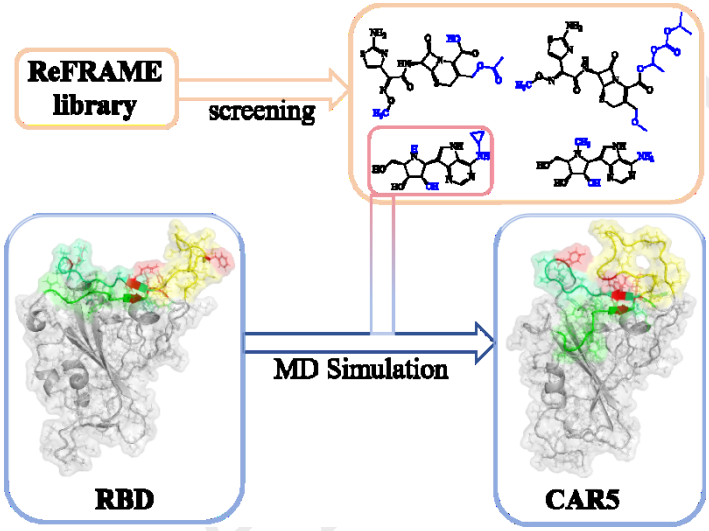
Mechanism of Irreversible Conformational Locking in the SARS-CoV-2 Spike Protein Induced by Transient Molecular Binding

Zijian Liu, Mengke Jia, Chuanbo Wang, Xiang Fan, Yvning Guan, and Hongqi Ai *

School of Chemistry and Chemical Engineering, University of Jinan, Jinan 250022, P.R. China

*To whom correspondence should be addressed, E-mail: *chm_aihq@ujn.edu.cn*

Graphic Abstract



Through structure-based virtual screening from the ReFRAME small-molecule drug library, we identified four small-molecule drugs enabling to be repurposed to induce closure and locking of the receptor-binding motif (RBM) region, thereby occluding the active site.

ABSTRACT:

This study investigates the molecular mechanism of allosteric deactivation of the SARS-CoV-2 spike protein through computational screening and molecular dynamics simulations. We focused on identifying compounds that induce long-lasting conformational changes by targeting non-receptor-binding motif (non-RBM) regions. Our simulations revealed that four small-molecule compounds can stably bind and induce deformations in the receptor-binding motif (RBM), which mask key residues and hinder ACE2 recognition. Most notably, we discovered a unique 'hit-and-run' mechanism: one compound, CA5, triggers an irreversible deformation of the RBM after only a transient contact (~1 ns). Crucially, the conformational change persisted long after the drug dissociates. Energetic and structural analysis identifies the cyclopropylamine group (R3 substituent) of CA5 as the key moiety that facilitates its dissociation while locking the protein in an inactive state. This work elucidated a novel biophysical principle for controlling protein conformation through transient interactions, providing a potential strategy for intervening in protein dynamics that is relevant to antiviral therapy.

Keywords: SARS-CoV-2 spike protein; Molecular dynamics simulations; Allosteric inhibition; Irreversible conformational change; Virtual screening; Hit-and-run mechanism

1 INTRODUCTION

The SARS-CoV-2 spike protein is a dynamic molecular machine whose conformational changes are critical for viral entry into host cells (Gorbalenya *et al.* 2020; Wu *et al.* 2020). Its receptor-binding domain (RBD) alternates between inactive ('down') and ACE2-accessible ('up') states, with the receptor-binding motif (RBM), a rich-mutated region (Fig.1b), serving as the essential interface for engaging the host angiotensin-converting enzyme 2 (ACE2) receptor (Ferrucci *et al.* 2021; Lan *et al.* 2020), while leaving less-mutated non-RBM, the more conserved region across different variants (Fig.1)(Gong *et al.* 2020; Song *et al.* 2020; Yu *et al.* 2022; Zhao *et al.* 2020). Furthermore, although frequent residue mutations in the RBM of various variants, especially of the Omicron variant (Fig.1b) can mediate immune escape, they generally do not affect the structure and function of the non-RBM . (Huang *et al.* 2020; Jiang *et al.* 2022; Khadka and Gupta 2021; Ngo *et al.* 2024; Robson 2020; Sultan Khan *et al.* 2024). Consequently, targeting these highly conserved non-RBM regions is emerging as a novel strategy, which holds promise for delivering broader inhibitory effects against continuously evolving viral variants. (Cao *et al.* 2021; Garg *et al.* 2024; Starr *et al.* 2020; Zhou *et al.* 2023).

A fundamental challenge, however, lies in understanding how molecular interventions at these alternative sites can allosterically perturb the RBM's structure and dynamics to achieve irreversible inactivation. Computational biophysics approaches, particularly molecular dynamics (MD) simulations, provide a powerful tool to dissect these mechanisms at atomic resolution. They allow us to move beyond static binding affinity predictions and directly observe the temporal evolution of drug-induced conformational changes—a capability critical for designing interventions that exploit the spike protein's intrinsic dynamics. This study employed such computational strategies to explore a novel inactivation mechanism: we investigate how transient, specific interactions with small molecules can induce persistent conformational changes that mechanically hinder ACE2 recognition, thereby proposing a new biophysical principle for antiviral intervention (Ashburn and Thor 2004).

2 COMPUTATIONAL METHODS

2.1 Selection and Processing of Viral Models

The SARS-CoV-2 Omicron variant spike glycoprotein was obtained from RCSB (Berman *et al.* 2000) (PDB ID: 7WVN, Fig.2(a1) (Hong *et al.* 2022)). To reduce computational resource requirements, the receptor-binding domain (RBD) portion was extracted using PyMOL

(Rigsby and Parker 2016).

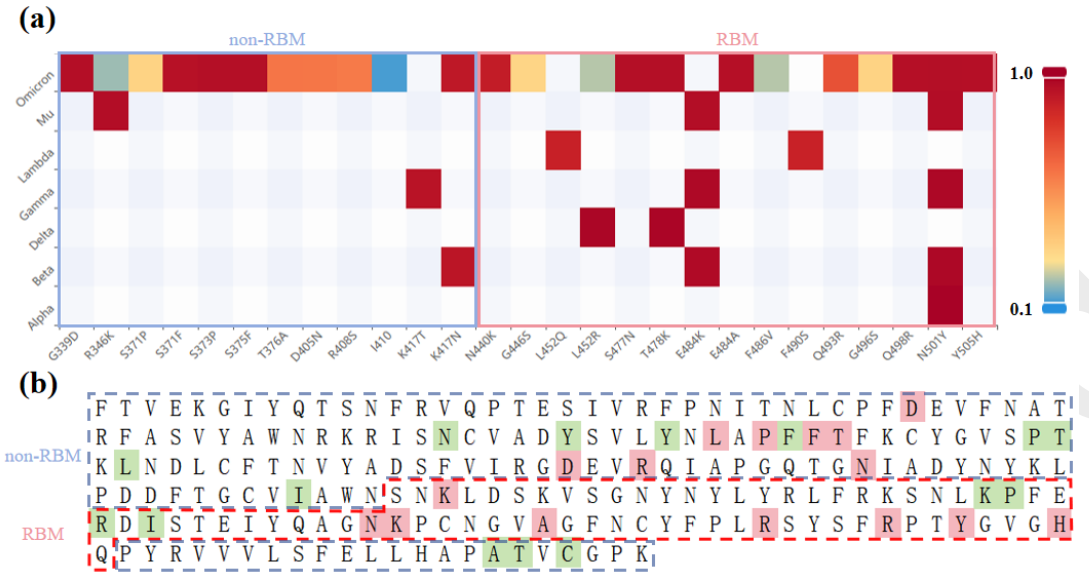


Figure 1 Mutation frequencies of the RBD regions in seven SARS-CoV-2 variants and labels of the mutation residues in the Omicron RBD. (a) Comparison of residue mutation frequencies among seven major SARS-CoV-2 variants, alpha, beta, …, Omicron, in which only the residues with mutation frequency > 0.1 are highlighted in bright color, and a large number of conservative residues were omitted; (b) The amino acid sequence of the RBD region in the spike protein of the Omicron variant. The RBD region includes the RBM and non-RBM areas outlined with the red and blue boxes, respectively. Residues with a mutation frequency greater than 0.5 are highlighted in light red, while small molecule binding sites used in subsequent experiments are marked in light green.

