Journal of Medicinal Chemistry

pubs.acs.org/jmc

Article

Probing the Allosteric Inhibition Mechanism of a Spike Protein Using Molecular Dynamics Simulations and Active Compound Identifications

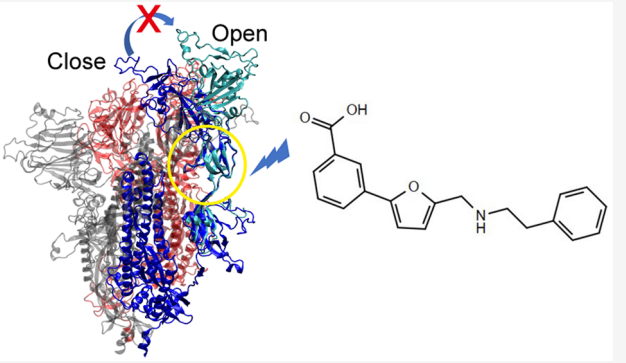
Qian Wang,*,▽ Lin Wang,▽ Yumin Zhang,▽ XiangLei Zhang, Leike Zhang, Weijuan Shang,* and Fang Bai*

Cite This: https://doi.org/10.1021/acs.jmedchem.1c00320

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The receptor recognition of the novel coronavirus SARS-CoV-2 relies on the "down-to-up" conformational change in the receptor-binding domain (RBD) of the spike (S) protein. Therefore, understanding the process of this change at the molecular level facilitates the design of therapeutic agents. With the help of coarse-grained molecular dynamic simulations, we provide evidence showing that the conformational dynamics of the S protein are globally cooperative. Importantly, an allosteric path was discovered that correlates the motion of the RBD with the motion of the junction between the subdomain 1 (SD1) and the subdomain 2 (SD2) of the S protein. Building on this finding, we designed non-RBD binding modulators to inhibit SARS-CoV-2 by prohibiting the conformational change of the S protein. Their inhibition effect and function stages at inhibiting SARS-CoV-2 were evaluated experimentally. In summary, our studies establish a molecular basis for future therapeutic agent design through allosteric effects.



INTRODUCTION

As we are witnessing, the coronavirus disease 2019 (COVID-19) pandemic, caused by a novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has quickly evolved into a tremendous global health threat.[1−4] By February of 2021, there were 107 million infection cases with over 2 million deaths.[5] Vaccines, antibodies, and small-molecule drugs are urgently needed to resolve this crisis.

SARS-CoV-2 invading host cells initiate from the spike (S) protein of SARS-CoV-2 and target its receptor angiotensin-converting enzyme 2 (ACE2),[6−11] which triggers the viral−host-cell membrane fusion.[12−14] The S protein has two functional domains: S1 and S2. S1 is sequentially comprised of the following four subdomains: the N-terminal domain (NTD, residues 1−318), the ACE2 binding domain (RBD, residues 319−531), and two other small subdomains named SD1 (residues 532−591) and SD2 (residues 591−685). Since ACE2 only binds to RBD, one straightforward strategy of neutralizing SARS-CoV-2 is to design antibodies that competitively bind to RBD.[15,16] Alternatively, due to multiple important structural analyses,[6−8,17,18] it has been widely accepted that the recognition of ACE2 requires a critical "down-to-up" conformational transition of the RBD. Therefore, another possible way of neutralizing SARS-CoV-2 is to design therapeutic agents such as antibodies or small-molecule modulators that lock RBD in the "down" conformation.[19] To specify the target region of therapeutic agents, the "down-to-up" conformational transition needs to be understood at the molecular level. If such transition is an isolated motion of RBD alone, then the target should be RBD. Otherwise, if such transition is coupled with a global cooperative motion of the whole S protein, then designing allosteric regulators becomes promising.

There have been several experimental findings implying the cooperative conformational dynamics of S protein. For the SARS-CoV-2 variant with the D614G mutation, the probability of RBD in the "up" conformation increases,[20] although residue 614 is located at the SD1−SD2 junction instead of RBD. Besides, multiple effective antibodies were found to target the NTD domain of the S protein.[21,22] Regardless of these pieces of indirect evidence, to date it is still challenging to experimentally trace the dynamic motion of each residue in the S protein so that the global cooperative dynamics can be directly analyzed. There are studies trying to tackle this

Special Issue: COVID-19

Received: March 2, 2021

problem using all-atomistic molecular dynamic simulations.[23−25] However, restrained by computational powers, the conformational change of the S protein cannot be observed in unbiased simulation trajectories. As a result, the kinetic information is still missing.

To solve this issue, we performed coarse-grained (CG) dynamic simulations to explore conformational transitions between the closed state (RBD-down) and the open state of the S protein (RBD-up). The usage of a CG model allows multiple conformational switches in a single unbiased trajectory so that the dynamics can be studied conveniently. Herein, direct pieces of evidence are provided to demonstrate that the conformational dynamics of the S protein are indeed cooperative. Specifically, we highlighted an allosteric propagating path from the SD1−SD2 junction to the RBD in a S protein trimer. Building upon that, we designed several allosteric regulators that inhibit the open state of the S protein, which was verified by experiments. In conclusion, our study paves a road for future therapeutic agent design through allosteric effects.

RESULTS

Conformational Dynamics of a Truncated S Protein Monomer. We first studied the conformational dynamics of a truncated S protein monomer by molecular dynamic simulations. A parameter *dist_1*, measuring the size of the pocket between RBD and SD1, was used to distinguish between the closed state (RBD-down, small *dist_1*) and the open state (RBD-up, large *dist_1*) of the S protein (Figure 1A). Owing to the advantage of coarse-grained simulations,

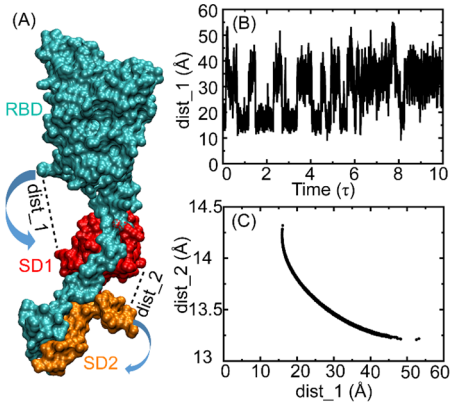


Figure 1. Dynamics of a truncated S protein (residue 296−626). (A) Structural illustration of the truncated S protein, including RBD (cyan), SD1 (residue 532−591, red), and SD2 (orange). *dist_1* represents the distance from residue 519 to 569, quantifying the size of the pocket between RBD and SD1. *dist_2* represents the distance from residue 536 to 618, quantifying the size of the pocket between SD1 and SD2. Blue arrows represent the motions of RBD and SD2 relative to SD1. (B) Two-state kinetics of *dist_1*. τ is equal to 10^7 simulation steps. (C) Correlation between *dist_1* and *dist_2*.

dist_1 shows clear two-state kinetics without any biasing potential (Figure 1B). The two-state kinetics of *dist_1* indicate that repeated conformational transitions between the closed state and the open state of S protein are well captured by simulations, allowing us to decipher the molecular details of such transitions.

