



OPEN Multi-structural molecular docking (MOD) combined with molecular dynamics reveal the structural requirements of designing broad-spectrum inhibitors of SARS-CoV-2 entry to host cells

Anqi Da^{1,3}, Meritxell Wu-Lu^{2,3}, Jovan Dragelj², Maria Andrea Mroginski²✉ & Kourosh H. Ebrahimi¹✉

New variants of SARS-CoV-2 that can escape immune response continue to emerge. Consequently, there is an urgent demand to design small molecule therapeutics inhibiting viral entry to host cells to reduce infectivity rate. Despite numerous *in silico* and *in situ* studies, the structural requirement of designing viral-entry inhibitors effective against multiple variants of SARS-CoV-2 has yet to be described. Here we systematically screened the binding of various natural products (NPs) to six different SARS-CoV-2 receptor-binding domain (RBD) structures. We demonstrate that Multi-structural Molecular Docking (MOD) combined with molecular dynamics calculations allowed us to predict a vulnerable site of RBD and the structural requirement of ligands binding to this vulnerable site. We expect that our findings lay the foundation for *in silico* screening and identification of lead molecules to guide drug discovery into designing new broad-spectrum lead molecules to counter the threat of future variants of SARS-CoV-2.

Abbreviations

ACE2	Angiotensin-converting enzyme 2
MD	Molecular dynamics
MOD	Multi-structural molecular docking
NP	Natural products
PKC	Protein kinase C
RBD	Receptor-binding domain
TMPRSS2	Transmembrane serine protease 2
WT	Wild-type

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) first appeared in December 2019 in Wuhan, Hubei province China¹. Until June 2022, there have been more than 550 million confirmed covid-19 infections worldwide, and the number of deaths is reportedly over six million individuals (WHO Coronavirus Statistics 2022). The global pandemic of the virus and the lack of therapeutics to counter severe disease put high pressure on the medical care system in many countries and have had substantial social, religious, economic, financial, and political burdens. An unprecedented joint effort of scientists around the world led to the development of vaccines and therapies inhibiting viral replication machinery, e.g. remdesivir, to protect the most vulnerable population and reduce mortality rates.

It is now clear that SARS-CoV-2 infects host cells using its homotrimeric surface spike glycoprotein (S protein)^{2,3}. This protein comprises two functional subunits: S1 subunit, responsible for the attachment of the virus

¹Institute of Pharmaceutical Science, King's College London, London, UK. ²Institute of Chemistry, Technische Universität Berlin, Berlin, Germany. ³These authors contributed equally: Anqi Da and Meritxell Wu-Lu. ✉email: andrea.mroginski@tu-berlin.de; Kourosh.Ebrahimi@kcl.ac.uk

to the host cell receptor, and S2 subunit, responsible for the fusion of the viral and cellular membranes⁴. The S protein is cleaved at the boundary of S1 and S2, remaining non-covalently bound in the prefusion conformation, then further cleaved by transmembrane serine protease 2 (TMPRSS2) inducing irreversible conformational changes for binding to the cellular receptor angiotensin-converting enzyme 2 (ACE2) (Fig. 1a). The distal domain of S1, also known as the receptor-binding domain (RBD) (Fig. 1b), is responsible for ACE2 recognition, similar to the mechanism described for SARS-CoV-1⁵. The pocket of RBD interacting with the ACE2 receptor can be divided into R1 and R2 pockets (Fig. 1b) based on the mutations occurring on different major variants affecting RBD-binding to the ACE2 receptor (Fig. 1c). Mutations of the S protein, particularly those in the R1 and R2 pocket of RBD (Fig. 1c), can lead to the rise of variants escaping antibody response and increased RBD binding affinity to the ACE2, undermining the efficacy of current therapies and vaccines^{6–9}. These mutations have led to the rise of new variants, namely Alpha, Beta or Omicron (BA.1) (Fig. 1c). The mutational landscape of the RBD is predicted to lead to new variants of concern with the ability to escape antibody response⁹. Hence, there is a growing interest in discovering broad-spectrum antiviral small molecules binding to R1, R2, or the R1–R2 interface and interfering with the viral entry process to counter the emergence of future variants of concern.

In silico screening, approaches have become an integral part of drug discovery pipelines to enable identification of the best leads in initial stages and reduce costs^{10–13}. Computational studies have largely been used to identify lead natural products (NPs) and other small molecules and nanobodies interfering with RBD binding to the ACE2 receptor^{12,14–23}. Multiple conformation-based virtual screening (ensemble docking) have been introduced to identify SARS-CoV-2 main protease (M^{Pro}) inhibitors among commercially available compounds²⁴. Some of these predictions are confirmed using biochemical studies suggesting that in silico studies using modern computational methods are suitable for discovering lead molecules and guiding drug discovery^{19,24–26}. Computational studies take advantage of the structure of RBD solved in the presence of a binding partner, i.e. ACE2, antibody, or nanobody. In these studies, the protein partner is removed for docking studies. Most computational studies have not tested how certain conformational changes induced in the RBD structure by the binding of protein-partner (ACE2, antibody, or nanobody) can affect in silico screening. Additionally, a vulnerable pocket to which a small molecule ligand can bind, irrespective of the RBD mutations, has yet to be reported. Such a vulnerable pocket is comparable to the vulnerable epitopes of broad-spectrum neutralization antibodies^{27–31}.

Here, we aimed to identify new natural products with the ability to interfere with RBD-binding to the ACE2 receptor. We used in silico multi-structural molecular docking (MOD) screening combined with molecular dynamics (MD) calculations. We systematically screened the binding of more than 70 NPs using three structures of wild-type (WT) RBD solved in the presence of ACE2, antibody, or nanobody and those of three major variants: alpha, beta, and omicron. We tested some of the already identified compounds (SI Table 1) to check if their bindings are affected by the structure of RBD. The results revealed a vulnerable pocket of RBD and the structural details of lead ligands targeting this pocket. The findings suggest key structural requirements for discovering and synthesizing new ligands as broad-spectrum inhibitors of SARS-CoV-2 entry to host cells.

Methods

Molecular docking

Molecular docking simulations were performed using PyRx open-source software, taking advantages of Auto Dock Vina program³². The specific operation method, including preparing input files for docking and running virtual screening using vina wizards are described step by step in Chemical Biology (Dalla Kyan S, Olson A, 2014)³³. It is worth noting that the ACE2, antibody and nanobody fragments in the RBD structures were removed

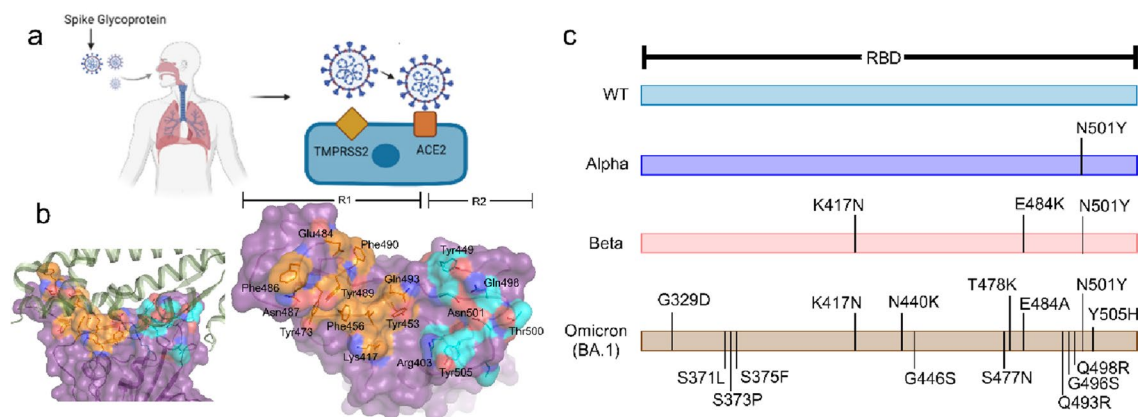


Figure 1. SARS-CoV-2 and its entry process. **(a)** SARS-CoV-2 entry into nasal epithelial cells is a complex process that requires the proteolytic processing by TMPRSS2 to induce conformational changes on spike protein and concerted action of RBD binding to ACE2 receptor. **(b)** Structure of RBD in complex with ACE2 receptor (green) (PDB code: 7C8D). The two regions of RBD encompassing the key amino acid residues interacting with the ACE2 receptor are coloured orange (R1) and cyan (R2). **(c)** Mutations of the RBD of three major variants are shown as compared to wild-type (WT) RBD.

by modifying the protein PBD text file to make the structures suitable for molecular docking. The water molecules in RBD structures were removed when docking since the structures were solved in the absence of a small molecule ligand and the binding pocket of a small molecule ligand could be occupied with water molecules. The Vina search space was performed to cover the top domain of RBD where it interacts with the ACE2 receptor (SI Fig. 1a). An example of the Vina search space defined for molecular docking is shown below¹¹ (PDB Code: 7F63). The Vina search space parameters were:

Center : X : 175.4453; Y : 199.3489; Z : 159.6444

Dimensions (Ångstrom) : X : 41.8118; Y : 50.7784; Z : 34.0094

After virtual screening is completed, PyRx automatically advances to Analyse Results page, where results of virtual screening computation can be viewed. Auto Dock Vina, by default, outputs the nine best binding modes for each ligand and mode with highest binding energy for each ligand was selected for further analysis. It should be emphasized that we ignore unreasonable binding models in which the ligand binds to the side or back of the top region of RBD (SI Fig. 1b), lacking the interaction with the ACE2 receptor.

