

# Sequence variation of SARS-CoV-2 spike protein may facilitate stronger interaction with ACE2 promoting high infectivity

**CURRENT STATUS:** POSTED



Masaud Shah

Department of Physiology, Ajou University School of Medicine, Suwon, Republic of Korea

Bilal Ahmad

Department of Molecular Science and Technology, Ajou University, Suwon, Republic of Korea

Sangdun Choi

Department of Molecular Science and Technology, Ajou University, Suwon, Republic of Korea

Hyun Goo Woo

Department of Physiology, Ajou University School of Medicine, Suwon, Republic of Korea

✉ hg@ajou.ac.kr *Corresponding Author*

## DOI:

10.21203/rs.3.rs-16932/v1

## SUBJECT AREAS

*Structural Biology*    *Immunology*

## KEYWORDS

*COVID-19, mAbs, Mutation, Neutralization, SARS-CoV-2, Spike, South Korea*

## Abstract

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which causes coronavirus disease (COVID-19), is a novel beta coronavirus emerged in China in 2019. Coronavirus uses spike glycoprotein to interact with host angiotensin-converting enzyme 2 (ACE2) and ensure cell recognition. High infectivity of SARS-CoV-2 raises questions on spike-ACE2 binding affinity and its neutralization by anti-SARS-CoV monoclonal antibodies (mAbs). Here, we observed Val-to-Lys417 mutation in the receptor-binding domains (RBD) of SARS-CoV-2, which established a Lys-Asp electrostatic interaction enhancing its ACE2-binding. Pro-to-Ala475 substitution and Gly482 insertion in the **AGSTPCNGV**-loop of RBD hindered neutralization of SARS-CoV-2 by anti-SARS-CoV mAbs. In addition, we identified unique and structurally conserved conformational-epitopes on RBDs, which can be potential therapeutic targets. Collectively, we provide new insights into the mechanisms underlying the high infectivity of SARS-CoV-2 and development of new effective neutralizing agents.

## Introduction

SARS-CoV-2 causes coronavirus disease (COVID-19), which was initially emerged in the Wuhan city of China that has transmitted rapidly to almost every corner of the world within two months (1, 2). Phylogenetic analysis has confirmed that SARS-CoV-2 is a novel beta coronavirus (3, 4) and apparently transmit from human-to-human way faster than previously known SARS-CoV and MERS-CoV (5, 6). Although epidemiological features of the SARS-CoV-2 are largely unknown, asymptomatic transmission and poor self-quarantine measurements of the irresponsible infected personals are thought as the most crucial reasons for the uncontrollable transmission. In current pandemic situation of SARS-CoV-2, there is an urgent need to develop effective therapeutics and vaccines. Some pre-existing anti-viral drugs are now under clinical trials (2).

Among the mode of action mechanisms of the viral infection, cell recognition and entry of virus is the most crucial step which determines viral infectivity and pathogenesis (7). Coronaviruses uses spike glycoprotein (S) to interact with the human respiratory and epithelial cells expressing angiotensin-converting enzyme-2 (ACE2) receptors (8, 9). The ectodomain of S protein is a ~1200 amino acid long trimeric class 1 fusion protein and normally exists as a metastable pre-fusion conformation “laying or

down”, which undergoes conformational rearrangement and acquires an ACE2-feasible conformation *i.e.* “up or standing” (9, 10). The “laying or down” and “up or standing” poses are differentiated due to the conformational rearrangement of the receptor binding domain (RBD, ~200 amino acid) in the S1 subunit of S. The RBD contains receptor-binding determining region (RBDR) that recognizes ACE2. The availability of RBDR is controlled by the hinge-like conformational motion of the RBD (9). Thus, S protein is indispensable for the virus survival and remained a priority target for antibodies to curb the viral entry. Although a recent report has shown crudely the arrangement of the SARS-CoV-2 S protein domains through Cryo-EM, the structure was not complete lacking many crucial loop regions that are responsible for receptor and antibodies binding (PDB ID: 6VSB) (9). In this study, we constructed a full-length model of the SARS-CoV-2 S protein in its pre-fusion monomeric and trimeric conformation, delineated its interaction with ACE2 and suggested its possible neutralization through epitopes investigation. By performing structural analyses of ACE2 to SARS-CoV RBD (sRBD) or SARS-CoV-2 RBD (cRBD) using molecular docking, molecular dynamics (MD) simulation, and molecular mechanics poisson-boltzmann surface area (MMPBSA) approaches, we investigated their relative dynamic interaction, stability, and binding affinities. In addition, we also identified conformational epitopes on the cRBD, which might be a novel neutralizing target for anti-sRBD mAbs.

## Methods

### SARS-CoV-2 spike protein and antibody modelling.

The full-length monomeric and trimeric Spike (S) models were modelled using multiple SARS-CoV structures as template (PDB ID: 5X5B, 6ACG, 5I08). The monomeric spike proteins were assembled into two conformational states based on the position of RBD (standing or laying). The amino acid sequence used in SARS-CoV-2 S modelling was retrieved from NCBI (accession # NC\_045512).

Modelling procedures for protein-protein docking and interface analyses were performed as previously described (11–13). Protein surface and patch analyses were performed in MOE suit (2019.0102) as described previously (14).

For the CR3014 and CR3022 mAbs modeling, a built-in MOE suit was used and the single chain variable fragments (scFv) were constructed (15). The complementarity-determining regions (CDR)

were annotated and numbered according to Chothia and Lesk numbering scheme (16, 17). Structure data of other mAbs including 80R, m396, F26G19 and s230 were obtained from PDB database (18–21). For mAbs docking, a built-in protein-protein docking procedure was used in MOE suit. In docking simulation, CDR regions of mAbs were considered as ligand-sites instead of entire scFv regions. Conformational epitopes of cRBD were predicted by using Epipred implemented in SAbPred web server (22, 23). Briefly, this tool utilizes CDR information of an input mAb and predict conformational epitopes on a target protein. By calculating geometric fitting and knowledge-based asymmetrical antibody-antigen scoring, the epitopes of the cRBD were predicted and ranked on the basis of combined conformational matching of the antibody-antigen structures. The score of the epitope is given by:

$$\text{Epitope Score} = \sum_{n \in G} d(n) \Pr(T_{ab}, T_{ag})$$

where  $T_{ab}$  and  $T_{ag}$  are the amino acid types of the antibody and antigen residues, respectively, which belong to node  $n$ .

Molecular dynamics (MD) simulations and binding free energy analysis.

