Predicted structural mimicry of spike receptor-binding motifs from highly pathogenic human coronaviruses

Christopher A Beaudoin[1*], Arian R Jamasb[1,2], Ali F Alsulami[1], Liviu Copoiu[1], Andries J van Tonder[3], Sharif Hala[4,5], Bridget P Bannerman[1], Sherine E Thomas[1], Sundeep Chaitanya Vedithi[1], Pedro H M Torres[6], and Tom L Blundell[1*]

[1]Department of Biochemistry, Sanger Building, University of Cambridge, Tennis Court Rd, Cambridge, CB2 1GA, UK
[2]Department of Computer Science & Technology, University of Cambridge, JJ Thomson Ave, Cambridge, CB3 0FD, UK
[3]Department of Veterinary Medicine, University of Cambridge, Madingley Rd, Cambridge, CB3 0ES, UK
[4]King Abdullah International Medical Research Centre - Ministry of National Guard Health Affairs, Jeddah, Saudi Arabia
[5]King Saud bin Abdulaziz University for Health Sciences, Jeddah, Saudi Arabia
[6]Laboratório de Modelagem e Dinâmica Molecular, Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil

*corresponding author and lead contact
Joint corresponding authors email addresses:
Tom Blundell: tlb20@cam.ac.uk
Chris Beaudoin: cab233@cam.ac.uk

Highlights

- Potential coronavirus spike protein mimicry revealed by structural comparison
- Human and non-human protein potential interactions with virus identified
- Predicted structural mimicry corroborated by protein-protein docking
- Epitope-based alignments may help guide vaccine efforts

Summary

Viruses often encode proteins that mimic host proteins in order to facilitate infection. Little work has been done to understand the potential mimicry of the SARS-CoV-2, SARS-CoV, and MERS-CoV spike proteins, particularly the receptor-binding motifs, which could be important in determining tropism of the virus. Here, we use structural bioinformatics software to characterize potential mimicry of the three coronavirus spike protein receptor-binding motifs. We utilize sequence-independent alignment tools to compare structurally known or predicted three-dimensional protein models with the receptor-binding motifs and verify potential mimicry with protein docking simulations. Both human and non-human proteins were found to be similar to all three receptor-binding motifs. Similarity to human proteins may reveal which pathways the spike protein is co-opting, while analogous non-human proteins may indicate shared host interaction partners and overlapping antibody cross-reactivity. These findings can help guide experimental efforts to further understand potential interactions between human and coronavirus proteins.
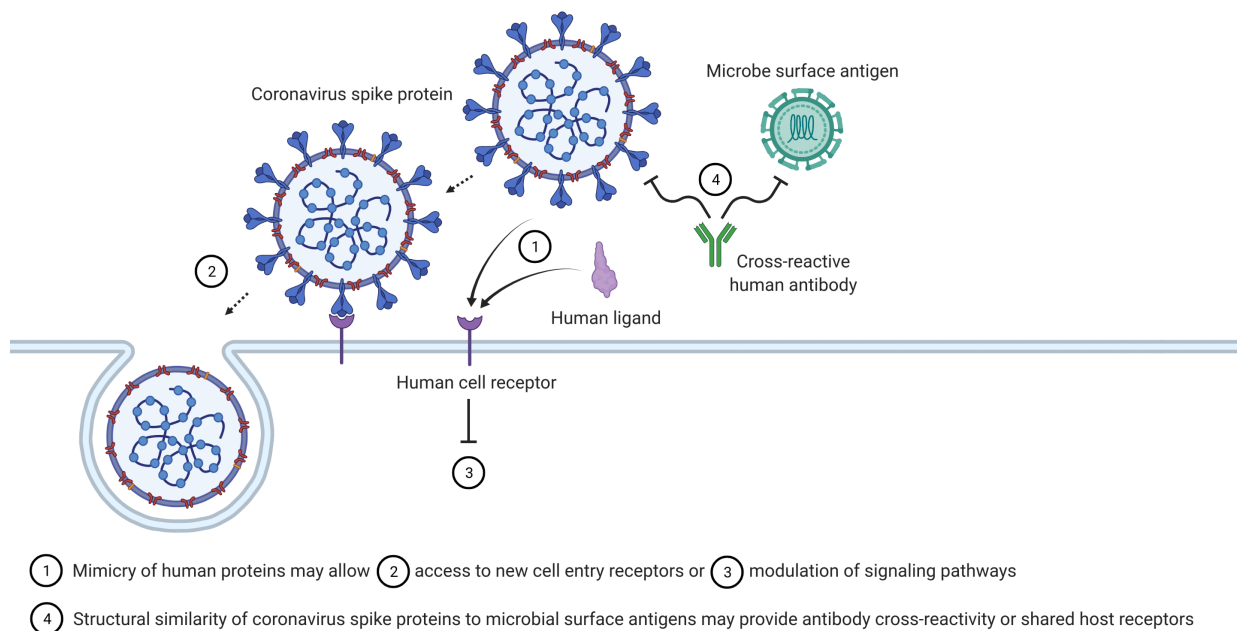
Keywords

SARS-CoV-2

SARS-CoV
MERS-CoV
coronavirus
coronavirus spike protein
COVID-19
structural alignment
viral host mimicry
infectious disease
structural bioinformatics

Graphical abstract

Potential molecular mimicry by coronavirus spike proteins

① Mimicry of human proteins may allow ② access to new cell entry receptors or ③ modulation of signaling pathways

④ Structural similarity of coronavirus spike proteins to microbial surface antigens may provide antibody cross-reactivity or shared host receptors

Introduction

Viruses have long been known to utilize molecular mimicry of host proteins to interrupt and exploit host biochemical pathways during infection (Alcami, 2003; Lasso et al., 2021). Alongside the need to employ host machinery for the viral replication cycle, the evolution of viral protein motifs that resemble host proteins can result in new virulence mechanisms, such as inducing inflammation and evading the immune system (Elde and Malik, 2009). Coronaviruses, in particular, have been suspected to have acquired human protein mimics throughout the long record of human coronavirus infections (Chew et al., 2003; Walls et al., 2019). As further evidence, the highly pathogenic human coronaviruses, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), SARS-CoV, and the Middle Eastern Respiratory Syndrome Coronavirus (MERS-CoV), have been shown to encode numerous short linear motifs across their genomes that are homologous to human proteins (Martínez et al., 2021). Although coronavirus infections are typically localized to the lungs, resulting in respiratory infections, viral material has also been found in other organs, such as the kidney, brain, and heart, resulting in more life-threatening infections (Renu et al., 2020). Furthermore, SARS-CoV-2 (the causative agent of the COVID-19 pandemic) infection has presented symptoms not previously seen in other coronavirus infections, such as conjunctival

discharge from the eyes (Cheong, 2020; Zheng et al., 2021). Investigations into coronavirus host mimicry may shed light on viral tropism and infection severity (Angileri et al., 2020a).

The structure of the receptor-binding motif (RBM) on the spike glycoprotein is particularly important for determining the tropism of the virus (Reguera et al., 2014). Host receptors that contain motif(s) that complement the electrochemical and spatial configurations of the viral RBM will interact and, thus, initiate viral entry (Li, 2015; Tortorici and Veesler, 2019). Angiotensin converting enzyme II (ACE2) has been established as the primary cell entry receptor for SARS-CoV-2 and SARS-CoV and dipeptidyl peptidase IV (DPP4) as the primary cell entry receptor for MERS-CoV. However, several reports, some preliminary, have proposed additional coronavirus cell entry receptors, such as transferrin receptor protein 1, kidney injury molecule-1, kremen protein 1, and αv integrins for SARS-CoV-2 (Gu et al., 2020; Ichimura et al., 2020; Li et al., 2003; Sigrist et al., 2020; Tang et al., 2020; Wang et al., 2013; Yang et al., 2020). Additionally, coronavirus spike proteins have been proposed to interact with host factors to facilitate infection aside from their role in cell entry (Patra et al., 2020). For instance, two studies found that the SARS-CoV-2 spike protein alone can interact with the blood brain barrier (Buzhdygan et al., 2020; Rhea et al., 2021). The importance in receptor-binding and low glycosylation surrounding the coronavirus RBM residues make it an attractive target for inhibition by small-molecule drugs, therapeutic peptides, and neutralizing antibodies (Hussain et al., 2020; Ling et al., 2020; Pandey et al., 2020).

To date, there has been limited investigation into the structural similarity of highly pathogenic coronavirus RBMs (Kanduc and Shoenfeld, 2020). Identifying structurally analogous human proteins may give insight into endogenous biochemical pathways that the virus is hijacking to facilitate infection or may help explain autoimmune disorders triggered by coronavirus infections (Angileri et al., 2020b; Drayman et al., 2013). Detecting similar microbial proteins may reveal shared host receptors or antibody cross-immunity (Huang et al., 2020). Short linear motifs on coronavirus spike RBMs have been shown to share high amino acid sequence identity with human proteins, which may indicate host mimicry (An and Park, 2020; Grifoni et al., 2020; Lin et al., 2020; Lucchese and Flöel, 2020). However, protein structure and fold similarity have been shown as more informative than amino acid sequence similarity in predicting molecular mimicry (Krishna and Grishin, 2004; Westall, 2006). Drayman *et al.* performed a structural similarity search using bacterial and viral motifs and experimentally validated the simian vacuolating virus 40 major capsid protein mimicry of Gas6 binding with TAM – Tyro3, Axl, and Mer – receptors, demonstrating that structural paralogs with low amino acid identity may still act as molecular mimics. Thus, to add to the understanding of host mimicry of highly pathogenic coronavirus RBMs, we used structural bioinformatics tools to model and map the extent to which the three-dimensional structures of the SARS-CoV-2, SARS-CoV, and MERS-CoV spike RBMs are potentially mimicking the interactions of experimentally-determined protein structures. We used structural alignment tools with distinct methodologies to perform a structural similarity screen between the RBMs and all known protein structures and, subsequently, tested potential RBM interactions with protein-protein docking simulations. Several cell signaling proteins, innate immune factors, snake and spider toxins, and microbial antigens are found to share structural features with the three RBMs. This information may help guide experimental efforts to elucidate spike RBM interactions, including that of vaccine design and cell entry receptor discovery.

Results and Discussion

Receptor-binding motif structural similarities and characteristics

Several models of the spike protein for each of the highly pathogenic coronaviruses have been experimentally determined; however, many of them are missing residues due to the difficulty in

resolving the structure of flexible protein motifs (Nwanochie and Uversky, 2019). To overcome this issue and obtain a representative three-dimensional model of each spike receptor-binding motif (RBM), we used ProtCHOIR, a recently developed pipeline to automate the modelling of homo-oligomers, to model each trimeric spike protein and, subsequently, manually selected the RBM residues for each coronavirus (Figure 1). All generated models were structurally aligned to experimental models using TM-align to determine modelling precision. On a scale from 0 to 1, a TM-score of over 0.5 between two proteins implies that they have the same fold, while below 0.2 suggests a random alignment. Each RBM alignment with the corresponding experimental structure reported a TM-score over 0.95, reflecting high-quality modelling. Although receptor-binding of coronavirus spike proteins has been shown to be an elaborate process that involves interactions with glycans and multiple protein domains, we selected the most interactive region of the spike RBD with primary receptors (i.e. ACE2 for SARS-CoV and SARS-CoV-2; DPP4 for MERS-CoV) from experimental models as the receptor-binding motif (RBM) (Wang and Xiang, 2020).

The structural similarity of the RBMs to one another was quantitatively assessed using TM-align before assessing their similarity to other known proteins. The SARS-CoV-2 and SARS-CoV RBMs were very similar with a TM-score of 0.71, while the TM-scores of MERS-CoV with the other two were both less than 0.25 (Figure 1). This level of divergence is also reflected at the amino acid sequence level for the RBD of SARS-CoV-2 and SARS-CoV at 64.6% sequence identity and MERS-CoV with SARS-CoV-2 and SARS-CoV at 19% and 21.6%, respectively.