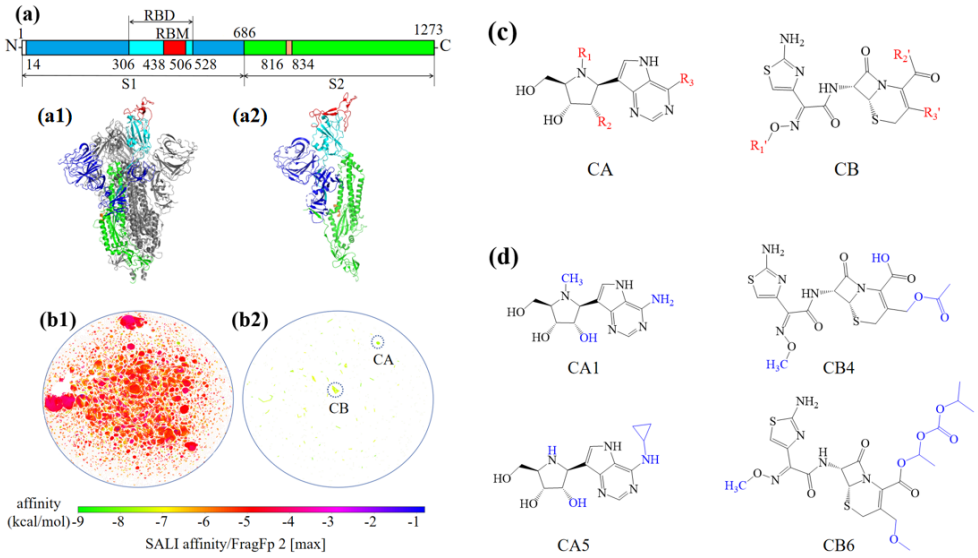


Figure 2 Conformation and sequence of SARS-CoV-2 Omicron monomer and six potential anti-inflammatory small molecules screened by clustering. (a) diagram illustrates the domain organization of the SARS-CoV-2 spike

protein monomer (see corresponding structure of the monomer below (a2)), along with the complete trimeric structure of the SARS-CoV-2 Omicron variant spike glycoprotein (a1). Functional residue regions of the monomer of the trimeric protomer is highlighted in the color consistent with that in (a); (b1) SALI map clustering of 12,640 anti-inflammatory small molecules based on binding energy and structural similarity to the small - molecule - RBD complex. Two most abundant clusters (CA and CB) of small molecules with binding energies below -7.0 kcal/mol are circled in (b2), and redisplayed on the right panel (c) in which two core structures are shown in black and three different substituents for each (R1, R2, and R3 for CA; R1', R2', and R3' for CB) are highlight in red; (d) Substitute for these substituents Ri/Rj' (i, j'=1, 2, 3) with their specific groups (shown in Tables S1. S2 of SI), 12 small molecules are obtained. Among them, four can bind to the non - RBM sequence in the RBD and induce conformational changes in the RBM region, causing it to close. The four small molecules are shown in (d).

2.2 Selection and Processing of Drug Molecules

The ReFRAME anti-inflammatory small-molecule library (Janes *et al.* 2018) was chosen as the source for screening. A python script was used with OpenBabelGUI (O'Boyle *et al.* 2011) to batch - convert 12,640 small - molecule drugs from sdf to pdbqt format. The ReFRAME library, supported by the Bill & Melinda Gates Foundation and others, was developed by the California Institute for Biomedical Research as a global drug repurposing platform. Of the 12,640 compounds (Fig.2(b1), approximately 35% are FDA-approved or registered drugs, around 58% are investigational new drugs in various stages of clinical development, and about 3% are preclinical compounds with available safety data, thus employed to accelerate the development of innovative therapies for infectious diseases (such as COVID-19) and cancer through high-throughput screening.

2.3 High-Throughput Screening Methods

Before high-throughput screening, the target protein was processed using MGL AutoDock Tools (Forli *et al.* 2016). After removing water molecules, polar hydrogen atoms were added, and Gasteiger charges were applied. The docking target was identified using the Grid function. To save computational resources, a grid box of $33 \times 50 \times 25 Å^3$ was created, with x, y, and z coordinates centered on the binding cavity. A Python script was built to perform batch docking with AutoDock Vina (Eberhardt *et al.* 2021), conducting five docking runs for each small molecule in the ReFRAME library. The best results for each molecule were compiled into a table. A SALI (Structure-Activity Landscape Index) clustering plot was generated via DataWarrior's (Sander *et al.* 2015) clustering analysis (clustering criterion: molecular structure; display parameter: binding energy) for small-molecule ligands exhibiting binding energies < 0 kcal/mol, clustering compounds with 80% structural similarity and coloring

them based on binding energy (Fig.2(b1)). Based on docking data in which 95% of compounds exhibited binding energies ≤ -7 kcal/mol against the viral protein, we selected the two largest clusters meeting this criterion for further testing. They were designated as Cluster A and Cluster B (CA and CB) (Figure 2 (b2)), respectively. The main structures of the two clusters characterize a scaffold and three substituents (R1, R2, R3 for CA and R1', R2', R3' for CB) (Fig.2(c)). Details of R1, R2, R3 and R1', R2', R3' are displayed in Tables S1 and S2 of the Supporting Information (SI). With these substituents, the two clusters are further subdivided into CA*i* and CB*j* (*i, j* =1, 2, …, 6), respectively, composed of 12 small molecules. Subsequently, AutoDock tools were employed again to assemble the 12 small molecules with RBD, respectively, to form 12 CA/CB-RBD (CAR/CBR) complexes.

2.4 Docking Methods

During system construction, the prepared target protein was imported into MGL AutoDock Tools. Nonpolar hydrogen atoms were merged, and rotatable bonds were detected for docking of all ligands. A grid box of 33 × 50 × 25Å^3 was created, with x, y, and z coordinates centered on the binding cavity. Other parameters were set to default values, and the Lamarckian genetic algorithm was used to calculate 200 possible conformations of the ligand and macromolecule. For each docking result, unreasonable docking sites were excluded, and the most reasonable conformations were selected based on binding energy intensity and structural distribution probability, with the results exported as PDB files.

2.5 Molecular Dynamics Simulations

The complex models were constructed via the CHARMM-GUI (Jo *et al.* 2008; Lee *et al.* 2016b) online platform. The complex systems were placed in a cube box with each face approximately 10 Å away from the protein, and surrounded by the TIP3P water molecules (Jorgensen *et al.* 1983) to mimic a solvent environment. 0.15 M NaCl solution was added to mimic physiological ion concentration.

For energy minimization, the steepest descent method was used with a convergence criterion of 1000 kJ/mol·nm and a maximum of 5000 steps. The constant temperature 310 K was maintained using Nose-Hoover (τT = 1 ns) (Bosko *et al.* 2005) in the NVT ensemble and the constant pressure was maintained at 1 bar using the Parrinello-Rahman (τP = 125 ps) (Bosko *et al.* 2005) pressure coupling algorithm in the NPT ensemble. Molecular dynamics

simulations were performed using the GROMACS (Abraham *et al.* 2015) 2020 software package with the Charmm36m (Huang *et al.* 2017) force field under periodic boundary conditions. After simulation, structures were checked using PyMOL (Brunger and Wells 2009) to ensure no disruption of protein-ligand interactions. Simulation times were extended for MD trajectory analysis using GROMACS tools.

Parallel Simulations: Each system underwent energy minimization, sequential 100 ps NVT and NPT equilibration, followed by initial production simulations of 1 ns, from which initial sampling with three snapshots at 1ns (Run 1), 0.3ns (Run 2), 0.6ns (Run 3), respectively, was extracted as starting structures. The methodological framework aligns well with Zeiske et al's approach (Zeiske *et al.* 2013) which demonstrated that multi-replicate sampling from early-equilibrium snapshots enhances conformational coverage without requiring extended simulation durations. Subsequently, 300 ns production simulations were performed for each system, with simulation time extended in 100 ns increments until convergence was achieved.

2.6 Residue Contact Map Analysis

2.6.1 RMSD/RMSF

Based on the models built from Charmm-GUI (Jo *et al.* 2008; Lee *et al.* 2016a) web interface (https://charmm-gui.org/), molecular dynamics (MD) simulations were performed by employing Gromacs 2020.4 pockage (Abraham *et al.* 2015). The system was considered to have reached equilibrium when the root mean square deviation (RMSD) of the protein backbone stabilized, fluctuating below 0.1 nm for a minimum consecutive period of 10 ns in the latter part of the trajectory (Fig. S1). Upon confirming stability, no further extension of simulation time was performed. Subsequently, the protein structures from the final 100 ns were subjected to clustering analysis using the gmx cluster command. To assess drug-induced residue flexibility, the Cα atoms from the most populated cluster were analyzed for their root mean square fluctuation (RMSF) with the gmx rmsf command.

2.6.2 MM/PBSA

The binding free energy between proteins and ligands was calculated using the molecular mechanics Poisson-Boltzmann (PB) surface area (MM/PBSA) method in GROMACS (Abraham *et al.* 2015). The method could be also used to solve the effect of electrostatic, van der Waals interaction and solvation change on binding affinity (Wang *et al.* 2016). The Debye-Huckel shielding method (Ding *et al.* 2021) was employed to consider the ionic strength so that the calculated binding free energy was more reasonable. The last 10 ns (equating to 100

frames) of the equilibrated molecular dynamics trajectory was used to calculate the average protein-ligand binding free energy (ΔGbinding) via the following equations, ensuring that the data represented a stable system and provided an accurate description of the complex's physicochemical properties(Ge *et al.* 2021):

$$\Delta G_{binding} = G_{complex} - (G_{protein} + G_{ligand}) \quad (2.1)$$

$$\Delta E_{MM} = \Delta E_{COU} + \Delta E_{VDW} \quad (2.2)$$

$$\Delta G_{solv} = \Delta G_{PB} + \Delta G_{SA} \quad (2.3)$$

$$\Delta G_{binding} = \Delta E_{MM} + \Delta G_{solv} - T\Delta S = \Delta H - T\Delta S \quad (2.4)$$

In these equations, ΔE_{MM} is the gas phase interaction energy between protein and ligand, which is composed of the change of electrostatic energy contribution ΔE_{COU} and the change of van der Waals contribution ΔE_{VDW}; T is the temperature and ΔS is the contribution of entropy change; ΔG_{solv} is the contribution of solvation free energy to binding, ΔG_{PB} and ΔG_{SA} are polar and nonpolar components of the solvation free energy, respectively. MM/PBSA decomposition analysis was also used to identify and compare the contribution of residue-ligand interaction pairs. Each residue-ligand interaction pair also consists of four parts mentioned above: ΔE_{VDW}, ΔE_{COU}, ΔG_{PB}, and ΔG_{SA}.

