular docking module, named Glide 8.5 [29]. The same protein structure (PBD code:6M17 [7]) used in hotspot searching was applied to carry out the molecular docking simulations. The structure was first prepared and refined by Protein Preparation Wizard [30] of Schrödinger 2019-4 with default parameter settings that included assigning bond orders, adding hydrogens, and assigning partial charges. A grid for docking simulations was generated with a size of $20 \times 20 \times 20$ Å^3 cubes by centering on the basis of the coordination of Q498 and Y489 on the RBD surface to cover the hotspot residues identified by Fd-DCA. During the virtual screening step, two docking methods, high throughput virtual screening (HTVS) and standard precision (SP) were used for preliminary rough docking and more accurate evaluation. The number of ligands retained from HTVS was set to 30% and the top 50% of SP docking ligands were eventually exported for further visual inspection.

Mass spectrometry

MS was performed by a Thermo Fisher fusion orbitrap with 20 µM protein and 1 mM compound in ammonium acetate buffer. Our protein sample of RBD (aa. 319–591) for MS analysis was provided by Dr. Guo from Nan Kai University (the detailed protein expression method is shown in supplementary materials).

Surface plasmon resonance

RBD was diluted in sodium acetate solution (pH 4.5) to a final concentration of 50 µg/mL and was then immobilized covalently on a CM5 sensor chip. The final immobilization level was 4430.3 resonance units. The measurement was run using PBS with
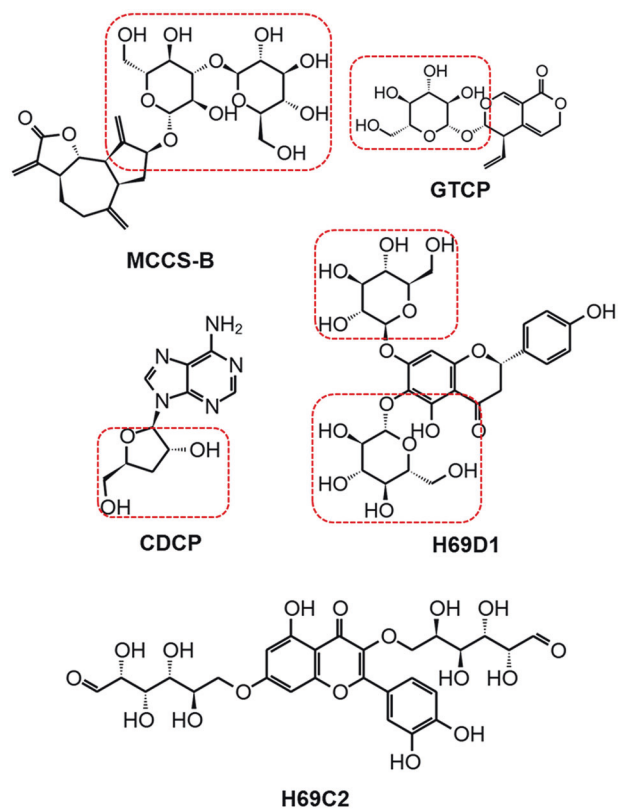

Fig. 3 Compound structures of the selected five candidate molecular binders of RBD. The groups of carbasugars were boxed by dashed red lines.

0.005% (v/v) surfactant P20 (pH 7.4) and 1% DMSO as the running buffer. The compound was diluted in the running buffer from the highest concentration. All compound measurements were performed at a flow rate of 30 µL/min. Data processing and analysis were performed using the BIA evaluation 1.1 software.

Anti-viral activity assay using native SARS-CoV-2 virus

To evaluate the anti-viral efficacy of these compounds, Vero cells were cultured overnight in 48-well plates at a density of 5×10^4 cells/well. The cells were pretreated with various doses of the indicated compounds for 1 h, and then the virus (MOI of 0.01) was added. At 48 h p.i., the culture supernatant was collected and treated with lysis buffer (Takara, Cat. no. 9766) for quantification as described in the previous study [31].

For cytotoxicity assays, Vero cells were suspended in a growth medium in 96-well plates at a density of 1×10^4 cells/well. The next day, appropriate concentrations of compounds were added to the medium. After 24 h, the relative numbers of surviving cells were measured by the CCK8 assay (Beyotime, China) in accordance with the manufacturer's instructions.

RESULTS

Virtual screening of the natural product database

Natural products are rich sources for drug development [32, 33]. Approximately 60% of drugs in the market have originated from natural sources [34]. In this study, we searched for lead compounds in an in-house natural product database from Dr. Ye Yang, which contained 2467 compounds. As a result, five candidate compounds were selected for experimental validation, as shown in Fig. 3.