Calculation of dissociation constant (K_d)

The protein–ligand–solvent mixture is a complex thermodynamic system containing complicated interactions and heat exchange. The driving forces that determine the association between RBD and ligands are a combination of various interactions and energy exchanges among the RBD, ligands and solvent. Gibbs free energy, known as a thermodynamic potential that could be used to measure the maximum amount of work in a reversible process performed by a thermodynamically closed system at constant temperature and pressure. As with all spontaneous processes, Gibbs free energy is minimized when a system reaches chemical equilibrium at constant pressure and temperature, the decrease of which equals to the work done by the system to its surroundings, minus the work of pressure forces. The accuracy of the PyRx for predicting the binding of a molecule was tested in the previous study³⁴. We estimated an error of ± 0.55 (kCal/mol) on the Gibbs free energy getting from PyRx and use a ΔG value of -8.3 ± 0.55 (kCal/mol)³⁴ as a benchmark for identifying NPs with reasonable binding affinity to RBD. ΔG is used to calculate the dissociation constant (K_d) (μM) using formula: $K_d = c^\ominus \times e \left(\frac{\Delta G}{R \times T} \right)$, where c^\ominus is the reference concentration (1 mol/l), R is the universal gas constant (1.98 cal/K mol) and T is the temperature in degrees of Kelvin (310 K). The equilibrium dissociation constant (K_d) is the basic parameter to evaluate the binding properties of the drug–receptor, which presents the extent of ligand leaving the protein. Hence, ligands having higher binding affinity to RBD should have smaller K_d .

Molecular modelling and dynamics

In this study we have set up 6 different systems corresponding to two ligands identified using MOD, i.e. *ligand viii* and *ligand ix*, complexed with glycosylated RBD of wild-type (PDB Code: 7C8D, 7F63 and 7KGJ) or mutated structures: Alpha (PDB Code: 7NEG), Beta (PDB Code: 7PS0) and Omicron (PDB code: 7WBP). The initial coordinates of the protein and ligand were taken after docking the ligand to the RBD. Initial modelling of RBDs was performed using CHARMM³¹ software and CHARMM 36 Force Field³². The protonation pattern of the protein (RBD) was determined Karlsberg2+³⁵, assuming a pH 7.0. Hence, all titratable residues were in their standard protonation state and His519 and His505 in the neutrally charged state (δ -tautomer).

The preparation of the ligand was first optimized quantum chemically using Gaussian16 as we described previously³⁶. The force field parameters of the ligand were generated with Ligand Modeler tool of CHARM-GUI Web Interface³⁷. The atomic partial charges of the ligands of interest were calculated from B3LYP/6-31G** optimized wave functions, using the electrostatic potential-fitted charge calculation protocol (ESP charges) using JAGUAR³⁸. To complete the molecular dynamics (MD) simulation setups, the structures of ligand and RBD complexes were neutralized with ions and solvated in a TIP3P water box with periodic boundary conditions (box size: 15 Å \times 15 Å \times 15 Å) as shown in SI Fig. 2. All MD simulations were performed using the same conditions with NAMDV2.14 software, with 2 fs time step using Shake algorithm to fix bond lengths of hydrogen atoms and Langevin dynamics at 300 K and flexible cell size. The resulting structures were then energy minimized and thermally equilibrated with harmonic constraints applied to all heavy atoms and sugar dihedrals angles. To ensure stability of isolated RBD constraints were applied in the production run as follows: sugar and heavy atoms retained harmonic constraints from equilibration, except for the selected residues present on the binding pockets of RBD in both WT and mutated structures within 3 Å as in our previous work³⁹. All MD simulations were repeated 3 times and were each 100 ns long. The thermal stability of the resulting structures was evaluated by computing the root-mean-square-deviations of heavy atoms (RMSD) with respect to initial structural models, the root-mean-square-fluctuations (RMSF) of heavy atoms of the ligands and their binding pocket, the radius of gyration (RG) of the protein–ligand complex and the solvent-accessible-surface-area (SASA) of the ligand and the protein. All these quantities were calculated using plugins of Visual Molecular Dynamics (VMD)⁵⁷. In addition, interaction energies between ligands and binding pocket were computed using INTER subroutine of CHARMM³¹ software using structures from time frames extracted every 100 ps from each MD trajectory. For this purpose, we have considered the electrostatic (Coulomb) and the Van der Waals components of the total interaction energies. VMD code was also employ for visualization of structures and generation of images.

Results and discussion

In silico screening predicts multiple NPs interfering with RBD-binding to the ACE2

Previous biochemical and in silico studies have revealed that some NPs, including bile acids like tauro- α -muricholic acid^{20,25}, β -amyrin⁴⁰, folic acid⁴¹, vitamin B6⁴², campesterol and stigmasterol⁴³ can bind to R1, R2, or at the interface of R1–R2 interfering with the interactions between the RBD and the ACE2 receptor. These natural products thus, are proposed to have therapeutic benefits. We aimed to expand previous studies and screen the binding of 70 NPs to the RBD using in silico molecular docking and identify those interfering with the RBD-binding to the ACE2 receptor. To perform docking studies, we first chose the structure of a WT-RBD solved in the presence of ACE2 receptor. To screen for the ligands that can bind to the RBD surface interacting with the ACE2 receptor (Figs. 1b and S1), we removed the ACE2 from the structure and defined the RBD-ACE2 interface including R1 and R2 pocket (Fig. 1b) as the docking pocket (Methods). We used PyRx open source software³³, which can predict the binding site of ligands accurately^{20,33}, and screened the binding of NPs from various sources (Supplementary Methods) (SI Table 1). The dissociation constant was calculated from the Gibbs free energy ΔG given by PyRx (Methods). Among all NPs tested, eight ligands have ΔG values close to the previously reported benchmarked value of -8.3 ± 0.55 (kCal/mol)²⁰ (Table 1). These ligands have some structural similarities: they all have aromatic and aromatic hydroxyl groups. They interact with the RBD via hydrogen bonds, dipole–dipole interactions, and hydrophobic interactions. The interaction includes amino acid residues participating in the binding of RBD to the ACE2 receptor (Fig. 1b) (SI Table 2), including Arg403, Tyr489, Phe456, Leu455, Gln493, Tyr505, Tyr449, Gln498, Arg439, and Asn487. Therefore, it is predicted that these ligands will interfere with RBD binding to the ACE2 receptor.