Our key finding is that the conformational dynamics of the S protein are cooperative. To investigate this cooperative motion, we applied principal component analysis to study the trajectories and filtered the motion along the first eigenvector. Next, we defined a parameter *dist_2* to quantify the size of the pocket between SD1 and SD2 (Figure 1A), then plotted the data point (*dist_1*, *dist_2*) at each snapshot (Figure 1C). It is shown that there is a strong negative correlation between *dist_1* and *dist_2*; when *dist_1* increases from 15 to 35 Å, *dist_2* decreases from 14 to 13 Å. This negative correlation proves that with respect to SD1 the relative motion of the RBD is coupled with that of SD2 during the conformational change of the S protein from the closed state to the open state. Specifically, SD2 moves against SD1 when the RBD moves toward SD1 (blue arrows in Figure 1A) and vice versa. One might think that the change of *dist_2* is relatively small. However, given the large size of the full-length S protein (127 Å × 140 Å × 160 Å), a small change at SD2 could be significantly amplified when propagated to other domains. This point will be further elaborated in the next section.

Conformational Dynamics of a S Protein Trimer. We next simulated a S protein trimer (Figure 2A) to check whether its conformational dynamics are also cooperative. First, only chain A was allowed to switch from the down conformation to the up conformation while the other two

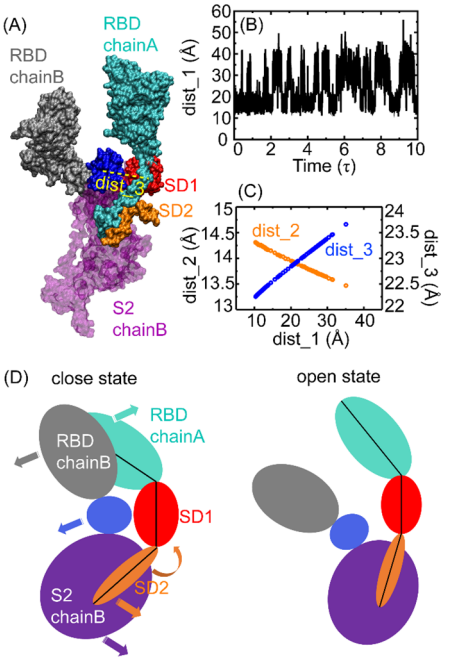


Figure 2. Dynamics of a full-length S protein trimer. (A) Structural illustration of the S protein trimer (for visualization purpose, here we only show domains important to the conformational change of chain A). RBD, SD1, and SD2 in chain A are colored in cyan, red, and orange, respectively. The S1 domain in chain B is colored in silver. The S2 domain in chain B is divided into two parts. The part (residues 745−755 and 975−1000) spatially close to SD1 in chain A is colored in blue. The other part of S2 is colored in purple. (B) Two-state kinetics of dist_1. (C) Correlation between *dist_1* and *dist_2* (orange circle) and between *dist_1* and *dist_3* (blue circle). (D) Schematic illustration of a cooperative conformational change of the S protein. *dist_1* and *dist_2* follow the same definition as that in Figure 1. *dist_3* (dashed yellow line) represents the distance between the center of mass (COM) of the blue domain in chain B and the COM of SD1 in chain A.

https://doi.org/10.1021/acs.jmedchem.1c00320
J. Med. Chem. XXXX, XXX, XXX−XXX

chains were kept in the down conformation. Like the monomer case described above, the *dist_1* of chain A in the trimeric molecule also shows two-state kinetics without any biasing potential (Figure 2B). The probability of the closed state increases compared to that of the monomer case (Figure 1B versus Figure 2B). This is not surprising because in the closed state the RBD of chain A forms interchain contacts with the RBD of chain B. Therefore, the transition to the open state requires the consumption of extra free energies to break the interchain interactions.

Similar to the monomer case, cooperativity was observed in the conformational dynamics of the S protein trimer. The motions of the RBD and SD2 in chain A are negatively correlated (Figure 2C, orange circle). Importantly, we found that the correlation between the RBD and SD2 is stronger in the trimer than that in the monomer, as the motion of SD2 can further couple with the motion of the RBD through interchain interactions. To elaborate this point, we defined another parameter *dist_3* to quantify the distance between SD1 in chain A and its interacting domain in chain B (the blue domain in Figure 2A), then plotted the data point (*dist_1*, *dist_3*) at each snapshot (Figure 2C, blue circle). It turns out that *dist_1* and *dist_3* are positively correlated, indicating that when the RBD in chain A moves against SD1, the blue domain in chain B also moves against SD1 in chain A. Thus, the conformational change of the S protein trimer from the close to the open state is a global cooperative process that requires not only the intrachain correlation (RBD and SD2 in chain A) but also the interchain correlation (chain A and chain B). This global cooperative process is illustrated in Figure 2D. Initiating from the closed state of S protein, when the RBD in chain A moves against SD1 (Figure 2D, cyan arrow), SD2 cooperatively rotates so that it moves toward SD1 (Figure 2D, orange arrow). The rotation of SD2 in chain A prompts S2 in chain B (Figure 2D, purple domain) to rotate as well (Figure 2D, purple arrow) due to the interchain attractions between them. Other domains in chain B (Figure 2D, blue and silver) sequentially move. Noticeably, just like a lever, a small change in the SD1−SD2 junction will be amplified when it propagates to RBD. As a result, the RBD in chain B (Figure 2D, silver domain) largely moves against the RBD in chain A (Figure 2D, cyan domain), which thermodynamically favors the contact breaking between the two, i.e., the prerequisite for forming the open state of S protein.

Next, all three chains were allowed to switch from the down to the up conformation. The main findings remain the same. Similar long-range cooperative conformational dynamics between the SD1−SD2 junction and the RBD domain were observed in such a model (Figure S1A) as those above (Figures 1 and 2). Interfering with such cooperation increases the probability of the closed state (Figure S1B, the black line versus the red line). In addition, we identified an extra interchain correlation where the up conformation of one chain prompts the up conformation of the other two chains, owing to the contact breaking between RBDs of different chains in the open state. As a result, when all three chains are allowed to switch from the down conformation to the up conformation, the overall population of the open state is larger than that in the case when two chains are kept in the down conformation (Figure S1B, the black line versus the green line).