As seen in Figure 1, the SARS-CoV-2 RBM is comprised almost exclusively of hydrophobic and polar non-charged amino acids, with the exception of one acidic glutamate and one basic lysine. SARS-CoV is similar to SARS-CoV-2 in that it is composed mostly of hydrophobic and polar non-charged residues with some exceptions as single amino acid differences, such as an acidic aspartate in the middle of the SARS-CoV RBM. The MERS-CoV RBM consists of more acidic and basic amino acids and contains fewer polar non-charged residues. Of note, the SARS-CoV-2 and SARS-CoV RBMs have 7 and 8 aromatic residues, respectively, exposed on the receptor-binding surface of the RBM. The recent discovery of the N501Y and E484K mutants add a potentially functional aromatic and basic residue, respectively, in the SARS-CoV-2 RBM – both of which have been proposed to increase binding to ACE2 (Nelson et al., 2021). Modelling of the mutants yielded very small structural changes in the SARS-CoV-2 RBM – TM-scores of the mutant RBMs aligned to the reference structure were above 0.9.

In terms of global architecture, the SARS-CoV-2 and SARS-CoV RBMs contain two anti-parallel beta-strands connecting three loops, although the SARS-CoV-2 RBM has two short beta-strands leading to a cystine disulfide loop (Figure 1). Both SARS-CoV and SARS-CoV-2 contain a similar cystine disulfide bond helping shape one end of the respective RBMs. The MERS-CoV RBM consists of three beta-strands connecting four loops. Because loop flexibility may affect overall structure, we submitted each RBD to the CABS-flex 2.0 web server and found that the cystine disulfide loop of both the SARS-CoV-2 and SARS-CoV RBMs displayed high flexibility (> 9 RMSF) - otherwise, the RBM residues on all three RBMs were predicted to exhibit low RMSF (< 6.5) (Supplementary Figure 1). The flexibility predictions from CABS-Flex 2.0 were supported by separate studies on coronavirus RBMs (Saputri et al., 2020; Spinello et al., 2020). The high flexibility of the cystine loops in the SARS-related RBMs motivated the use of two additional models provided by CABS-Flex 2.0 for the structural similarity screen. The added models reported surprisingly low TM-scores compared to the references (0.42 and 0.65 for SARS-CoV-2 and 0.45 and 0.41 for SARS-CoV), revealing the high flexibility in these loops (Supplementary Figure 1). Overall, SARS-CoV-2 and SARS-CoV were found to share higher structural homology with one another than in comparison with MERS-CoV.
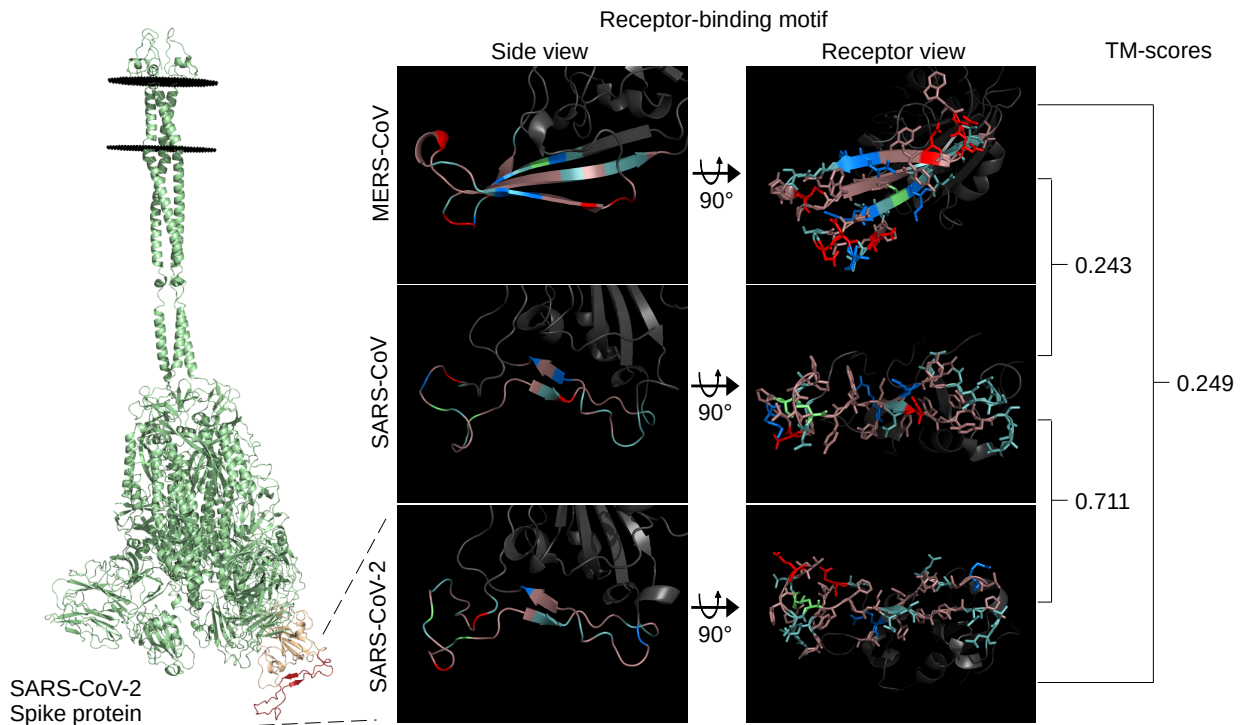
Figure 1. Spike receptor-binding motif comparison
The full-length SARS-CoV-2 spike protein (green), left, modelled using ProtCHOIR is shown with the receptor-binding domain (yellow) and receptor-binding motif (red) marked. The RBMs from the side view are shown, middle, with the amino acids labelled by color: red for acidic (D,E), blue for basic (H,R,K), light teal for polar non-charged (S,N,T,Q), dirty violet for hydrophobic (A,V,I,L,M,F,W,P,G,Y), and lime green for cysteine residues. RBMs from the host cell receptor side are shown, right, with amino acid stick configurations.

Structural similarity screen

After RBM model generation, we performed a structural similarity screen for each RBM. Four sequence-independent 3D-structure alignment tools with different methodologies were used to quantify the structural similarity between the RBMs and known 3D protein structures in order to better understand shared structural features between the RBMs and potential mimics. Notably in this study, although spike may engage in interactions within human cells, we focused on protein structures that would be found in the extracellular matrix (excluding antibodies, due to their structural diversity) to gain more insight into potential cell entry receptors, immunopathies, and shared antigenicity with other microorganisms (Versteeg et al., 2007).

The PDBeFold, RUPEE, and HMI-PRED web servers were used, and TM-align was locally-installed and run pairwise against the downloaded PDB database clustered at 100% sequence identity. The TM-score distributions between SARS-CoV-2 and SARS-CoV were quite similar, while MERS-CoV was more similar to a greater number of proteins (Figure 2A). The MERS-CoV RBM returned 3,954 structures with a TM-score of over 0.5 (~ top 1% of TM-scores) out of 245,055 total RBM-chain alignments and an average TM-score of 0.33. The SARS-CoV-2 and SARS-CoV RBMs had lower average TM-scores, 0.298 and 0.297 respectively, and the top 1% corresponded roughly to the 0.4 TM-score line. Thus, structures with a TM-score of > 0.4 were selected for further analysis for the SARS-related viruses: 4,025 for SARS-CoV-2 and 3,561 for SARS-CoV. PDBeFold returned 621-806 and 1,163 structures for the SARS-related and MERS-CoV RBM models, respectively. The top 1,000 hits from each RUPEE run were recorded. HMI-PRED outputs ranged from 20-50 mimicked PDB templates per RBM. All alignments of interest were manually inspected to validate the potential for structural mimicry. Returned aligned proteins

from each tool were linked to their corresponding PDB and UniProt codes. Shared UniProt codes between two or more tools were regarded as high-confidence hits. Biologically relevant structural alignments specific to each tool were also inspected and considered. Structural alignments that would not make sense biologically, such as when the RBM is facing the inside of the protein, were discarded, while alignments that were logical but found outside of protein-protein interfaces were included on a case-by-case basis. Returned structures not shown to be found in the extracellular matrix were removed. All tools returned their respective spike structures, confirming their validity.

A total of 62 UniProt codes, excluding 28 toxins, were considered as biologically relevant, which were comprised of 35, 19, 19, and 8 selected alignments from RUPEE, HMI-PRED, PDBeFold, and TM-align, respectively. When comparing tools (Figure 2B), RUPEE and PDBeFold web servers shared 7 UniProt codes for at least one RBM, while TM-align shared 1 with PDBeFold and 0 with RUPEE. HMI-PRED shared 2 structures with RUPEE, 1 with PDBeFold, and 0 with TM-align. Little overlap was shown between most of the tools, which is consistent with structural similarity-based studies on HIV and human proteins (Doolittle and Gomez, 2010). The combined returned UniProt codes, excluding toxins, from all four tools totalled 39, 23, and 29 for the SARS-CoV-2, SARS-CoV, and MERS-CoV RBMs, respectively. The top alignments consisted of cytokines, chemokines, and growth factors and their receptors; structures containing EGF-like domains; complement activation proteins; cystine disulfide-rich toxins derived from snakes and spiders; and antigenic microbial proteins. A Venn diagram showing some shared hits between the three RBMs can be seen in Figure 2C, and a full listing of the hits, alignment values, and tools can be found in Tables 1 and 2. The SARS-CoV and SARS-CoV-2 RBMs shared more structural domains, while MERS-CoV returned more unique hits compared to the other two. Altogether these results indicate that proteins from completely different protein families may interact with coronavirus spike RBMs.
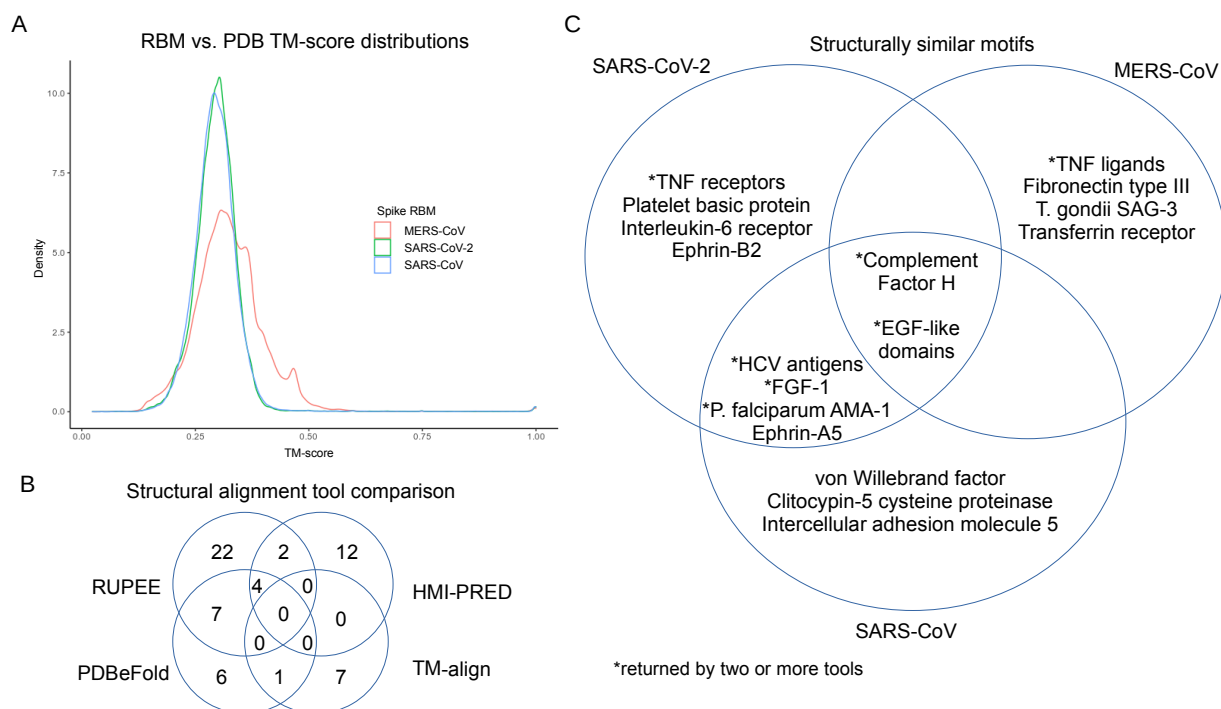

Figure 2. Summary of structural similarity screen
The TM-scores generated from the in-house TM-align screen are displayed as a density plot for each RBM (A). Number of shared proteins from the tools used in the structural similarity screen compared and contrasted (B). Structurally similar motifs, common between coronavirus receptor-binding motifs, compared and contrasted (C).