MOD revealed new ligands of interest

The RBD structure is widely used for molecular docking studies and discovering ligands interfering with the RBD-binding to the ACE2 receptor^{12,14–22}. The structures determined both using X-ray diffraction and Electron microscopy have been used. The structure of RBD by itself has never been solved due to its instability and the structure is always solved in the presence of one of its binding-protein partners namely ACE2 receptor, antibody, or nanobody. The binding partner is removed to prepare the structure for in silico studies. However, it has never been tested if solving the structure in the presence of different protein partners will affect the results of in silico screening. We first compared the structure of WT-RBD solved in the presence of ACE2, antibody, or nanobody (Fig. 2a, b). The results showed conformational changes and differences between these structures, specifically in the docking surface used for in silico screenings. Although the backbone structure is practically conserved, the orientation of side chains at the RBD surface are different. We hypothesized that these conformational changes could affect the outcome of in silico screening and that if the binding of a ligand was not affected by the conformational changes, that ligand would be ideal as a therapeutic lead molecule. Therefore, we performed molecular docking studies using all the 70 NPs, and the structure of WT-RBD solved in the presence of ACE2, antibody or nanobody (Methods) (SI Table 1). As an example, the results for the top eight ligands (smallest K_d) are summarised in Table 1. For most of the ligands including those tested previously such as β -amyrin⁴⁰, either the binding pocket shifted between R1 and R2 or the dissociation constant increased (binding affinity reduced). This observation is likely due to the conformational changes induced by ACE2, antibody, or nanobody binding to the RBD (Fig. 2a, b). Surprisingly, the binding energy and the binding pocket of two ligands (*ligands vi* and *viii*) were not affected by the conformational changes of RBD (Table 1). *Ligand viii* has the smallest dissociation constant and, thus, the highest binding affinity (Table 1). In all the structures, *ligand viii* continues to interact with amino acid residues in the R1 pocket of RBD (Fig. 2c–e). The amino acid residues of this pocket (SI Table 2) are

Ligand	WT-RBD with ACE2 ΔG (Kd) PDB code 7c8d			WT-RBD with nanobody ΔG (Kd) PDB code 7kgj			WT-RBD with antibody ΔG (Kd) PDB code 7f63		
	R1	R1-2	R2	R1	R1-2	R2	R1	R1-2	R2
<i>i</i>	–	–	– 8.2 (1.6)	–	–	– 7.1 (9.5)	–	–	– 8.3 (1.3)
<i>ii</i>	–	– 7.9 (2.6)	–	–	– 7.2 (8.0)	–	–	– 7.2 (8.0)	–
<i>iii</i>	–	– 7.8 (3.0)	–	–	– 7 (11.2)	–	–	– 7.2 (8.0)	–
<i>iv</i>	–	– 8.1 (1.6)	–	–	– 7.7 (3.6)	–	–	– 7.6 (4.2)	–
<i>v</i>	–	– 7.3 (6.8)	–	– 7.4 (5.8)	–	–	–	– 7.5 (4.9)	–
<i>vi</i>	– 7.1 (9.5)	–	–	– 7.4 (5.8)	–	–	– 6.9 (13.1)	–	–
<i>vii</i>	–	–	– 7.3 (6.8)	– 7.3 (6.8)	–	–	– 6.8 (15.4)	–	–
<i>viii</i>	– 8 (2.2)	–	–	– 8 (2.2)	–	–	– 8.6 (0.8)	–	–

Table 1. The binding of ligand (*viii*) is not affected by structural changes of RBD induced by the protein-binding partner. ΔG values are in kCal/mol and K_d values are in μM . (i) (1R,5R,6R,13R,21R)-16-[(1R,5R,6R,7R,13S,21R)-5,13-bis(3,4-dihydroxyphenyl)-6,9,17,19,21-pentahydroxy-4,12,14-trioxapentacyclo[11.7.1.02,11.03,8.015,20]henicosa-2(11),3(8),9,15,17,19-hexaen-7-yl]-5,13-bis(3,4-dihydroxyphenyl)-4,12,14-trioxapentacyclo[11.7.1.02,11.03,8.015,20]henicosa-2(11),3(8),9,15,17,19-hexaene-6,9,17,19,21-pentol; (ii) CinnamtanninB2; (iii) Escin Ia; (iv) Escin IIa; (v) Escin IIB; (vi) α -amyrin; (vii) β -amyrin; (viii) (1R)-1,6,7-trimethoxy-9-(4-methoxyphenyl)-8-[(4S)-1,4,9-trimethoxy-3-(4-methoxyphenyl)-5,6-dihydro-4H-phenalen-2-yl]-2,3-dihydro-1H-phenalene.

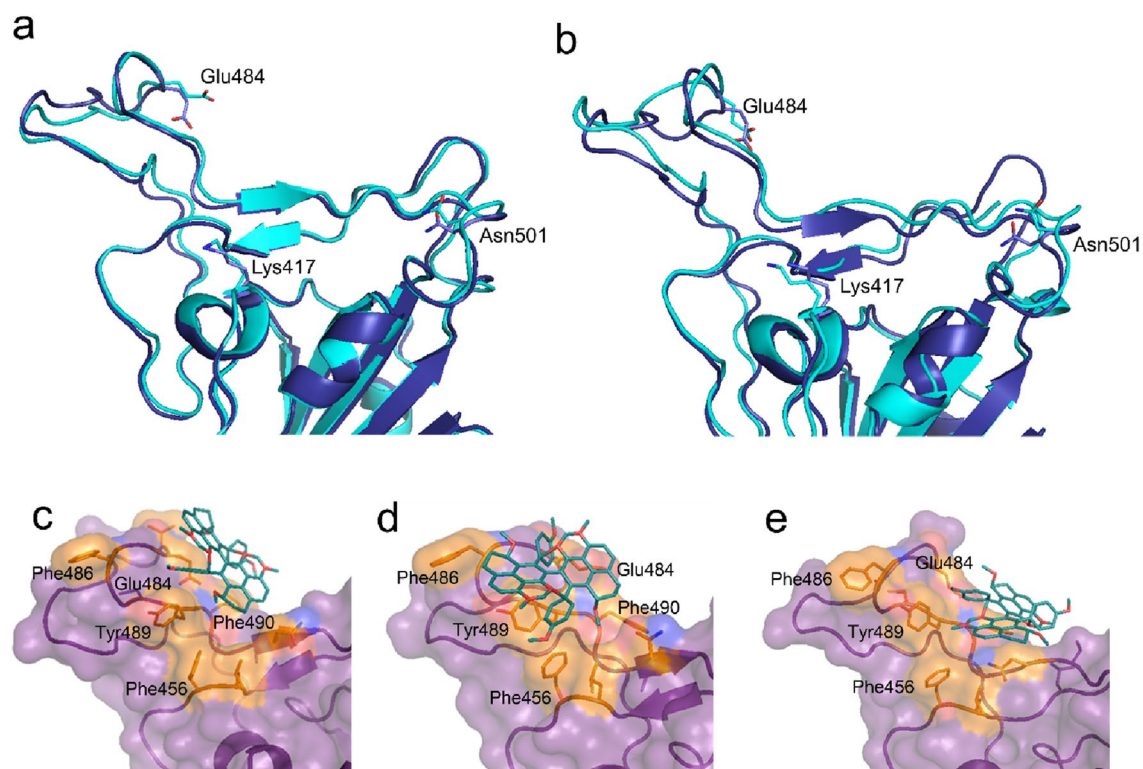


Figure 2. Predicting the effect of structural changes of RBD induced by protein partner reveals a ligand of interest. **(a)** Pairwise alignment of the structure of RBD solved in the presence of ACE2 (dark blue) (PDB Code: 7C8D) with that of RBD solved in the presence of nanobody (PDB Code: 7KGJ) (cyan) and **(b)** antibody (cyan) (PDB Code: 7F63). **(c)** The predicted binding pocket of ligand 8 using the RBD structure solved in the presence of ACE2, **(d)** nanobody, or **(e)** antibody.

among the key residues interacting with the ACE2 receptor (Fig. 1b). These results strongly suggest that *ligand viii* will interfere with the RBD-binding to the ACE2 and is a lead molecule for designing and synthesizing new inhibitors of viral entry to host cells.

Screening using multiple RBD variants revealed its vulnerable pocket

Some of the emerging variants of SARS-CoV-2 can escape antibodies generated by the immune response. This escape phenomenon is confirmed by several previous studies^{44–47}. We showed that mutations like N501Y on RBD abolish the binding of NPs of commensal microbiota like bile acids and minimize the interference of these NPs with the RBD binding to the ACE2 receptor²⁰. We proposed a new mechanism of selection and evolution of the virus²⁰. These observations highlight the growing need for discovering broad-spectrum small molecule inhibitors of RBD-binding to the ACE2 receptor. To predict which of the eight NPs of interest can bind to the R1 or R2 pocket of all major variants (alpha, beta, and omicron) and inhibit RBD-binding to ACE2, we used