MD simulations were performed using GROMACS 2019.3. The RBD-ACE2 complexes were solvated with TIP3P water cubic box of dimension boundaries extended to 10 Å from protein atoms. To neutralize the charge of the simulation system the  $\text{Na}^+/\text{Cl}^-$  counter ions were added and energy minimization was performed using CHARMM37 force field (24) and steep descent algorithm. After temperature and pressure equilibration, MD simulation were carried out for 30ns for each system. Detailed procedure has been described in our previous studies (12, 14).

Free energy of RBD-ACE2 complexes was calculated using molecular mechanics of Poisson-Boltzmann surface area (MM-PBSA) approach (25). The enthalpy of the system was calculated using molecular mechanics of MM-PBSA. The effect of polar and non-polar part of the solvent effect on free energy was determined using Poisson-Boltzmann equation and calculating the surface area. The basic equation is given by:

$$\Delta G_{\text{bind}} = \Delta E_{\text{MM}} + \Delta \Delta G_{\text{sol}} - T \Delta S$$



In GROMACS the built-in tool g\_mmpbsa and APBSA were called for the MMPBSA calculations, and the last 10ns of MD simulations trajectory of each complex with 1000 frames in each trajectory were taken for energy calculations. For g\_mmpbsa analysis, the dielectric constant of the aqueous solvent was set to 80, and the interior dielectric constant was set to 4; the surface tension constant  $\gamma$  was set to 0.022 kJ/mol.

## Results And Discussion

### Structural modelling of the SARS-CoV-2 spike and ACE2 interaction.

A full-length S protein is composed of S1 and S2 subunits, which are further divided into sub-domains with distinct functions. Based on the hinge-like motion of the RBD of S1 subunit, the trimeric S protein exists as transiently symmetric (RBD down) or asymmetric conformation (RBD standing) (*Figure 1A*)..

Recent studies with Cryo-EM analysis revealed consistent results that the cRBD, like other coronaviruses, exhibited stochastic breathing-like movement, facilitating receptor binding to the exposed RBD and subsequent shedding of the S1 subunit (9, 10). However, this trimeric structure may be not helpful in understanding the receptor binding mechanisms, because structural information on the residues in RBDR and mAbs-binding loops are missing in the 3D structure from protein data bank (PDB) (*Supplementary Figure 1*, PDB ID: 6VSB). Thus, by using SARS-CoV S proteins structures, we constructed monomeric and trimeric models of the SARS-CoV-2 S protein with different conformational states. The trimeric models are provided in the supplementary data. To understand the structure more clearly and validate our structure modelling, the Cryo-EM structure and our models were superimposed in standing (in the presence of ACE2) and laying conformation, respectively. Total root mean square deviation (RMSD) of the backbone atoms of S protein in standing pose was 2.93Å while the RMSD of RBD region was 1.45Å (*Supplementary Figure 1*).. Whereas these RMSD values in laying pose were 2.56Å and 1.38, respectively. This validation supports the reliability of our model and further suggests that the structure modelled by Cryo-EM may be not able to capture the loop structure from the stochastic motion of RBD in standing pose, which is critical for binding of receptors and antibodies. During preparation of our manuscript, Yan R *et al.* has demonstrated the interaction between SARS-CoV-2 RBD and ACE2 interaction by performing Cryo-EM analysis (26). Even though,

the findings of this study were consistent with our results and supported the reliability of our computational model, the structure related information were not yet deposited in PDB. Therefore, we suggest that our computational model is better than recently reported pre-fusion trimeric Cryo-EM structure of S protein.

By performing protein patch analysis, we demonstrated that the standing cRBD exposes Lys417 that establishes transient but strong electrostatic interaction with Asp30 of the ACE2, although this patch remains buried in the laying position of the cRBD (*Figure 1B*, patch analysis). This finding indicates that Lys417 mutation in cRBD plays a crucial role in ACE2 recognition that is otherwise substituted by hydrophobic valine in sRBD (*Figure 1C*). In addition, sequence and structure analyses revealed that RBDR of SARS-CoV-2 is substantially variable compared to that of SARS-CoV which harboured some conserved motifs. The average RMSD for the whole cRBD and sRBD was  $\sim 1.1\text{\AA}$ , whereas the average RMSD for RBDR was deviated by  $\sim 2\text{--}3\text{\AA}$  owing to the glycine insertion and other mutations. Both cRBD and sRBD established same contacts with ACE2 residues (*Table 1*), although their RBDR sequences were highly variable. Initial docking analysis revealed that the electrostatic contact between Arg426 and Glu329 in sRBD-ACE2 is analogous to that of Lys417 and Asp30 contact in cRBD-ACE2 (see *Figure 1D*). There 3D structure of cRBD-ACE2 complex has been provided in the supplementary data. However, this interaction was transient and break after the Asp30 of ACE2 established an intrachain contact with nearby His34.

**Mutation of Lys417 in cRBD may facilitate stronger interaction with ACE2.** Differing from SARS-CoV and SARS-related CoVs, SARS-CoV-2 had furin cleavage site at the S1/S2 boundary and similar binding affinity to cRBD of ACE2, which might be responsible for the efficient spread of SARS-CoV-2 (27). In addition to these, we next sought sequence mutations in cRBD which play critical roles in ACE2 interaction. As the static conformation of a protein complex provides limited information about the changes of the binding interface in physiological condition, we simulated the complex structure of cRBD-ACE2 and compared with sRBD-ACE2 complex. The distances between interface residues were monitored as function of time to trace the shifting, breaking, or formation of new bonds. Previously, surface plasmon resonance (SPR) and bio-layer interferometry (BLI) analyses