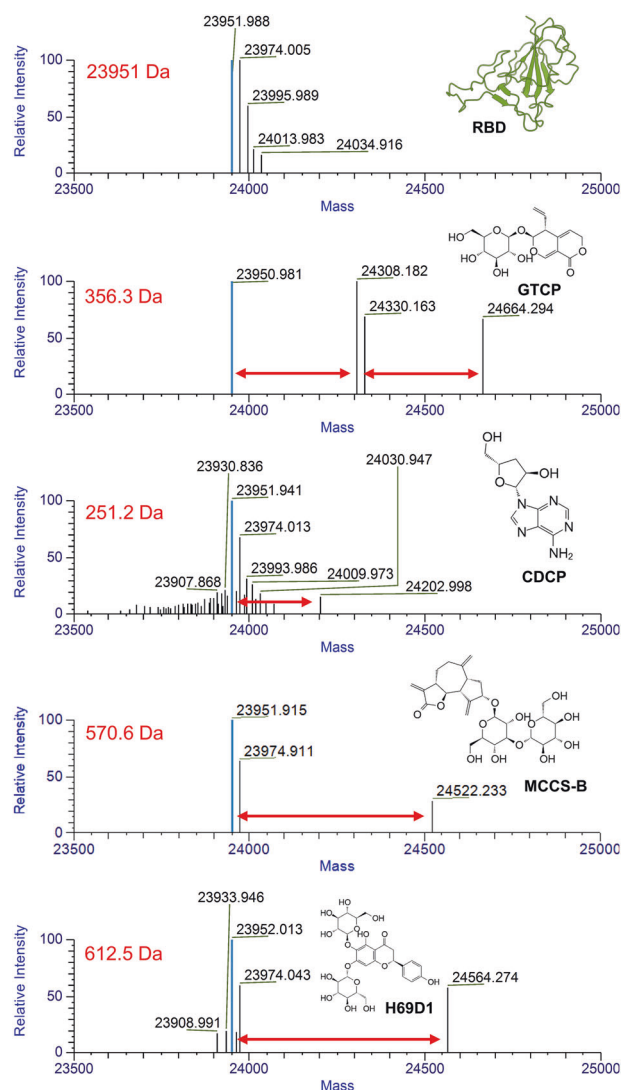Among the five compounds, four contained groups of carbasugars. Gentiopicrin (GTCP), is a precursor of gentiogenal

Fig. 4 Mass spectrometric determination of the interaction between RBD and our compounds. Differences in molecular weights between RBD protein and complexes are indicated by red arrows.

[35, 36]. Cordycepin (CDCP) is an excellent anti-cancer lead compound because of its various types of bioactivities, i.e., AMP-activated protein kinase (AMPK) agonist activity [37], inhibiting the activity of mTORC1 function [38], downregulating HIF-1α expression in tumor cells [38] and activating autophagy [39]. Potential bioactivities of the other three compounds, MCCS-B, H69D1, and H69C2, have not been reported to our knowledge.

Mass spectrometry of candidate molecular binders
To determine whether the compounds shown in Fig. 3 bound to the S protein as predicted, MS and SPR were employed for the binding evaluation. As a result, four of the five candidates presented signals of binding to RBD, as shown in Fig. 4. H69C2 was not examined because of its poor solubility in ammonium acetate buffer.

Surface plasmon resonance of candidate molecular binders
Next, we performed SPR experiments with a Biacore T200 (GE Healthcare) to determine the binding affinities and kinetics of the five compounds for RBD. The protein sample of RBD (aa. 319–591) used in the experiments was obtained from the National Protein Center of China (the method of recombinant protein preparation is described in the supplementary materials).

As shown in Fig. 5, all compounds exhibited concentration-dependent binding to the RBD of the S protein. We next obtained dissociation constant K_D values. Compared with the other three compounds, CDCP and GTCP showed weaker binding affinities (CDCP was 20.460 μM and GTCP was 29.559 μM). H69D1, MCCS-B, and H69C2 showed obviously stronger binding affinities than CDCP and GTCP, with a dissociation constant K_D values of 2.154, 3.560, and 0.0947 μM, respectively. Binding kinetics data of these compounds are shown in Fig. 5.

Binding model analysis
Interestingly, in accordance with the MS results shown in Fig. 4, GTCP was bound to RBD in the 2:1 model, i.e., one RBD molecule accommodated two GTCP molecules. To further determine how the compound interacted with RBD, we compared all predicted possible binding poses from docking simulations. As a result, there may be two potential binding sites in RBD of ACE2 to attach two GTCP molecules simultaneously (Fig. 6). While the rest of the compounds bound to RBD in the 1:1 model, the lowest-energetic binding confirmation for each of them was maintained. To refine the predicted binding conformation and relax the predicted binding complexes from the molecular docking simulations, the Molecular Mechanics Generalized Born and Solvent Accessibility module with the OPLS_2005 force field of Schrödinger 2019-4 was used for energy minimization.

By analyzing the refined binding conformations, as shown in Fig. 6, we found that the binding sites of these five small molecular blockers may be distributed on three subregions of the large binding site of RBD. CDCP, H69D1, and one of the binding poses of GTCP bound in the same subregion of RBD. A common hydrogen bond was observed between the E484 of RBD and CDCP, H69D1, and GTCP. Excluding this, a series of hydrophobic contacts formed with aromatic residues, such as Y351, Y449, and F490, or alkyl chain residues, i.e., L492 and L452. Our hotspot residue identification results showed that these residues i.e., Y449, F490, and L452, which strongly interacted with the compounds, were also potential hotspot residues in protein–protein interactions between S protein and ACE2. This indicated that these compounds competitively interacted with three hotspot residues and interfered with the interaction between them and ACE2. Excluding the common binding interaction features, an additional hydrogen bond or network was identified between H69D1 and the residues, such as T470 and Q493, which may make H69D1 bind much stronger to RBD (around 10 times stronger in accordance with SPR experiments) than CDCP or GTCP.