Ligand	WT ΔG (K_d)			Alpha ΔG (K_d)			Beta ΔG (K_d)			Omicron ΔG (K_d)		
	R1	R1-2	R-2	R1	R1-2	R-2	R1	R1-2	R-2	R1	R1-2	R-2
<i>i</i>	–	–	–8.2 (1.6)	–	–	–7.2 (8.0)	–	–7.1 (9.5)	–	–	–6.8 (15.4)	–
<i>ii</i>	–	–7.9 (2.6)	–	–	–6.7 (18.2)	–	–	–7.1 (9.5)	–	–	–7.2 (8.1)	–
<i>iii</i>	–	–7.8 (3.0)	–	–	–	–7.1 (9.5)	–	–7 (11.2)	–	–	–	–7.1 (9.5)
<i>iv</i>	–	–8.1 (1.6)	–	–	–	–7.4 (5.8)	–	–7.2 (8.1)	–	–	–	–7.4 (5.8)
<i>v</i>	–	–7.3 (6.8)	–	–	–	–7.4 (5.8)	–	–6.9 (13.1)	–	–	–	–7.3 (6.8)
<i>vi</i>	–7.1 (9.5)	–	–	–	–	–6.8 (15.4)	–7.7 (3.6)	–	–	–	–6.9 (13.1)	–
<i>vii</i>	–	–	–7.3 (6.8)	–7 (11.2)	–	–	–	–	–7.9 (2.3)	–	–7.1 (9.5)	–
<i>viii</i>	–8 (2.2)	–	–	–8.5 (0.97)	–	–	–8.9 (0.5)	–	–	–8 (2.2)	–	–

Table 2. The ligand (*viii*) binding pocket and dissociation constant are not affected by mutations of RBD. ΔG values are in kcal/mol and K_d values are in μM . PDB codes for WT, Alpha, Beta and Omicron variants are 7C8D, 7NEG, 7PS0, and 7WBP, respectively.

molecular docking studies taking advantage of RBD structures from all major variants. The ΔG values and the binding pocket of each ligand obtained from PyRx are given in Table 2. Mutations of RBD significantly increased the dissociation constant (reduced binding affinity) and shifted the binding pocket of all ligands except that of *ligand viii*. This ligand continued to interact with amino acid residues of R1 pocket with high affinity (Fig. S3). Therefore, R1 pocket is potentially an Achilles heel of RBD and ideal for structural-guided design of small molecule inhibitors of RBD binding to the ACE2 receptor and viral entry. This predicted pocket is different than the previously reported vulnerable epitopes of RBD targeted by various neutralizing antibodies^{27–31}. This pocket consists of four short peptides (Fig. 3a, b). We used ProtScale server and Hphob.OHH scale⁴⁸ and analyzed the hydrophobicity of R1 pocket encompassing these peptides (Fig. 3c). The results revealed that S-2 and S-4 peptides are in two hydrophobic regions, and there is no or little variability among the major variants in the amino acid sequences of these hydrophobic peptides.

Predicting the molecular structure of ligands binding to the vulnerable region of RBD

Based on in silico screening, we found that the presence of aromatic/aromatic hydroxyl groups is essential for strong binding to the RBD. By comparing the chemical structure of *ligand viii* with other ligands we hypothesized that the flat surface of the *ligand viii* is vital for its binding to the vulnerable R1 site. To test this prediction, we used density functional theory to optimize the structure of *ligand viii* (Fig. S3). The energy minimized structure of *ligand viii* bends around a central C7-C9 bond (Fig. S3). Next, we performed molecular docking using the optimized structure of *ligand viii*. The results show that after optimization the dissociation constant of the optimized *ligand viii* increased between 6- and 85-fold (Table 3). Additionally, the binding pocket shifted towards the R2 pocket on the RBD surface (Fig. 4a). To confirm that the significant increase in the dissociation constant and the shift in the binding pocket result from the bent conformation of the ligand, we virtually designed and optimized a new ligand (*ligand ix*) (Figs. 4b and S3) with a flat surface as in the original structure of *ligand viii*. The molecular structure of *ligand ix* is very similar to that of the hypericin⁴⁹, a NP with antiviral activity against SARS-CoV-2. We then performed molecular docking using *ligand ix*. The results show that our designed ligand *ix* binds to the predicted vulnerable site of RBD (Fig. 4b) in all the structures with a dissociation constant at least ten-fold (Table 3) less than those obtained for the *ligand viii* in its original conformation (Table 2). In all the structures tested, we found that *ligand ix* binds to the R1 and slides over this pocket (Fig. S4). *Ligand ix* has a flat surface and multiple benzene and phenol rings. Similarly, in silico and biochemical screening have shown that several drugs with multiple benzene and phenol rings, like Hydroxycamptothecine⁵⁰, KT203⁵¹, BMS195614⁵¹,

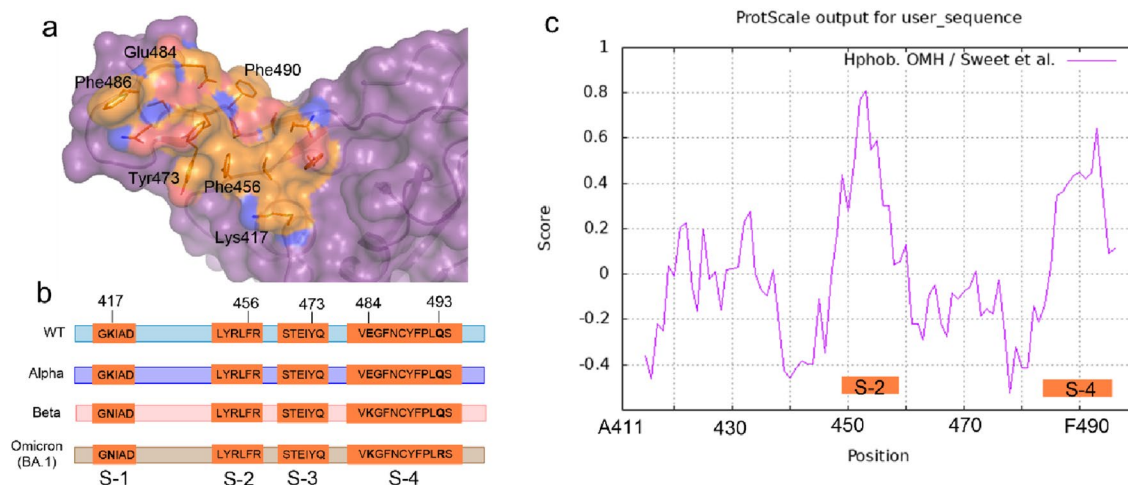


Figure 3. The vulnerable site of RBD has two conserved hydrophobic segments. (a) The structure of the vulnerable R1 region of RBD. (b) The vulnerable region consists of four short peptides (S-1 to S-4). These peptides are conserved. (c) Analysis of the peptides shows that peptide S-2 and S-4 have highly conserved hydrophobic regions.

Ligand	WT (ACE2) ΔG (K_d)			WT (Nanobody) ΔG (K_d)			WT (antibody) ΔG (K_d)			Alpha ΔG (K_d)			Beta ΔG (K_d)			Omicron ΔG (K_d)		
	R1	R1-2	R-2	R1	R1-2	R-2	R1	R1-2	R-2	R1	R1-2	R-2	R1	R1-2	R-2	R1	R1-2	R-2
Optimized <i>viii</i>	-	-	-6.9 (13.1)	-	-5.4 (151)	-	-	-5.3 (178)	-	-	-5.6 (109)	-	-	-6.1 (48)	-	-	-	-5.6 (109)
Optimized <i>ix</i>	-10.5 (0.04)	-	-	-9.8 (0.12)	-	-	-9.2 (0.31)	-	-	-9.3 (0.26)	-	-	-9.3 (0.26)	-	-	-9.6 (0.16)	-	-

Table 3. A designed ligand with tight binding to the vulnerable site of RBD. ΔG values are in kCal/mol and K_d values are in μM .