have shown that cRBD-ACE2 interaction is stronger than sRBD-ACE2 interaction (9, 27, 28). Supporting this, we also observed that the total number of hydrogen bonds remained similar throughout the simulation time in both sRBD-ACE2 and cRBD-ACE2 complexes (*Figure 2A*).. This result may imply that the stronger binding affinity of cRBD towards ACE2 might be attributed to stronger Lys417-Asp30 interaction compared to Arg426-Glu329 interaction in sRBD-ACE2. Interestingly, when we monitored the minimum interaction distances with respect to the simulation time, we observed that Lys417-Asp30 pair was more compact as compared to Arg426-Glu329 pair. Initially the residues in both pairs were  $\sim 1.4\text{\AA}$  apart; however, the Arg426-Glu329 pair separated by  $2.6\text{\AA}$ , but the Lys417-Asp30 pair remained intact till the midpoint of the simulation. The bonds between both pairs broke at the same time point and remained separated by  $\sim 5\text{\AA}$  till the end of simulation (*Figure 2A*).. This strong yet transient electrostatic contacts can partly explain the phenomena of receptor recognition and S1 shedding. S protein transiently utilizes the RBD of S1 subunit for receptor recognition and shed them during cell internalization. Thus, the faster SARS-CoV-2 transmission as compared to SARS-CoV is, at least in part, facilitated by the robust Lys417-Asp30. In addition, we also observed that Tyr449, Tyr489, Gln493, and Asn501 in cRBD established strong hydrogen bonds with the interface residues of ACE2 and remain intact along the course of simulation (*Figure 2B*).. These results indicate that these residues are equally responsible for the relatively stronger interaction with ACE2. To demonstrate our results more clearly, we captured the motions of these interface residues in animations, and calculated binding free energies for each complex along the simulation time (*Supplementary movies 1 and 2*).. The polar solvation energy of cRBD-ACE2 is almost half of the sRBD-ACE2, which may compensate the difference of the electrostatic energies of these complexes, resulting in overall similar total binding free energies (see *Figure 2A*).. Collectively, by performing structural modelling, we could demonstrate stronger cRBD-ACE2 interaction compared to sRBD-ACE2 interaction. In addition, we demonstrated that the mutation of Lys417 in RBDR in cRBD allowed to bind ACE2 more readily, which may facilitate the rapid transmission of SARS-CoV-2 compared to SARS-CoV.

Peptide vaccine can block SARS-CoV-2\_ACE2 interaction.

In past the interface information of sRBD-ACE2 had been utilized to design peptide vaccine that were able to block the receptor binding of the virus. The peptide, S<sub>471-503</sub>, derived from the ACE2 binding region of the sRBD was able to hinder the ACE2-RBD interaction and thus viral entry into the cell, as confirmed *in vitro* (29). Another peptide, constructed by the glycine linkage of two separate segments of ACE2 was able to exhibit efficient antiviral activity (IC<sub>50</sub> = 0.1μM) (30). By comparing the cRBD region corresponding to the S<sub>471-503</sub> (ALNCYWPLNDYGFYTTTGIGYQPYRVVLSFEL) peptide, we found that the N-terminus (bold letters) of this peptide and corresponding cRBD region are considerably different but the C-terminus portion (non-bold letters) was 100% identical with cRBD (Figure 1C).. Owing to the difference in the N-terminal half, S<sub>471-503</sub> may not hinder the SARS-CoV-2 cell entry as it exhibited in SARS-CoV. Alternatively, we suggest that peptide, HW1, comprising (not disclosed yet) may have the potential to abrogate the cRBD-ACE2 interaction and subsequent cell entry. Thus far we have observed that cRBD and sRBD interact with the overall same helical peptide of the ACE2 with some differing interface residues from the RBDs (see Figure 2B and Table 1). Hence we suggest that P6 peptide (EEQAKTFLDKFNHEAEDLFYQSS-G-LGKGDFR), reported by Han DP *et al.*, may have the ability to block SARS-CoV-2 cell entry similar to SARS-CoV. Overall, peptides-based vaccines are easy to design, cheaper to synthesise and faster to investigate as compared to mAbs, which take longer time.

### Identification of epitopes on cRBD that bind to SARS-CoV-2 mAbs.

As SARS-CoV-2 and SARS-CoV belong to the *betacoronavisus* genus of coronaviridae family and share ~97% of sequence similarity in the RBD region of S protein (5, 27), researchers have tried to delineate the cross-reactivity of anti-sRBD mAbs with cRBD (9, 28). Although the data for RBD structure is limited yet, SARS-CoV-2 had some conserved RBDR motifs with SARS-CoV (see Figure 1C).. Thus, we could predict the interaction of cRBD with the previous known anti-sRBD mAbs including 80R (18), m396 (20), F26G19 (19), s230 (21), CR3014, and CR3022 (31). To estimate potential of the anti-sRBD mAbs to bind with and neutralize cRBD, we obtained their structures from PDB [for 80R (18), m396 (20), F26G19 (19), s230 (21)] or modelled [for CR3014, and CR3022 (30)](for details see Methods; the 3D structures of these mAbs have been provided in supplementary data).

The variable heavy (VH) and variable light (VL) chains of scFv regions in these mAbs were aligned and their CDRs were annotated (*Figure 3A*).. From these models, we could observe that the VL-CDR1 of CR3022 and s230 were relatively longer and similar as compared to the CDR1 of the other mAbs in their VL chain. In addition, the VH-CDR3 of s230 was more expanded than other mAbs (*Figure 3B*).. Differences in the sequence and length of the CDRs indicate that these mAbs recognise distinct epitopes on the RBD and may not overlap thoroughly.

In addition, we predicted conformational epitopes of cRBD by using structure information of the mAbs. To ensure the authenticity of the epitope prediction, the co-crystal structure of sRBD-F26G19 was used as control. Among the three epitopes we predicted, epitope 1 was completely overlapped with the experimental results, supporting the reliability of our analysis (*Figure 3C*).. Similarly, we also predicted three cRBD epitopes, of which the residues in epitope 2 were mainly composed of RBDR which was highly variable between sRBD and cRBD. By contrast, the residues of the epitope 1 and epitope 3 were significantly conserved respectively between sRBD and cRBD (epitope 1, 93%; epitope 3 and 100%, see *Figure 3D*). The high conservation of structures and sequences between cRBD and sRBD epitopes may indicate that anti-SARS-CoV mAbs can also interact with SARS-CoV-2. Thus, we postulated that the anti-cRBD mAbs that recognize the predicted epitope 1 or epitope 3 might be able to neutralize cRBD. While the epitope 2 region was highly variable between cRBD and sRBD, therefore the epitope 2-binding anti-sRBD mAbs were not expected to bind to cRBD.

### Highly conserved epitope 3 of cRBD is a potential new target for mAbs against SARS-CoV-2

A recent study using SPR and BLI analyses has demonstrated that sRBD mAbs such as m396, 80R, s230, and CR3014 could not recognize cRBD (9, 28). However, these studies did not show whether the interface residues of the these mAbs can hinder their binding to cRBD. To evaluate this, we placed or docked the scFv regions of these mAbs into cRBD and identified their interface residues (*Table 2*).. 80R and s230 interacted with a part of the overlapping residues at the hypervariable RBDR region of cRBD, while m396 and F26G19 were partly overlapped with the residues at non-epitope regions. (*Figure 4A*).. These findings indicate that m396 and F26G19 can bind and neutralize cRBD. The

binding affinity of F26G19 with cRBD has not been studied yet by SPR or BLI analyses, requiring further evaluation in near future. In addition, we found that epitope 2 on cRBD harbored a mutated loop region (475-AGSTPCNGV-483) that has been reported to abrogate the binding of CR3014 mAb to sRBD (31). Taken together, we suggest that these m396, 80R, s230, and F26G19 mAbs recognize non-conserved or non-epitope regions of cRBD, therefore might not be able to neutralize the S protein of SARS-CoV-2.