For GTCP, its second potential binding site was located around Q498 and Y453, as shown in Fig. 6. Multiple hydrogen bonds formed with the side chain of Y453 and the backbone of G502 and G496 were identified. Although two possible binding conformations of GTCP were found, its binding affinity measured by SPR was still relatively lower than that of the other small molecular blockers.

In accordance with the predictions, the binding site of H69C2 (Fig. 6) was adjacent to the binding site of H69D1 and the other two. Several hydrogen bonds around the molecule formed between the side chains of R403, D405, E406, Q493, and S494. Moreover, its flavonoid skeleton played a role in stabilizing the binding via its hydrophobic interaction with hydrophobic residues, such as L455, Y495, and Y505 (Fig. 6). Among these critical residues, several of them, i.e., R403, Q493, L45, and Y505, were predicted as hotspot residues. These important molecular interactions conferred H69C2 with the best binding affinity and binding kinetics (around 20 to 300 times stronger in accordance with SPR experiments) compared with the other four small molecular blockers. This further confirmed the importance of designing molecules to interact with hotspot residues and interfere with protein–protein interactions.

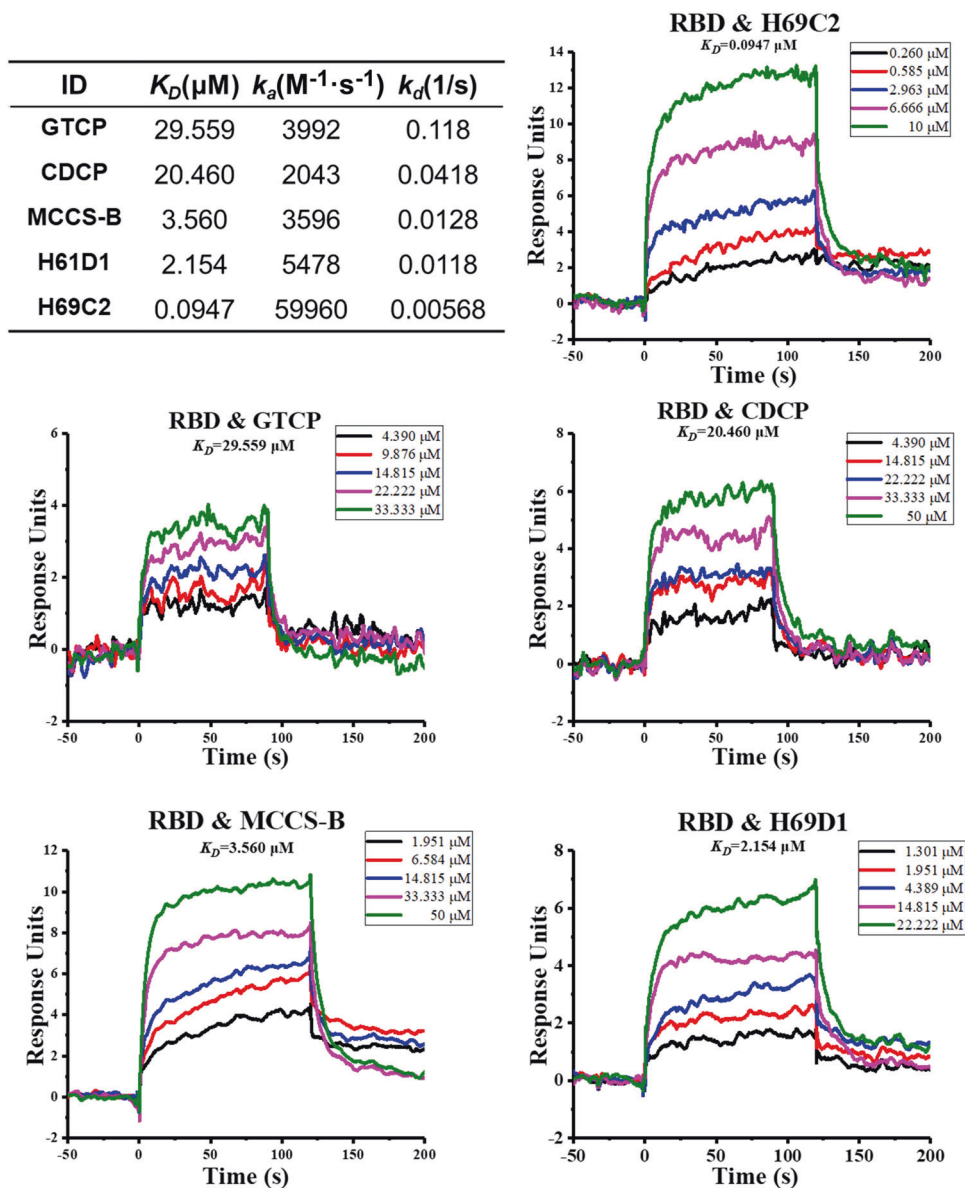| ID	K_D(μM)	k_a(M^{-1}·s^{-1})	k_d(1/s)
GTCP	29.559	3992	0.118
CDCP	20.460	2043	0.0418
MCCS-B	3.560	3596	0.0128
H61D1	2.154	5478	0.0118
H69C2	0.0947	59960	0.00568

Fig. 5 Binding kinetics and affinity analysis of the five compounds for binding to the RBD domain of S protein of SARS-CoV-2 using a Biacore T200. Multiple concentrations of the compounds were injected into the equipment to fit the binding data, which are represented by different colors in each subgraph.