The updated binding free energy calculation script from https://jerkwin.github.io/2021/03/16/gmx_mmpbsa was employed, providing enhanced consideration of screening effects and entropic contributions. This implementation accounts for ionic strength while utilizing the Debye length for ΔE_{cou} calculations, exponentially attenuating electrostatic interactions (Colwill *et al.* 2022; Genheden and Ryde 2015; Sun *et al.* 2014).

$$\Delta G_{fit} = 0.05402\ (\Delta E_{cou} + \Delta G_{pb}) + 0.14852\ \Delta E_{vdw} + 0.05584\ \Delta G_{np} + 0.11351\ (-T\Delta S) - 4.77148 \quad (2.5)$$

Final calibration of binding energies used Huang et al.'s free energy estimator (Huang *et al.* 2020) for closer approximation to experimental values (2.5).

2.6.3 Residue Contact Map Analysis

In this study, pairwise residue distances were calculated using the gmx mindist tool in GROMACS 2020, and data was converted into images and plotted automatically using gnuplot. The command recorded and calculated the covariance of backbone C atom (C_α) position fluctuations in equilibrated trajectories. A time interval (-dt) was set to obtain conformations periodically, and C_{ij} (dynamic cross-correlation matrix) was calculated using a

specific formula (2.5).

$$C_{ij}=\frac{(\Delta r_i \times \Delta r_j)}{\sqrt{(\Delta r_i^2 \times \Delta r_j^2)}} \quad (2.5)$$

3 RESULTS AND DISCUSSION

3.1 Distinctions of the binding region (BR) for the 12 small molecules

MD results indicated that among the 12 small molecules studied, only CA2 and CA6 can directly bind to the RBM region (termed as BR1, Fig.S2 in SI), therefore block SARS-CoV-2 entry into host cells (Wang *et al.* 2020) to interact with human ACE2 protein, having the potential as targeted therapeutics. Another six small molecules (CA3, CA4, CB1, CB2, CB3, and CB5) were found to primarily bind to the non-RBM regions of the RBD (termed as BR2, Fig.S3 in the SI), no significant structural changes were observed in the viral architecture, however. CA3 is reported to inhibit viral RNA replication by suppressing RNA polymerase activity, thus inhibiting the virus (Julander *et al.* 2021). Notably, CA4 has a binding energy of -10 kcal/mol, which is far more than the -7 kcal/mol of other small molecules, suggesting a greater research potential (Cairo *et al.* 2002) than other five from the view of binding strength point. The remaining four small molecules, CA1, CA5, CB4, and CB6 (Fig.2(d)), also exhibit preferentially binding to BR2 (Fig.2(a)). However, this interaction induces significant conformational deformation of the RBM, occluding its receptor-binding residues and thereby impairing ACE2 engagement, while concurrently circumventing immune escape (Garg *et al.* 2024; Starr *et al.* 2020; Zhou *et al.* 2023). Consequently, the subsequent discussion will focus primarily on the binding characteristics of these four molecules to BR2.

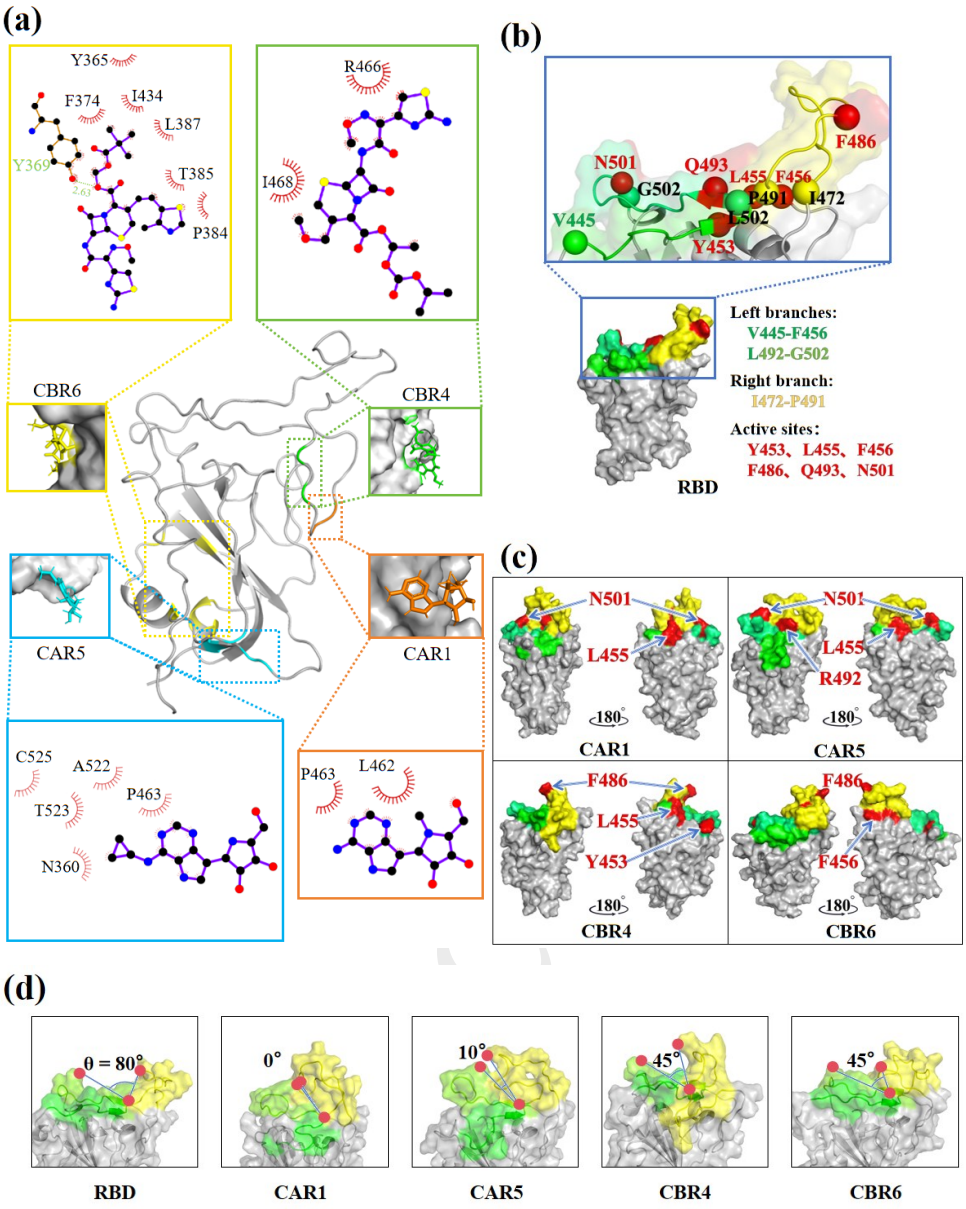


Figure 3 Effects of selected small molecules on the binding sites and conformation of the SARS-CoV-2. (a) Small molecule binding sites within the viral RBD domain (highlighted in corresponding colors in the viral cartoon representation) and specific residues involved in binding CA1, CA5, CB4, and CB6, respectively; (b) Division of the primary binding residues (marked in red) and surrounding residues into left and right branches based on their spatial positions on the RBM, with the left branch (residues V445-L455 and R492-G502) colored in green and the right branch (residues I472-P491) colored in yellow, respectively. Regions excluding the two branches are marked in gray; (c) Persistently exposed key residues (in red) after the closure of the mouth derived from the right branch close over the left branches upon contact of small molecule CA1, CA5, CB4, or CB6 to the non-RBM regions of the RBD; (d) A change in angle θ characterizes the conformational deformation of the RBM region induced by small-molecule binding. The angle θ is defined as the angle between the Cα-Cα vectors P491→N501 and P491→F486, representing the left and right branches, respectively.

3.2 Binding sites and conformational deformation defined by two branches

Fig.3(a) illustrates the binding sites and involved residues within the BR2 of the small molecules CA1, CA5, CB4, and CB6. As demonstrated by the Root Mean Square Fluctuation (RMSF) analysis (Fig.4(a)), the viral residues V445-L455, A475-G485, and F490-T500 exhibited enhanced conformational flexibility under the influence of the small molecules, a characteristic that aligns with the conformational flexibility features observed in Dhami et al.'s experimental study (Dhami *et al.* 2024). In the Omicron variant, the residues primarily binding to ACE2 are Y453, L455, F456, F486, Q493, and N501 (Quitté *et al.* 2024; Yang *et al.* 2023), dominant to be impacted by any of the four small molecules if these molecules are present. Considering the spatial distribution of these residues within the RBM region, closely proximal residues and their adjacent residues will be considered as a part. Consequently, residues V445-L455 and R492-G502 were designated as the left branches (LB) (branches 1 and 2, LB1 and LB2) of the RBM, while residues I472-P491 were assigned as the right branch (RB, Fig.3(b)). As shown in Fig.3(c) and Fig.3 (d), upon binding, small molecules induce the left and right branches to approach each other, masking at least half of the key residues and thereby obstructing RBM-ACE2 binding (e.g., only L455 and N501 remain exposed, while other four sites are already obscured in CAR1). This revealed a novel interaction pattern and a potentially unprecedented mechanism. The following discussion will focus on the four small molecules in preclinical optimization and their corresponding RBD complexes: CAR1, CAR5, CBR4, and CBR6.