It is surprising that how cRBD escapes from the anti-sRBD mAbs even though cRBD can bind to ACE2 with higher affinity. This could be explained in part by the structural differences of binding regions between them. Anti-sRBD mAbs had CDR which are very specific and recognize conformational epitopes, while ACE2 utilize a long helix binding longitudinally to RBD.

Interestingly, we observed a mutation Ala475 in cRBD, which was corresponded to Pro462 in sRBD (Figure 1C). Indeed, a previous study has shown that CR3014 mAb is not effective to the mutant Pro462Leu viruses, although it can prevent lung damage and SARS-CoV shedding in ferrets (31). Moreover, we also found a glycine insertion mutation in the same loop (475-AGSTPCNGV-483), lengthening the loop RMSD to 2-3Å (see *Figure 1C*). This might be the reason why the previous BLI study could not observe the binding of cRBD with CR3014 (28).

Thus, to further evaluate the binding of CR3014, we performed epitope mapping and protein ligand interaction fingerprints (PLIF) analyses, which revealed a CR3014 cluster around the same AGSTPCNGV-loop, a part of the epitope 2 (*Figure 4B*). This finding implies that the sRBD mAbs targeting the epitope 2 in RBD cannot neutralize cRBD. Also, as the previously known mAbs recognize the epitopes in variable RBDR regions, they may not be able to neutralize cRBD. From these results, we suggest that only the mAbs that can bind to a conserved epitope or the peptides that can mimic the cRBD-ACE2 interface might be able to neutralize cRBD.

A novel mAb, CR3022 has been reported to completely neutralize the CR3014 escape mutants (*i.e.*, Pro462Leu), and synergizes the neutralizing effect of CR3014 without competing to its epitopes (30). This study also has shown that CR3022 did not compete with cRBD showing strong binding affinity in the BLI analysis, while the other mAbs including CR3014, m396, and MERS-CoV neutralizing mAb

m336 could not bind to cRBD. These results imply that CR3022 binds to a conserved epitope of RBD in SARS-CoV-2 and SARS-CoV. To answer this, we performed an antibody docking procedure and calculated PLIF based on 100 docked poses of the CR3022-cRBD complex. Differing from CR3014, CR3022 was clustered over Arg24 and Arg26, and interacted with Glu19 in cRBD, which we designated as epitope 3 (*Figure 4B, C*).

Tamina Park *et al.* has performed computational analyses to demonstrate whether the previously known anti-MERS-CoV and anti-SARS-CoV mAbs can bind and neutralize cRBD (32). They suggested that binding of CR3022 and s230 were overlapped in the docking analysis, neutralizing cRBD. However, this study did not consider the fact that the spike protein particularly the RBD of SARS-CoV and MERS-CoV have significant variations, which may preclude neutralization of cRBD by anti-sRBD mAbs (33). In addition, Xiaolong Tian *et al.* (28) and we have demonstrated that there is no overlap between CR3022 and the ACE2-binding region of cRBD (see *Supplementary Figure 2A*). A crystal structure analysis also revealed that s230 can bind to the ACE2-binding region of sRBD (21). To validate these results, we superimposed the structures of s230-sRBD and ACE2-sRBD complexes, which revealed that ACE2 and s230 were overlapped with the same interface of sRBD (*Supplementary Figure 2B*). We also found that CR3022 did not compete with ACE2 and CR3014 interface of the cRBD in consistent with the previous studies for SARS-CoV (31) and recent SARS-CoV-2 (28) (*Supplementary Figure 2A*). Strikingly, we found that CR3022 bind to a highly conserved region, partly overlapping with epitope 3 (*Figure 4C*). This result strongly indicates that CR3022 has a potential, at least in part, to neutralize cRBD, and raise a possibility that the conserved epitopes of RBD can be promising targets for the broad-range or cross-reactive mAbs engineered for SARS-CoV and SARS-CoV-2.

In summary, we suggest that Lys417 mutation in cRBD acquires stronger electrostatic interaction with ACE2, which may facilitate faster receptor-recognition of cells. This interaction is strengthened by hydrophobic and Van Der Waals contacts at the cRBD-ACE2 interface. Moreover, we demonstrated that CR3022, but not the other mAbs for SARS-CoV, can recognize a conserved epitope on RBD, implying its potential to neutralize SARS-CoV-2. We suggest that our findings can provide new

insights on the underlying mechanisms on the high infectivity of SARS-CoV-2, which may be helpful for developing new effective neutralizing agents against SARS-CoV-2.

## Declarations

**Acknowledgements** This research was supported by grants from the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (MSIT) (NRF-2017M3C9A6047620, NRF-2019R1A5A2026045, NRF-2017M3A9B6061509), Republic of Korea.

## Author contributions

M.S. S.C, and H.W. contributed toward conceptualization and methodology; M.S. and B.A. performed the investigation and formal analysis; M.S. and H.W. wrote the original manuscript draft; H.W and S.C. supervised the study; H.W. was responsible for funding acquisition; all the authors contributed to editing and reviewing the manuscript.

## Competing interests

All authors declare that there is no competing interest.

## Materials & Correspondence

Correspondence and material requests should be addressed to Hyun Goo Woo.

## References

- 1.WHO. Coronavirus disease 2019 (COVID-19) Situation Report—46 6 March 2020.
- 2.Behzadi MA, Leyva-Grado VH. Overview of Current Therapeutics and Novel Candidates Against Influenza, Respiratory Syncytial Virus, and Middle East Respiratory Syndrome Coronavirus Infections. *Front Microbiol.* 2019;10:1327.
- 3.Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. *Nature.* 2020.
- 4.Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* 2020.
- 5.Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med.* 2020;382(8):727–33.
- 6.Paraskevis D, Kostaki EG, Magiorkinis G, Panayiotakopoulos G, Sourvinos G, Tsiodras S. Full-genome



evolutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event. *Infect Genet Evol.* 2020;79:104212.

7.Gallagher TM, Buchmeier MJ. Coronavirus spike proteins in viral entry and pathogenesis. *Virology.* 2001;279(2):371–4.

8.Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol.* 2020.

9.Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science.* 2020.