Compound MCCS-B, which was identified to bind to the area adjacent to H69D1 (Fig. 6), also had a moderate binding affinity for RBD with a KD value of 3.560 μM. This strong binding affinity may be attributed to molecular interactions, such as hydrogen bonds between it and the hotspot residues of Q493, N501, and Q498, as well as hydrophobic interactions with hotspot residues of L452, as an example.

H69C2 for SARS-CoV-2 treatment as an entry inhibitor
By measuring viral RNA in culture supernatant by quantitative RT-PCR, we found that H69C2 inhibited SARS-CoV-2 replication in a dose-dependent manner (Fig. 7a), with an IC50 value of 85.75 μM. We also performed cytotoxicity assays using H69C2 and found that the CC50 value of H69C2 was above 250 μM (Fig. 7b). The anti-SARS-CoV-2 activity of H69C2 was also evaluated by monitoring the intracellular SARS-CoV-2 NP level using immunofluorescence. We found that the intracellular SARS-CoV-2 NP level was decreased after treatment with H69C2 (Fig. 7c), which indicates that H69C2 inhibited SARS-CoV-2 replication in vitro.

DISCUSSION
As shown in Fig. 8, the small-molecular blockers, derived from natural products, bound to the surface of RBD, which hindered the binding of RBD and ACE2 and thus protected healthy cells and prevented further deterioration in the early stage of infection. By combining computational methods and experiments, we successfully identified five small molecular binders of the RBD of S protein, which bound on the RBD to block binding of ACE2 and prevent the infection by viruses. These compounds may competitively interact with hotspot residues, that were identified to contribute to the majority binding affinity of RBD to ACE2. Two of them showed low micromolar binding and H69C2 showed the strongest binding affinity of 0.0947 μM. In addition, H69C2 was validated to block viral infection in vitro with an IC50 of 85.75 μM.

Apart from designing anti-viral agents by targeting S protein, there are several highly potent anti-virus inhibitors that target Mpro and RdRp. Inhibitors of Mpro and RdRp have been proposed for combining to the treatment of COVID-19. Multi-target drug combination therapy has become a promising treatment strategy.
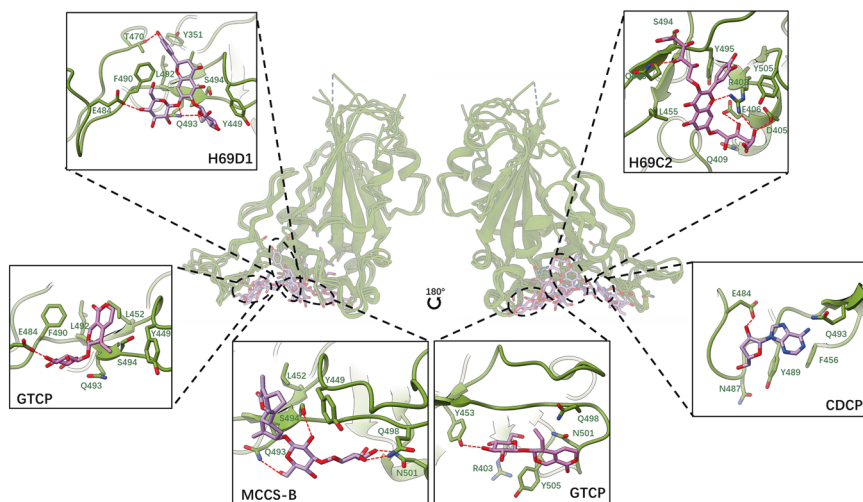
Fig. 6 Binding model of the five compounds. Global overview of compounds on RBD and details of the five compound binding models are linked by the dashed line. The hydrogen bonds between the RBD sidechain and small molecular blockers are shown as a dashed line.
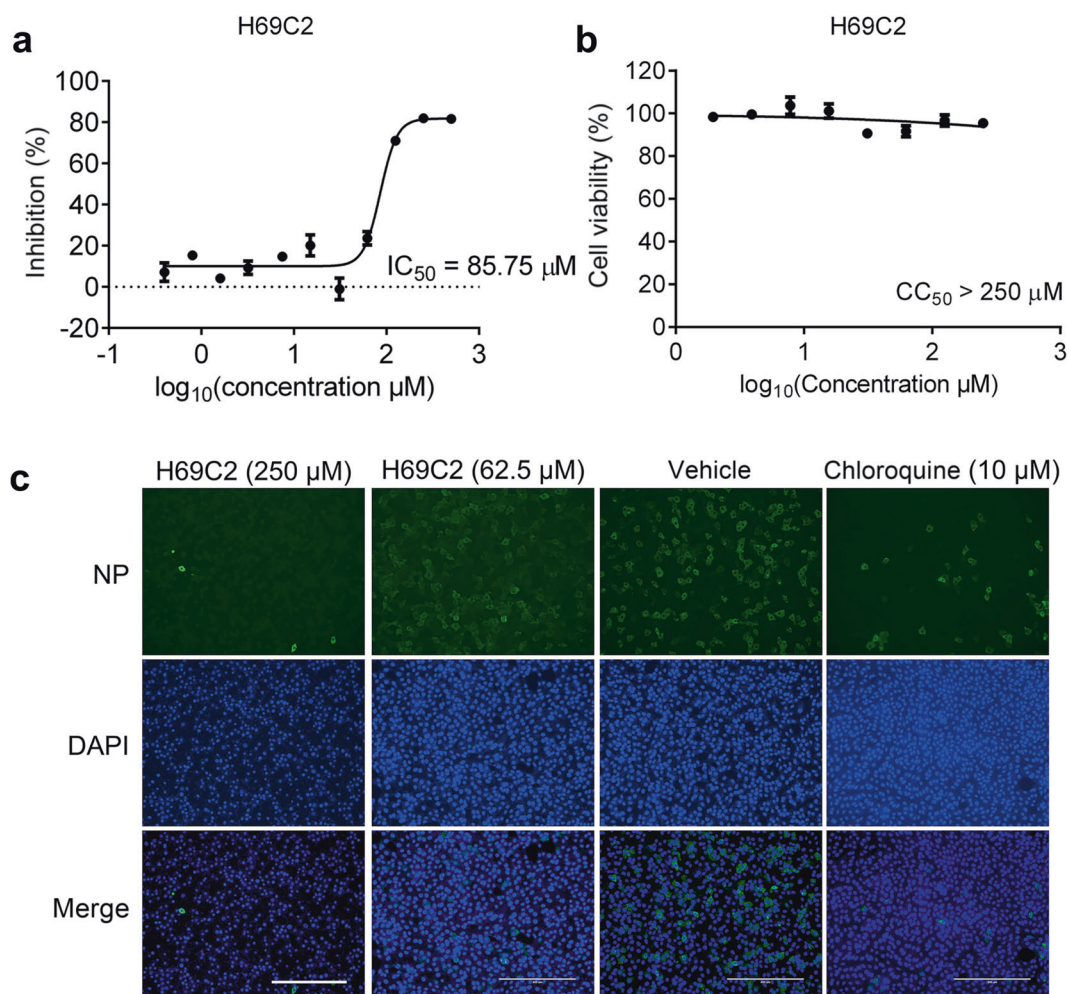