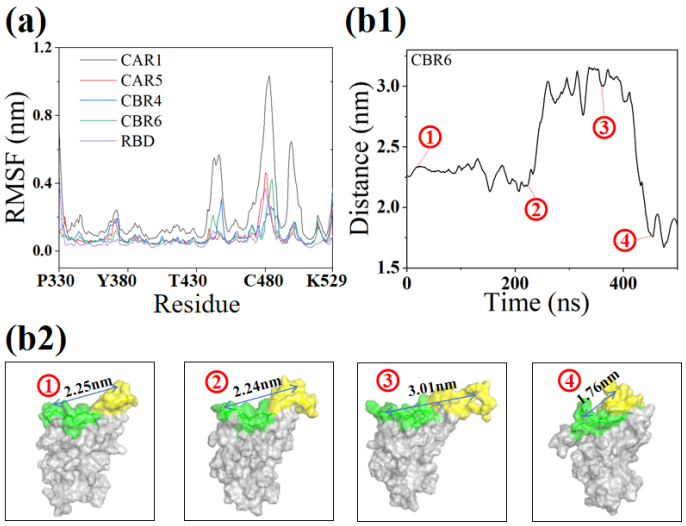


Figure 4 Deformation progress marked by RMSF and distance changes. (a) RBD RMSF plots: free state, RBD, and complexed states, CAR1, CAR5, CBR4, and CBR6; (b1) Time-dependent minimum distance between left branches and right branch in the CBR6 system; (b2) Detailed structural changes of four key

snapshots, with the left branches marked in green and the right branch marked in yellow. The distance between the Cα atoms of N501 and F486 is measured as a proxy for the inter-branch separation.

3.3 Revelation of Allosteric Shielding Mechanisms

Based on their distinct binding outcomes, the four compounds are categorized into two groups: Group 1 comprises CA1, CB4, and CB6 (capable of stable binding at the target site), while Group 2 exclusively contains CA5 (which initiates spontaneous dissociation after transient binding).

3.3.1 *Changes of secondary structure of two branches induced by group 1 molecules*

CA1, CB4, and CB6 consistently induced similar dynamic changes of the branch conformations: initial slight extension of the viral left and right branches followed by subsequent mutual approach. Comparative analysis of secondary structure (Fig.5) and viral branch residue contact maps (Fig.6) before and after four small-molecule treatments revealed that four complexes exhibited slight reduction in β-sheet content followed by significant increases in Bend/Turn structures, relative to free RBD. This structural reorganization consequently drove the sequential extension and mutual approximation of the two branches, leading to novel interfacial contacts that occluded key functional residues to contact the ACE2 (Fig.3 (b, c)). Take CBR6 system as an example. The binding of CB6 reduced β-sheet content from 0.19 to 0.18, accompanied by extension of the LB-RB distance from ~2.3 nm to ~3.0 nm (Fig.4(b)), which was temporarily sustained. Concurrently, Bend/Turn proportion increased from 0.23 in RBD to 0.30 in CBR6 (Fig.5), triggering the right branch close progressively toward the left branch (distance reduced to ~1.8 nm; Fig.4(b)). New interfacial contacts are formed between residues L452 (LB1) and A484 (LB2), and between I472 (RB) and R498 (LB2), wrapping four major functional residue sites (Y453, L455, Q493, N501) and remaining F456 and F486 open in CBR6 (Figures 2(c) and 5), thereby greatly inhibiting the spike protein to infect human ACE2.

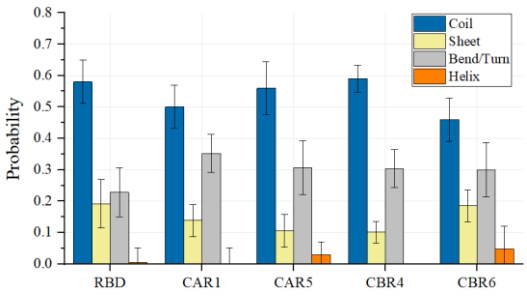


Figure 5 Secondary structure compositions of both left and right branches in the viral RBM region of RBD and four complexes CAR1, CAR5, CAR4, and CAR6.

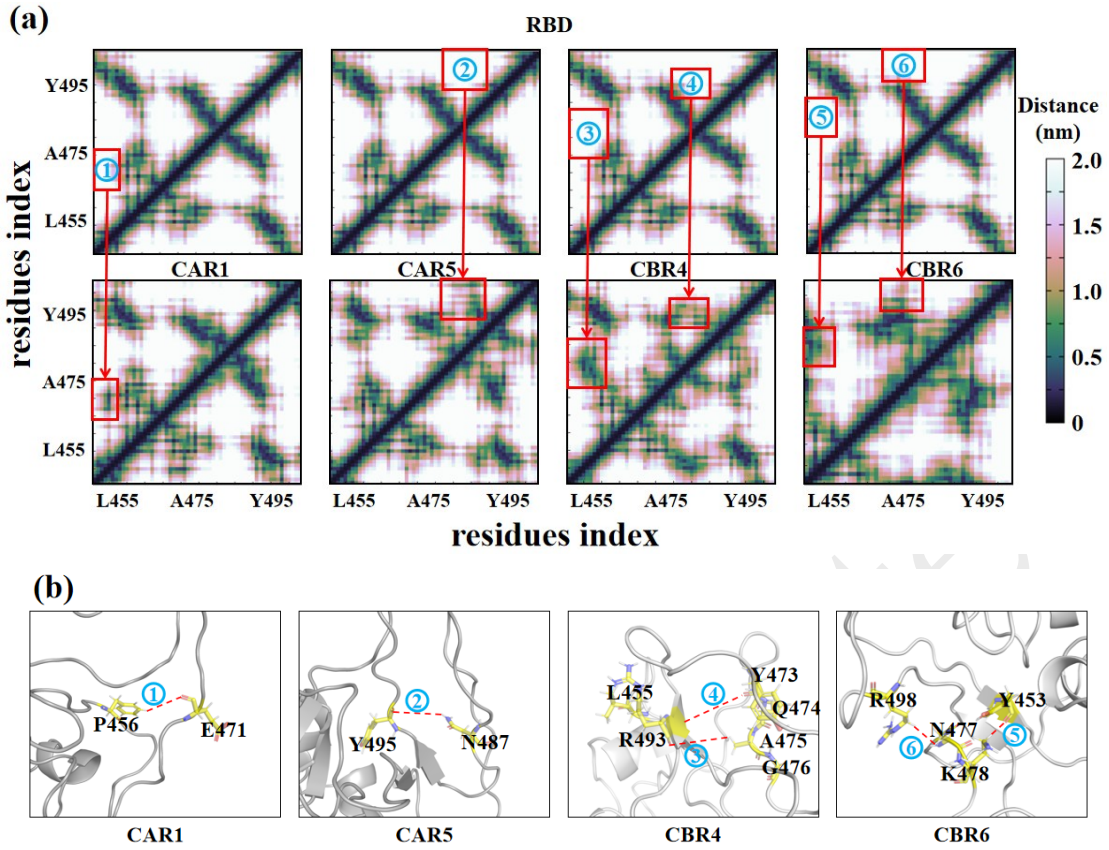


Figure 6 Inter-branch residue contacts in the RBD induced by small-molecule binding. (a) Residue contact maps between the left and right branches in the RBD before (panels in 1st row) and after small compounds bound (panels in 2nd row). New contact regions, highlighted with red boxes, include those between F456 (LB1) and E471 (RB) in CAR1; N487 (RB) and Y495 (LB2) in CAR5; L455 (LB1) and Y473-A475 (RB), and G476 (RB) and R493 (LB2) in CAR4; Y453 (LB1) and K478 (RB), and N477 (RB) and R498 (LB2) in CAR6; (b) 3D contact structures, in which six key interactions (primarily hydrogen bonds between critical residues) across the four complex systems, are marked in ①-⑥.

3.3.2 *Validation of irreversible conformational locking*

Given the relatively weak binding (Tables S1 and S2 in SI) between the three small compounds in group 1 and the viral target and apt to dissociation, additional simulations were conducted following the ligand removal from the deformed complexes to evaluate the reversibility of small molecule-induced conformational changes. Superimposition of equilibrated RBM structures before and after ligand removal revealed high structural similarity with minimal conformational alterations (Fig.7), indicating an irreversible structure deformation. These findings demonstrate that across all three complex systems, the deformed region failed to reopen and persistently blocks the active site required for host binding after small-molecule-induced closure and subsequent dissociation from the RBM region.