10.Yuan Y, Cao D, Zhang Y, Ma J, Qi J, Wang Q, et al. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat Commun.* 2017;8:15092.

11.Park S, Shin HJ, Shah M, Cho HY, Anwar MA, Achek A, et al. TLR4/MD2 specific peptides stalled in vivo LPS-induced immune exacerbation. *Biomaterials.* 2017;126:49–60.

12.Junaid M, Li CD, Shah M, Khan A, Guo H, Wei DQ. Extraction of molecular features for the drug discovery targeting protein-protein interaction of *Helicobacter pylori* CagA and tumor suppressor protein ASSP2. *Proteins.* 2019;87(10):837–49.

13.Achek A, Shah M, Seo JY, Kwon HK, Gui X, Shin HJ, et al. Linear and Rationally Designed Stapled Peptides Abrogate TLR4 Pathway and Relieve Inflammatory Symptoms in Rheumatoid Arthritis Rat Model. *J Med Chem.* 2019;62(14):6495–511.

14.Shah M, Anwar MA, Park S, Jafri SS, Choi S. In silico mechanistic analysis of IRF3 inactivation and high-risk HPV E6 species-dependent drug response. *Sci Rep.* 2015;5:13446.

15.Narayanan A, Sellers BD, Jacobson MP. Energy-based analysis and prediction of the orientation between light- and heavy-chain antibody variable domains. *J Mol Biol.* 2009;388(5):941–53.

16.Chothia C, Lesk AM. Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol.* 1987;196(4):901–17.

17.Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol.* 1995;247(4):536–40.

18.Hwang WC, Lin Y, Santelli E, Sui J, Jaroszewski L, Stec B, et al. Structural basis of neutralization by

- a human anti-severe acute respiratory syndrome spike protein antibody, 80R. *J Biol Chem.* 2006;281(45):34610–6.
- 19.Pak JE, Sharon C, Satkunarajah M, Auperin TC, Cameron CM, Kelvin DJ, et al. Structural insights into immune recognition of the severe acute respiratory syndrome coronavirus S protein receptor binding domain. *J Mol Biol.* 2009;388(4):815–23.
- 20.Prabakaran P, Gan J, Feng Y, Zhu Z, Choudhry V, Xiao X, et al. Structure of severe acute respiratory syndrome coronavirus receptor-binding domain complexed with neutralizing antibody. *J Biol Chem.* 2006;281(23):15829–36.
- 21.Walls AC, Xiong X, Park YJ, Tortorici MA, Snijder J, Quispe J, et al. Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell.* 2019;176(5):1026–39 e15.
- 22.Dunbar J, Krawczyk K, Leem J, Marks C, Nowak J, Regep C, et al. SAbPred: a structure-based antibody prediction server. *Nucleic Acids Res.* 2016;44(W1):W474–8.
- 23.Krawczyk K, Liu X, Baker T, Shi J, Deane CM. Improving B-cell epitope prediction and its application to global antibody-antigen docking. *Bioinformatics.* 2014;30(16):2288–94.
- 24.Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods.* 2017;14(1):71–3.
- 25.Kumari R, Kumar R, Open Source Drug Discovery C, Lynn A. g\_mmpbsa—a GROMACS tool for high-throughput MM-PBSA calculations. *J Chem Inf Model.* 2014;54(7):1951–62.
- 26.Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of the SARS-CoV-2 by full-length human ACE2. *Science.* 2020.
- 27.Alexandra C. Walls1 Young-Jun Park M. Alejandra Tortorici AW ATMaDV. Structure, function and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell press.* 2020.
- 28.Tian X, Li C, Huang A, Xia S, Lu S, Shi Z, et al. Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. *Emerg Microbes Infect.* 2020;9(1):382–5.
- 29.Hu H, Li L, Kao RY, Kou B, Wang Z, Zhang L, et al. Screening and identification of linear B-cell epitopes and entry-blocking peptide of severe acute respiratory syndrome (SARS)-associated

coronavirus using synthetic overlapping peptide library. *J Comb Chem.* 2005;7(5):648-56.

30.Han DP, Penn-Nicholson A, Cho MW. Identification of critical determinants on ACE2 for SARS-CoV entry and development of a potent entry inhibitor. *Virology.* 2006;350(1):15-25.

31.Ter Meulen J, van den Brink EN, Poon LL, Marissen WE, Leung CS, Cox F, et al. Human monoclonal antibody combination against SARS coronavirus: synergy and coverage of escape mutants. *PLoS Med.* 2006;3(7):e237.

32.Tamina Park. Sang LY. Seil Kim MJ, Hong GK, Sangmi Jun, Seung Kim, Bum TK, Edmond CP, Daeui Par. Spike protein binding prediction with neutralizing antibodies of SARS-CoV-2. *bioRxiv.* 2020.

33.Chen Y, Rajashankar KR, Yang Y, Agnihothram SS, Liu C, Lin YL, et al. Crystal structure of the receptor-binding domain from newly emerged Middle East respiratory syndrome coronavirus. *J Virol.* 2013;87(19):10777-83.

## Tables

Due to technical limitations, the tables are only available as a download in the supplemental files section.