Inhibiting the Conformational Transition of the S Protein by Allosteric Regulators. The global cooperative conformational dynamics of the S protein trimer imply the possibility of inhibiting the open state through allosteric regulations. To test this idea, we increased the rigidity of the SD1−SD2 junction of chain A in simulations. By doing so, the allosteric propagating path from the SD1−SD2 junction to the RBD (Figure 2D) was hindered from the very beginning. Indeed, the probability of the open state decreased by 27% (Figure 3). Inspired by this result, we designed ligands for allosteric regulation and tested them experimentally.

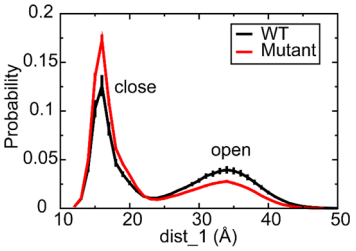


Figure 3. Computed probability distributions of *dist_1* for a wide-type (WT) S protein and a designed S protein mutant. In this mutant, the rigidity of the SD1−SD2 junction was increased to interfere the global cooperative conformational dynamics of the S protein (please see details in the last paragraph of the Supporting Information). *dist_1* follows the same definition as that in Figure 1.

To identify potential allosteric small-molecule modulators against our identified binding pocket, a synergistic workflow that combined virtual screening and experimental validation was used. As indicated above, there is a significant correlation in the process of conformational change between the RBD and the SD1−SD2 junction of the S protein. Hence, designing a molecular binder to insert between SD1 and SD2 of the S protein may impede the conformational motion of the protein, maintaining the RBD in the "down" conformation. However, by analyzing the snapshots of conformations of the S protein that were obtained from our simulations, it was found that the pocket formed between SD1 and SD2 is too shallow to accommodate small molecules (as shown in Figure 4A). Therefore, an extended region around 5 Å from SD1 and SD2 was considered to identify other candidate pockets using a binding site identification method named SiteMap (Schrödinger, Inc.).[26] An exposed candidate site between SD1 and the S2 domain of another monomer of the S protein trimer was discovered, as shown in Figure 4A. This pocket may confer biological activity to small-molecule compounds that bind to and potentially stabilize the closed state conformation of S protein by gluing the red and blue domains in Figure 2D together. Such a hypothesis was supported by coarse-grained simulations, as shown in Figure S4.

Based on the identified binding site, a large cube grid with a size of 35 × 35 × 35 Å^3, covering both the SD1−SD2 shallow pocket and the identified binding site, was designed with the center according to the coordinates of the residues of E619 and N536. Then, molecular docking-based virtual screening was conducted, and top-ranked molecules were reserved for visual inspection. Finally, six molecules were chosen as candidates (as shown in Table S1, labeled as CPD1−CPD6), and their compound samples were ordered from Chemdiv (https://www.chemdiv.com/) and ChemBridge (https://www.chembridge.com/) to perform experimental validations.

These six molecules present distinct differences in their structures and various binding modes with the S protein. As shown in Figure S2, these molecules generally may bind to the

https://doi.org/10.1021/acs.jmedchem.1c00320
J. Med. Chem. XXXX, XXX, XXX−XXX

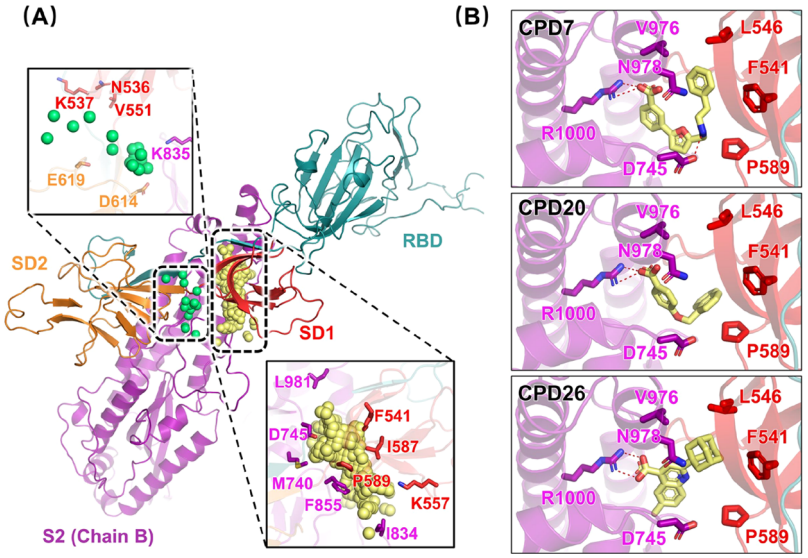


Figure 4. (A) Elucidation of the two binding pockets identified for designing molecular modulators of the S protein. The green spheres indicate the area of the shallow pocket between SD1 with SD2, and the yellow spheres represent the region of the other identified binding sites near SD1. (B) Three discovered potential molecular modulators and their binding modes.

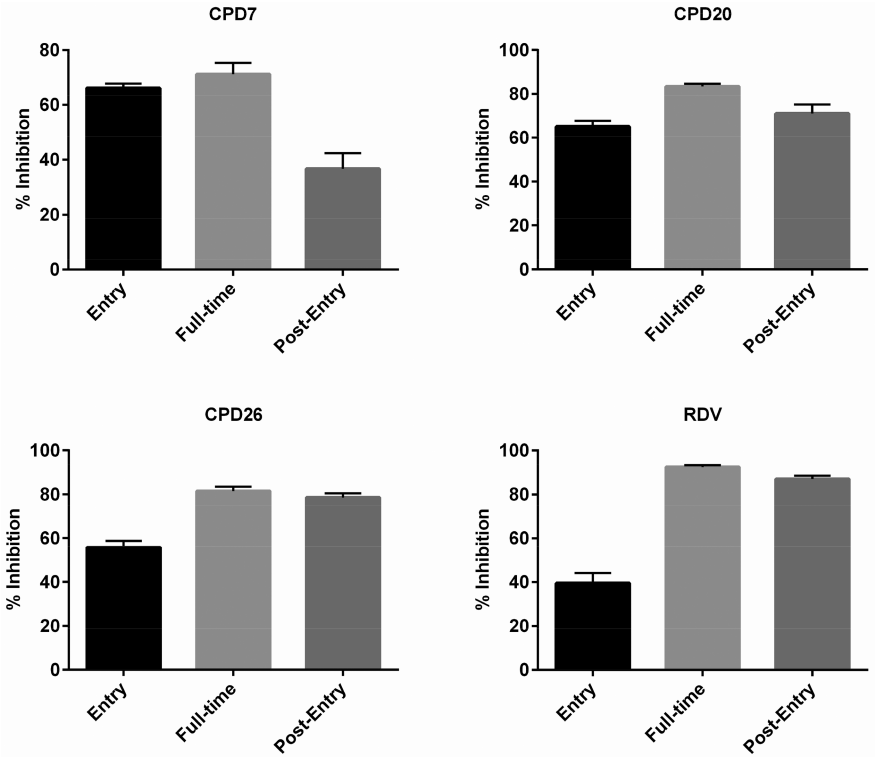


Figure 5. Time-of-addition experiment to study the inhibitory stage of three discovered allosteric modulators against SARS-CoV-2: CPD7, CPD20, and CPD26 in Vero E6 cells. The data are representative of two independent experiments, and error bars represent the mean ± SD of triplicate experiments.

different sites. Among of them, CPD1 and CPD3−CPD6 are located in a similar region, while CPD2 is located in a different area. By analyzing the binding mode, a common hydrogen bond can be observed between the residue N978 and the first group of these five molecules. CPD6 also forms another hydrogen bond with the residue of R1000. In contrast with the other compounds, CPD2 forms a cation−π interaction with K557. Beyond this, a series of hydrophobic contacts that interact with P589, I587, F541 of SD1, and F855 of the S2 domain can be identified (as shown in Figure S2), which may contribute to the majority of the binding affinities for these molecules.