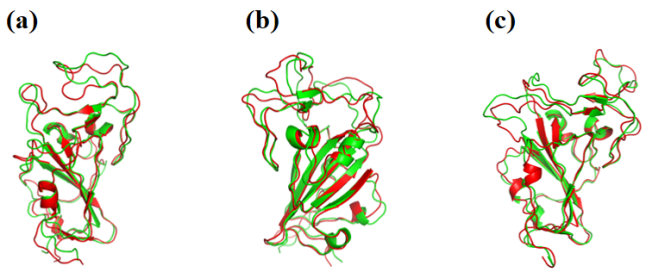


Figure 7 Overlap comparison of RBD conformations in CAR1(a), CBR4 (b) and CBR6 (c) before (green) and after (red) ligand removal.

3.4 Aberrant Dissociation Behavior and Transient Induction Effect of CA5

3.4.1 *The dissociation dynamics and features of CA5 from its complex*

Unlike the metastable binding of the three small molecules to the RBM (≤ -5.37 kcal/mol, Tables S1 and S2 in SI), the small molecule CA5 in the CAR5 system fails to bind stably to the virus (Fig.8). Initially, it dissociated from the original docking site of RMD and migrated to the surface of residues E516-L518, establishing contact with the virus at 85 nanoseconds (Fig.8). After a pause of ~1 ns, the small molecule begins to detach from the RMD. Correspondingly, the right branch of the RBM starts to extend but gradually recovers as the small molecule dissociates. At 191 ns, the small molecule makes the second transitory contact (~1 ns) with residues V367 and P373, then detaches again, inducing contact between the left and right branches and formation of a stable interaction interface.

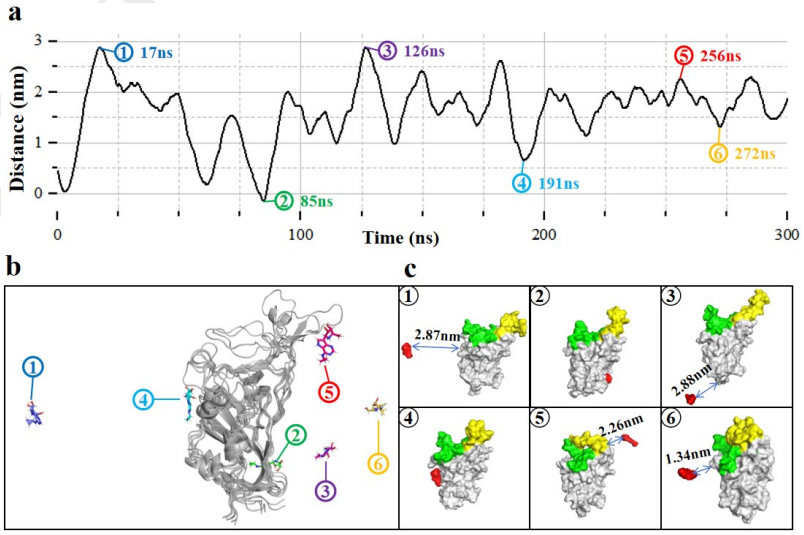


Figure 8 Spatial evolution of CA5 relative to RBD in complex CAR5. (a) Minimum distance between CA5 and RBD over time, with six key snapshots selected; (b) 3D spatial views of CA5 (color-coded) against the

RBD (gray) at the six time points mentioned in (a); (c) The spatial orientation of CA5 across six conformational snapshots and the variation in the minimum distance between the small molecule and the protein are presented. The ligand is highlighted in red, with the left and right branches colored in green and yellow, respectively, to facilitate comparison of structural changes.

To ensure reproducibility, we constructed two additional sets of parallel simulations (CAR5-1 and CAR5-2) from sampling new CA5-RBM structures (see details of the method in section 1.5 of SI). When the simulations were extended to 500 nanoseconds, CAR5-2 exhibited similar behavior to the original CAR5, with the small molecule inducing RBM closure before dissociating (Fig.S4 in SI). However, in CAR5-1, after the small molecule initially contacted the virus, the right branch of the virus extended and attempted to approach the left branches. The proximity was unstable—before a new β-turn form between the two branches, the small molecule dissociated following brief contact. This caused the virus's right branch to reinitiate extension (Fig.S5 in SI) instead of forming stable interfacial interactions with the left branches. Three parallel experiments collectively demonstrated that the CA5 compound consistently dissociated from the virus and probably induce RBM closure.

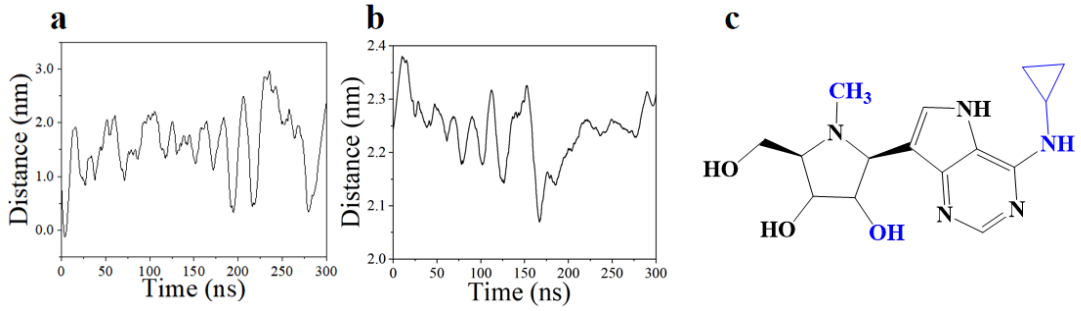


Figure 9 Structure and distance changes of CA1s. (a) Shortest distance between CA1s and the virus over time. (b) Time-dependent minimum distance between the left and right branches of RBM under CA1s's influence over time. (c) Structural formula of CA1s.

3.4.2 *Verification of cyclopropylamine moiety-mediated dissociation propensity*

To elucidate structural features underlying CA5's distinct behavior, we performed comparative structural analysis against other CA cluster small molecules. We found that CA5's R3 substituent is cyclopropylamine, different from other CA small molecules (e.g., the amino group in CA1), which may confer distinct properties to CA5. To verify this, the amino group at the R3 position of CA1 was substituted with cyclopropylamine, to yield a new small

molecule CA1s (Fig.9c). Then we conducted three parallel simulations for its complexes CAR1s-n (n=1, 2, 3) (Fig.S6 in SI). The minimum distance plot (Fig.9a) revealed that CA1s dissociated shortly after binding and intermittently approaching the virus, maintaining a relative distance of 1-3 nm.

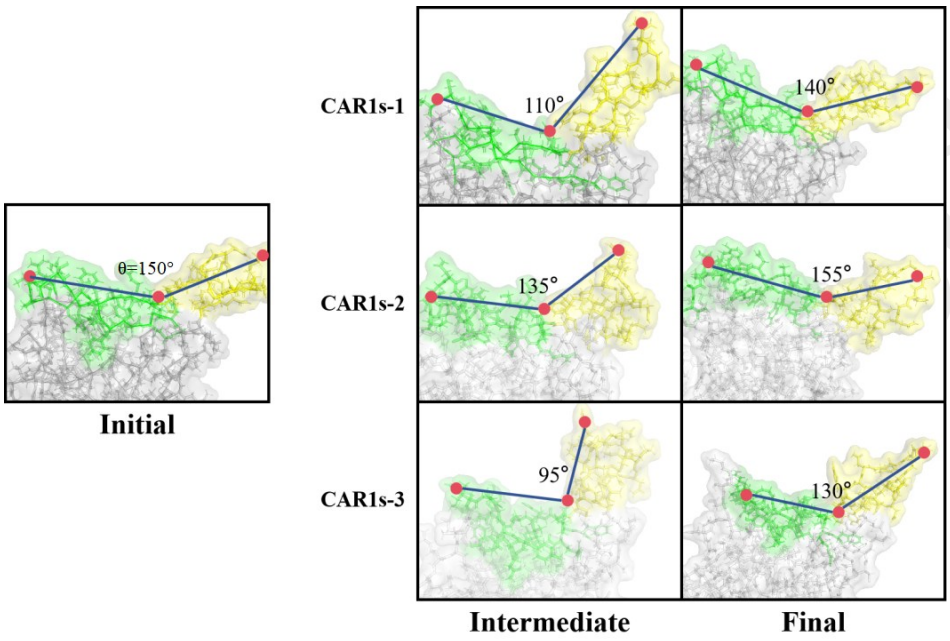


Figure 10 Schematic representations illustrating the LB-RB angles in the initial, intermediate, and final structures from three independent replicates of CAR1s-n (n=1, 2, 3) simulations. See θ definition in Figure 3d.

Three replicate MD simulations uniformly showed that CA1s binding to the RBD triggered transient convergence of the RBM's left and right branches. Subsequent gradual divergence of these branches results in the restoration (open) of the space between BR1 and BR2 throughout the simulation period (Fig.10). For instance, the initial angle between the branches in the CAR1s-3 measured 150°. After CA1s bound to the RBD, this angle reduced to 95°. Upon dissociation of CA1s from the viral target at 140 ns, the angle progressively returned to 130°. Dissociation of the small molecule from the viral target was also observed in both the CAR1s-1 and CAR1s-2 simulations. In the three CAR1s-n simulations, CA1s always reflected a high propensity for dissociation as its binding energy only maintained at -5.40±1.59 kcal/mol (Table S3 in SI), corresponding to large dissociation constant (1.04×10^5nM). Therefore, these parallel experiments confirmed the cyclopropylamine moiety plays a key role in promoting the dissociation of active ligands, such as CA5. This finding provides a clear direction for drug design. However, elucidating the dominant functional group

driving RBM (RNA-binding motif) closure is challenging, due to intricate intermolecular interactions between them.