## Figures

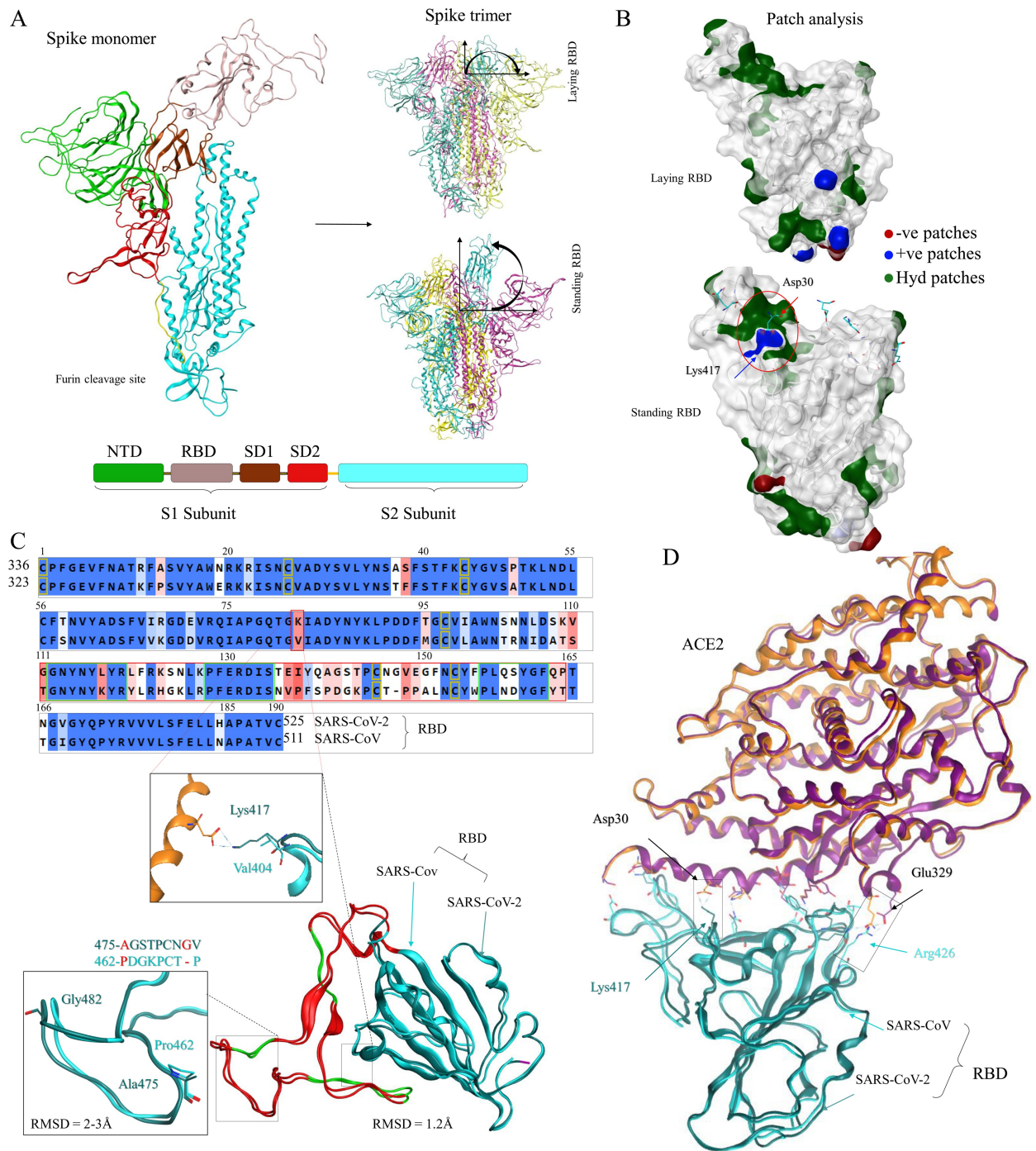


Figure 1

Modelling of the monomeric and trimeric SARS-CoV-2 S and receptor binding of the RBD. A) Domains architecture of the monomeric full-length S protein and its trimeric states are shown. The RBD are displayed in receptor accessible “standing” and non-accessible “Laying” conformation. B) The surface patch analysis suggests the exposure of Lys417 in

standing conformation (Asp30 belongs to ACE2 in the patch analysis). The color cods are displayed according to the charge of the patch. C) RBD alignment of the SARS-CoV-2 (accession # NC\_045512) and SARS-CoV (PDB ID: 6ACG). Residues number at the start are according to the full-length S model and the top numbering are for the RBD region only. The numbering on the start are used in RBD-ACE2 docking analysis while the top numbering are used in RBD-mAbs docking analysis. The lower panel shows the superimposition of sRBD and cRBD. Enhanced box shows crucial mutations responsible for high receptor binding (Lys417) and antibody resistance (Ala475, Gly482). D) The superimposed RBD-ACE2 complexes in both SARS-CoV and SARS-CoV-2 are shown and the interface residues are depicted as sticks. Electrostatic contacts are highlighted in boxes.