Analysis of predicted structural mimicry

Further examination of the structural alignments and their relevance to biological activity was performed to elucidate potential mechanisms of molecular mimicry by the SARS-CoV-2, SARS-CoV, and MERS-CoV spike RBMs. The UniProt and STRING databases were used to link the predicted mimics with potential interaction partners, and the PDB provided template structures to determine whether the alignments were found in ligand-binding regions. Selected high-confidence potential interactions were further evaluated using protein-protein docking with ClusPro PIPER in order to better understand electrochemical, in addition to structural, complementarity considering the low amino acid sequence identity. The docked models were then analyzed with the FoldX AnalyseComplex program to determine the complex interaction energy. Docking of the natural ligand to the receptor was performed to obtain a control interaction energy. The energy of the original PDB protein complex was also predicted as an experimental control. The exploration of these interactions with structural alignment visualization and protein-protein docking may help explain their potential roles in infection.

The potential mimics were split into two categories: endogenous vs. exogenous, or human vs. non-human, to more effectively describe the results in the context of infection. Mimicry of endogenous proteins may reveal which human pathways, specifically, the viral RBM is hijacking; structurally similar exogenous proteins may exhibit shared interference of human interaction pathways or antigenicity with the coronavirus RBMs. Endogenous hits, both discovered by single and multiple structural alignment tools, are summarized in Table 1 and exogenous hits in Table 2.

Endogenous

Several proteins containing EGF-like domains were found to be similar to all three RBMs. EGF-like domains are evolutionarily conserved domains that share homology to the epidermal growth factor and have been shown to function primarily in tissue organization and repair (Engel, 1989; Tombling et al., 2020). Both the cystine disulfide loop and the central beta-strand sub-motif structures in the SARS-CoV-2 and SARS-CoV RBMs and the MERS-CoV beta-strands were found to mimic EGF-like domains.

The EGF-like domain of the urokinase-type plasminogen activator (uPa) in complex with its receptor, urokinase plasminogen activator receptor (uPAR), (PDB: 2fd6) was found to be similar to both SARS-CoV-2 and SARS-CoV RBMs using RUPEE. Interestingly, the uPa/uPAR system has been implicated in SARS-CoV-2 pathogenesis with uPAR as an early predictor of severe respiratory failure (D'Alonzo et al., 2020; Rovina et al., 2020). Although the RBMs protrude into the receptor in the structural alignments, the alignments suggest that the RBMs might bind to uPAR (Supplementary Figure 2A).

The neurogenic locus notch homolog protein 1 (NOTCH1) EGF-like domain was returned for the SARS-CoV-2 RBM central beta-strands and MERS-CoV RBM by RUPEE and PDBeFold. NOTCH1 is involved in developmental, innate immunity, and inflammation signaling pathways, and natural ligands of the NOTCH1 EGF-like domains include jagged-1, jagged-2, delta-like 1 (DLL1), DLL3, and DLL4 (Shang et al., 2016). Alignment of the SARS-CoV-2 and MERS-CoV RBMs with the EGF-like domain of NOTCH1 bound to DLL4 (PDB: 4xl1) shows potential for molecular mimicry, i.e. the coronavirus RBMs may bind to DLL4 (Supplementary Figure 2B) (Luca et al., 2015). The SARS-CoV-2 RBM was also found similar to NOTCH2 by RUPEE, but no PDB complex models were available for further inspection. No direct interactions with the NOTCH1 pathway have been revealed, but its inhibition has been proposed to help fight SARS-CoV-2 infection (Rizzo et al., 2020).

All three RBMs were found to potentially mimic the EGF-like domain of coagulation factor VIIa. Further inspection of the alignment in complex with tissue factor (PDB: 1dan) showed potential for mimicry (Supplementary Figure 2C) (Banner et al., 1996). Interestingly, tissue factor expression has been shown to be up-regulated in severe SARS-CoV-2 infections, although there are several plausible theories (Bautista-Vargas et al., 2020; Eslamifar et al., 2020). The cystine disulfide loops of SARS-CoV and SARS-CoV-2 were found to resemble the EGF-like domains of coagulation factors X and IX and fibrillin, which are known to bind calcium (Handford et al., 1995; Stenflo et al., 2000). However, there is no evidence for calcium binding to the RBMs.

All three RBMs were found to mimic the EGF-like domain of thrombomodulin, specifically in the region that binds thrombin (PDB: 1dx5), by RUPEE and PDBeFold, while the SARS-CoV-2 similarity was also detected by HMI-PRED (Fuentes-Prior et al., 2000). Studies have shown that both thrombin and thrombomodulin blood concentrations are correlated with SARS-CoV-2 infection severity (Goshua et al., 2020; Ranucci et al., 2020). The verification by three tools and relevance to the literature led us to explore the potential mimicking of thrombomodulin binding to thrombin by the SARS-CoV-2 RBM using protein-protein docking (Figure 3A). Calculation of the interaction energies revealed that the reference docking and experimental controls showed similar affinities of -6.12 and -6.88 kJ/mol, and the SARS-CoV-2 RBM bound at a slightly lower affinity of -1.96 kJ/mol. The similarity to thrombomodulin might help explain the prothrombotic coagulopathy presented in SARS-CoV-2 infections (Bongiovanni et al., 2021).

The central beta-strands of the SARS-CoV-2 RBM were found to be structurally similar to the transforming growth factor alpha, epiregulin, and epigen EGF-like domains using RUPEE; however, alignment of the RBM with the proteins in complex with the epidermal growth factor receptor (EGFR) (PDBs: 1mox, 5wb7, 5wb8) showed that the RBM was just out of ligand-binding range (Supplementary Figure 2E) (Freed et al., 2017; Garrett et al., 2002). There is no evidence for interaction of the SARS-CoV-2 RBM with the extracellular domain of EGFR; however, the alignments were included for potential off-target effects related to the EGF-like domains.

Structural mimicry of chemokine and cytokine signaling has been reported for several viruses (Alcami, 2003). Viral proteins can mimic the chemokine, as in the case of HIV gp120 and CCL5, or they can mimic the receptor and bind directly to the cytokine (inhibiting its function), such as the vaccinia virus B15R protein that mimics the IL-1B receptor and binds to IL-1B (Ahuja and Murphy, 1999; Alcami and Smith, 1992).

Several cell signaling ligands and receptors were found similar to the coronavirus RBMs. The SARS-CoV-2 and SARS-CoV RBMs were both found to mimic IL-8 like chemokines, fibroblast growth factor 1, C-C motif chemokines 3, interleukin-18, and ephrins; they also individually mimic BMP2 and von Willebrand factor, respectively. The MERS-CoV RBM structurally resembled C-X-C motif chemokines 2 and 4 and growth/differentiation factor 5. The alignments of the RBMs with IL-8 like chemokines, C-C motif chemokines 2 (CCL2), 3, and 4, and IL-18 in complex with their respective receptors shows only partial alignment with the ligand-binding regions (Supplementary Figure 2F). Interestingly, however, expression levels of these cytokines have all been shown as correlating with SARS-CoV-2 infection, although other explanations have been proposed (Buszko et al., 2020; Chen et al., 2010a; Jamilloux et al., 2020; Younes et al., 2020). For example, IL-33 release by damaged lower respiratory cells during SARS-CoV-2 has been demonstrated to trigger inflammation, increasing CCL2 and CCL3 expression (Zizzo and Cohen, 2020).

Fibroblast growth factor 1 (FGF1) was shown to be similar to the SARS-CoV-2 and SARS-CoV RBMs using PDBeFold and HMI-PRED. A transcriptomic profiling revealed that FGF1 was upregulated in coronavirus infections (Alsamman and Zayed, 2020). Thus, to look more closely at

potential interference, we docked the SARS-CoV RBM with the fibroblast growth factor receptor 2 (FGFR2) (PDB: 3OJ2), which was predicted by HMI-PRED (Figure 3B) (Beenken et al., 2012). The RBM-FGFR2 docking analysis predicted a potentially favourable affinity of -2.49 kJ/mol, although not as high as the experimental complex (-9.78 kJ/mol) (Figure 3D). The FGF1 signaling pathway may, thus, be modulated by the coronavirus spike RBMs.

HMI-PRED predicted that the SARS-CoV-2 RBM mimics ephrin-A5 and ephrin-B2 binding to the ephrin type 4a receptor (EPHA4), and SARS-CoV mimics ephrin-A5 binding to the ephrin receptor type 3a. EPHA4 is unique among known class A ephrin receptors in that it binds both ephrin a and b ligands (Bowden et al., 2009; Xu et al., 2013). The structural similarity of two ligands for the same receptor for the SARS-CoV-2 RBM motivated further testing with protein-protein docking with EPHA4. Although there is no evidence for ephrin receptor involvement in coronavirus infections, other viral surface proteins have been shown to utilize ephrin receptors for cell entry, such as the rhesus r virus (Wang et al., 2020). The docking revealed similar affinities between the SARS-CoV-2 RBM-EPHA4, ephrin-A5-EPHA4, and experimental (PDB: 4m4r) complexes: -1.41, -3.32, -0.45 kJ/mol, respectively (Figure 3A).

The platelet glycoprotein Ib (GP-Ib) binding domain of von Willebrand factor (VWF) was found to be similar to SARS-CoV by HMI-PRED. VWF-GP-Ib interaction has been shown as critical in modulating thrombosis and inflammation (Denorme et al., 2019). Although there is no literature on VWF and SARS-CoV infection, blood concentration levels of VWF have been shown as correlated with SARS-CoV-2 infection severity, which may indicate potential pathway interference (Klok et al., 2020).

SARS-CoV-2 was predicted to mimic bone morphogenetic protein 2 binding to activin receptor type-2B and MERS-CoV to mimic growth/differentiation factor 5 (GDF5) binding to bone morphogenetic protein receptor type-1A (BMPRT1A) by HMI-PRED; however, no experimental evidence is available for either case. To explore the potential involvement of the MERS-CoV RBM in cell signaling, we docked the MERS-CoV RBM to the BMPRT1A in the GDF5-binding region (PDB: 3qb4) (Klammert et al., 2015). The docking of the RBM and GDF5 displayed similar affinities to BMPRT1A with -6.73 and -6.99 kJ/mol, respectively, while the experimental complex bound with -10.58 kJ/mol.

RUPEE detected structural resemblance between the SARS-CoV-2 RBM and IL-6 receptor alpha and beta chains, both of which show mimicry of the IL-6 binding sites (Supplementary Figure 2G). IL-6 has been reported as an overexpressed cytokine in SARS-CoV-2 infections, which can lead to induction of a hyper-innate inflammatory response (Chen et al., 2020; Magro, 2020). Mimicry of the IL-6 receptors by the RBM could result in binding and, thus, interference of IL-6 related interactions. However, several alternative theories have been proposed to explain the increases in IL-6 during severe infection; for example, the SARS-CoV nucleocapsid protein has been shown to activate IL-6 expression (Zhang et al., 2007). HMI-PRED additionally predicted MERS-CoV RBM mimicry of the binding of the T cell receptor beta chain to the major histocompatibility complex class I-related gene protein and interferon lambda receptor 1 binding to the beta subunit of the interleukin-10 receptor, both of which could have implications in immunosurveillance and inflammatory pathways (Corbett et al., 2014; Mendoza et al., 2017).