4 CONCLUSIONS

In this study, we employed molecular dynamics simulations to elucidate a novel biophysical mechanism for the allosteric inactivation of the SARS-CoV-2 spike protein. Experimental studies have established that SARS-CoV-2 infection is primarily initiated by the binding of the viral receptor-binding motif (RBM) within the spike protein to the angiotensin-converting enzyme 2 (ACE2) receptor on host cell surfaces—a key interaction governing viral infectivity, pathogenicity, and host range (Shang *et al.* 2020; Wang *et al.* 2020; Yang *et al.* 2020). Through MD simulations, we demonstrate that by targeting conserved non-RBM regions, small molecules can induce an irreversible conformational lock in the RBM, effectively masking key residues required for RBM-ACE2 binding, thereby inhibiting SARS-CoV-2 infection.

Our simulations revealed two distinct pathways to achieve this inhibitory state, including (1) stable binding-induced occlusion for Compounds CA1, CB4, and CB6, that stabilize the RBD in a closed conformation through persistent binding, creating a stable interface between the RBM's structural elements; and (2) a 'hit-and-run' mechanism for compound CA5, that represents a more dynamic and previously underappreciated mechanism. It triggers a complete conformational rearrangement of the RBM after only a transient (~1 ns) association. Crucially, the inactivation persists long after the ligand dissociates, suggesting that the molecule acts as a catalyst for conformational change rather than a static stabilizer. More importantly, CA5's unique strengths demonstrate its active dissociation from the target, conferring dual therapeutic advantages: reduced risk of resistance (through lowered selective pressure) and an extended therapeutic window (via sustained plasma concentrations)(Dai *et al.* 2020).

We also observe that the "hit-and-run" action is facilitated by specific molecular interactions, particularly those involving the cyclopropylamine (R3) group of CA5, which enable efficient dissociation after initiating the deformation. The induced conformational shift is characterized by a significant increase in β-turn/bend content and a decrease in random coil, indicating a transition towards a more structured, locked state.

Beyond proposing specific antiviral candidates, this work establishes a general biophysical principle: transient, specific molecular interactions can be sufficient to drive persistent, large-scale conformational changes in proteins. This mechanism, which exploits the intrinsic dynamics of the protein, could inform the design of a new class of molecular

interventions that are less susceptible to resistance. Our findings provided a foundational framework for understanding and manipulating protein allostery and conformational dynamics, with implications that extend beyond antiviral development to fundamental biophysics and therapeutic design.

Author contributions

Z. L. performed experiments and analysis of results and wrote the original draft. K. J. and C.W. helped with MD simulations. X.F. and Y.G performed experiments. H. A. put forward the idea and reviewed the manuscript.

Acknowledgements

This work was supported by the Shandong Provincial Natural Science Foundation (ZR2022MB073) of China and the High-performance Computing Platform at University of Jinan.

Compliance with Ethics Guidelines

Conflict of interest

Zijian Liu, Hongqi Ai, Mengke Jia, Chuanbo Wang, Xiang Fan and Yvning Guan declare that they have no conflict of interest.

Human and animal rights and informed consent

This article does not contain any studies with human or animal subjects performed by the any of the authors.

References

Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1-2:19-25

Ashburn TT, Thor KB (2004) Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discov 3(8):673-683

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. Nucleic Acids Res 28(1):235-242

Bosko JT, Todd BD, Sadus RJ (2005) Molecular simulation of dendrimers and their mixtures under shear: comparison of isothermal-isobaric (NpT) and isothermal-isochoric (NVT) ensemble systems. J Chem Phys 123(3):34905

Brunger AT, Wells JA (2009) Warren L. DeLano 21 June 1972–3 November 2009. Nat Struct Mol Biol 16(12):1202-1203

Cairo CW, Strzelec A, Murphy RM, Kiessling LL (2002) Affinity-based inhibition of beta-amyloid toxicity. Biochemistry 41(27):8620-8629

Cao Y, Wang J, Jian F, Xiao T, Song W, Yisimayi A, Huang W, Li Q, Wang P, An R, Wang J, Wang Y, Niu X, Yang S, Liang H, Sun H, Li T, Yu Y, Cui Q, Liu S, Yang X, Du S, Zhang Z, Hao X, Shao F, Jin R, Wang X, Xiao J, Wang Y, Xie XS (2021) Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. Nature 602(7898):657-663

Colwill K, Galipeau Y, Stuible M, Gervais C, Arnold C, Rathod B, Abe KT, Wang JH, Pasculescu A, Maltseva M, Rocheleau L, Pelchat M, Fazel-Zarandi M, Iskilova M, Barrios-Rodiles M, Bennett L, Yau K, Cholette F, Mesa C, Li AX, Paterson A, Hladunewich MA, Goodwin PJ, Wrana JL, Drews SJ, Mubareka S, McGeer AJ, Kim J, Langlois MA, Gingras AC, Durocher Y (2022) A scalable serology solution for profiling humoral immune responses to SARS-CoV-2 infection and vaccination. Clin Transl Immunology 11(3):e1380

Dai W, Zhang B, Jiang XM, Su H, Li J, Zhao Y, Xie X, Jin Z, Peng J, Liu F, Li C, Li Y, Bai F, Wang H, Cheng X, Cen X, Hu S, Yang X, Wang J, Liu X, Xiao G, Jiang H, Rao Z, Zhang LK, Xu Y, Yang H, Liu H (2020) Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. Science 368(6497):1331-1335

Dhami LS, Dahal P, Thapa B, Gautam N, Pantha N, Adhikari R, Adhikari NP (2024) Insights from in silico study of receptor energetics of SARS-CoV-2 variants. Phys Chem Chem Phys 26(11):8794-8806

Ding H-m, Yin Y-w, Ni S-d, Sheng Y-j, Ma Y-q (2021) Accurate Evaluation on the Interactions of SARS-CoV-2 with Its Receptor ACE2 and Antibodies CR3022/CB6. Chin Phys Lett 38(1):018701

Eberhardt J, Santos-Martins D, Tillack AF, Forli S (2021) AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. J Chem Inf Model 61(8):3891-3898

Ferrucci V, Kong DY, Asadzadeh F, Marrone L, Boccia A, Siciliano R, Criscuolo G, Anastasio C, Quarantelli F, Comegna M, Pisano I, Passariello M, Iacobucci I, Monica RD, Izzo B, Cerino P, Fusco G, Viscardi M, Brandi S, Pierri BM, Borriello G, Tiberio C, Atripaldi L, Bianchi M, Paolella G, Capoluongo E, Castaldo G, Chiariotti L, Monti M, De Lorenzo C, Yun KS, Pascarella S, Cheong JH, Kim HY, Zollo M (2021) Long-chain polyphosphates impair SARS-CoV-2 infection and replication. Sci Signal 14(690):eabe5040

Forli S, Huey R, Pique ME, Sanner MF, Goodsell DS, Olson AJ (2016) Computational protein-ligand docking and virtual drug screening with the AutoDock suite. Nat Protoc 11(5):905-919

Garg P, Hsueh SCC, Plotkin SS (2024) Testing the feasibility of targeting a conserved region on the S2 domain of the SARS-CoV-2 spike protein. Biophys J 123(8):992-1005

Genheden S, Ryde U (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov 10(5):449-461

Gong Z, Zhu JW, Li CP, Jiang S, Ma LN, Tang BX, Zou D, Chen ML, Sun YB, Song SH, Zhang Z, Xiao JF, Xue YB, Bao YM, Du ZL, Zhao WM (2020) An online coronavirus analysis platform from the National Genomics Data Center. Zool Res 41(6):705-708

Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA, Haagmans BL, Lauber C, Leontovich AM, Neuman BW, Penzar D, Perlman S, Poon LLM, Samborskiy DV, Sidorov IA, Sola I, Ziebuhr J (2020) The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol 5(4):536-544

Hong Q, Han W, Li J, Xu S, Wang Y, Xu C, Li Z, Wang Y, Zhang C, Huang Z, Cong Y (2022) Molecular basis of receptor binding and antibody neutralization of Omicron. Nature 604(7906):546-552

Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, Grubmüller H, MacKerell AD (2017) CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods 14(1):71-73

Huang Y, Yang C, Xu X-f, Xu W, Liu S-w (2020) Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. Acta Pharmacol Sin 41(9):1141-1149

Janes J, Young ME, Chen E, Rogers NH, Burgstaller-Muehlbacher S, Hughes LD, Love MS, Hull MV, Kuhen KL, Woods AK, Joseph SB, Petrassi HM, McNamara CW, Tremblay MS, Su AI, Schultz PG, Chatterjee AK (2018) The ReFRAME library as a comprehensive drug repurposing library and its application to the treatment of cryptosporidiosis. Proc Natl Acad Sci USA 115(42):10750-10755

Jiang S, Wu S, Zhao G, He Y, Guo X, Zhang Z, Hou J, Ding Y, Cheng A, Wang B (2022) Identification of a promiscuous conserved CTL epitope within the SARS-CoV-2 spike protein. Emerg Microbes Infect 11(1):730-740