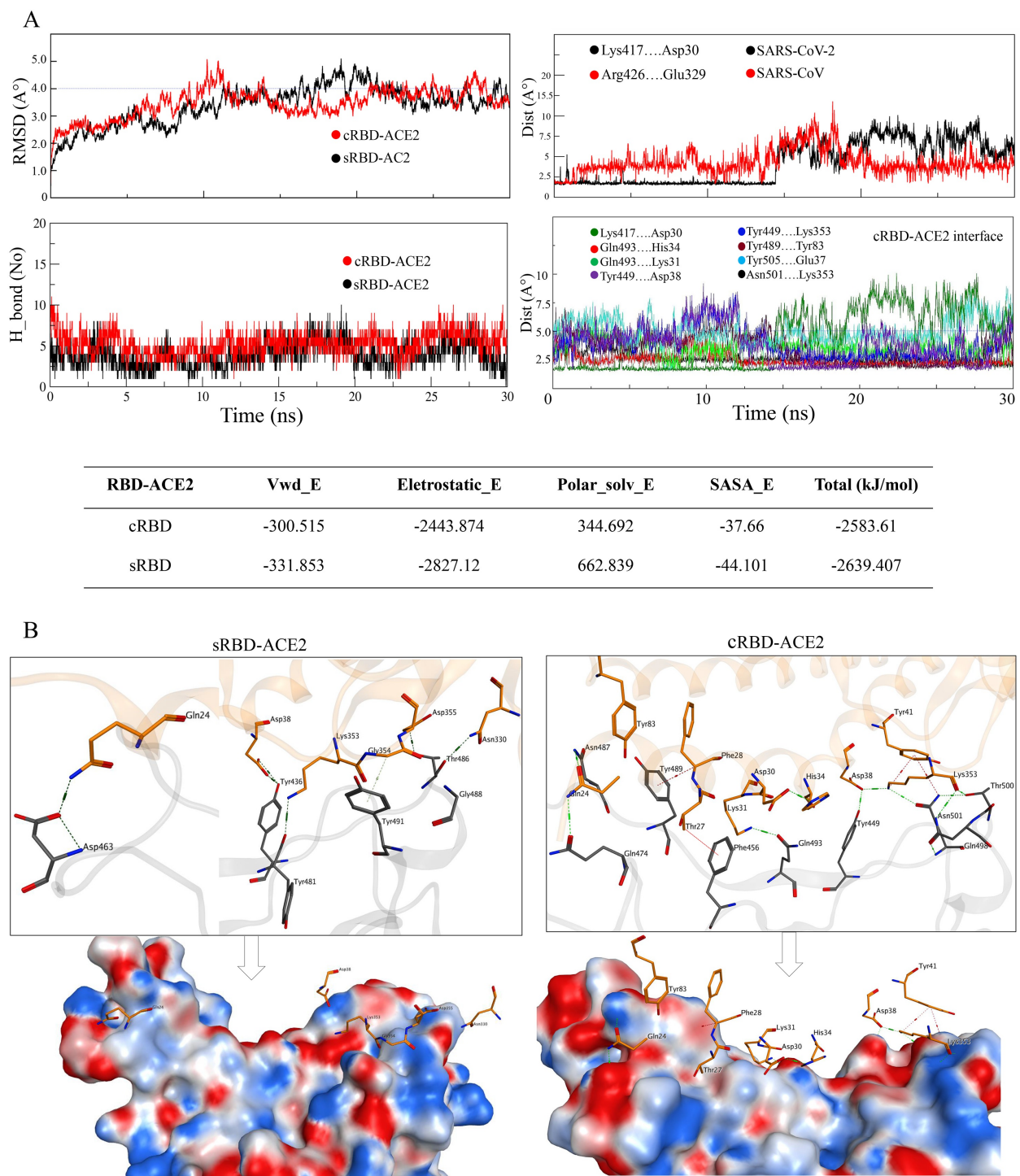


Figure 2

Binding affinity and interface analysis of the RBD-ACE2. A) The RMSD of the backbone atoms of the RBD-ACE2 complexes with reference to the 0 ns conformation are shown. Total number of hydrogen bonds between RBD-ACE2 interfaces are shown in the bottom left panel. The minimum contact distances, as a function of time, between the oppositely



charged residues at the RBD-ACE2 interface are shown in the right top panel. The contact distances between the interface residues of cRBD-ACE2 complex are shown. The bottom panel show the comparative binding free energies for the RBD-ACE2 complexes. B) The interfaces (PDB frames extracted at 20 ns of simulation) of ACE2 with sRBD (left) and cRBD (right) and their electrostatic surface maps, respectively.

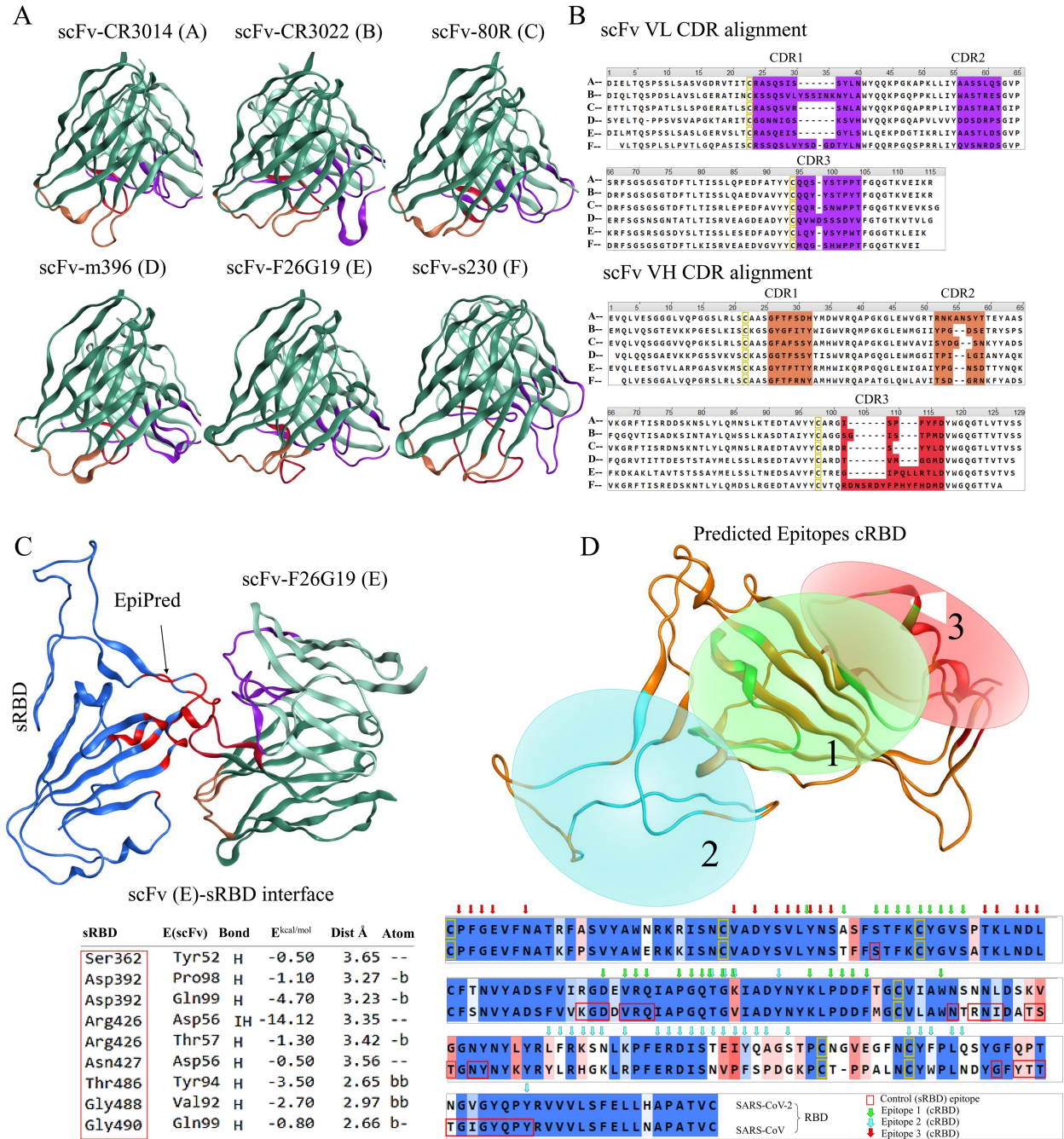
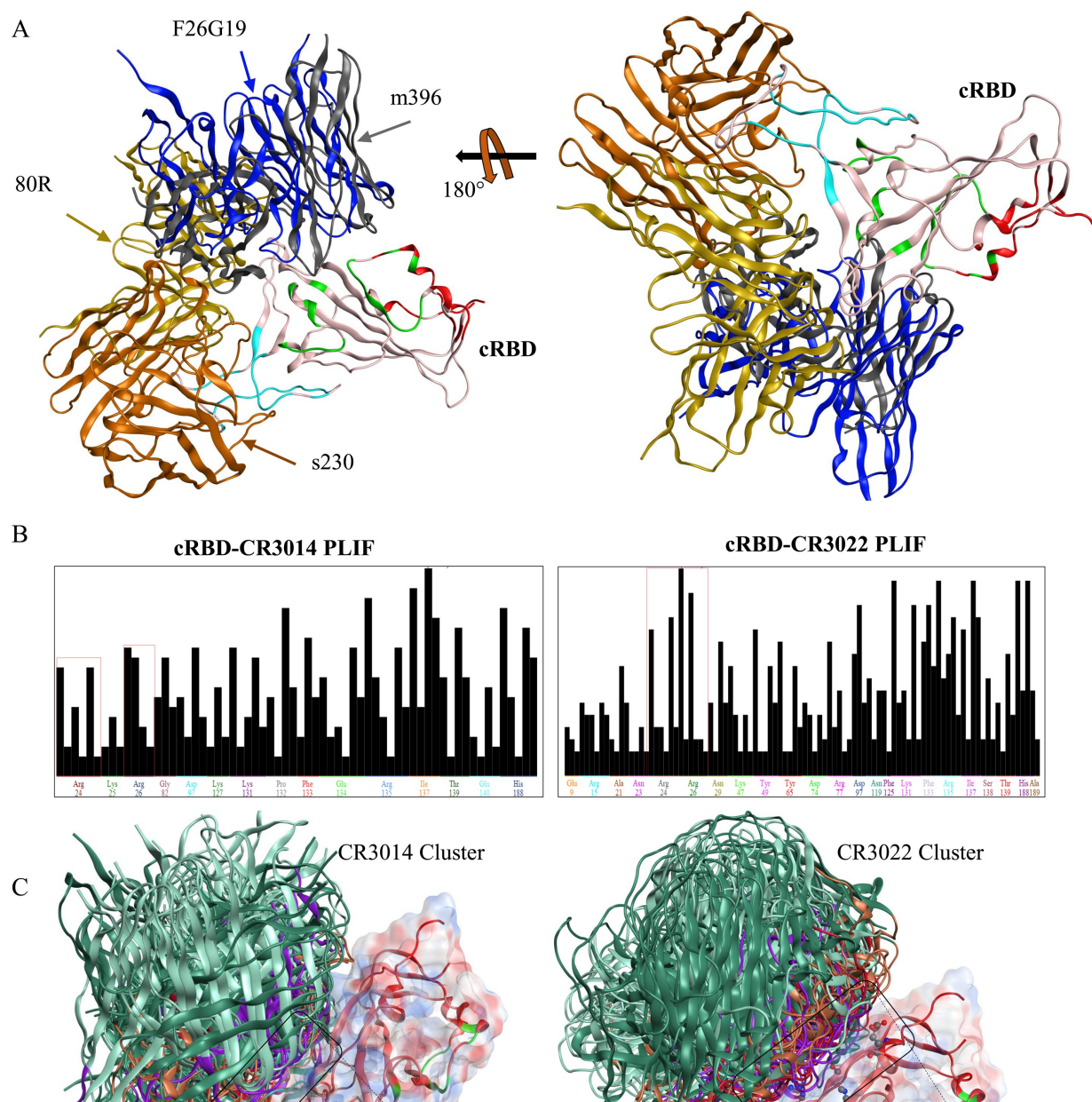