To validate our computational prediction, surface plasmon resonance (SPR) experiments were first used to determine the binding affinities and kinetics of these six compounds to the S protein using Biacore K8000 (GE Healthcare) equiptment. As shown in Figure S3, all these compounds exhibit concentration-dependent binding to the S protein.

https://doi.org/10.1021/acs.jmedchem.1c00320
J. Med. Chem. XXXX, XXX, XXX−XXX

To further investigate the antiviral properties of these molecules, CPD1, CPD3, CPD4, and CPD6, which showed better binding affinities with the trimeric S protein, were sent to be evaluated on their anti-SARS-CoV-2 effect. Vero E6 cells were treated with 100 μM CPD1, CPD3, CPD4, and CPD6 and then infected with SARS-CoV-2 at an multiplicity of infection (MOI) of 0.01. Twenty-four hour postinfection, the supernatant was collected, and the intracellular level of viral RNA was measured with quantitative RT-PCR. As shown in Figure S6, CPD1, CPD3, CPD4, and CPD6 can decrease the intracellular level of viral RNA, suggesting that CPD1, CPD3, CPD4, and CPD6 can inhibit SARS-CoV-2 replication.

Based on the structure−activity relationship of the above six compounds identified in the first round of identification, we performed a second round of virtual screening to search for more potent compounds. Another 20 compounds were discovered, and their compound structures are attached in Table S1 (names starting from CPD7 to CPD26). The ADMET properties of these picked molecules have been computationally evaluated using the QikProp module of Schrödinger2021-2 and the admetSAR Web server.[27] All these compounds were passed by both the above methods (shown in Table S2). Then, similar experimental validations were performed, including a binding affinity evaluation with SPR experiments and anti-SARS-CoV-2 assays. Interestingly, 19 of the 20 compounds discovered in this round were shown to be able to interact with the S protein, and their binding affinities are given in the Table S1. For the antiviral assay of these 19 compounds, Vero E6 cells were preseeded in 48-well plates at a density of 6 × 104 cells per well and cultured overnight. Afterward, cells were pretreated with the medium with compounds at a final concentration of 100 μM for 1 h, and SARS-CoV-2 was subsequently added (MOI of 0.01) to allow infection for 1 h. Next, the virus/drug mixture was removed, and cells were further cultured with a fresh drug-containing medium. At 24 h p.i. (post infection), the cell supernatant was collected, and the amount of viral RNA in the supernatant was determined by qRT-PCR analysis. Compounds were stored as 40 mM stock solutions in DMSO at 4 °C until use. The top three compounds, i.e., CPD7, CPD20, and CPD26, were proven to be able to inhibit SARS-CoV-2 at a concentration of 100 μM (Figure S7) and showed inhibitory effects that were selected for further research. Moreover, the SPR sensorgrams for the interactions of these three top compounds with the full length of the S protein are given in Figure S8.

To further investigate which stage of the life cycle of SARS-CoV-2 was inhibited by CPD7, CPD20, and CPD26, a time-of-addition experiment was performed using the method reported previously.[28] As shown in Figure 5, remdesivir can affect SARS-CoV-2 replication at 2 h p.i., indicating that viral RNA replication or transcription may happen at this time point since it has been reported that remdesivir can inhibit viral RNA replication or transcription but not entry. Compared with remdesivir, all these two compounds, especially the CPD7, show a stronger inhibitory effect, suggesting that these compounds may inhibit viral entry, However, a shorter time of drug incubation is needed to confirm this result. On the other hand, it has been reported that the S protein of SARS-CoV-2 is subjected to cotranslational and post-translational processing in the RER and Golgi complex, including signal peptide removal, trimerization, extensive glycosylation, and subunit cleavage.[29] Hence, the conformational change of the S protein may also play important function in this "post-entry" stage. It indicates that the compounds that are bound to the identified allosteric site may inhibit the virus by blocking the infection in the early stage and inhibiting the release of new virions in the late stage of the life circle.

DISCUSSION AND CONCLUSIONS

In this work, we utilized coarse-grained molecular dynamic simulations to study the conformational changes of a S protein. The motion of the RBD was found to correlate with that of the SD1−SD2 junction in the same chain as follows: RBD in chain A → SD1−SD2 junction in chain A. Moreover, in a S protein trimer, the latter can influence the former in turn through an interchain allosteric propagating path as follows: SD1−SD2 junction in chain A → S2 in chain B → RBD in chain B → RBD in chain A, which forms a positive feedback loop. Consequentially, the conformational dynamics of a S protein trimer are highly cooperative. Noticeably, D614G mutation, which has been found to increase the population of the open state of the S protein experimentally,[20] happens to be located at SD1/SD2 junction. We hypothesize that glycine can add extra flexibility to this junction,[30] which enhances the global conformational correlations of the S protein. Therefore, we reduced the interaction between residue 614 and other parts of the S protein to mimic the effect of D614G. The computational result (Figure S5) shows that the probability of the open state increases due to this mutation, which is consistent with the experimental findings.

Non-RBD-binding therapeutic agents have been proposed to inhibit SARS-CoV-2 via the direct perturbation of its fusion core.[31] Our work establishes a molecular basis for the design of new non-RBD-binding therapeutic agents via allosteric regulations. Although a variety of SARS-CoV-2 antibodies[15,21,22,32] have been developed, such a design is still needed for at least three reasons. First, a combination therapy with multiple therapeutic agents shows an additive inhibition effect.[16,19] Second, SARS-CoV-2 undergoes diverse mutations,[33] which cause "antibody escape" to certain existing antibodies.[34] Third, the sequence of the receptor interaction domain in RBD is not well-conserved during evolution. Therefore, RBD-binding therapeutic agents are often virus-specific[35,36] while non-RBD-binding ones could show a cross-neutralization effect.