Different tumor necrosis factor-related ligands and receptors were found to be structurally analogous to the MERS-CoV RBM and SARS-CoV-2 RBM, respectively. Tumor necrosis factor receptor superfamily (TNFRSF) 1A, 4, 13C, and 14 were returned for the cystine disulfide loop for the SARS-CoV-2 RBM by RUPEE, while the tumor necrosis factor ligand superfamily (TNFSF) 13B and 14 were found to resemble the MERS-CoV RBM by RUPEE and HMI-PRED. Similarity of the SARS-CoV-2 RBM to TNFRSF 13C was also found by PDBeFold. These signaling pathways

normally promote B-cell and the T-cell survival and maturation (Tamada et al., 2000; Yu et al., 2000). The structural similarity of this family of ligands and receptors to the SARS-CoV-2 and MERS-CoV RBMs led us to further inspect the interactions with protein-protein docking: mimicry of SARS-CoV-2 to TNFRSF 13C and MERS-CoV to TNFSF 13B. Thus, we simulated the binding of the SARS-CoV-2 RBM to TNFSF 13B and MERS-CoV to TNFRSF 13C (PDB: 3v56) (Smart et al., 2012). Both cases revealed that the RBM is predicted to dock at a higher affinity than the natural ligand (Figure 3D).

The complement system comprises a series of protein cascades that form an integral part of the innate immune response to viruses (Nonaka and Yoshizaki, 2004). Viruses are generally susceptible to the complement system; however, viral proteins can utilize complement proteins through molecular mimicry in a variety of ways, such as using complement receptors for viral entry or evading detection by the immune system (Bernet et al., 2003). Infections from all three highly pathogenic coronaviruses have been reported to activate the complement system, enhancing pathogenicity, although the exact mechanisms remain unclear (Java et al., 2020). The spike protein of SARS-CoV-2 has been shown to localize near C4d and C5b-9 in lung vasculature, and mutations in several complement activation proteins, such as complement factors H, I, and III, have been found to correlate with infection severity (Magro et al., 2020; Ramlall et al., 2020). The structural similarity screen yielded three motifs from the complement system that potentially mimic RBMs: complement factor I (CFI) binding domain of CFH for all three RBMs and both the complement C3d binding domain of complement receptor 2 (CR2) and the complement C1r binding domain of complement C1s for the MERS-CoV RBM. Interestingly, CFH and the SARS-CoV-2 spike protein have been proposed to compete for heparan sulfate binding (Yu et al., 2020). The SARS-CoV RBM, however, was predicted to be similar to CFH by RUPEE, PDBeFold, and HMI-PRED; thus, we docked the SARS-CoV RBM to CFI (PDB: 5o32) and found that the natural ligand was predicted to bind at a slightly higher affinity than the SARS-CoV RBM: -1.60 vs. -0.14 kJ/mol, respectively (Xue et al., 2017). The C3d-binding domain of CR2 for the MERS-CoV RBM was also identified by RUPEE, PDBeFold, and HMI-PRED and was, thus, explored with docking of the MERS-CoV RBM to C3 (PDB: 3oed) (Figure 3C) (van den Elsen and Isenman, 2011). The MERS-CoV RBM was predicted to bind at a higher affinity than both the control docking and experimental complexes: -5.10 vs. -0.93 and -1.20 kJ/mol, respectively (Figure 3D). Additionally, HMI-PRED found that the MERS-CoV RBM also mimics the complement C1r binding site of complement C1s. Additional experimental efforts are needed to validate the relationship between coronavirus spike proteins and the complement activation pathway.

Other endogenous hits included several unrelated proteins, such as protease inhibitors and serotransferrin. The MERS-CoV RBM resembled the fibronectin type III (FNIII) domains of mouse myosin-binding protein C and tenascin-X using RUPEE. Although myosin-binding protein C is intracellular, FNIII domains are found across the domains of life and function in diverse ways, from cell adhesion to cell signaling (Campbell and Spitzfaden, 1994). Drayman *et al.* found that the West Nile virus envelope glycoprotein E resembles the structural architecture of the FN10 domain of fibronectin, which is a natural ligand for integrin αvβ3. Thus, we checked and found that the MERS-CoV RBM shares structural properties with other FNIII domains, such as those from fibronectin and neural cell adhesion molecule 1 (PDBs: 2haz and 1fnf, respectively) (Figure 3C) (Leahy et al., 1996; Mendiratta et al., 2006). The MERS-CoV RBM was also found to mimic part of the jagged-2, DLL1, and DLL4 proteins; however, the alignment was largely out of ligand-binding range when compared to jagged-1 in complex with NOTCH1 (PDB: 5uk5) – although the alignment may be relevant in other scenarios (Supplementary figure 2E) (Luca et al., 2017). Protease inhibitors included neuroserpin for the SARS-CoV-2 RBM and clitocypin-5 cysteine protease for the SARS-CoV RBM. The alignment of the SARS-CoV with clitocypin-5 cysteine protease showed potential binding to cathepsin L2 (PDB: 3h6s) (Renko et al., 2010). The role of cathepsins in coronavirus cell entry has been described as helping process the spike protein for viral and host

membrane fusion (Pišlar et al., 2020). To investigate the potential for additional interactions between coronavirus RBMs and cathepsins, we performed protein-protein docking. The binding of the SARS-CoV RBM to cathepsin L2 was predicted to be more favourable than the docking and experimental controls (Figure 3D). Experimental evidence is required to validate this interaction, however. Both the SARS-related RBMs resembled motifs of serotransferrin using TM-align, and, interestingly, the transferrin receptor protein 1 has been proposed as a potential cell entry receptor for SARS-CoV-2 (Tang et al., 2020). However, the alignments were generally out of ligand-binding range (PDB: 1suv) (Supplementary Figure 2H); since no binding mode was apparent, it was not considered for docking (Cheng et al., 2004). HMI-PRED predicted that the SARS-CoV-2 RBM mimics the dimerization domain of C-type lectin domain family 5 and that the SARS-CoV RBM mimics intercellular adhesion molecule 5 binding to integrin alpha-L (Watson et al., 2011; Zhang et al., 2008). Integrins have been proposed to bind to the SARS-CoV-2 spike protein, although that is due to a new RGD motif in the RBD – of note, the RGD motif is not included in the selected residues for this study's SARS-CoV-2 RBM since it does not interact with ACE2 in experimental models (Sigrist et al., 2020). Because integrin binding has not been hypothesized outside of the discussion of the SARS-CoV-2 RGD motif, docking was not pursued. Both the SARS-CoV-2 and SARS-CoV RBMs mimicked the nicotine-binding domain of the nicotinic acetylcholine receptor by RUPEE, which may have implications in the 'nicotinic hypothesis' (Changeux et al., 2020).

D Protein-protein docking analysis for selected coronavirus RBM interactions

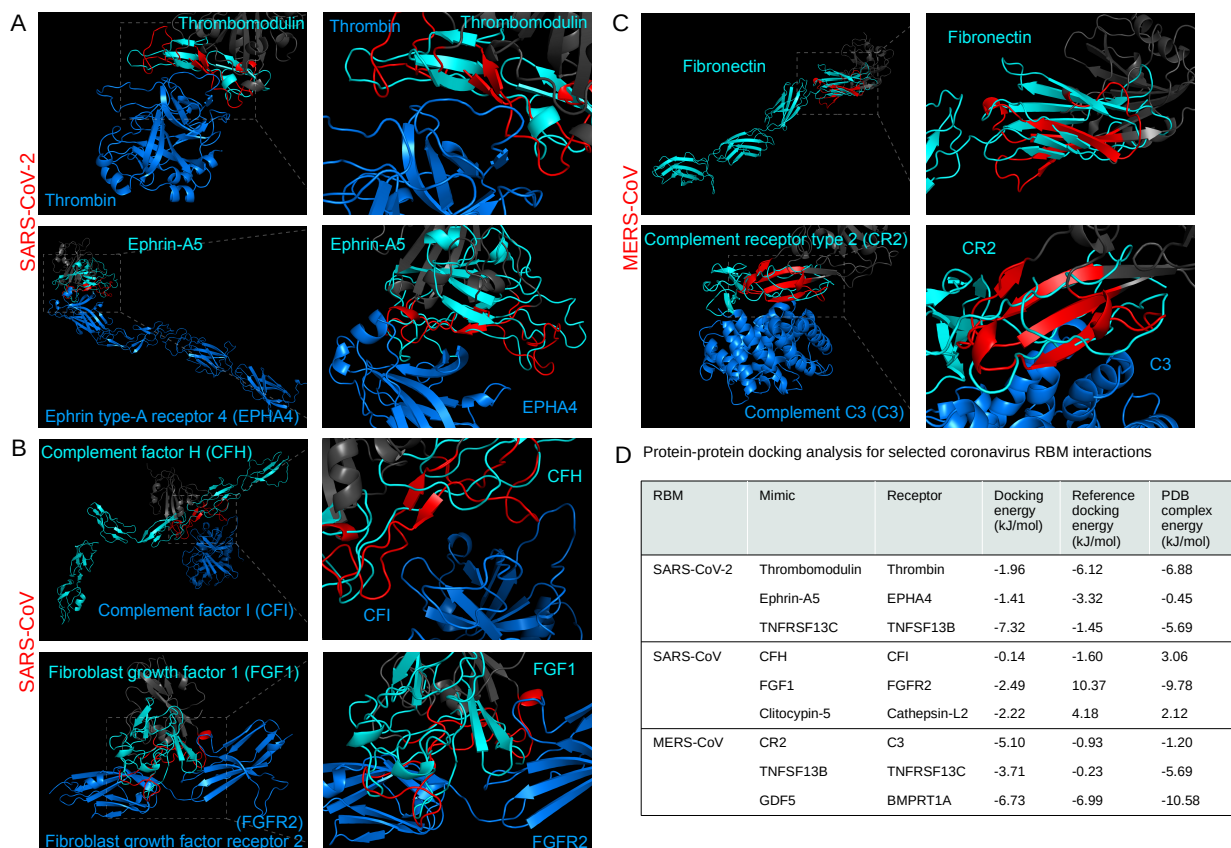| RBM	Mimic	Receptor	Docking energy (kJ/mol)	Reference docking energy (kJ/mol)	PDB complex energy (kJ/mol)
SARS-CoV-2	Thrombomodulin	Thrombin	-1.96	-6.12	-6.88
	Ephrin-A5	EPHA4	-1.41	-3.32	-0.45
	TNFRSF13C	TNFSF13B	-7.32	-1.45	-5.69
SARS-CoV	CFH	CFI	-0.14	-1.60	3.06
	FGF1	FGFR2	-2.49	10.37	-9.78
	Clitocypin-5	Cathepsin-L2	-2.22	4.18	2.12
MERS-CoV	CR2	C3	-5.10	-0.93	-1.20
	TNFSF13B	TNFRSF13C	-3.71	-0.23	-5.69
	GDF5	BMPRT1A	-6.73	-6.99	-10.58

Figure 3. Analysis of endogenous structural alignments
Protein-protein docking was performed using ClusPro PIPER to test the potential interactions between the coronavirus RBMs and potential interaction partners. The following alignments are shown: between the SARS-CoV-2 RBM and thrombin (A, top) and ephrin type-A receptor 4 (A, bottom), the SARS-CoV RBM and complement factor I (B, top) and fibroblast growth factor receptor 2 (B, bottom), and the MERS-CoV RBM with complement C3 (C, bottom). The MERS-CoV RBM aligned to fibronectin type III domain (C, top). RBMs are labelled red, the remainder of the RBDs are dark gray, mimicked proteins are cyan, and potential interaction partners are marine blue. Interaction energy scores predicted using FoldX on docked and experimental complexes (D).