Figure 3

Epitope mapping of the cRBD and CDRs annotation of the mAbs. A) Anti-sRBD mAbs (single-

chain variable fragment scFv are shown) and their complementarity-determining regions (CDRs) are shown. B) The variable light (VL) and variable heavy (VH) chains of the scFv regions are aligned and the CDR regions are annotated according to Chothia and Lesk numbering scheme. C) The epitope prediction was validated through sRBD-F26G19 complex (PDB ID: 3BGF). The tabular interface is reported in the crystal structure while red boxes in the aligned sequences show the EpiPred predicted epitope. D) conformational epitopes predicted with reference to six known anti-sRBD mAbs are highlighted and encircled. Residues participating in epitopes are indicated with arrows (the arrow colors correspond to their respective epitopes).





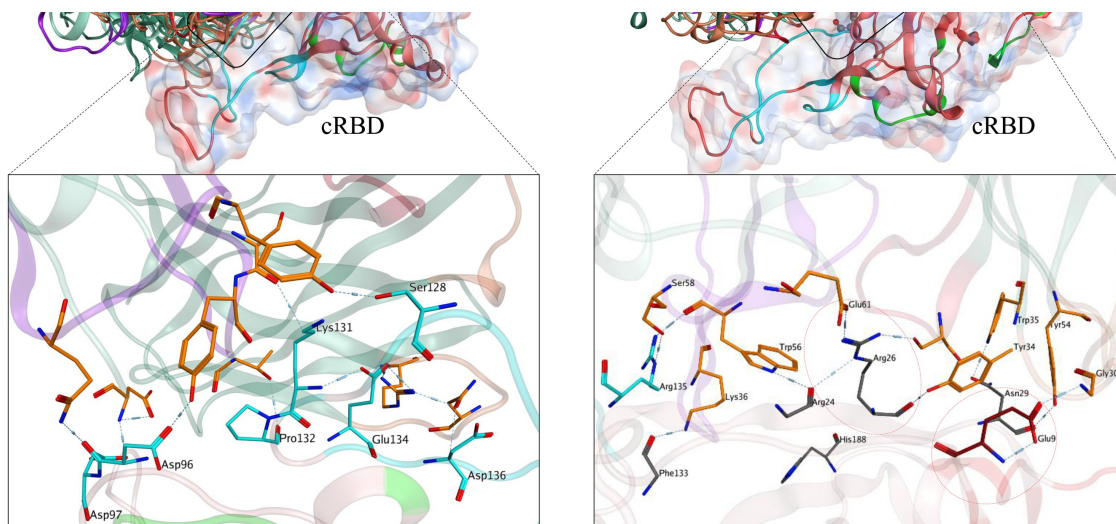


Figure 4

The interface and protein ligand interaction fingerprints (PLIF) analysis. A) The binding interface of mAbs are displayed with respect to the predicted epitopes on cRBD (Green = Epitope 1, Cyan = Epitope 2, Red = Epitope 3; this has been further explained in Figure 4). B) PLIF analysis of the CR3014 and CR3022 mAbs with cRBD. C) CR3014 cluster around the epitope 2, which is highly variable between sRBD (left) and cRBD (right). CR3022 cluster near epitope 3, which is conserved between cRBD and sRBD.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

Supplementary\_movie\_2.m4v  
 SARS-CoV-2\_Spike\_ACE2\_Supplementary.pdf  
 SARS-CoV-2\_Spike\_ACE2\_TABLES.pdf  
 Supplementary\_movie\_1.m4v  
 Spike\_trimer.pdb  
 Spike\_trimer\_ACE2\_complex.pdb  
 SARS-CoV-2-RBD\_ACE2 Complex.pdb  
 CR3014-mAb.pdb  
 CR3022\_mAb.pdb