We discovered three potential allosteric regulators by two rounds of virtual screening. Although their potencies are weaker than those of CQ and RVD (Figures 5, S6, and S7), these compounds were confirmed to bind to the spike protein (Figure S8 and Table S1) and inhibit the virus in both stages of "pre-entry" and "post-entry" (Figure 5). Herein, we emphasized that the major role of these antiviral experiments is to support our key finding on the cooperative conformational dynamics of the S protein and to demonstrate the potential application of designing allosteric regulators. Further molecular modifications are still needed for real therapeutic purposes because the binding affinities of our designed compounds to the spike protein are relatively weak in general. In addition, the virtual screening process can be further improved by considering the glycan shield of the spike glycoprotein.[37] According to the predicted binding poses, all these compounds can partially form several hydrophobic contacts with a bunch of hydrophobic residues, i.e., I587, F541, P589, F855, etc., together with a hydrogen bond with N978, leading to a relatively stable but not strong binding interaction with the

https://doi.org/10.1021/acs.jmedchem.1c00320
J. Med. Chem. XXXX, XXX, XXX−XXX

spike protein. Additionally, there are two polar residues, that is, R1000 and D745, that could strengthen the interactions by forming ionic bonds or hydrogen bonds with some of these compounds. To improve the inhibitor potency, a further modification for these compounds is needed to fully utilize the hydrophobic property of the pocket.

As one of the major structural proteins of SARS-CoV-2, the S protein is responsible for the attachment and penetration of the virus into the host cell; therefore, is the direct target for designing anti-infection drugs. However, more and more mutations have been identified on the the S protein, especially occurring on its binding site to the host cell receptor, i.e., ACE2. Hence, the binding site between the S protein and ACE2 is not an ideal drug design site. To overcome this potential drug resistance induced by evolutionary mutations, we designed such a work to investigate the allosteric mechanism of the S protein and tried to identify the potential allosteric sites, which are relatively conserved and able to be targeted for the design of allosteric modulators that indirectly interfere with the binding between the S protein and host cells. Certainly, there are several targets, such as Mpro, PL, and RdRp, that are responsible for the different biological functions in the life circle of SARS-CoV-2. Several potent inhibitors targeted on these targets have been reported, even in the clinical trials. These targets are relatively conserved and more promising for drug design. Drawing from the anti-HIV "drug cocktail" therapy, combining multiple drugs that can inhibit the virus in the different stages of the life circle of SARS-CoV-2 may be a more promising therapy to treat the COVID-19.

EXPERIMENTAL SECTION

Computational Model of SARS-CoV-2 S Protein. We built a coarse-grained structure-based model[38] to simulate the conformational transitions of the S protein between its closed state and its open state. The initial structures of these two states[17] were taken from the protein data bank (PDB ID 6VXX and 6VYB, respectively), with all missing loops filled in by the SWISS-MODEL server.[39] Next, the structural information was used to generate the Cα-only model by SMOG[40] for the following three systems: (I) a truncated S protein monomer (residues 296−626), (II) a full-length wild-type (WT) S protein trimer, and (III) a full-length mutated S protein trimer. The Hamiltonian of these systems follows our previous work[41,42] with slight modifications. The details can be found in the Supporting Information. Simulations were performed by an in-house version of Gromacs.[40,43,44] The Langevin equations of motion were integrated to the low-friction limit.[45] Each system was simulated 10^8 steps and repeated five times. Data from all trajectories were collected and analyzed.

Virtual and Visual Inspection. Compounds in the Chinese National Compound library were first prepared for docking using LigPrep (Schrödinger, Inc.), which is a ligand preparation module that can generate multiple minimized conformations as well as protonation or tautomerization states with default settings. Next, virtual screening was performed using a molecular docking module named Glide 8.5.[46] The receptor structure used for docking is the same one used to modeling the closed state of the trimeric S protein in the above-mentioned simulations. This structure was refined by the Protein Preparation Wizard[47] of Schrödinger 2019-4 with default parameter settings, including assigning bond orders, adding hydrogens, treating charge, etc. A grid for the docking simulations was generated with a size of a 35 × 35 × 35 Å^3 cube[46] by centering it based on the coordinates of the residues of E619 and N536. During the virtual screening step, two different docking methods, that is, high throughput virtual screening (HTVS) and standard precision (SP), were used for the preliminary rough docking and a more accurate evaluation, respectively. The number of ligands retained from HTVS was set to 10%, and the top 50% of SP docking ligands were eventually exported for further visual inspection.

Compounds in the Specs library were used as the compound database to perform the second round of virtual screening. In this round of computational simulations, the grid for the docking simulations was generated with a size of a 20 × 20 × 20 Å^3 cube by centering it based on these six compounds (as shown in Figure S2), which were the best compounds of the first virtual screening round. The first round of the identification process indicated that the allosteric site identified was reasonable, and a smaller size of box was focused on in the second round of computational simulation. The other parameters of LigPrep and the docking simulation in this round are same as those for the last round.

Surface Plasmon Resonance. All SPR experiments were performed at 25 °C in HBS-EP buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 3.0 mM EDTA, and 0.005% (v/v) surfactant P20) containing 5% DMSO with Biacore 8K (GE Healthcare). The data were analyzed using Biacore 8K evaluation software. The sensorgrams were reference, and the blank was subtracted. Running buffer was used for the blank injections, and bulk effects were corrected using eight cycles of solvent correction.

Prior to determining the binding affinity of each compound, the SARS-CoV-2 full-length spike protein was diluted to a concentration of 20 μg/mL in sodium acetate, pH 4.5, and immobilized on a CM5 chip using standard amine coupling to a level of 15 000 response units (RU). Tested compounds in twofold dilutions with concentrations ranging from 100 to 0.78 μM with final a DMSO concentration of 5% were prepared with running buffer. Three startup cycles with running buffer were performed first, and the analyte was injected and run through the chip with an association time of 120 s and a disassociation time of 100 s, then washed with 50% DMSO in each cycle. Solvent correction with different DMSO concentrations varying from 4.5% to 5.8% was performed every 48th cycle. Raw data were reduced, double-referenced, and solvent-corrected using Biacore 8K evaluation software, and K_D values of each compound were calculated using the steady-state affinity model with a constant R_{max} value.

Cell Cultures and Virus Species Information. Vero E6 was maintained in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS), 100 IU/mL penicillin, and 100 μg/mL streptomycin. Cells were maintained at 37 °C with 5% CO_2. Cells were infected with SARS-CoV-2 (nCoV-2019BetaCoV/Wuhan/WIV04/2019, GISAID EPI_ISL_402124) isolated from a clinical patient in Wuhan as described in our previous study.[28] The virus strain was obtained from the National Virus Resource and was propagated in Vero E6 cells. Cells were infected at a multiplicity of infection (MOI) of 0.01. The virus titer was determined by using a tissue culture infectious dose-50 (TCID50) end-point dilution assay.