Jo S, Kim T, Iyer VG, Im W (2008) CHARMM-GUI: a web-based graphical user interface for CHARMM. J Comput Chem 29(11):1859-1865

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79(2):926-935

Julander JG, Demarest JF, Taylor R, Gowen BB, Walling DM, Mathis A, Babu YS (2021) An update on the progress of galidesivir (BCX4430), a broad-spectrum antiviral. Antiviral Res 195:105180

Khadka B, Gupta RS (2021) Conserved molecular signatures in the spike protein provide evidence indicating the origin of SARS-CoV-2 and a Pangolin-CoV (MP789) by recombination(s) between specific lineages of Sarbecoviruses. Peer J 9:e12434

Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L, Wang X (2020) Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature 581(7807):215-220

Lee J, Cheng X, Swails JM, Yeom MS, Eastman PK, Lemkul JA, Wei S, Buckner J, Jeong JC, Qi Y, Jo S, Pande VS, Case DA, Brooks CL, 3rd, MacKerell AD, Jr., Klauda JB, Im W (2016a) CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. J Chem Theory Comput 12(1):405-413

Lee J, Cheng X, Swails JM, Yeom MS, Eastman PK, Lemkul JA, Wei S, Buckner J, Jeong JC, Qi Y, Jo S, Pande VS, Case DA, Brooks CL, III, MacKerell AD, Jr., Klauda JB, Im W (2016b) CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. J Chem Theory Comput 12(1):405-413

Ngo V-N, Winski DP, Aho B, Kamath PL, King BL, Waters H, Zimmerberg J, Sodt A, Hess ST (2024) Conserved sequence features in intracellular domains of viral spike proteins. Virology 599:110198

O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR (2011) Open Babel: An open chemical toolbox. J Cheminform 3:33

Quitté L, Leclercq M, Prunier J, Scott-Boyer MP, Moroy G, Droit A (2024) A Machine Learning Approach to Identify Key Residues Involved in Protein-Protein Interactions Exemplified with SARS-CoV-2 Variants. Int J Mol Sci 25(12):6535

Rigsby RE, Parker AB (2016) Using the PyMOL application to reinforce visual understanding of protein structure. Biochem Mol Biol Educ 44(5):433-437

Robson B (2020) COVID-19 Coronavirus spike protein analysis for synthetic vaccines, a peptidomimetic antagonist, and therapeutic drugs, and analysis of a proposed achilles' heel conserved region to minimize probability of escape mutations and drug resistance. Comput Biol Med 121:103749

Sander T, Freyss J, von Korff M, Rufener C (2015) DataWarrior: an open-source program for chemistry aware data visualization and analysis. J Chem Inf Model 55(2):460-473

Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, Geng Q, Auerbach A, Li F (2020) Structural basis of receptor recognition by SARS-CoV-2. Nature 581(7807):221-224

Song S, Ma L, Zou D, Tian D, Li C, Zhu J, Chen M, Wang A, Ma Y, Li M, Teng X, Cui Y, Duan G, Zhang M, Jin T, Shi C, Du Z, Zhang Y, Liu C, Li R, Zeng J, Hao L, Jiang S, Chen H, Han D, Xiao J, Zhang Z, Zhao W, Xue Y, Bao Y (2020) The Global Landscape of SARS-CoV-2 Genomes, Variants, and Haplotypes in 2019nCoVR. Genomics Proteomics Bioinformatics 18(6):749-759

Starr TN, Greaney AJ, Hilton SK, Ellis D, Crawford KHD, Dingens AS, Navarro MJ, Bowen JE, Tortorici MA, Walls AC, King NP, Veesler D, Bloom JD (2020) Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. Cell 182(5):1295-1310.e1220

Sultan Khan M, Shakya M, Kumar Verma C, Mukherjee R (2024) Identification of highly conserved surface-exposed peptides of spike protein for multiepitope vaccine design against emerging omicron variants: An immunoinformatic approach. Hum Immunol 85(6):111117

Sun H, Li Y, Tian S, Xu L, Hou T (2014) Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set. Phys Chem Chem Phys 16(31):16719-16729

Wang C, Nguyen PH, Pham K, Huynh D, Le TB, Wang H, Ren P, Luo R (2016) Calculating protein-ligand binding affinities with MMPBSA: Method and error analysis. J Comput Chem 37(27):2436-2446

Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, Lu G, Qiao C, Hu Y, Yuen KY, Wang Q, Zhou H, Yan J, Qi J (2020) Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. Cell 181(4):894-904.e899

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 EC, Zhang Y-Z (2020) A new coronavirus associated with human respiratory disease in China. Nature 579(7798):265-269

Yang J, Petitjean SJL, Koehler M, Zhang Q, Dumitru AC, Chen W, Derclaye S, Vincent SP, Soumillion P, Alsteens D (2020) Molecular interaction and inhibition of SARS-CoV-2 binding to the ACE2 receptor. Nat Commun 11(1):4541

Yang K, Liu S, Yan H, Lu W, Shan X, Chen H, Bao C, Feng H, Liao J, Liang S, Xu L, Tang H, Yuan JX, Zhong N, Wang J (2023) SARS-CoV-2 spike protein receptor-binding domain perturbates intracellular calcium homeostasis and impairs pulmonary vascular endothelial cells. Signal Transduct Target Ther 8(1):276

Yu D, Yang X, Tang B, Pan YH, Yang J, Duan G, Zhu J, Hao ZQ, Mu H, Dai L, Hu W, Zhang M, Cui Y, Jin T, Li CP, Ma L, Su X, Zhang G, Zhao W, Li H (2022) Coronavirus GenBrowser for monitoring the transmission and evolution of SARS-CoV-2. Brief Bioinform 23(2):bbab583

Zeiske T, Stafford KA, Friesner RA, Palmer AG, 3rd (2013) Starting-structure dependence of nanosecond timescale intersubstate transitions and reproducibility of MD-derived order parameters. Proteins 81(3):499-509

Zhao WM, Song SH, Chen ML, Zou D, Ma LN, Ma YK, Li RJ, Hao LL, Li CP, Tian DM, Tang BX, Wang YQ, Zhu JW, Chen HX, Zhang Z, Xue YB, Bao YM (2020) The 2019 novel coronavirus resource. Yi Chuan 42(2):212-221

Zhou J, Liu Z, Zhang G, Xu W, Xing L, Lu L, Wang Q, Jiang S (2023) Development of variant-proof severe acute respiratory syndrome coronavirus 2, pan-sarbecovirus, and pan-β-coronavirus vaccines. J Med Virol 95(1):e28172

Supporting Information

Differences in substituent groups of the small molecules, Binding energies, RMSDs of RBD in 12 complexes, Schematic diagrams of binding structures, Variations in distance between the CAR5-1/2 and RBD, RMSDs and Distances of CAR1s.

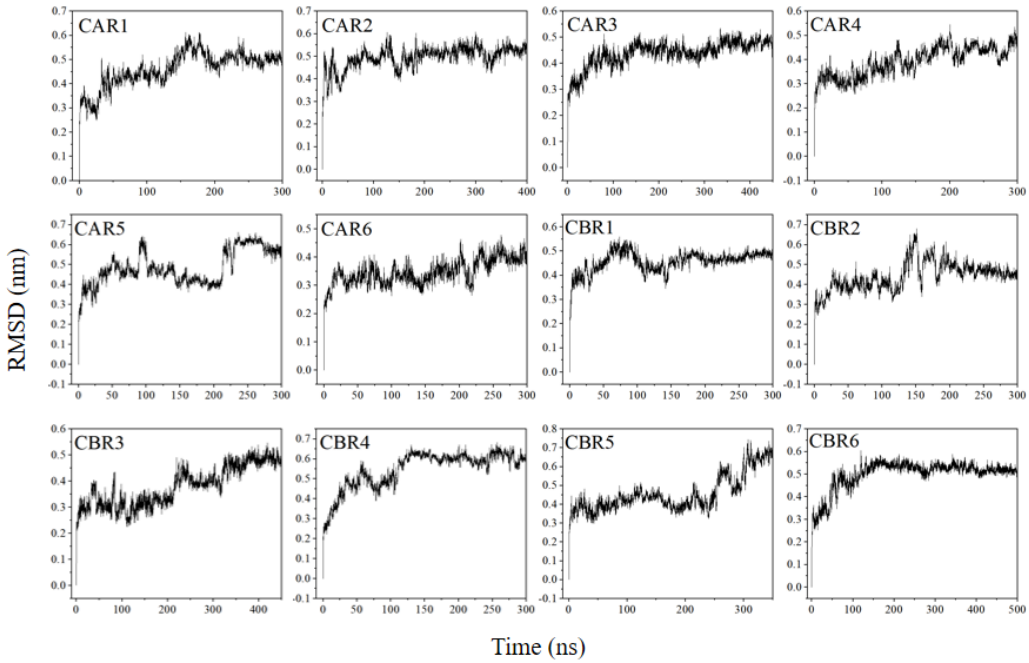


Figure S1 RMSD plots of RBD in 12 complexes.