Mimic Representative PDB			Mimic UniProt Name	Alignment Length	Alignment Sequence Identity (%)	TM-score	RMSD	Tools
MERS-CoV	SARS-CoV	SARS-CoV-2						
		1mox	Transforming growth factor alpha	25	4	0.25	3.53	R
1dx5	1zaq	1dqb	Thrombomodulin	21;20;25	4.8;5;4	0.31074;0.20501;0.2719	2.76;3.02;3.08	R,P and H (SARS-CoV-2)
5ky0		1toz	Neurogenic locus notch homolog protein 1	18;23	5.6;4.3	0.29958;0.26322	2.04;2.67	R,P
		5mwb	Neurogenic locus notch homolog protein 2	19	0	0.23683	2.81	R
1xka	2vwl	1fjs	Coagulation factor X	23;21	17.4;0	0.29888;0.24978	2.67;2.88	R,P
3th4	5ky2	1bf9	Coagulation factor VII	25;22	8;4.5	0.27601;0.25062	3.00;2.97	R,P
5f85		1edm	Coagulation factor IX	17;20	5.9;5	0.33731;0.25615	1.27;2.91	R,P
	3bt1	2fd6	Urokinase-type plasminogen activator	23	8.7	0.25522	3.13	R
		5wb7	Epiregulin	19;25	5.3;4	0.30636;0.23605	2.39;3.52	R
		5wb8	Epigen	22	13.6	0.34781	2.67	R
	1emn	1emn	Fibrillin	26;25;27	3.8;8;11.1	0.24425;0.26079	2.67;3.40;3.59	R
4hsv	1rhp	1nap	Platelet factor 4 (Platelet basic protein)	26;23;25	3.8;4.3;8	0.33673;0.19566;0.28516	2.67;3.77;2.97	R, H and P (SARS-CoV-2)
	2r3z	1o7z	Small inducible cytokine B10	27;28	0;7.1	0.28217;0.30758	3.38;2.9	R
	3oj2	3cu1	Fibroblast growth factor 1	37;21	5.4;14.3	0.29384;0.27446	3.54;2.74	H,P
	4c2b		von Willebrand factor	31	3.2	0.27715	3.53	H
	1b3e	1b3e	Serotransferrin	54;50	7.4;10	0.42052;0.41965	3.86;3.83	T
	4l0p	4m4r	Ephrin-A5	39;37	10.3;2.7	0.27482;0.31208	4.36;3.68	H
		2wo2	Ephrin-B2	34	8.8	0.3341	3.08	H
	1b53	1b50	C-C motif chemokine 3	26;22	0;0	0.24893;0.21513	3.13;2.97	R
		2h62	Bone morphogenetic protein 2	32	3.1	0.28716	3.69	H
	4xfu	4r6u	Interleukin-18	31;46	0;2.2	0.33910;0.34459	2.88;4.37	R
1qnk			C-X-C motif chemokine 2	27	7.4	0.2786	3.53	P
1je4			C-C motif chemokine 4	24	8.3	0.32333	3.12	P
3qb4			Growth/differentiation factor 5	39	17.9	0.35895	3.78	H
5mw5			Jagged-2	38	15.8	0.54617	2.56	T
4xbm			Delta-like protein 1	37	5.4	0.51041	2.32	T
5mvx			Delta-like protein 4	38	5.3	0.50604	2.47	T
		5fuc	Interleukin-6 receptor subunit alpha	36	0	0.34661	2.87	R
		1pvh	Interleukin-6 receptor beta chain	34	5.9	0.32838	3.51	R
4nqc			TCR beta chain	28	10.7	0.40641	2.42	H
5t5w			Interferon lambda receptor 1	27	7.4	0.37707	2.65	H
		2hey	Tumor necrosis factor receptor superfamily member 4	32	6.2	0.31459	3.33	R
		1ncf	Tumor necrosis factor receptor superfamily member 1A	24	4.2	0.29911	2.94	R
		2aw2	Tumor necrosis factor receptor superfamily member 14	25	12	0.23837	3.42	R
		1oqe	Tumor necrosis factor receptor superfamily member 13C	23	4.3	0.25858	3.29	R,P
3v56			Tumor necrosis factor ligand superfamily member	40	12.5	0.39446	3.31	H,R
4en0			Tumor necrosis factor ligand superfamily member	37	16.2	0.44547	2.86	H,R
1hfh	5o32	1hfi	Complement factor H	33;35;28	9.1;2.9;3.6	0.3487;0.31744;0.27510	3.39;3.82;3.57	R,P (no SARS-CoV),H (SARS-CoV)
3oed			Complement receptor type 2 (CR2)	31	6.5	0.3175	3.23	H,R,P
6f1c			Complement C1s	23	8.7	0.38981	2.02	H
1x5y			Fibronectin type-III domain of mouse myosin-binding protein C	28	14.3	0.42255	2.45	R
2cum			Fibronectin type III domain of human Tenascin-X	29	3.4	0.40251	2.8	R
		3f5n	Neuroserpin	41	12.2	0.36905	3.58	H
	3h6s		Clitocypin-5 cysteine proteinase	34	2.9	0.32861	3.45	H
		2yhf	C-type lectin domain family 5 member A	30	3.3	0.30137	3.42	H
	3bn3		Intercellular adhesion molecule 5	39	3.34	0.36204	3.34	H
	3sq9	3u8m	Acetylcholine-binding protein	38;38	13.2;10.5	0.35462;0.38651	3.69;3.14	R

Table 1. Structural alignment values and data for endogenous hits

Exogenous

We classified the exogenous hits by the pathogen type. There were motifs from apicomplexan parasites, viruses, one bacterial protein, and snake and spider toxins found to resemble the coronavirus RBMs.

The EGF-like domains from merozoite surface protein 1 (MSP1) of several *Plasmodium* species were found to be structurally similar to all three RBMs using RUPEE. Compared to the other two, the SARS-CoV-2 RBM was found to be similar to the most *Plasmodium* species: *falciparum, yoelii,*

cynomolgi, knowlesi, vivax. The SARS-CoV RBM returned *P. yoelii* MSP1 and the MERS-CoV RBM returned *P. falciparum* MSP1. A closer inspection at the *P. falciparum* MSP1 alignments revealed that two EGF-like domains on the same PDB structure (1ob1) were found to resemble the SARS-CoV-2 RBM (Figure 4A) (Pizarro et al., 2003). The PDB structure is originally modelling the antibody-binding epitope of the EGF-like domain of MSP1; however, the antibody epitope is located on a loop just outside of the EGF-like domain. Thus, antibody-binding to the SARS-CoV-2 RBM could not be verified, but the presence of two EGF-like domains near an epitope may motivate experimental testing. The *P. falciparum* apical membrane antigen 1 epitope (PDB: 2j5l) was also found to resemble the SARS-CoV-2 and SARS-CoV RBMs, although, as in the case of MSP1, both RBMs aligned to a region outside of the antibody-interacting residues (Figure 4B) (Igonet et al., 2007). These EGF-like domains from *Plasmodium* parasites may provide structural epitope scaffolding for cross-reactivity against the coronavirus spike RBMs (Craig et al., 1998). Recent studies have pointed to a potential protective effect of *P. falciparum* infections against SARS-CoV-2 infection, although direct experimental evidence is yet to be established (Iesa et al., 2020; Kalungi et al., 2021; Panda et al., 2020; Raham, 2021). The MERS-CoV RBM was also found to resemble the rhoptry neck protein 2 and thrombospondin-related anonymous protein from *P. falciparum*. The surface antigen 3 of *Toxoplasma gondii* was found to be similar to the MERS-CoV RBM (Figure 4C). Although there are no data on MERS-CoV and *T. gondii* co-infections, SARS-CoV-2 has been shown to have negative covariation with toxoplasmosis, which may indicate a protective effect from *T. gondii* (Jankowiak et al., 2020).

The coronavirus RBMs were found to structurally mimic several motifs on the HIV and Influenza spike proteins; however, they were found either facing inwards or buried inside the mimicked protein and were, therefore, discarded. PDBeFold and TM-align indicated that the SARS-CoV-2 and SARS-CoV RBMs structurally mimic several hepatitis C virus (HCV) antibody epitopes. The SARS-CoV-2 and SARS-CoV RBMs were found to be similar to 10 and 6 PDB HCV E2 protein epitopes structures, respectively (Supplementary Table 2). The HCV E2 protein is implicated in host entry, which has been explored as an inhibitory target with neutralizing antibodies (Heile et al., 2000; Ploss and Evans, 2012). A closer inspection of the mapping of the epitopes to the SARS-CoV-2 RBM show that they are distributed across the RBM (Figure 4D). Some studies have suggested that HCV may be negatively correlated with SARS-CoV-2 infection (Mirzaie et al., 2020; Reddy, 2020; Richardson et al., 2020). Since several of the epitopes were aligned in ways that were accessible to antibodies in the original PBD, we selected three epitopes, one at each region of the RBM, and docked the respective antibody to the SARS-CoV-2 RBM using the ClusPro PIPER 'antibody' mode (Figure 4D). As shown in Figure 4E, the RBM-antibody docking results were compared to docking and experimental controls – the antibodies bound in a similar way to docking controls in all three cases, while the experimental complexes were predicted to bind more tightly. These structural similarities may take part in potential cross-reactivity between HCV and coronavirus infections. Of note, two recently proposed cell entry receptors for the SARS-CoV-2 spike protein, ASGR1 and APOA4, have been shown as potentially implicated in mediating HCV viral entry (Brockbank et al., 2021). In an interesting case, the MERS-CoV RBM was found to structurally mimic both the Machupo virus glycoprotein polyprotein GP complex RBM (TM-score: 0.47) and its receptor, transferrin receptor protein 1 (0.52) (PDB: 3kas) using TM-align, although the transferrin receptor scores slightly higher (Supplementary Figure 3A) (Abraham et al., 2010).

Only one bacterial protein was selected in the structural similarity screen. The adhesin-binding fucosylated histo-blood group antigen of *Helicobacter pylori* was found to be similar to the MERS-CoV RBM by TM-align. The structure (PDB: 5f7l) shows binding of the bacterial protein to a nanobody; however, the RBM alignment is just outside of the nanobody binding site (Supplementary Figure 3B) (Moonens et al., 2016). No studies have detailed any connections between MERS-CoV and *H. pylori*.

Motifs from snake, spider, and cone snail toxins were found to be similar to all three RBMs using PDBeFold and RUPEE. The SARS-CoV-2 and SARS-CoV RBMs shared similarity to four toxins, and the MERS-CoV RBM only returned unique proteins. The two SARS-related viruses mimicked three-finger bungarotoxins and inhibitor cystine-knot toxins, such as psalmotoxin-1, while MERS-CoV RBM resembled other three-finger toxins, like cytotoxin 4 (Supplementary Figures 3C,3D) (Corfield et al., 1989; Dellisanti et al., 2007; Lee et al., 2014; Saez et al., 2011; Volpon et al., 2004). In total, these toxins may bind to several receptors involved in nociception, e.g. ASIC1 and Nav1.7, which may be relevant to the taste and pain perception changes experienced during SARS-CoV-2 infection (McFarland et al., 2021). Importantly, and perhaps confoundingly, a recent study found no changes in depolarization for Nav1.7 and Cav2.2 upon exposure to the SARS-CoV-2 RBD (Moutal et al., 2020). Thus, further experimental work is necessary to validate these interactions.