Antiviral Efficacy in the Vero Cell Line. Vero E6 cells were cultured in T75 flasks and incubated at 37 °C and 5% CO2. Cells were passaged at a ratio of 1:3 every 48 h. Media formula is as follows: 90% DMEM (Gibco Invitrogen) and 10% fetal bovine serum (Gibco Invitrogen).

In our study, Vero E6 cells were advance-prepared to 48-well plates (50 000 cells/well) for 16−18 h. Then, cells were treated with a medium containing a gradient concentration of the compound at 200 μL/well for 1 h. DMSO was used in the control. Cells in the board were infected at a multiplicity of infection (MOI) of 0.01 for 24 h. The cells incubated at 37 °C for 24 h with a 5% CO_2 atmosphere.

Twenty-four hours after infection, the supernatant was collected. The inactivated sample was immersed in disinfectant and removed from the biosafety level 3 laboratory. RNA was extracted from the inactivated supernatant, reverse transcription occurred, and copies were determined using real-time fluorescence-based quantitative PCR. After determining the copy numbers, the DMSO group was used as the control well to calculate the inhibition rate of viral replication by the compound.

qRT-PCR. The collected supernatant was extracted by an RNA/DNA extraction kit (TaKaRa MiniBEST Viral RNA/DNA Extraction Kit ver. 5.,0 TaKara 9766). RNA was reverse transcribed using the RT reagent kit (PrimeScript RT reagent kit with gDNA eraser, TaKaRa

https://doi.org/10.1021/acs.jmedchem.1c00320
J. Med. Chem. XXXX, XXX, XXX−XXX

RR047A). Quantitative RT−PCR was performed on a SYBR Green I Taq (TaKaRa SYBR Premix Ex Taq II, TaKaRa RR820A) system .

Primers used were as follows: RBD-qF1:50-CAATGGTTTA-ACAGGCACAGG-30 and RBD-qR1:50-CTCAAGTGTCTGTGG-ATCACG-30.

Time-of-Addition Experiment. To evaluate which stage of the life cycle of SARS-CoV-2 was inhibited by CPD7, CPD20, and CPD26 at a concentration of 100 μM, a time-of-addition experiment was performed as previously described.[28] Remdesivir (10 μM) was used in parallel as a positive control for the post-entry-stage inhibition. For all the treatment groups, Vero E6 cells were infected with SARS-CoV-2 at an MOI of 0.01. For "full-time" treatment, cells were pretreated with the drugs for 1 h. The virus was added to attach to cells for 2 h, then the virus/drug mixture was removed, and the cells were cultured with the drug-containing medium until the end of the experiment. For the "entry" treatment, the drugs were added to the cells for 1 h before viral attachment for 2 h, then the virus/drug mixture was replaced with fresh culture medium and maintained until the end of the experiment. For the "post-entry" experiment, drugs were added 2 h postinfection, then maintained until the end of the experiment. All the treatment groups were done in triplicate, and the virus yield in the infected cell supernatants was quantified by qRT-PCR at 24 h p.i.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jmedchem.1c00320.

Additional computational modeling details and predicted effects of various mutations and 2D structures, binding affinities, and inhibition effects of allosteric regulators identified by virtual screening (PDF)

Molecular formula strings for final compounds (CSV)

Mass spectrometry data for the studied compounds (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Qian Wang − *Hefei National Laboratory for Physical Sciences at the Microscale and Department of Physics, University of Science and Technology of China, Hefei, Anhui 230026, China;* orcid.org/0000-0002-9908-4317; Email: wqq@ustc.edu.cn

Weijuan Shang − *State Key Laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, Hubei 430071, China*; Email: shangweijuan@wh.iov.cn

Fang Bai − *School of Life Science and Technology and Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, Shanghai 201210, China;* orcid.org/0000-0003-1468-5568; Email: baifang@shanghaitech.edu.cn

Authors

Lin Wang − *School of Life Science and Technology and Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, Shanghai 201210, China*

Yumin Zhang − *State Key Laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, Hubei 430071, China*

XiangLei Zhang − *School of Life Science and Technology and Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, Shanghai 201210, China*

Leike Zhang − *State Key Laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, Hubei 430071, China;* orcid.org/0000-0002-2593-2571

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jmedchem.1c00320

Author Contributions

$^{\nabla}$Q.W., L.W., and Y.Z. contributed equally to this work

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Q.W. acknowledges the funding support from the National Natural Science Foundation of China (32000882) and "USTC Research Funds of the Double First-Class Initiative" (YD2030002006). F.B. was supported by the Science and Technology Commission of Shanghai Municipality grants (Grants 20430780300, 20431900102, and 20431900100), Shanghai Science and Technology Development Funds (Grant 20QA1406400), National Natural Science Foundation of China (no. 82003654), and start-up package from ShanghaiTech University. L.-K.Z. was supported by the Youth Innovation Promotion Association CAS (Grant 2018367). The numerical calculations in this paper have been done on the supercomputing system in the Supercomputing Center of the University of Science and Technology of China.

■ ABBREVIATIONS

ACE2, angiotensin-converting enzyme 2; CG, coarse-grained; COVID-19, coronavirus disease 2019; CPD, compound; CQ, chloroquine; DMEM, Dulbecco's modified Eagle's medium; FBS, fetal bovine serum; HTVS, high throughput virtual screening; MOI, multiplicity of infection; NTD, N-terminal domain; RBD, receptor-binding domain; RVD, Remdesivir; S protein, spike protein; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SD1, subdomain 1; SD2, subdomain 2; SP, standard precision; SPR, surface plasmon resonance; TCID50, tissue culture infectious dose-50; WT, wide-type

■ REFERENCES

(1) Wang, C.; Horby, P. W.; Hayden, F. G.; Gao, G. F. A novel coronavirus outbreak of global health concern. *Lancet* **2020**, *395*, 470−473.

(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. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **2020**, *382*, 727−733.

(3) Wu, F.; Zhao, S.; Yu, B.; Chen, Y. M.; Wang, W.; Song, Z. G.; Hu, Y.; Tao, Z. W.; Tian, J. H.; Pei, Y. Y.; Yuan, M. L.; Zhang, Y. L.; Dai, F. H.; Liu, Y.; Wang, Q. M.; Zheng, J. J.; Xu, L.; Holmes, E. C.; Zhang, Y. Z. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, *579*, 265−269.