Fig. 7 Inhibition of SARS-CoV-2 by H69C2 in Vero E6 cells. Vero E6 cells infected with SARS-CoV-2 at an MOI of 0.01 were treated with various concentrations of H69C2. (**a**) Quantitative RT-PCR assays were performed to measure the viral copy number in the cellular supernatant. The y-axis indicates percentage inhibition of the virus relative to the sample treated with DMSO (vehicle). (**b**) Cell viability assay. The y-axis represents the percentage of cell viability relative to the sample treated with DMSO (vehicle). (**c**) Immunofluorescence images of intracellular NP. At 24 h post-infection, cells were fixed, and intracellular NP levels were monitored by immunofluorescence. Scale bars, 400 μm. Data are shown as the mean ± s.e.m., $n = 3$.
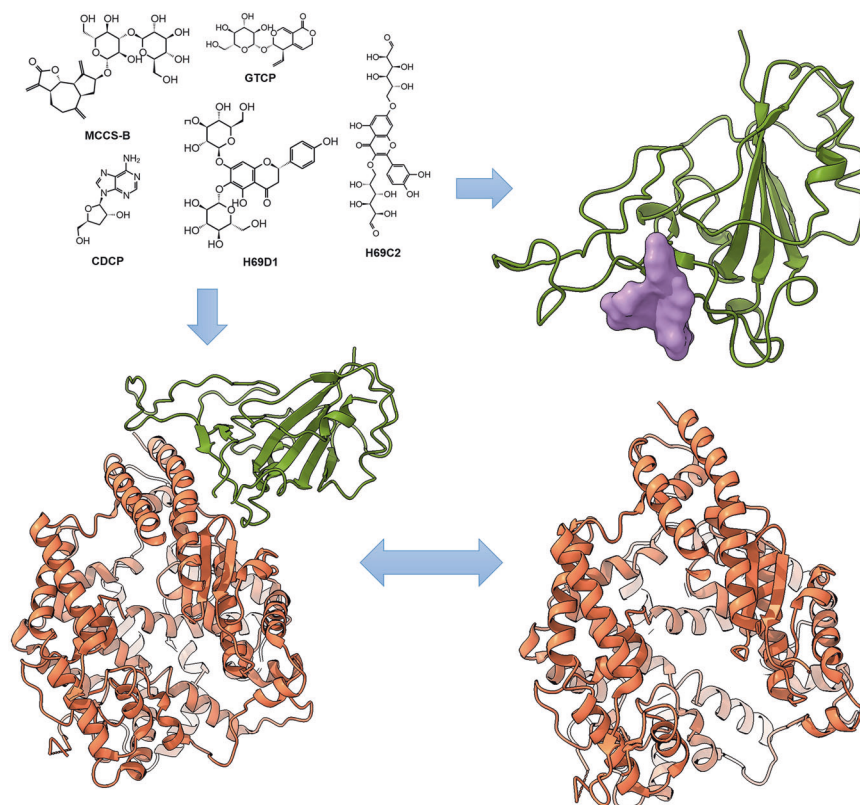
Fig. 8 Competitive binding mechanism of the identified small molecules over ACE2 to bind to the RBD of S protein. Small molecule blockers bind to the RBD and interfere with RBD binding to ACE2.