Protein-protein docking analysis for SARS-CoV-2 RBM and selected HCV E2 antibodies

HCV PDB	HCV E2 residues	Docking energy (kJ/mol)*	Reference docking energy (kJ/mol)*	PDB complex energy (kJ/mol)*
4g6a	412-423	-2.70 / -1.68	0.33 / 0.86	3.44 / -29.16
5fgc	412-423	2.02 / -1.54	-0.06 / -0.20	-5.65 / -25.83
5nph	532-540	4.28 / -1.38	-0.86 / -0.39	-3.23 / -41.16

*Antibody heavy chain / antibody light chain
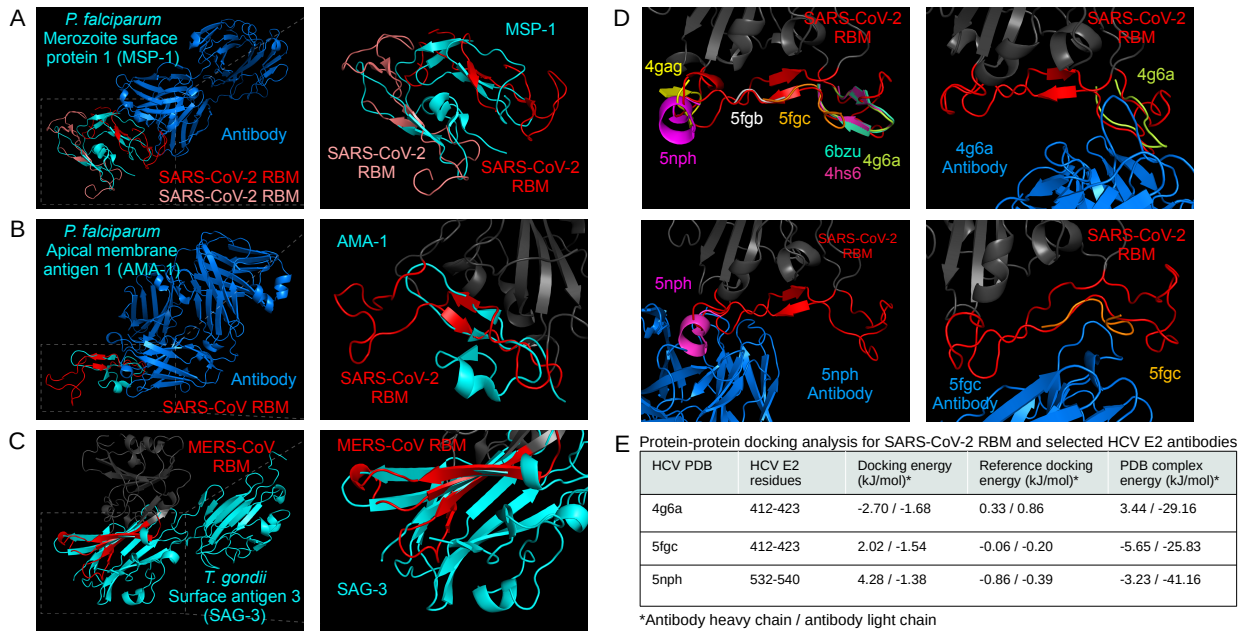
Figure 4. Analysis of exogenous structural alignments
Plasmodium falciparum merozoite surface protein 1 (A) and apical membrane antigen 1 (B) structurally aligned with the SARS-CoV-2 RBM and SARS-CoV RBM, respectively. The *Toxoplasma gondii* surface antigen 3 aligned with the MERS-CoV RBM (C). RBMs are labelled red, mimicked proteins are cyan, and potential interaction partners are marine blue (A,B,C). Hepatitis C virus epitopes structurally aligned to the SARS-CoV-2 RBM, and the respective antibody structures from PDBs 5fgc, 5nph, 4g6a docked to the RBM (D) using ClusPro PIPER "antibody" mode. Interaction energy scores predicted using FoldX on docked and experimental complexes (E).

Representative PDB									
MERS-CoV	SARS-CoV	SARS-CoV-2	Mimic UniProt Name	Species	Alignment Length	Alignment Sequence Identity (%)	TM-score	RMSD	Tools
		1n1i	Merozoite surface protein-1	Plasmodium knowlesi	22	4.5	0.25874	3.65	R
	2mgp	2mgp	Merozoite surface protein 1	Plasmodium yoelii	25;24	4;4.2	0.25287;0.27536	2.98;3.33	R
1cej		1ob1	Merozoite surface protein 1	Plasmodium falciparum	31;23	6.5;0	0.30361;0.24401	3.29;3.09	R
		1b9w	Merozoite surface protein 1	Plasmodium cynomolgi	24	4.2	0.26388	3.48	R
		2npr	Merozoite surface protein 1	Plasmodium vivax	24	4.2	0.24929	3.62	R
	1hn6	2j5l	Apical membrane antigen 1	Plasmodium falciparum	40;39	0;7.7	0.25857;0.26931	4.62;4.43	R,P
	1hky	1hky	Micronemal protein MIC4, related	Eimeria tenella	31;33	9.7;3	0.25771;0.31278	3.53;3.19	P
4yiz	4yiz	4yiz	Rhoptry neck protein 2, putative	Eimeria tenella	18;24;21	11.1;4.2;0	0.22386;0.28855;0.21409	2.79;3.29;3.36	P
2j4w			Apical membrane antigen 1	Plasmodium vivax	26	7.7	0.29687	2.84	R,P
2bbx			Thrombospondin-related anonymous protein	Plasmodium falciparum	25	4.8;13.5	0.30614	2.83	P
3sri			Rhoptry neck protein 2	Plasmodium falciparum	17	5.9	0.28299	2.39	P
5wa2			Surface antigen 3	Toxoplasma gondii	36	8.3	0.50524	3.02	T
	4xvj	4g6a	E2 envelope glycoprotein	Hepatitis C virus	10;12	10;2.9	0.45252;0.37438	1.73;3.82	T,P
3kas			Transferrin receptor protein 1 / glycoprotein polyprotein GP complex	Machupo virus	38	10.5	0.51761	2.52	T
5f7l			Adhesin binding fucosylated histo-blood group antigen	Helicobacter pylori	37	2.7	0.53448	2.77	T
		2a2v	Kappa-theraphotoxin-Cg1a 1	Chilobrachys guangxiensis	18	0	0.23611	2.46	R
	2kni	2kni	Psalmotoxin-1	Psalmopoeus cambridgei	21;22	14.3;13.6	0.25651;0.25962	2.89;2.79	R,P
		1oma	Omega-agatoxin-Aa4b	Agelenopsis aperta	25	4	0.24137	3.07	R
		1g1p	Delta-conotoxin EVIA	Conus ermineus	17	5.9	0.22626	3.01	R
	1qdp	1qdp	Delta-hexatoxin-Ar1a	Atrax robustus	26;23	7.7;0	0.25966;0.22613	3.76;3.7	R
		1la4	Kappa-theraphotoxin-Scg1a	Stromatopelma calceatum griseipes	22	13.6	0.22418	3.76	R
		2jtb	Hainantoxin-III 1	Haplopelma hainanum	18	5.6	0.22565	2.57	R
		1i26	Toxin Ptu1	Peirates turpis	21	9.5	0.25106	3.07	R
		1eit	Mu-agatoxin-Aa1a	Agelenopsis aperta	21	0	0.26268	2.73	R
		2mpq	Mu-theraphotoxin-Hd1a	Cyriopagopus doriae	18	5.6	0.27363	2.73	R
	1abt	2qc1	Alpha-bungarotoxin	Bungarus multicinctus	31;28	12.9;10.7	0.27217;0.33380	3.83;3.54	R
	4lft	4lft	Alpha-elapitoxin-Dpp2a	Dendroaspis polylepis polylepis	26	0	0.25472;0.32330	3.38;2.63	R
		2jqp	Weak toxin 1	Bungarus candidus	26;25	7.7;4	0.30834	3.36	R
		2nbt	Kappa-bungarotoxin	Bungarus multicinctus	28	3.6	0.27782	3.74	R
		2ctx	Long neurotoxin 3	Naja naja	28	3.6	0.32545	3.61	R
	1lxh		Long neurotoxin 1	Naja kaouthia	23	0	0.27252	2.85	R
	1txa		Long neurotoxin 2	Ophiophagus hannah	28	3.6	0.29226	3.28	R
	1c6w		Maurocalcin	Scorpio palmatus	18	0	0.27771	2.6	R
4om5			Cytotoxin 4	Naja atra	31	9.7	0.40943	2.71	R
4om4			Cytotoxin 2	Naja atra	29	10.3	0.40317	2.31	R
1onj			Cobrotoxin-b	Naja atra	30	10	0.38853	2.56	R
1era			Erabutoxin b	Laticauda semifasciata	34	17.6	0.3819	3.14	R
5ebx			Erabutoxin a	Laticauda semifasciata	36	16.7	0.38112	3.49	R
2mj4			Short neurotoxin 1	Naja oxiana	35	17.1	0.36641	3.33	R
1cod			Cobrotoxin homolog	Naja atra	34	11.8	0.35821	3.29	R
3hh7			Haditoxin	Ophiophagus hannah	34	8.8	0.37387	3.42	R
1jgk			Candoxin	Bungarus candidus	33	6.1	0.37726	3.47	R
1je9			Cobrotoxin-c	Naja kaouthia	31	16.1	0.34998	3.02	R

Table 2. Structural alignment values and data for exogenous hits

Conclusions

This study involved the structural bioinformatics characterization of potential molecular mimicry by highly pathogenic coronavirus spike protein RBMs. Using protein homology modelling, we built representative models of the spike RBMs and tested structural changes in the SARS-CoV-2 RBM induced by recently recorded mutations, which had little effect on overall RBM structure. Comparison of the RBMs revealed that the SARS-CoV and SARS-CoV-2 RBMs share higher structural homology than with MERS-CoV, which was underlined by the number of common returned proteins in the structural similarity screen using four structural alignment tools. The flexibility of the cystine disulfide loop in the SARS-related RBMs was found to permit large global changes in RBM structure; however, since most of the predicted mimicry was mapped to the RBM central beta-strands, which are quite rigid, the models of different conformations did not return significantly different proteins from the structural alignment tools. The structural alignment screen highlighted the similarity of the RBMs to evolutionarily unrelated human and non-human proteins. Further validation of the alignments with protein-protein docking revealed that all tested coronavirus RBM-endogenous protein interactions were predicted to be energetically favourable, confirming that the structural similarity screen may be useful in identifying potential molecular mimics.

The predicted endogenous mimicry comprised of proteins in cell signaling, adhesion, and complement pathways. Potential mimicry of several microbial antigenic proteins and exogenous

toxins was also discovered. The EGF-like domains of both endogenous and exogenous proteins structurally resemble all three RBMs. Predicted mimicked endogenous interactions include the EGF-like domain of thrombomodulin binding to thrombin, NOTCH1 binding to DLL4, and coagulation factor VIIa binding to tissue factor. Interference in these pathways may partially explain coagulopathies in coronavirus infections (Iba et al., 2020). Exogenous EGF-like domains of MSP1 from different *Plasmodium* species, on the other hand, may provide a structural epitope scaffold for cross-reactivity between coronavirus and *Plasmodium* infections (Panda et al., 2020). Epitope similarity was further explored among the several antibody-bound hepatitis C virus E2 protein motifs that were structurally analogous to the SARS-related RBMs. Structural similarity to antigenic proteins from other microbes may confer cross-immunity and, thus, also potentially guide vaccine design. Cell signaling pathway proteins, such as TNF-related and ephrin ligands, were also found as potential mimics of the coronavirus RBMs, which may lead to use of alternative co-receptors for viral entry or modulation of signaling cascades. Complement factor H was returned for all three RBMs and has also been implicated in coronavirus infections (Yu et al., 2020). The mimicry of complement proteins is widespread among viruses, and the spike RBM may have secondary roles interfering in these pathways (Mastellos et al., 2003). Many snake and spider toxins were also found similar to the coronavirus RBMs, which implies the potential usage of receptors involved in pain, muscle contraction, cell adhesion, and coagulation pathways (Dongol et al., 2019; Rowan, 2001; Wu et al., 2006). The prediction of evolutionarily unrelated, yet structurally similar, potential protein mimics reveals that previously unidentified pathways could be altered by the spike RBMs. The structural variation between coronavirus RBMs and their resulting molecular mimics can possibly be connected to differences in tropism, infection severity, and immune system reactivity between coronaviruses.

Although experimental verification of the predicted interactions is required to take these results further, the findings presented in this study provide insight into the potential molecular mimicry utilized by highly pathogenic coronavirus RBMs. The data can be used to support inhibitory drug, peptide, and antibody design efforts in order to prevent viral cell entry and virulence mechanisms related to coronavirus RBMs. Additional work is needed to better understand how coronaviruses co-opt host machinery to enhance fitness.