(4) Huang, C. L.; Wang, Y. M.; Li, X. W.; Ren, L. L.; Zhao, J. P.; Hu, Y.; Zhang, L.; Fan, G. H.; Xu, J. Y.; Gu, X. Y.; Cheng, Z. S.; Yu, T.; Xia, J. A.; Wei, Y.; Wu, W. J.; Xie, X. L.; Yin, W.; Li, H.; Liu, M.; Xiao, Y.; Gao, H.; Guo, L.; Xie, J. G.; Wang, G. F.; Jiang, R. M.; Gao, Z. C.; Jin, Q.; Wang, J. W.; Cao, B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan. *Lancet* **2020**, *395*, 497−506.

(5) World Health Organization Report. Weekly epidemiological and operational updates February 2021. *Coronavirus disease (COVID-19) Weekly Epidemiological Update and Weekly Operational Update.* https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.

(6) Lan, J.; Ge, J. W.; Yu, J. F.; Shan, S. S.; Zhou, H.; Fan, S. L.; Zhang, Q.; Shi, X. L.; Wang, Q. S.; Zhang, L. Q.; Wang, X. Q.

Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020**, *581*, 215−220.
(7) Shang, J.; Ye, G.; Shi, K.; Wan, Y. S.; Luo, C. M.; Aihara, H.; Geng, Q. B.; Auerbach, A.; Li, F. Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020**, *581*, 221−224.
(8) Wang, Q. H.; Zhang, Y. F.; Wu, L. L.; Niu, S.; Song, C. L.; Zhang, Z. Y.; Lu, G. W.; Qiao, C. P.; Hu, Y.; Yuen, K. Y.; Wang, Q. S.; Zhou, H.; Yan, J. H.; Qi, J. X. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell* **2020**, *181*, 894−904.
(9) Zhou, P.; Yang, X. L.; Wang, X. G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H. R.; Zhu, Y.; Li, B.; Huang, C. L.; Chen, H. D.; Chen, J.; Luo, Y.; Guo, H.; Jiang, R. D.; Liu, M. Q.; Chen, Y.; Shen, X. R.; Wang, X.; Zheng, X. S.; Zhao, K.; Chen, Q. J.; Deng, F.; Liu, L. L.; Yan, B.; Zhan, F. X.; Wang, Y. Y.; Xiao, G. F.; Shi, Z. L. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *579*, 270−273.
(10) Liu, K.; Tan, S.; Niu, S.; Wang, J.; Wu, L.; Sun, H.; Zhang, Y.; Pan, X.; Qu, X.; Du, P.; Meng, Y.; Jia, Y.; Chen, Q.; Deng, C.; Yan, J.; Wang, H.-W.; Wang, Q.; Qi, J.; Gao, G. F. Cross-species recognition of SARS-CoV-2 to bat ACE2. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, e2020216118.
(11) Yan, R. H.; Zhang, Y. Y.; Li, Y. N.; Xia, L.; Guo, Y. Y.; Zhou, Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020**, *367*, 1444−1448.
(12) Gallagher, T. M.; Buchmeier, M. J. Coronavirus spike proteins in viral entry and pathogenesis. *Virology* **2001**, *279*, 371−374.
(13) Simmons, G.; Zmora, P.; Gierer, S.; Heurich, A.; Pohlmann, S. Proteolytic activation of the SARS-coronavirus spike protein: Cutting enzymes at the cutting edge of antiviral research. *Antiviral Res.* **2013**, *100*, 605−614.
(14) Shang, J.; Wan, Y. S.; Luo, C. M.; Ye, G.; Geng, Q. B.; Auerbach, A.; Li, F. Cell entry mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 11727−11734.
(15) Lv, Z.; Deng, Y.-Q.; Ye, Q.; Cao, L.; Sun, C.-Y.; Fan, C.; Huang, W.; Sun, S.; Sun, Y.; Zhu, L.; Chen, Q.; Wang, N.; Nie, J.; Cui, Z.; Zhu, D.; Shaw, N.; Li, X.-F.; Li, Q.; Xie, L.; Wang, Y.; Rao, Z.; Qin, C.-F.; Wang, X. Structural basis for neutralization of SARS-CoV-2 and SARS-CoV by a potent therapeutic antibody. *Science* **2020**, *369*, 1505−1509.
(16) Koenig, P.-A.; Das, H.; Liu, H.; Kummerer, B. M.; Gohr, F. N.; Jenster, L.-M.; Schiffelers, L. D. J.; Tesfamariam, Y. M.; Uchima, M.; Wuerth, J. D.; Gatterdam, K.; Ruetalo, N.; Christensen, M. H.; Fandrey, C. I.; Normann, S.; Todtmann, J. M. P.; Pritzl, S.; Hanke, L.; Boos, J.; Yuan, M.; Zhu, X.; Schmid-Burgk, J. L.; Kato, H.; Schindler, M.; Wilson, I. A.; Geyer, M.; Ludwig, K. U.; Hallberg, B. M.; Wu, N. C.; Schmidt, F. I. Structure-guided multivalent nanobodies block SARS-CoV-2 infection and suppress mutational escape. *Science (Washington, DC, U. S.)* **2021**, *371*, eabe6230.
(17) 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.
(18) Xu, C.; Wang, Y.; Liu, C.; Zhang, C.; Han, W.; Hong, X.; Wang, Y.; Hong, Q.; Wang, S.; Zhao, Q.; Wang, Y.; Yang, Y.; Chen, K.; Zheng, W.; Kong, L.; Wang, F.; Zuo, Q.; Huang, Z.; Cong, Y. Conformational dynamics of SARS-CoV-2 trimeric spike glycoprotein in complex with receptor ACE2 revealed by cryo-EM. *Science Advances* **2021**, *7*, eabe5575.
(19) Toelzer, C.; Gupta, K.; Yadav, S. K. N.; Borucu, U.; Davidson, A. D.; Williamson, M. K.; Shoemark, D. K.; Garzoni, F.; Staufer, O.; Milligan, R.; Capin, J.; Mulholland, A. J.; Spatz, J.; Fitzgerald, D.; Berger, I.; Schaffitzel, C. Free fatty acid binding pocket in the locked structure of SARS-CoV-2 spike protein. *Science* **2020**, *370*, 725−730.
(20) Gobeil, S. M. C.; Janowska, K.; McDowell, S.; Mansouri, K.; Parks, R.; Manne, K.; Stalls, V.; Kopp, M. F.; Henderson, R.; Edwards, R. J.; Haynes, B. F.; Acharya, P. D614G mutation alters SARS-CoV-2 spike conformation and enhances protease cleavage at the S1/S2 junction. *Cell Rep.* **2021**, *34*, 108630.
(21) Chi, X. Y.; Yan, R. H.; Zhang, J.; Zhang, G. Y.; Zhang, Y. Y.; Hao, M.; Zhang, Z.; Fan, P. F.; Dong, Y. Z.; Yang, Y. L.; Chen, Z. S.; Guo, Y. Y.; Zhang, J. L.; Li, Y. N.; Song, X. H.; Chen, Y.; Xia, L.; Fu, L.; Hou, L. H.; Xu, J. J.; Yu, C. M.; Li, J. M.; Zhou, Q.; Chen, W. A neutralizing human antibody binds to the N-terminal domain of the spike protein of SARS-CoV-2. *Science* **2020**, *369*, 650−655.
(22) Liu, L. H.; Wang, P. F.; Nair, M. S.; Yu, J.; Rapp, M.; Wang, Q.; Luo, Y.; Chan, J. F. W.; Sahi, V.; Figueroa, A.; Guo, X. Z. V.; Cerutti, G.; Bimela, J.; Gorman, J.; Zhou, T. Q.; Chen, Z. W.; Yuen, K. Y.; Kwong, P. D.; Sodroski, J. G.; Yin, M. T.; Sheng, Z. Z.; Huang, Y. X.; Shapiro, L.; Ho, D. D. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* **2020**, *584*, 450−456.
(23) Qiao, B.; de la Cruz, M. O. Enhanced binding of SARS-CoV-2 spike protein to receptor by distal polybasic cleavage sites. *ACS Nano* **2020**, *14*, 10616−10623.
(24) Gur, M.; Taka, E.; Yilmaz, S. Z.; Kilinc, C.; Aktas, U.; Golcuk, M. Conformational transition of SARS-CoV-2 spike glycoprotein between its closed and open states. *J. Chem. Phys.* **2020**, *153*, 075101.
(25) Casalino, L.; Gaieb, Z.; Goldsmith, J. A.; Hjorth, C. K.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; McLellan, J. S.; Fadda, E.; Amaro, R. E. Beyond Shielding: The roles of glycans in the SARS-CoV-2 spike protein. *ACS Cent. Sci.* **2020**, *6*, 1722−1734.
(26) Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377−389.
(27) Yang, H. B.; Lou, C. F.; Sun, L. X.; Li, J.; Cai, Y. C.; Wang, Z.; Li, W. H.; Liu, G. X.; Tang, Y. admetSAR 2.0: web-service for prediction and optimization of chemical ADMET properties. *Bioinformatics* **2019**, *35*, 1067−1069.
(28) Wang, M. L.; Cao, R. Y.; Zhang, L. K.; Yang, X. L.; Liu, J.; Xu, M. Y.; Shi, Z. L.; Hu, Z. H.; Zhong, W.; Xiao, G. F. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res.* **2020**, *30*, 269−271.
(29) Vinod, N. Exploring COVID-19: Relating the spike protein to infectivity, pathogenicity and immunogenicity. *ACS Central Science* **2021**, *5*, 001−010.
(30) Yan, B. X.; Sun, Y. Q. Glycine residues provide flexibility for enzyme active sites. *J. Biol. Chem.* **1997**, *272*, 3190−3194.
(31) Martin, W. R.; Cheng, F. X. Repurposing of FDA-Approved toremifene to treat COVID-19 by blocking the spike glycoprotein and NSP14 of SARS-CoV-2. *J. Proteome Res.* **2020**, *19*, 4670−4677.
(32) Wang, C.; Li, W.; Drabek, D.; Okba, N. M. A.; van Haperen, R.; Osterhaus, A. D. M. E.; van Kuppeveld, F. J. M.; Haagmans, B. L.; Grosveld, F.; Bosch, B.-J. A human monoclonal antibody blocking SARS-CoV-2 infection (vol 11, 2251, 2020). *Nat. Commun.* **2020**, *11*, 2251.
(33) van Dorp, L.; Acman, M.; Richard, D.; Shaw, L. P.; Ford, C. E.; Ormond, L.; Owen, C. J.; Pang, J.; Tan, C. C. S.; Boshier, F. A. T.; Ortiz, A. T.; Balloux, F. Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. *Infect., Genet. Evol.* **2020**, *83*, 104351.
(34) Li, Q.; Wu, J.; Nie, J.; Zhang, L.; Hao, H.; Liu, S.; Zhao, C.; Zhang, Q.; Liu, H.; Nie, L.; Qin, H.; Wang, M.; Lu, Q.; Li, X.; Sun, Q.; Liu, J.; Zhang, L.; Li, X.; Huang, W.; Wang, Y. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. *Cell* **2020**, *182*, 1284−1294.
(35) Tian, X.; Li, C.; Huang, A.; Xia, S.; Lu, S.; Shi, Z.; Lu, L.; Jiang, S.; Yang, Z.; Wu, Y.; Ying, T. Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. *Emerging Microbes Infect.* **2020**, *9*, 382−385.
(36) Menachery, V. D.; Yount, B. L.; Debbink, K.; Agnihothram, S.; Gralinski, L. E.; Plante, J. A.; Graham, R. L.; Scobey, T.; Ge, X. Y.; Donaldson, E. F.; Randell, S. H.; Lanzavecchia, A.; Marasco, W. A.; Shi, Z. L. L.; Baric, R. S. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat. Med.* **2015**, *21*, 1508−1513.
(37) Martin, W. R.; Cheng, F. X. 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**.
(38) Clementi, C.; Nymeyer, H.; Onuchic, J. N. Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein