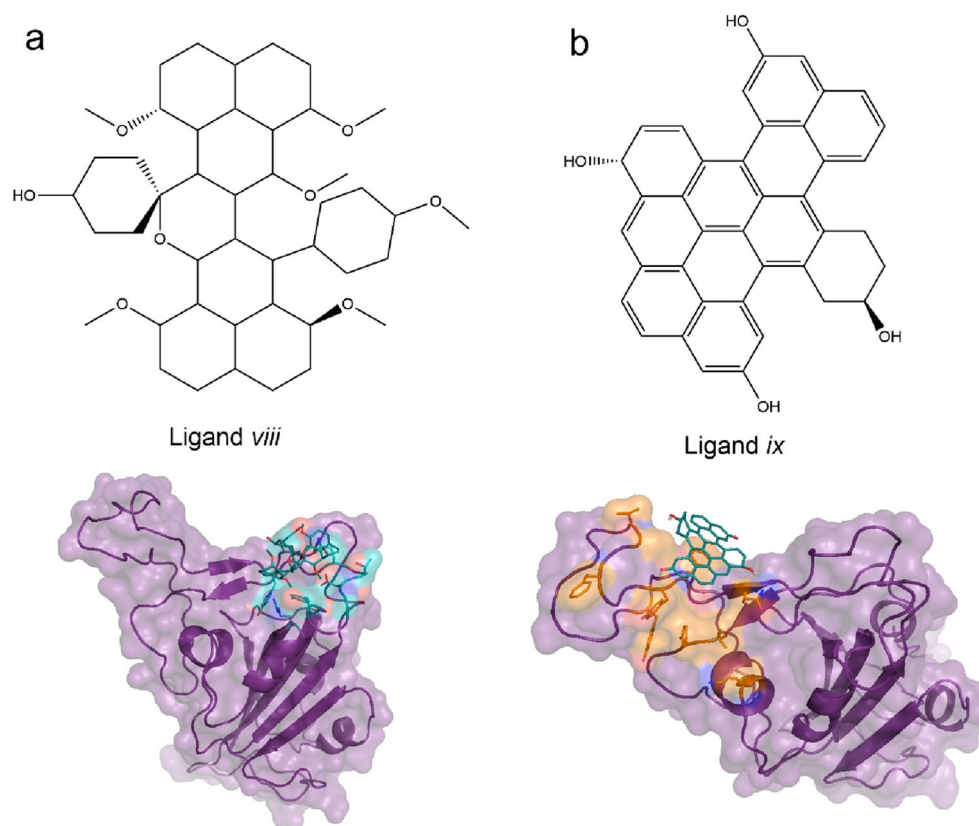


Figure 4. Structural requirement for designing ligands targeting vulnerable site of RBD. (a) The 2D chemical structure of ligand (viii) (top panel). After optimization the ligand folds and binds to R2 region with increased dissociation constant. (b) The chemical structure of the designed ligand (ix) (top panel). The optimized ligand has a flat surface, and it binds to the vulnerable R1 region of RBD. The docking results are from WT-RBD (PDB code 7C8D).

and rilapladi⁵², interact with RBD and interfere with its binding to the ACE2 receptor. The binding energies of *ligand ix* (Table 3) are significantly smaller than those of drugs, such as KT203 (– 8.73 kCal/mol) BMS195614 (– 8.25 kCal/mol), obtained using PyRx⁵¹.

Molecular dynamics revealed conformational stability of ligands

To obtain a molecular understanding of the ligand docking site and its dynamic properties, we carried out a series of classical nanosecond-scale molecular dynamics (MD) simulations. Specifically, we investigated the conformational stability and flexibility of the two docked ligands (*ligands viii* and *ix*) in the R1 and R2 binding pockets of WT RBDs and the variants, as well as the nature and strength of the intermolecular interactions. In total, six models were considered for these computations (see “Methods”). First, the RBD binding pocket accommodates both *ligands viii* and *ix*, irrespective of the initial starting structure used for the molecular simulations and the mutations of the RBD in the SARS-CoV-2 variants. The binding of these ligands is not restricted to the specific sites in R2 and R1 (Fig. 4), respectively, but to various regions of R1 and R2 (Figs. 5 and S6). While *ligand viii* binds to both R2 and R1, *ligand ix* slides over R1 pocket (Fig. S6). This observation is consistent with the results of rigid docking showing that *ligand viii* binding pocket changes while *ligand ix* binds to different sites of pocket R1 in different structures (Fig. S6). The amino acid residues of RBD interacting with *ligand viii* and *ix* are summarized in supplementary table 6. Among these residues, Arg403, Tyr449, Gln498, Asn501, Glu484, Phe490, Gln493, Tyr453 directly interact with ACE2 receptor. Therefore, consistent with the MOD studies *ligands viii* and *ix* are predicted to interfere with RBD binding to the ACE2.

The extent of conformational changes which undergo the structural models upon thermal relaxation can be estimated by monitoring the evolution of root-mean-square deviations of atomic positions (RMSD) experienced by all heavy atoms with respect to the initial reference structure during the simulation. In our analysis, the initial reference structure is represented by the corresponding crystal structure with the DFT geometry optimized docked ligand in the pose as predicted by the in silico docking experiments (vide supra). Another measure for protein compactness and stability is the radius of gyration (RG). The flexibility of the ligands and their binding environment can be accessed by evaluating the root-mean-square fluctuations (RMSF) of each heavy atom around their average position during the MD trajectory. Furthermore, the solvent-accessible protein surface (SASA) is used to determine whether the ligand is retained inside the binding pocket, or it expels out

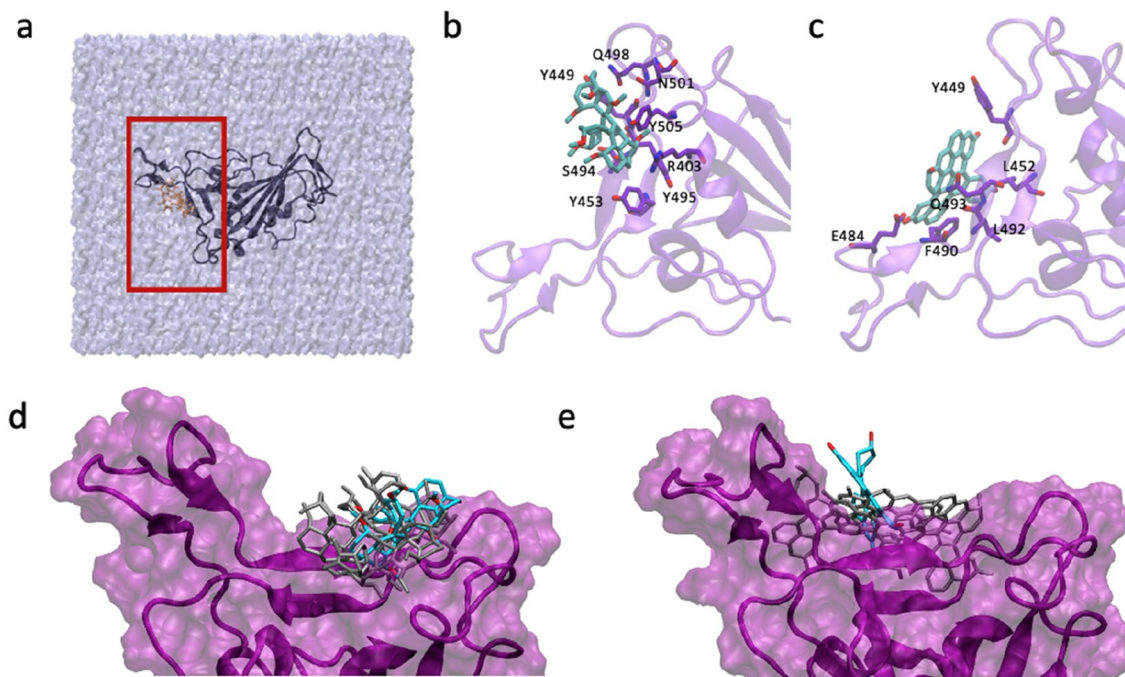


Figure 5. Results from classical molecular dynamics simulations for WT-RBD model using 7c8d.pdb structure. (a) Simulation box consisting of RBD protein, ligand (*ix* depicted in orange) and 78,310 water molecules (ice-cube representation), red box indicate the RBM of the RBD of the SARS-CoV-2 Spike protein. (b) Close up into the binding of ligand *viii* to R2 site of WT-RBD in conformation with strongest interaction (snapshot after 99.0 ns, -33.9 kcal/mol), (c) binding of ligand *ix* to R1 site of RBD in conformation with strongest interaction (snapshot after 93.3 ns, -30.9 kcal/mol). Ligands are shown as cyan sticks while relevant RBD residues interacting with the ligands are highlighted as purple sticks. Conformational space of ligand *viii* (d) and ligand *ix* (e) sampled over the MD simulation. Only representative ligand conformations characterized by highest interaction energies with WT-RBD are depicted as grey stick.

from the binding cavity. These quantities were computed over the last 10 ns of MD simulations (considering all three repetitions), when all systems reached a thermal equilibrium (Fig. S5). The average RMSD and RMSF values are listed in Table 4.