Table 1. Predicted ADMET properties of identified natural products.

ADMET properties	H69C2	H69D1	MCCS-B	CDCP	GTCP
Human intestinal absorption	+	+	−	+	−
Blood–brain barrier	−	−	+	−	−
P-glycoprotein substrate and inhibition	+−	−−	−−	−−	−−
CYPs substrate and inhibition	+−−−−	+−−−−	+−−−−	−−−−−−	+−−−−
Ames mutagenesis	+	−	−	−	+
Hepatotoxicity	+	+	−	+	−

+/− indicate positive/negative. Symbols on the P-glycoprotein substrate and inhibition from left to right indicate properties of the substrate and inhibition. Symbols on the CYPs substrate and inhibition from left to right indicate the properties of CYP3A4, CYP2C9, CYP2D6, CYP3A4, CYP2C19, and CYP1A2, respectively.

The strategy is applicable to RBD molecular blockers, such as H69C2, which have different anti-virus effects from Mpro and RdRp inhibitors. There are two phases in the cell infection of SARS-CoV-2, namely cell entry, and proliferation. Inhibitors of Mpro, RdRp, and host proteases function in the stage of proliferation, but our molecular blockers can inhibit the stage of virus infection. Hence, it is expected that the combinations of RBD molecular blockers and other inhibitors will enhance the anti-viral effect. The cytokine release syndrome is the major cause of the fatal outcomes of severe COVID-19 patients [40, 41]. As discussed above, multi-target drug combination therapy is becoming a promising treatment. In addition to targets of SARS-CoV-2, targets of the host can also be considered, such as melatonin [42], and other anti-inflammatory agents.

In addition to designing RBD blockers that directly disrupt the binding between ACE2 and S protein, another approach is to identify the allosteric binding site and discover allosteric modulators, such as toremifene, to indirectly destabilize the binding between two proteins [43]. Recently, many advances have been made in the design of vaccine epitopes [44, 45], monoclonal antibodies [46, 47], and peptides [48] that target the S protein. Small molecular blockers that also bind to the epitopes of the spike are promising for COIVD-19 treatment.

So far, there are 930 naturally occurring missense mutations in the SARS-CoV-2 S protein reported in the GISAID database. Four common mutations among them are located on the interface between S protein and ACE2, which include K417N, L452R, E484K/Q, and N501Y (https://bigd.big.ac.cn/ncov/). Several studies have shown that these mutations allow escape from vaccines, monoclonal antibodies, and convalescent plasma [49–52]. In accordance with our predicted binding mode of H69C2 to S protein, we found that L452 and E484 were not involved in the interaction with H69C2, and thus may not affect the binding of H69C2. We performed residue scanning simulations using Schrödinger 2021-1 with the OPLS4 force field to evaluate the effect of K417N and N501Y on the binding affinities of H69C2 and ACE2 to the RBD domain of the S protein, respectively. The mutation of K417N may minimally affect the binding between ACE2 and RBD ($\Delta\Delta G = \Delta G_{RBD_{K417N}-ACE2} - \Delta G_{RBD_{WT}-ACE2} = 0.16$), but N501Y increased their binding to a certain extent($\Delta\Delta G = \Delta G_{RBD_{N501Y}-ACE2} - \Delta G_{RBD_{WT}-ACE2} = -1.42$). Similar studies have been performed to monitor the binding affinities between the mutants of RBD and H69C2. These computationally obtained observations agreed well with an experimental study [53]. Interestingly, we found that both mutations increased the binding between H69C2 and RBD, with higher binding affinities ($\Delta\Delta G = \Delta G_{RBD_{K417N}-H69C2} - \Delta G_{RBD_{WT}-H69C2} = -1.60$ and

$\Delta\Delta G = \Delta G_{RBD_{N501Y}-H69C2} - \Delta G_{RBD_{WT}-H69C2} = -5.67$), respectively. Hence, H69C2 was also effective against the mutants.

This study also predicted the ADMET properties of the identified natural products using the admetSAR webserver [54]. As shown in Table 1, H69C2, H69D1, and CDCP were predicted to be absorbed by the intestines. MCCS-B could pass the blood–brain barrier. In terms of metabolic properties, CDCP had an easier metabolic reaction than the other products. H69C2 inhibited P-glycoprotein and CYP3A4. Unfortunately, H69C2 and H69D1 might cause Ames mutagenesis and hepatotoxicity, which indicates that they should be chemically modified.