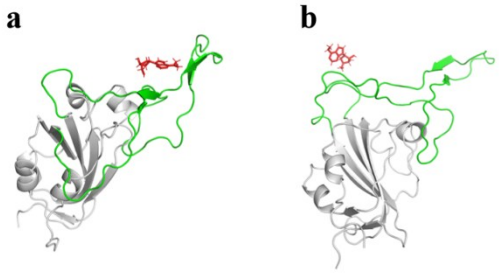


Figure S2 Structural schematic diagram of the binding of small molecules (in red) to the RBD (RBM in green, other regions in gray) in the CAR2(a) and CAR6(b) systems.

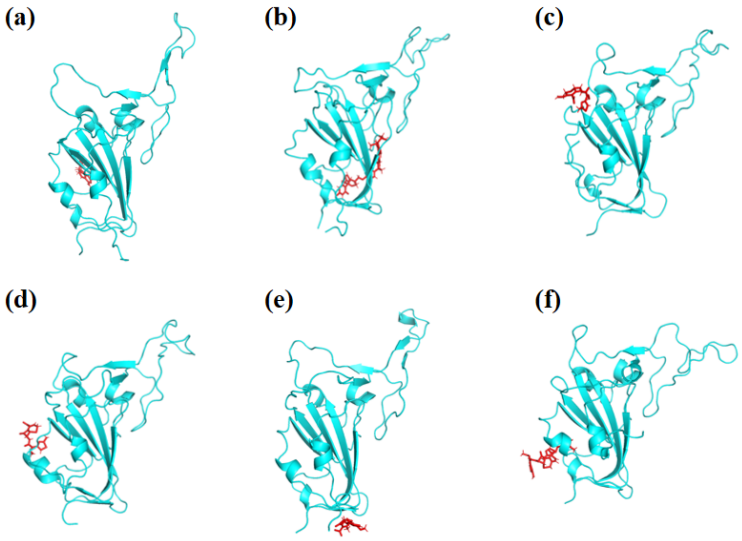


Figure S3 Structural schematic diagram of the binding of small molecules (red) to the viral RBD region (cyan) in the CAR3(a), CAR4(b), CBR1(c), CBR2(d), CBR3(e), and CBR5(f) systems.

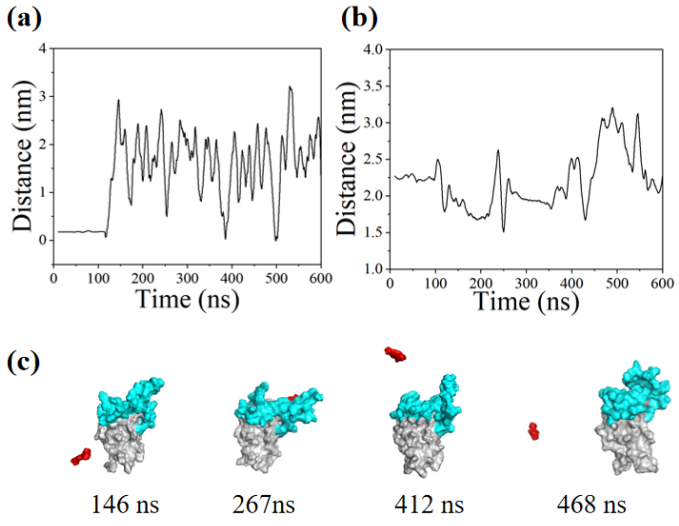


Figure S4 Molecular dynamics results of CAR5-2. (a) The shortest distance between the CA5 and RDB in Run 2;(b) The centroid distance between the left and right branches of CAR5-2;(c) Four selective snapshots of CAR5-2, where CA5 locates at different distance from the RBD.

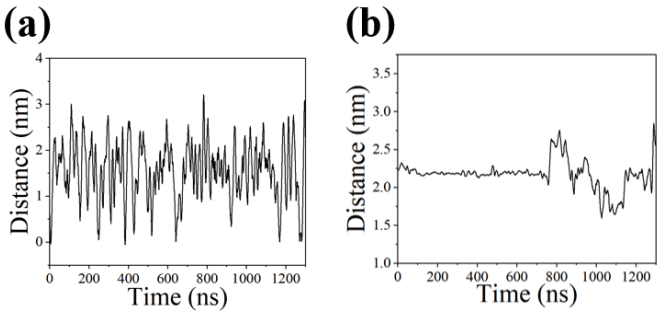


Figure S5 (a) The shortest distance between the CA5 and the virus in CAR5-1 as the reaction evolves;(b) Changes in the centroid distance between the left and right branches of CAR5-1 as the reaction evolves

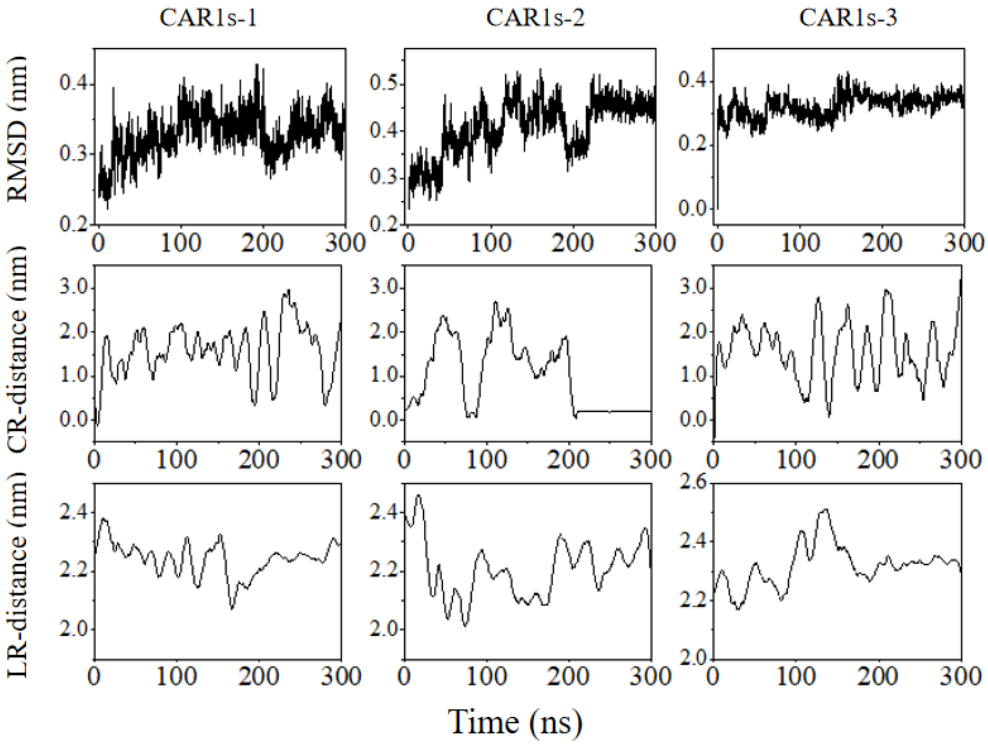


Figure S6 Three parallel experimental datasets for the improved compound: RMSD values; Minimum atomic distance between the compound and viral target (CR-distance); Centroid distance between viral left/right branches under ligand influence (LR-distance).

Table S1 Differences in substituent group of small molecules within the CA cluster and their binding results (kcal/mol).

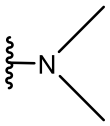
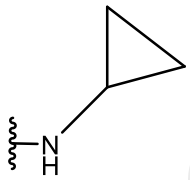
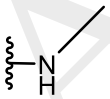
ID	R_1	R_2	R_3	Binding site	Induced deformation	ΔG_{fit}
CA1				non-RBM	**Yes**	-5.37
CA2				RBM	No	-4.95
CA3				non-RBM	No	-4.57
CA4				non-RBM	No	-10.10
CA5				non-RBM	**Yes**	-4.75
CA6				RBM	No	-4.80

Table S2 Differences in substituent group of small molecules within the CB cluster and their binding results (kcal/mol).

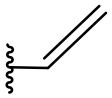
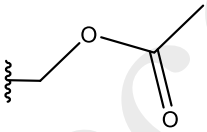
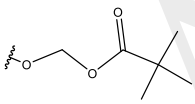
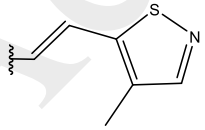
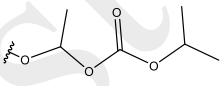
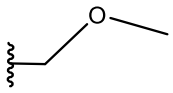
ID	R_1	R_2	R_3	Binding site	Induced deformation	ΔG_{fit}
CB1				non-RBM	No	-5.67
CB2				non-RBM	No	-7.38
CB3				non-RBM	No	-5.56
CB4				non-RBM	**Yes**	-4.43
CB5				non-RBM	No	-7.05
CB6				non-RBM	**Yes**	-4.76

Table S3 Binding energy between engineered small molecules (CA1s) and viral RBD in two of three parallel simulations (kcal/mol).

ID	ΔE_{pb}	ΔE_{sa}	ΔE_{ele}	ΔE_{vdw}	$-T\Delta S$	$\Delta G_{binding}$	ΔG_{fit}
CA1s-1	17.52	-2.26	-12.23	-7.01	-18.30	14.26	-3.58
CA1s-2	31.20	-4.78	-5.14	-30.09	2.77	-6.02	-7.78
CA1s-3	-0.10	-0.93	0.04	-0.01	-0.01	-0.98	-4.83
CA1s-ave							-5.40±1.59