The relatively low (around 1.0 Å) RMSD and RG values predicted for the six protein–ligand structural models suggest minor conformational changes of the RBD upon ligand binding, regardless of the nature of the ligand. Interestingly a slightly higher RMSD value of 1.61 Å is predicted when *ligand viii* binds the crystal structure of WT RBD complexed to the antibody. This may be a consequence of the initial arrangement of the side chain atoms in 7f63.pdb structure, in particular the Gln493 side chain, which points away from Ser494. In addition, the RMSF values of the RBD residues at the ligand docking site is relatively low (c.a 0.5 Å) for all models suggesting that the RBD mutations in Alpha-, Beta- and Omicron- SARS-CoV-2 variants practically do not affect the conformational mobility at the receptor binding site. The most flexible residues at the RBD were identified as Gln498 and Phe497. These residues flip their side chains towards the ligands upon docking. Other flexible side chains interacting with the ligands were Arg403, TYr505 and Gln493 (Fig. 5).

The dynamic properties of the two ligands, however, are significantly different. While *ligand viii* shows high flexibility as reflected by RMSD values up to 2.5 Å, and RMSF values below 0.3 Å, *ligand ix* is very rigid (RMSD < 0.3 Å, RMSF < 0.2 Å) (Fig. S6). In particular, the methoxycyclohexane-moiety in *ligand viii* shows the highest mobility as reflected by the different conformations it adopts when docking onto the protein surface (Table S4) and by the fluctuations of the dihedral angle $\tau_{C2-C10-C22-C31}$ (Table S4 and Fig. S3) of more than 5° predicted during the MD simulations. The opposite is observed for *ligand ix*, which practically slides over the R1 surface as a rigid body without changing its conformation (Fig. S6). Thus, *ligand viii* is more prone to adapt its form to the RBD topology compared to *ligand ix*. Indeed, analysis of atomic contact for neighbouring atoms of binding pocket supports such a statement showing that *ligand viii* remains in very close contact with the S-protein during the course of all MD simulations while *ligand ix* interacts with the protein less than 95% of the simulation time. Moreover, the number of contacts established between ligand and RBD are correlated with ligand flexibility: the flexible *ligand viii* forms in average 18 contacts with side chains of RBD whereas the rigid *ligand ix* only 14 contacts with it. Specifically, *ligand viii* presents short-range contacts of hydrophobic and H-bond nature with Arg403, Tyr495, Gln498, Asn501, Tyr505, Tyr453, Ser494 and Tyr449. Ligand *ix* mainly interacts with Tyr449, Leu452, Leu492, Phe490, Gln493 and Glu484 (Fig. 5).

The strength of the interactions between *ligands viii* and *ix* with RBD and their nature has been quantified by computing the Van der Waals and electrostatic contributions of the average total interaction energy (Table 5). Among the two ligands, *ligand ix* shows the strongest interaction with R1 site of wild-type RBD as well as with

Ligand <i>viii</i>						
	WT-RBD with ACE2 PDB code 7c8d	WT-RBD with nanobody PDB code 7kgj	WT-RBD with antibody PDB code 7f63	RBD Alpha variant PDB code 7neg	RBD Beta variant PDB code 7ps0	RBD Omicron variant PDB code 7wbp
Protein						
RMSD	1.08 ± 0.06	0.94 ± 0.11	1.61 ± 0.12	1.13 ± 0.08	1.21 ± 0.06	1.01 ± 0.07
RMSF	0.47 ± 0.28	0.48 ± 0.34	0.60 ± 0.39	0.48 ± 0.32	0.46 ± 0.31	0.47 ± 0.33
RG	18.2 ± 0.02	18.2 ± 0.02	19.2 ± 0.02	17.6 ± 0.02	18.2 ± 0.02	18.2 ± 0.02
Ligand						
RMSD	0.86 ± 0.08	1.85 ± 0.06	2.09 ± 0.23	1.88 ± 0.06	2.53 ± 0.07	1.86 ± 0.06
RMSF	0.28 ± 0.20	0.27 ± 0.18	0.27 ± 0.17	0.27 ± 0.18	0.29 ± 0.18	0.27 ± 0.17
SASA	870.5 ± 12.2	879.5 ± 13.9	892.5 ± 13.0	887.0 ± 12.0	892.6 ± 11.0	884.7 ± 12.6
Protein–Ligand						
Number of contacts	28.2 ± 5.5	15.7 ± 5.3	10.8 ± 3.1	15.1 ± 5.2	17.5 ± 5.4	17.1 ± 4.6
SASA	11,024.8 ± 66.1	10,978.9 ± 90.1	12,623.8 ± 98.4	10,733.0 ± 96.9	11,433.1 ± 92.2	11,367.1 ± 83.3
Ligand <i>ix</i>						
	WT-RBD with ACE2 PDB code 7c8d	WT-RBD with nanobody PDB code 7kgj	WT-RBD with antibody PDB code 7f63	RBD Alpha variant PDB code 7neg	RBD Beta variant PDB code 7ps0	RBD Omicron variant PDB code 7wbp
Protein						
RMSD	1.33 ± 0.11	1.02 ± 0.11	1.16 ± 0.10	1.05 ± 0.08	0.81 ± 0.07	0.95 ± 0.12
RMSF	0.48 ± 0.33	0.47 ± 0.33	0.51 ± 0.34	0.45 ± 0.28	0.44 ± 0.27	0.46 ± 0.32
RG	18.1 ± 0.02	18.2 ± 0.02	19.2 ± 0.02	17.6 ± 0.02	18.2 ± 0.02	18.2 ± 0.02
Ligand						
RMSD	0.29 ± 0.06	0.28 ± 0.05	0.25 ± 0.06	0.27 ± 0.07	0.31 ± 0.06	0.26 ± 0.06
RMSF	0.17 ± 0.08	0.17 ± 0.08	0.17 ± 0.08	0.18 ± 0.09	0.17 ± 0.08	0.18 ± 0.09
SASA	682.5 ± 10.6	704.6 ± 7.9	722.5 ± 7.9	718.4 ± 10.8	724.3 ± 8.4	715.5 ± 9.9
Protein–Ligand						
Number of contacts	15.1 ± 3.8	13.7 ± 3.6	19.8 ± 4.0	13.5 ± 3.6	10.0 ± 3.6	12.5 ± 3.9
SASA	10,992.1 ± 72.9	10,680.5 ± 63.3	12,125.6 ± 85.1	10,467.5 ± 72.0	11,258.3 ± 73.4	11,304.9 ± 73.6

Table 4. Structural properties of the proteins, the ligands *viii* and *ix*, and the protein–ligand complexes: average root mean square deviations (RMSD) of heavy atoms of ligands *viii* and *ix* and their corresponding docking sites with respect to the initial structural models of WT-RBD and the variant, average root mean square fluctuations (RMSF) of heavy atoms of the ligands and their docking sites, radius of gyration (RG) of the proteins, solvent accessible surface area (SASA) of the ligands and the protein–ligand complex and average number of contacts between them. All these quantities were evaluated over the last 10 ns of the corresponding MD trajectories (including repetitions). The number of contacts was computed using Numpy 1.20.3, and MDAnalysis 2.0.0¹³ with Python 3.8 compatible packages. A contact is defined as an event where the distance between ligand and protein is below 3 Å. SASA and RG were computed using VMD⁵⁷ tools. The units for RMSD, RMSF and RG values is Å, while SASA is given in Å².

Ligand	WT (ACE2)			WT (Nanobody)			WT (antibody)			Alpha			Beta			Omicron		
	MD1	MD2	MD3	MD1	MD2	MD3	MD1	MD2	MD3	MD1	MD2	MD3	MD1	MD2	MD3	MD1	MD2	MD3
Optimized <i>viii</i>	− 30.0 ±1.9	− 29.6 ±2.0	− 29.8 ±2.3	− 17.1 ±4.6	− 16.1 ±2.5	− 12.1 ±4.1	− 12.7 ±3.6	− 17.7 ±2.3	− 15.8 ±3.6	− 12.9 ±4.2	− 17.3 ±3.8	− 17.2 ±3.3	− 22.2 ±2.0	− 22.1 ±2.1	− 20.7 ±3.4	− 19.7 ±1.7	− 14.2 ±3.2	− 13.5 ±2.8
Optimized <i>ix</i>	− 24.5 ±3.1	− 21.3 ±2.6	− 25.8 ±3.1	− 23.8 ±2.3	− 23.7 ±2.0	− 25.0 ±3.7	− 32.7 ±2.5	− 32.0 ±2.3	− 30.0 ±2.3	− 22.2 ±2.2	− 14.9 ±2.4	− 22.7 ±3.9	− 24.7 ±2.7	− 26.6 ±2.5	− 25.1 ±2.8	− 26.3 ±2.8	− 24.0 ±2.3	− 27.3 ±2.2

Table 5. Average total interaction energy (kcal/mol) between optimized ligands *viii* and *ix* with binding sites of wild type RBD and the variant of SARS-CoV-2 virus. All energy values are evaluated over the last 10 ns of the MD trajectories (3 repetitions).

alpha, beta and omicron variants of the SARS-CoV-2 as reflected by the predicted low average interaction energies of − 26.5 kcal/mol, − 19.9 kcal/mol, − 25.5 kcal/mol and − 25.9 kcal/mol, respectively. These values are followed by ligand *viii*, which also shows moderate interaction energy with the R2 site. Furthermore, the computations clearly indicate that Van der Waals forces are the primary contributor to the interaction between the ligands and the RBD. Electrostatics contribute to less than 1% of ligand *viii* binding to the RBD and only 15% to the binding of ligand *ix* to the RBD. Glu484 is responsible for the attractive electrostatic interactions in R1.