Overall, this study computationally identified critical druggable binding sites on the interface between ACE2 and S protein and identified five potential small molecular blockers that target this site to disrupt the binding of these two proteins, which provides a type of chemical scaffold for further optimization and evaluation.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

FB, LKZ, YY, WJS, LLZ, and HLJ designed research; LW, PXR, HQL, MAA, and FB performed calculations and analyzed data; SY and YY prepared compound samples; YG and PXM prepared protein samples; WZC and WZ performed Mass Spectrometry experiments and analyzed the results; YZ performed Bio-Layer Interferometry analysis and analyzed data; HG, HYW, KC, XLZ, YZ, WQX, and LLZ performed Surface Plasmon Resonance and analyzed data; YW, LKZ, and WJS performed anti-viral activity assay and analyzed data. The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41401-021-00735-z.

REFERENCES

1. Zhou P, Yang X, Lou, Wang XG, Hu B, Zhang L, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020;579:270–3.
2. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. Nature. 2020;579:265–9.
3. Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS Coronavirus. J Virol. 2020;94:7–20.
4. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet. 2020;395:565–74.
5. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature. 2020;583:459–68.
6. Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire A, Veesler D. Structure, function and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell. 2020;181:281–92.
7. 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;367:1444–8.
8. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature. 2020;581:215–20.
9. Hwang SS, Lim J, Yu Z, Kong P, Sefik E, Xu H, et al. MRNA destabilization by BTG1 and BTG2 maintains T cell quiescence. Science. 2020;367:1255–60.
10. Gui M, Song W, Zhou H, Xu J, Chen S, Xiang Y, et al. Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. Cell Res. 2017;27:119–29.
11. Shang J, Wan Y, Luo C, Ye G, Geng Q, Auerbach A, et al. Cell entry mechanisms of SARS-CoV-2. Proc Natl Acad Sci USA. 2020;117:11727–34.
12. Towler P, Staker B, Prasad SG, Menon S, Tang J, Parsons T, et al. ACE2 X-Ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. J Biol Chem. 2004;279:17996–8007.
13. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell. 2020;181:271–280.e8.
14. Wu C, Liu Y, Yang Y, Zhang P, Zhong W, Wang Y, et al. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. Acta Pharm Sin B. 2020;10:766–88.
15. Yang L, Pei R, Li H, Ma X, Zhou Y, Zhu F, et al. Identification of SARS-CoV-2 entry inhibitors among already approved drugs. Acta Pharmacol Sin. 2020;42:1–7.
16. Clausen TM, Sandoval D, Spliid C, Pihl, J, Painter C, Thacker B, et al. SARS-CoV-2 infection depends on cellular Heparan sulfate and ACE2. bioRxiv. 2020;183:1043–57.
17. Benítez-Cardoza CG, Vique-Sánchez JL. Potential inhibitors of the interaction between ACE2 and SARS-CoV-2 (RBD), to develop a drug. Life Sci. 2020;256:117970.
18. Tai W, Zhang X, He Y, Jiang S, Du L. Identification of SARS-CoV RBD-targeting monoclonal antibodies with cross-reactive or neutralizing activity against SARS-CoV-2. Antiviral Res. 2020;179:104820.
19. Pan X, Zhou P, Fan T, Wu Y, Zhang J, Shi X, et al. Immunoglobulin fragment F(ab') 2 against RBD potently neutralizes SARS-CoV-2 in vitro. bioRxiv. 2020;182:104868.
20. Huo J, Le Bas A, Ruza RR, Duyvesteyn HME, Mikolajek H, Malinauskas T, et al. Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. Nat Struct Mol Biol. 2020;27:846–54.
21. Hanke L, Perez LV, Sheward D, Das H, Schulte T, Moliner-Morro A, et al. An alpaca nanobody neutralizes SARS-CoV-2 by blocking receptor interaction. bioRxiv. 2020;11:1–9.
22. Mores A, Matziari M, Beau F, Cuniasse P, Yiotakis A, Dive V. Development of potent and selective phosphinic peptide inhibitors of angiotensin-converting enzyme 2. J Med Chem. 2008;51:2216–26.
23. Yap JL, Chen L, Lanning ME, Fletcher S. Expanding the cancer arsenal with targeted therapies: disarmament of the antiapoptotic Bcl-2 proteins by small molecules. J Med Chem. 2017;60:821–38.
24. Souers AJ, Leverson JD, Boghaert ER, Ackler SL, Catron ND, Chen J, et al. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. Nat Med. 2013;19:202–8.
25. Arkin MR, Tang Y, Wells JA. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. Chem Biol. 2014;21:1102–14.
26. Gonzalez-Ruiz D, Gohlke H. Targeting protein-protein interactions with small molecules: challenges and perspectives for computational binding epitope detection and ligand finding. Curr Med Chem. 2006;13:2607–25.
27. Guo W, Wisniewski JA, Ji H. Hot spot-based design of small-molecule inhibitors for protein-protein interactions. Bioorg Med Chem Lett. 2014;24:2546–54.
28. Bai F, Morcos F, Cheng RR, Jiang H, Onuchic JN. Elucidating the druggable interface of protein-protein interactions using fragment docking and coevolutionary analysis. Proc Natl Acad Sci USA. 2016;113:E8051–E8058.
29. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, et al. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J Med Chem. 2004;47:1750–9.
30. Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J Comput Aided Mol Des. 2013;27:221–34.
31. Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res. 2020;30:269–71.
32. Tayeh Z, Dudai N, Schechter A, Chalifa-Caspi V, Barak S, Ofir R. Molecular mode of action of asteriscus graveolens as an anticancer agent. Int J Mol Sci. 2018;19:2162.
33. Müller H, Brackhagen O, Brunne R, Henkel T, Reichel F. Natural products in drug discovery. Ernst Schering Res Found Workshop. 2000;32:205–16.
34. Newman DJ, Cragg GM. Natural products as sources of new drugs over the 30 years from 1981 to 2010. J Nat Prod. 2012;75:311–35.
35. Van der Nat J, Van der Sluis WG, Labadie RP. Gentiogenal, a new antimicrobial iridoid derived from gentiopicrin (gentiopicroside). Planta Med. 1982;45:161–2.
36. Van der Sluis WG, Van der Nat JM, Spek AL. Gentiogenal, a conversion product of gentiopicrin (gentiopicroside). Planta Med. 1983;49:211–5.
37. Hawley SA, Ross FA, Russell FM, Atrih A, Lamont DJ, Hardie DG. Mechanism of activation of AMPK by Cordycepin. Cell Chem Biol. 2020;27:214–222.e4.
38. Wu WD, Hu ZM, Shang MJ, Zhao DJ, Zhang CW, Hong DF, et al. Cordycepin down-regulates multiple drug resistant (MDR)/HIF-1α through regulating AMPK/