https://doi.org/10.1021/acs.jmedchem.1c00320
J. Med. Chem. XXXX, XXX, XXX−XXX

folding? An investigation for small globular proteins. *J. Mol. Biol.* **2000**, *298*, 937−953.
(39) 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.
(40) Noel, J. K.; Whitford, P. C.; Sanbonmatsu, K. Y.; Onuchic, J. N. SMOG@ctbp: simplified deployment of structure-based models in GROMACS. *Nucleic Acids Res.* **2010**, *38*, W657−W661.
(41) Wang, Q.; Diehl, M. R.; Jana, B.; Cheung, M. S.; Kolomeisky, A. B.; Onuchic, J. N. Molecular origin of the weak susceptibility of kinesin velocity to loads and its relation to the collective behavior of kinesins. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E8611−E8617.
(42) Wang, Q.; Jana, B.; Diehl, M. R.; Cheung, M. S.; Kolomeisky, A. B.; Onuchic, J. N. Molecular mechanisms of the interhead coordination by interhead tension in cytoplasmic dyneins. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 10052−10057.
(43) Lammert, H.; Schug, A.; Onuchic, J. N. Robustness and generalization of structure-based models for protein folding and function. *Proteins: Struct., Funct., Genet.* **2009**, *77*, 881−891.
(44) Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(45) Veitshans, T.; Klimov, D.; Thirumalai, D. Protein folding kinetics: timescales, pathways and energy landscapes in terms of sequence-dependent properties. *Folding Des.* **1997**, *2*, 1−22.
(46) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750−1759.
(47) Sastry, G. M.; 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*, 221−234.

https://doi.org/10.1021/acs.jmedchem.1c00320
J. Med. Chem. XXXX, XXX, XXX−XXX