From the standpoint of SASA, *ligand ix* shows a SASA value of about 700 Å², which is almost 200 Å² lower than the value predicted for *ligand viii*, indicating that *ligand ix* is less exposed to water and a larger part of its surface is protected by the receptor protein, compared to *ligand viii*. Thus, the highly mobile *ligand viii* is stabilized via short, weak interactions with several side chains in the entire RBD. On the other hand, the rigid *ligand ix* establishes strong hydrophobic interactions with specific residues of the vulnerable R1 pocket. Furthermore, the sequence mutations detected in alpha, beta and omicron variants of the SARS-CoV-2 practically do not affect the binding properties of *ligands viii* and *ix* to RBD. These results support the predictions from the in silico rigid docking experiments described above.

Conclusions

In silico screening of small molecule ligands is a powerful tool to identify the suitable lead molecules helping the process of drug discovery^{10,11,53}. These studies take advantage of the structure(s) of a protein. In the case of SARS-CoV-2 RBD, the structure is always solved in the presence of a binding partner like ACE2, antibody, or nanobody. Thus, the binding partner is removed from the structure before in silico screening. However, how the interaction of the binding partner with RBD affects the outcome of molecular docking has not been investigated^{12,14–22}. To address this question, we used the structure of WT-RBD solved in the presence of ACE2, antibody, or nanobody (Table 1). We used MOD and screened the binding of more than 70 NPs. Our findings show that the structural changes of WT-RBD induced by the binding of ACE2, antibody, or nanobody significantly affect the binding of most ligands. These studies enabled us to identify a vulnerable binding pocket (R1) of RBD, where binding of a ligand was not affected by the structural changes of RBD and mutations observed in the emerging variants. A combination of MOD and MD simulations allowed us to define the structural requirement of ligands targeting the R1 pocket. We found that hydroxyl groups, the flat surface and the rigid structure of the ligand are essential for binding to R1 pocket. To confirm this, we designed *ligand ix*. Molecular docking and dynamics studies confirmed that *ligand ix* binds to the R1 pocket irrespective of the structural changes induced by ACE2, antibody, or nanobody and the mutations present in alpha, beta, and omicron variants. The *ligand ix* structure is very similar to the antiviral and anticancer NP hypericin. Recent biochemical studies revealed that hypericin reduces SARS-CoV-2 infectivity via different mechanisms⁵⁴ and inhibits the S-glycoprotein binding to the ACE2 receptor²⁶.

In silico studies have limitations. Virtual screening methods can generate false positives/negatives. These false outcomes can arise from inaccuracies of the structural models used, inclusion/exclusion of water molecules in the models, and approximation of the binding sites⁵⁵. Our data suggest that MOD reduces the generation of false positive/negatives outcomes. Although we used the structure of multiple variants of SARS-CoV-2, new variants of the virus continue to emerge. Given the adaptability of the virus and its continuous evolution and possible mutational landscapes, it is not possible to screen the binding of ligands against all possible variants. Therefore, further experimental studies and screening using multiple variants of the virus are required to corroborate the findings of our in silico screening. A limitation of our study is the sample size of the structures used. We selected one structure either determined using X-ray diffraction or electron microscopy (SI Methods). Screening using more structures, at least 3 structures solved using X-ray and 3 structure solved using electron microscopy for each RBD will extend the confidence of MOD studies and potentially identifying broad-spectrum lead molecules for drug discovery and development. Finally, in silico studies cannot predict off-target and toxicity of newly identified compounds. It is possible that like hypericin, which has some cytotoxicity and can inhibit protein kinase C (PKC)⁵⁶, our predicted *ligand ix* will have some cellular toxicity and off-target activity. These activities and IC50 value of the ligand must be determined experimentally to validate its potential for further drug development.

In summary, we demonstrate that the application of MOD improves identification of ligands binding to various RBD variants and structures irrespective of mutations in the RBD and structural variants. The MOD findings combined with MD calculations reveal new insight into the structural requirement of designing SARS-CoV-2 viral entry inhibitors. Further experiments using a larger number of structures of RBD should be conducted to corroborate our findings. We suggest that MOD should be used to help structural-guided design and synthesis of new lead molecules for drug discovery and development.

Data availability

All data generated or analysed during this study are included in this published article and its supplementary information files.

Received: 3 May 2023; Accepted: 4 September 2023

Published online: 29 September 2023

References

1. Ciotti, M. *et al.* The COVID-19 pandemic. *Crit. Rev. Clin. Lab. Sci.* **57**, 365–388 (2020).
2. Hoffman, M. *et al.* SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280 (2020).
3. Yan, R. *et al.* Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **367**, 1444–1448 (2020).
4. Jackson, C. B., Farzan, M., Chen, B. & Choe, H. Mechanisms of SARS-CoV-2 entry into cells. *Nat. Rev. Mol. Cell Biol.* **23**, 3–20 (2022).
5. Li, W. *et al.* Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450–454 (2003).
6. Ferraz, M., Moreira, E., Coêlho, D. F., Wallau, G. & Lins, R. Immune evasion of SARS-CoV-2 variants of concern is driven by low affinity to neutralizing antibodies. *Chem. Comm.* **57**, 6094–6097 (2021).
7. Gupta, R. K. Will SARS-CoV-2 variants of concern affect the promise of vaccines?. *Nat. Rev. Immunol.* **21**, 340–341 (2021).
8. Hu, J. *et al.* Increased immune escape of the new SARS-CoV-2 variant of concern Omicron. *Cell. Mol. Immunol.* **19**, 293–295 (2022).