mTORC1 signaling in GBC-SD gallbladder cancer cells. Int J Mol Sci. 2014;15:12778–90.

39. Marcelo A, Brito F, Carmo-Silva S, Matos CA, Alves-Cruzeiro J, Vasconcelos-Ferreira A, et al. Cordycepin activates autophagy through AMPK phosphorylation to reduce abnormalities in Machado-Joseph disease models. Hum Mol Genet. 2019;28:51–63.

40. Mangalmurti N, Hunter CA. Cytokine storms: understanding COVID-19. Immunity. 2020;53:19–25.

41. Xiao N, Nie M, Pang H, Wang B, Hu J, Meng X, et al. Integrated cytokine and metabolite analysis reveals immunometabolic reprogramming in COVID-19 patients with therapeutic implications. Nat Commun. 2021;12:1–13.

42. Zhou Y, Hou Y, Shen J, Kallianpur A, Zein J, Culver DA, et al. A network medicine approach to investigation and population-based validation of disease manifestations and drug repurposing for COVID-19. ChemRxiv. 2020;18:e3000970.

43. Martin WR, Cheng F. Repurposing of FDA-approved toremifene to treat COVID-19 by blocking the spike glycoprotein and NSP14 of SARS-CoV-2. ChemRxiv. 2020;19:4670–7.

44. Bangaru S, Ozorowski G, Turner HL, Antanasijevic A, Huang D, Wang X, et al. Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. Science. 2020;370:1089–94.

45. Martin WR, Cheng F. A rational design of a multi-epitope vaccine against SARS-CoV-2 which accounts for the glycan shield of the spike glycoprotein. J Biomol Struct Dyn. 2021; 1–15. https://doi.org/10.1080/07391102.2021.1894986.

46. Chen X, Li R, Pan Z, Qian C, Yang Y, You R, et al. Human monoclonal antibodies block the binding of SARS-CoV-2 spike protein to angiotensin converting enzyme 2 receptor. Cell Mol Immunol. 2020;17:647–9.

47. Wang C, Li W, Drabek D, Okba NMA, van Haperen R, Osterhaus ADME, et al. A human monoclonal antibody blocking SARS-CoV-2 infection. Nat Commun. 2020;11:1–6.

48. Karoyan P, Vieillard V, Gómez-Morales L, Odile E, Guihot A, Luyt CE, et al. Human ACE2 peptide-mimics block SARS-CoV-2 pulmonary cells infection. Commun Biol. 2021;4:197.

49. Wibmer CK, Ayres F, Hermanus T, Madzivhandila M, Kgagudi P, Oosthuysen B, et al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. Nat Med. 2021;27:622–5.

50. Doç Y, Vaizoglu SA, Akça IO, Akdag IA, Akpinar IA, Fişne M, et al. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. Elife. 2020;9: e61312.

51. Andreano E, Piccini G, Licastro D, Casalino L, Johnson NV, Paciello I, et al. SARS-CoV-2 escape in vitro from a highly neutralizing COVID-19 convalescent plasma. bioRxiv. 2020.12.28.424451.

52. Wang Z, Schmidt F, Weisblum Y, Muecksch F, Barnes CO, Finkin S, et al. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. Nature. 2021;592:616–22.

53. Tian F, Tong B, Sun, L, Shi S, Zheng B, Wang Z, et al. Mutation N501Y in RBD of spike protein strengthens the inter-action between COVID-19 and its receptor ACE2. bioRxiv. 2021.02.14.431117.

54. Yang H, Lou C, Sun L, Li J, Cai Y, Wang Z, et al. AdmetSAR 2.0: web-service for prediction and optimization of chemical ADMET properties. Bioinformatics. 2019;35:1067–9.