Acknowledgements

Author contributions

CAB performed analyses and wrote the paper; ARJ, AFA, and LC performed analyses and provided expertise on bioinformatics tools and statistical analysis; AvT, SH, BPB, and SET provided expertise on sequence and structural investigations; SCV helped acquire funding and provided expertise on the structural mutation analysis; PHMT and TLB supervised and provided overall direction for the study.

Declaration of interests

The authors declare no competing interests.

Methods

Spike receptor-binding motif model generation and characterization
Amino acid sequences of the SARS-CoV-2 (NCBI code: NC_045512), SARS-CoV (NC_004718), and MERS-CoV (NC_038294) spike proteins were extracted as FASTA files from the NCBI Viral Genomes Resource (Brister et al., 2015). Each amino acid sequence corresponds to one of three identical protomers of the full homo-oligomeric spike trimer. Due to the high number of available experimentally-resolved structures for each spike protein, representative models were generated using ProtCHOIR – a recently developed bioinformatic tool to automate 3D homology modelling of homo-oligomers (Torres and Blundell, manuscript in preparation https://doi.org/10.5281/zenodo.3384945). ProtCHOIR builds homo-oligomeric assemblies by searching for homolog templates on a locally created homo-oligomeric protein database using PSI-BLAST, performing a series of structural analyses on the input protomer structure or sequence using Molprobity, PISA, and GESAMT (all three tools as part of the CCP4 Molecular Graphics package), and comparative homology modelling using MODELLER (version 9.24) with molecular dynamics-level optimization and refinement (Altschul et al., 1997; Chen et al., 2010b; Krissinel, 2012; Krissinel and Henrick, 2007; Šali, 2019; Šali and Blundell, 1993).

The residues for the receptor-binding domains (RBD) and RBMs of each spike model were manually selected based on experimental structures with primary receptors (residues defined in Supplementary Table 1) and made into sub-structures during manual inspection of full-length models on PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). Global amino acid sequence alignment of RBDs was performed with EMBOSS Needle (Madeira et al., 2019). The full-length SARS-CoV-2 spike protein modelled with the lipid bilayer displayed in Figure 1 was retrieved from the SARS-CoV-2 3D database (Alsulami et al., 2021). No records exist of N-linked or O-linked glycosylation motifs near the three RBMs, which was supported by NetNGlyc 1.0 and NetOGlyc 4.0 predictions (Gupta et al., 2004; Steentoft et al., 2013). To determine the flexibility of each residue in the RBDs, we used CABS-flex 2.0, a web server that offers fast simulations and resulting data of protein structure flexibility (Kuriata et al., 2018). Default values were used and residue flexibility was reported as root mean squared flexibility (RMSF) (Kmiecik et al., 2016).

Structure similarity screen
Several web servers and stand-alone tools have become available to perform pairwise or multiple sequence-independent protein structure alignments, such as DALI, FATCAT, iSARST, MADOKA, PDBeFold, TM-align, and RUPEE (Ayoub and Lee, 2019; Deng et al., 2019; Holm and Laakso, 2016; Krissinel and Henrick, 2004; Li et al., 2020; Lo et al., 2009; Zhang and Skolnick, 2005). After testing each tool, the PDBeFold web server, RUPEE web server, and a locally-installed

version of TM-align were selected due to the diversity of structural alignment methodologies, ease-of-use, data accessibility, and widespread-usage. A newly published web server for structural prediction of host-microbe interactions based on interface mimicry, HMI-PRED, was also included in the analysis.

Of note, mTM-align (the web server version of TM-align) was considered, but no non-spike proteins were shown – restricting the downstream analysis (Dong et al., 2018). Thus, all 3D models in the PDB database clustered at 100% sequence identity were downloaded, and TM-align was run in a pairwise manner, using GNU parallel, between each RBM model and each chain of every downloaded PDB file (O. Tange (2018): GNU Parallel, March 2018, https://doi.org/10.5281/zenodo.1146014.). TM-align works by, first, combining secondary structure similarity alignments, defined by DSSP (Define Secondary Structure of Proteins), and TM-score-based structural alignments (Kabsch and Sander, 1983; Zhang and Skolnick, 2004). A structure rotation matrix is applied to the alignments in order to maximize the TM-score, which was used to rank the alignments for each RBM.

The PDBeFold web server utilizes SSM, a graph-matching algorithm that superimposes *PROMOTIF*-defined secondary structures and, subsequently, maps backbone carbon atoms of, first, matched and, second, unmatched secondary structures (Hutchinson and Thornton, 1996; Krissinel and Henrick, 2004, 2005). The hits are ranked by their Q-Score, which is calculated to achieve a lower root mean squared deviation (RMSD) and an increased number of aligned residues. Since the highest percentage (%) of secondary structure matches for the SARS-CoV RBM was found at 67% (while SARS-CoV-2 and MERS-CoV returned hits with 100% secondary structure matches), we set the PDBeFold search parameters for 65% structural similarity at "highest precision" for each RBM.

The RUPEE web server performs structural similarity comparisons using a purely geometric approach: 1) a linear encoding of the protein structure is defined to identify separable regions of permissible torsion angles for DSSP secondary structure assignments; 2) the encoding is converted into a bag of features; 3) a protein structure indexing method is established using min-hashing and locality sensitive hashing; 4) the top 8,000 matches are sorted based on adjusted Jaccard similarity scores; and 5) if running in "Top-Aligned" mode (used in this analysis), the alignments are re-scored using TM-align (Ayoub and Lee, 2019). RUPEE allows the specific comparison of a query protein to the CATH, SCOP, PDB, and ECOD databases (Andreeva et al., 2020; Burley et al., 2019; Cheng et al., 2014; Sillitoe et al., 2019). Additional settings are also offered, such as the "contains" (finding query protein inside database protein) and "contained in" options (small protein motif detection in query protein) – both of which were used in this analysis.

HMI-PRED combines TM-align with NACCESS to search through template host protein-protein complexes (Guven-Maiorov et al., 2020; Hubbard and Thornton, 1993). The structural alignment model of the ligand and putative receptor are refined with RosettaDock to quantify electrochemical complementarity, which is not included in a strict structural alignment screen (Wang et al., 2007). Alignments for both the RBM and RBD of each of the three coronaviruses were collected from HMI-PRED.

Since the flexibility of the cystine disulfide loop on the SARS-CoV-2 and SARS-CoV RBMs may affect the global structure of the RBMs and, thus, search outcome, we used two additional models provided by CABS-flex 2.0 for both RBMs, making a total of three conformations for each SARS-CoV-2 and SARS-CoV. Results of the different conformations provided were pooled together for both SARS-related RBMs. The top-scoring alignments from all four tools for each RBM were matched with their corresponding PDB and, subsequently, UniProt accession code (Bateman, 2019). The UniProt accession codes were then compared across tools to identify shared top hits.

Protein-protein docking and interaction energy prediction
A local installation of ClusPro PIPER (version 1.1.5) was used for protein-protein docking (Desta et al., 2020; Kozakov et al., 2013, 2017; Vajda et al., 2017). Annotations informing potential protein-protein interactions were obtained from the PDB, STRING, and UniProt databases (Szklarczyk et al., 2019). The ClusPro PIPER "antibody" docking mode was used to dock RBDs with the hepatitis C antibody PDB structures (PDBs: 4g6a, 5fgc, 5nph), and the "others" mode was used for all non-antibody docking. In order to minimize non-biologically relevant binding during the docking runs, residues outside of the RBM on the spike RBDs and outside of the ligand-binding region on the predicted interaction partners were masked. The docked models were minimized using CHARMM22. To gain better insight into the binding strength of the potential RBD-receptor complex interactions, we used the FoldX (version 4.0) AnalyseComplex program, which predicts the interaction energy by finding the difference in stability between the individual unfolded molecules and the overall complex (Schymkowitz et al., 2005). The original PDB ligands were also docked to the receptor in order to obtain a "Reference docking energy" when compared with the predicted RBD-receptor energy. The "PDB complex energy" was obtained from the original PDB containing the ligand and receptor to understand binding resolution of the experimental complex.

Data analysis and visualization
A full representation of the pipeline and tools used can be found in Figure 5. Data were analyzed and plots were generated using R version 3.6.3 (2020-02-29). Protein structural alignments were visualized with PyMOL (version 1.8.4.0) (Stout, 2004). Pdb-tools was used to manipulate and organize PDB files (Rodrigues et al., 2018). The graphical abstract was adapted from the "SARS-CoV-2 Spike Protein Conformations" template on BioRender. Figure 5 was created on draw.io. Raw data and alignment models are made available at https://github.com/tlb-lab.
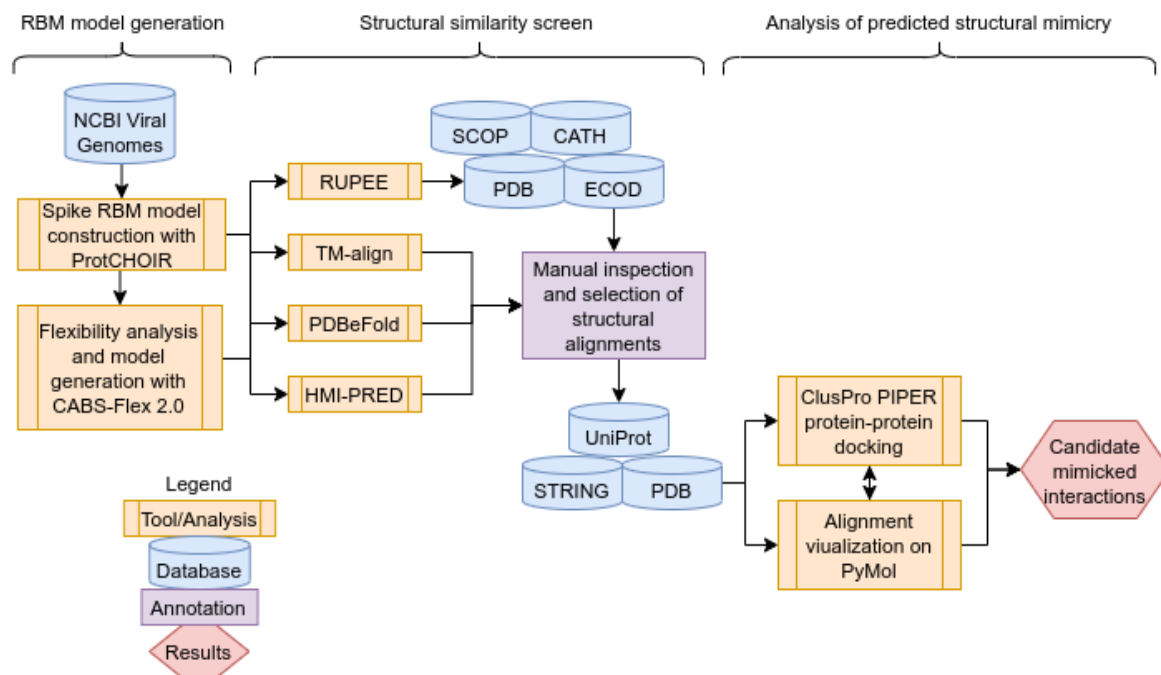

Figure 5. Pipeline flow
A flow chart of the analyses performed in this study.

References

Abraham, J., Corbett, K.D., Farzan, M., Choe, H., and Harrison, S.C. (2010). Structural basis for receptor recognition by New World hemorrhagic fever arenaviruses. Nat. Struct. Mol. Biol. *17*, 438–444.

Ahuja, S.K., and Murphy, P.M. (1999). Viral Mimicry of Chemokines and Chemokine Receptors BT - Chemokines in Disease: Biology and Clinical Research. C.A. Hébert, ed. (Totowa, NJ: Humana Press), pp. 235–251.

Alcami, A. (2003). Viral mimicry of cytokines, chemokines and their receptors. Nat. Rev. Immunol. *3*, 36–50.