9. Maher, M. C. *et al.* Predicting the mutational drivers of future SARS-CoV-2 variants of concern. *Sci. Transl. Med.* **14**, eabk3445 (2022).
10. Terstappen, G. C. & Reggiani, A. In silico research in drug discovery. *Trends Pharmacol. Sci.* **22**, 23–26 (2001).
11. Seifert, M. H. J., Wolf, K. & Vitt, D. Virtual high-throughput in silico screening. *Biosilico* **1**, 143–149 (2003).
12. Nag, A., Paul, S., Banerjee, R. & Kundu, R. In silico study of some selective phytochemicals against a hypothetical SARS-CoV-2 spike RBD using molecular docking tools. *Comput. Biol. Med.* **137**, 104818 (2021).
13. Muratov, E. N. *et al.* A critical overview of computational approaches employed for COVID-19 drug discovery. *Chem. Soc. Rev.* **50**, 9121–9151 (2021).
14. Singh, R., Bhardwaj, V. K., Sharma, J., Kumar, D. & Purohit, R. Identification of potential plant bioactive as SARS-CoV-2 Spike protein and human ACE2 fusion inhibitors. *Comput. Biol. Med.* **136**, 104631 (2021).
15. Rout, J., Swain, B. C. & Tripathy, U. In silico investigation of spice molecules as potent inhibitor of SARS-CoV-2. *J. Biomol. Struct. Dyn.* **40**, 860–874 (2022).
16. Behloul, N. *et al.* In silico identification of strong binders of the SARS-CoV-2 receptor-binding domain. *Eur. J. Pharmacol.* **890**, 173701 (2021).
17. Muhseen, Z. T., Hameed, A. R., Al-Hasani, H. M. H., Tahir-ul-Qamar, M. & Li, G. Promising terpenes as SARS-CoV-2 spike receptor-binding domain (RBD) attachment inhibitors to the human ACE2 receptor: Integrated computational approach. *J. Mol. Liq.* **320**, 114493 (2020).
18. Deganutti, G., Prischi, F. & Reynolds, C. A. Supervised molecular dynamics for exploring the druggability of the SARS-CoV-2 spike protein. *J. Comput. Aided. Mol. Des.* **35**, 195–207 (2021).
19. Gangadevi, S. *et al.* Kobophenol A inhibits binding of host ACE2 receptor with spike RBD domain of SARS-CoV-2, a lead compound for blocking COVID-19. *J. Phys. Chem. Lett.* **12**, 1793–1802 (2021).
20. Dragelj, J., Mroginski, M. A. & Ebrahimi, K. H. Hidden in plain sight: Natural products of commensal microbiota as an environmental selection pressure for the rise of new variants of SARS-CoV-2. *ChemBioChem* **21**, 2946–2950 (2021).
21. Honarmand Ebrahimi, K. SARS-CoV-2 spike glycoprotein-binding proteins expressed by upper respiratory tract bacteria may prevent severe viral infection. *FEBS Lett.* **594**, 1651–1660 (2020).
22. Yepes-Pérez, A. F., Herrera-Calderon, O. & Quintero-Saumeth, J. Uncaria tomentosa (cat's claw): A promising herbal medicine against SARS-CoV-2/ACE-2 junction and SARS-CoV-2 spike protein based on molecular modeling. *J. Biomol. Struct. Dyn.* **40**, 2227–2243 (2022).
23. Yang, J. *et al.* Computational design and modeling of nanobodies toward SARS-CoV-2 receptor binding domain. *Chem. Biol. Drug Des.* **98**, 1–18 (2021).
24. Yang, J. *et al.* Structure-based discovery of novel nonpeptide inhibitors targeting SARS-CoV-2 Mpro. *J. Chem. Inf. Model.* **61**, 3917–3926 (2021).
25. Carino, A. *et al.* Hijacking SARS-CoV-2/ACE2 receptor interaction by natural and semi-synthetic steroidal agents acting on functional pockets on the receptor binding domain. *Front. Chem.* **8**, 572885 (2020).
26. Mohamed, F. F. *et al.* Hypericum perforatum and its ingredients hypericin and pseudohypericin demonstrate an antiviral activity against SARS-CoV-2. *Pharmaceuticals* **15**, 530 (2022).
27. Li, T. *et al.* Uncovering a conserved vulnerability site in SARS-CoV-2 by a human antibody. *EMBO Mol. Med.* **13**, e14544 (2021).
28. McCallum, M. *et al.* N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell* **184**, 2332–2347 (2021).
29. Joyce, M. G. *et al.* A cryptic site of vulnerability on the receptor binding domain of the SARS-CoV-2 spike glycoprotein. *BioRxiv* **2020**, 15 (2020).
30. Makdasi, E. *et al.* Neutralizing monoclonal anti-SARS-CoV-2 antibodies isolated from immunized rabbits define novel vulnerable spike-protein epitope. *Viruses* **13**, 566 (2021).
31. Burnett, D. L. *et al.* Immunizations with diverse sarbecovirus receptor-binding domains elicit SARS-CoV-2 neutralizing antibodies against a conserved site of vulnerability. *Immunity* **54**, 2908–2921 (2021).
32. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
33. Dallakyan, S. & Olson, A. J. Small-molecule library screening by docking with PyRx. In *Chemical Biology* 243–250 (Springer, 2015).
34. Dragelj, J., Mroginski, M. A. & Ebrahimi, K. H. Hidden in plain sight: Natural products of commensal microbiota as an environmental selection pressure for the rise of new variants of SARS-CoV-2. *ChemBioChem* **21**, 2946–2950 (2021).
35. Kieseritzky, G. & Knapp, E.-W. Optimizing pKa computation in proteins with pH adapted conformations. *Proteins Struct. Funct. Bioinforma.* **71**, 1335–1348 (2008).
36. Wolf, A. *et al.* The redox-coupled proton-channel opening in cytochrome c oxidase. *Chem. Sci.* **11**, 3804–3811 (2020).
37. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).
38. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).
39. Dragelj, J., Mroginski, M. A. & Ebrahimi, K. H. Hidden in plain sight: Natural products of commensal microbiota as an environmental selection pressure for the rise of new variants of SARS-CoV-2. *ChemBioChem* **22**, 2946–2950 (2021).
40. Wahyuni, D. K. *et al.* Molecular simulation of compounds from n-hexane fraction of *Sonchus arvensis* L. leaves as SARS-CoV-2 antiviral through inhibitor activity targeting strategic viral protein. *J. Pharm. Pharmacogn. Res.* **10**, 1126–1138 (2022).
41. Eskandari, V. Repurposing the natural compounds as potential therapeutic agents for COVID-19 based on the molecular docking study of the main protease and the receptor-binding domain of spike protein. *J. Mol. Model.* **28**, 153 (2022).
42. Pandya, M. *et al.* Unravelling Vitamin B12 as a potential inhibitor against SARS-CoV-2: A computational approach. *Inform. Med. Unlocked* **30**, 100951 (2022).
43. Dinata, R. *et al.* Repurposing immune boosting and anti-viral efficacy of *Parkia* bioactive entities as multi-target directed therapeutic approach for SARS-CoV-2: exploration of lead drugs by drug likeness, molecular docking and molecular dynamics simulation methods. *J. Biomol. Struct. Dyn.* **2023**, 1–39 (2023).
44. Harvey, W. T. *et al.* SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* **19**, 409–424 (2021).
45. Cao, Y. *et al.* Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* **602**, 657–663 (2022).
46. Hoffmann, M. *et al.* SARS-CoV-2 variants B.1.351 and P.1 escape from neutralizing antibodies. *Cell* **184**, 2384–2393 (2021).
47. Zhou, D. *et al.* Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine induced sera. *Cell* <https://doi.org/10.1016/j.cell.2021.02.037> (2021).
48. Sweet, R. M. & Eisenberg, D. Correlation of sequence hydrophobicities measures similarity in three-dimensional protein structure. *J. Mol. Biol.* **171**, 479–488 (1983).
49. Liu, Q. *et al.* Hypericin: Single molecule spectroscopy of an active natural drug. *J. Phys. Chem. A* **124**, 2497–2504 (2020).
50. Tsegay, K. B. *et al.* A repurposed drug screen identifies compounds that inhibit the binding of the COVID-19 spike protein to ACE2. *Front. Pharmacol.* **12**, 685308 (2021).
51. Choudhary, S., Malik, Y. S. & Tomar, S. Identification of SARS-CoV-2 cell entry inhibitors by drug repurposing using in silico structure-based virtual screening approach. *Front. Immunol.* **11**, 1664 (2020).

52. Alexpandi, R., De Mesquita, J. F., Pandian, S. K. & Ravi, A. V. Quinolines-based SARS-CoV-2 3CLpro and RdRp inhibitors and Spike-RBD-ACE2 inhibitor for drug-repurposing against COVID-19: an in silico analysis. *Front. Microbiol.* **11**, 1796 (2020).
53. Davies, J. W., Glick, M. & Jenkins, J. L. Streamlining lead discovery by aligning in silico and high-throughput screening. *Curr. Opin. Chem. Biol.* **10**, 343–351 (2006).
54. Delcanale, P. *et al.* The interaction of hypericin with SARS-CoV-2 reveals a multimodal antiviral activity. *ACS Appl. Mater. Interfaces* **14**, 14025–14032 (2022).
55. Phatak, S. S., Stephan, C. C. & Cavasotto, C. N. High-throughput and in silico screenings in drug discovery. *Expert Opin. Drug Discov.* **4**, 947–959 (2009).
56. Kocanova, S. *et al.* Characterization of the interaction of hypericin with protein kinase C in U-87 MG human glioma cells. *Photochem. Photobiol.* **82**, 720–728 (2006).
57. Humphrey, W., Dalke, A. & Schulten, K. V. M. D. Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

Acknowledgements

KHE thanks the Royal Society for an International Exchanges Grant (IES\R1\221117) and for a Research Grant (RGS\R1\231135). MAM, JD and MWL thank the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC 2008–390540038 (UniSysCat). This article/publication is based upon work from COST Action FeSImmChemNeT (CA21115) supported by COST (European Cooperation in Science and Technology). For the purposes of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Accepted Author Manuscript version arising from this submission.

Author contributions

K.H.E conceived the project; A.D. performed screening using molecular docking, M.W.L conducted molecular modelling and dynamics; A.D., M.W.L, J.D., M.A.M, and K.H.E analysed data and prepared figures; M.A.M and K.H.E supervised the project and wrote the manuscript with contributions from A.D., M.W.L, and J.D; All authors reviewed the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-42015-2>.

Correspondence and requests for materials should be addressed to M.A.M. or K.H.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023