Alcami, A., and Smith, G.L. (1992). A soluble receptor for interleukin-1β encoded by vaccinia virus: A novel mechanism of virus modulation of the host response to infection. Cell *71*, 153–167.

Alsamman, A.M., and Zayed, H. (2020). The transcriptomic profiling of SARS-CoV-2 compared to SARS, MERS, EBOV, and H1N1. PLoS One *15*, e0243270.

Alsulami, A.F., Thomas, S.E., Jamasb, A.R., Beaudoin, C.A., Moghul, I., Bannerman, B., Copoiu, L., Vedithi, S.C., Torres, P., and Blundell, T.L. (2021). SARS-CoV-2 3D database: understanding the coronavirus proteome and evaluating possible drug targets. Brief. Bioinform. *22*, 769–780.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.

An, H., and Park, J. (2020). Molecular Mimicry Map (3M) of SARS-CoV-2: Prediction of potentially immunopathogenic SARS-CoV-2 epitopes via a novel immunoinformatic approach. BioRxiv 2020.11.12.344424.

Andreeva, A., Kulesha, E., Gough, J., and Murzin, A.G. (2020). The SCOP database in 2020: Expanded classification of representative family and superfamily domains of known protein structures. Nucleic Acids Res. *48*, D376–D382.

Angileri, F., Legare, S., Marino Gammazza, A., Conway de Macario, E., JL Macario, A., and Cappello, F. (2020a). Molecular mimicry may explain multi-organ damage in COVID-19. Autoimmun. Rev. *19*, 102591.

Angileri, F., Légaré, S., Marino Gammazza, A., Conway de Macario, E., Macario, A.J.L., and Cappello, F. (2020b). Is molecular mimicry the culprit in the autoimmune haemolytic anaemia affecting patients with COVID-19? Br. J. Haematol. *190*, e92–e93.

Ayoub, R., and Lee, Y. (2019). Rupee: A fast and accurate purely geometric protein structure search. PLoS One *14*, 1–17.

Banner, D.W., D'Arcy, A., Chène, C., Winkler, F.K., Guha, A., Konigsberg, W.H., Nemerson, Y., and Kirchhofer, D. (1996). The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. Nature *380*, 41–46.

Bateman, A. (2019). UniProt: A worldwide hub of protein knowledge. Nucleic Acids Res. *47*, D506–D515.

Bautista-Vargas, M., Bonilla-Abadía, F., and Cañas, C.A. (2020). Potential role for tissue factor in the pathogenesis of hypercoagulability associated with in COVID-19. J. Thromb. Thrombolysis *50*, 479–483.

Beenken, A., Eliseenkova, A. V, Ibrahimi, O.A., Olsen, S.K., and Mohammadi, M. (2012). Plasticity in Interactions of Fibroblast Growth Factor 1 (FGF1) N Terminus with FGF Receptors Underlies Promiscuity of FGF1. J. Biol. Chem. *287*, 3067–3078.

Bernet, J., Mullick, J., Singh, A.K., and Sahu, A. (2003). Viral mimicry of the complement system. J Biosci. *28*, 249–264.

Bongiovanni, D., Klug, M., Lazareva, O., Weidlich, S., Biasi, M., Ursu, S., Warth, S., Buske, C., Lukas, M., Spinner, C.D., et al. (2021). SARS-CoV-2 infection is associated with a pro-thrombotic platelet phenotype. Cell Death Dis. *12*, 50.

Bowden, T.A., Aricescu, A.R., Nettleship, J.E., Siebold, C., Rahman-Huq, N., Owens, R.J., Stuart, D.I., and Jones, E.Y. (2009). Structural Plasticity of Eph Receptor A4 Facilitates Cross-Class Ephrin Signaling. Structure *17*, 1386–1397.

Brister, J.R., Ako-Adjei, D., Bao, Y., and Blinkova, O. (2015). NCBI viral Genomes resource. Nucleic Acids Res. *43*, D571–D577.

Brockbank, S.M. V, Soden, J., Faba-Rodriguez, R., Ribeiro, L.R., Geh, C., Thomas, H., Delight, J., Coverley, L., Abbott, W.M., and Freeth, J. (2021). SARS-CoV-2 comprehensive receptor profiling: mechanistic insight to drive new therapeutic strategies. BioRxiv 2021.03.11.434937.

Burley, S.K., Berman, H.M., Bhikadiya, C., Bi, C., Chen, L., Di Costanzo, L., Christie, C., Dalenberg, K., Duarte, J.M., Dutta, S., et al. (2019). RCSB Protein Data Bank: Biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. Nucleic Acids Res. *47*, D464–D474.

Buszko, M., Park, J.-H., Verthelyi, D., Sen, R., Young, H.A., and Rosenberg, A.S. (2020). The dynamic changes in cytokine responses in COVID-19: a snapshot of the current state of knowledge. Nat. Immunol. *21*, 1146–1151.

Buzhdygan, T.P., DeOre, B.J., Baldwin-Leclair, A., Bullock, T.A., McGary, H.M., Khan, J.A., Razmpour, R., Hale, J.F., Galie, P.A., Potula, R., et al. (2020). The SARS-CoV-2 spike protein alters barrier function in 2D static and 3D microfluidic in-vitro models of the human blood–brain barrier. Neurobiol. Dis. *146*, 105131.

Campbell, I.D., and Spitzfaden, C. (1994). Building proteins with fibronectin type III modules. Structure *2*, 333–337.

Changeux, J.-P., Amoura, Z., Rey, F.A., and Miyara, M. (2020). A nicotinic hypothesis for Covid-19 with preventive and therapeutic implications. C. R. Biol. *343*, 33–39.

Chen, I.-Y., Chang, S.C., Wu, H.-Y., Yu, T.-C., Wei, W.-C., Lin, S., Chien, C.-L., and Chang, M.-F. (2010a). Upregulation of the Chemokine (C-C Motif) Ligand 2 via a Severe Acute Respiratory Syndrome Coronavirus Spike-ACE2 Signaling Pathway. J. Virol. *84*, 7703–7712.

Chen, L.Y.C., Hoiland, R.L., Stukas, S., Wellington, C.L., and Sekhon, M.S. (2020). Confronting the controversy: Interleukin-6 and the COVID-19 cytokine storm syndrome. Eur. Respir. J. 2003006.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010b). MolProbity: All-atom structure validation for macromolecular crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 12–21.

Cheng, H., Schaeffer, R.D., Liao, Y., Kinch, L.N., Pei, J., Shi, S., Kim, B.H., and Grishin, N. V. (2014). ECOD: An Evolutionary Classification of Protein Domains. PLoS Comput. Biol. *10*.

Cheng, Y., Zak, O., Aisen, P., Harrison, S.C., and Walz, T. (2004). Structure of the Human Transferrin Receptor-Transferrin Complex. Cell *116*, 565–576.

Cheong, K.X. (2020). Systematic Review of Ocular Involvement of SARS-CoV-2 in Coronavirus Disease 2019. Curr. Ophthalmol. Rep. *8*, 185–194.

Chew, F.T., Ong, S.Y., and Hew, C.L. (2003). Severe acute respiratory syndrome coronavirus and viral mimicry. Lancet *361*, 2081.

Corbett, A.J., Eckle, S.B.G., Birkinshaw, R.W., Liu, L., Patel, O., Mahony, J., Chen, Z., Reantragoon, R., Meehan, B., Cao, H., et al. (2014). T-cell activation by transitory neo-antigens derived from distinct microbial pathways. Nature *509*, 361–365.

Corfield, P.W., Lee, T.J., and Low, B.W. (1989). The crystal structure of erabutoxin a at 2.0-A resolution. J. Biol. Chem. *264*, 9239–9242.

Craig, L., Sanschagrin, P.C., Rozek, A., Lackie, S., Kuhn, L.A., and Scott, J.K. (1998). The role of structure in antibody cross-reactivity between peptides and folded proteins. J. Mol. Biol. *281*, 183–201.

D'Alonzo, D., De Fenza, M., and Pavone, V. (2020). COVID-19 and pneumonia: a role for the uPA/uPAR system. Drug Discov. Today *25*, 1528–1534.

Dellisanti, C.D., Yao, Y., Stroud, J.C., Wang, Z.-Z., and Chen, L. (2007). Crystal structure of the extracellular domain of nAChR alpha1 bound to alpha-bungarotoxin at 1.94 A resolution. Nat. Neurosci. *10*, 953–962.

Deng, L., Zhong, G., Liu, C., Luo, J., and Liu, H. (2019). MADOKA: An ultra-fast approach for large-scale protein structure similarity searching. BMC Bioinformatics *20*, 1–10.

Denorme, F., Vanhoorelbeke, K., and De Meyer, S.F. (2019). von Willebrand Factor and Platelet Glycoprotein Ib: A Thromboinflammatory Axis in Stroke. Front. Immunol. *10*, 2884.

Desta, I.T., Porter, K.A., Xia, B., Kozakov, D., and Vajda, S. (2020). Performance and Its Limits in Rigid Body Protein-Protein Docking. Structure *28*, 1071-1081.e3.

Dong, R., Pan, S., Peng, Z., Zhang, Y., and Yang, J. (2018). MTM-align: A server for fast protein structure database search and multiple protein structure alignment. Nucleic Acids Res. *46*, W380–W386.

Dongol, Y., Cardoso, F.C., and Lewis, R.J. (2019). Spider Knottin Pharmacology at Voltage-Gated Sodium Channels and Their Potential to Modulate Pain Pathways. Toxins (Basel). *11*.

Doolittle, J.M., and Gomez, S.M. (2010). Structural similarity-based predictions of protein interactions between HIV-1 and Homo sapiens. Virol. J. *7*, 82.

Drayman, N., Glick, Y., Ben-Nun-Shaul, O., Zer, H., Zlotnick, A., Gerber, D., Schueler-Furman, O., and Oppenheim, A. (2013). Pathogens use structural mimicry of native host ligands as a mechanism for host receptor engagement. Cell Host Microbe *14*, 63–73.

Elde, N.C., and Malik, H.S. (2009). The evolutionary conundrum of pathogen mimicry. Nat. Rev. Microbiol. *7*, 787–797.

van den Elsen, J.M.H., and Isenman, D.E. (2011). A Crystal Structure of the Complex Between Human Complement Receptor 2 and Its Ligand C3d. Science (80-.). *332*, 608–611.

Engel, J. (1989). EGF-like domains in extracellular matrix proteins: Localized signals for growth and differentiation? FEBS Lett. *251*, 1–7.

Eslamifar, Z., Behzadifard, M., Soleimani, M., and Behzadifard, S. (2020). Coagulation abnormalities in SARS-CoV-2 infection: overexpression tissue factor. Thromb. J. *18*, 38.

Freed, D.M., Bessman, N.J., Kiyatkin, A., Salazar-Cavazos, E., Byrne, P.O., Moore, J.O., Valley, C.C., Ferguson, K.M., Leahy, D.J., Lidke, D.S., et al. (2017). EGFR Ligands Differentially Stabilize Receptor Dimers to Specify Signaling Kinetics. Cell *171*, 683-695.e18.

Fuentes-Prior, P., Iwanaga, Y., Huber, R., Pagila, R., Rumennik, G., Seto, M., Morser, J., Light, D.R., and Bode, W. (2000). Structural basis for the anticoagulant activity of the thrombin–thrombomodulin complex. Nature *404*, 518–525.

Garrett, T.P.J., McKern, N.M., Lou, M., Elleman, T.C., Adams, T.E., Lovrecz, G.O., Zhu, H.-J., Walker, F., Frenkel, M.J., Hoyne, P.A., et al. (2002). Crystal Structure of a Truncated Epidermal Growth Factor Receptor Extracellular Domain Bound to Transforming Growth Factor. Cell *110*, 763–773.

Goshua, G., Pine, A.B., Meizlish, M.L., Chang, C.H., Zhang, H., Bahel, P., Baluha, A., Bar, N., Bona, R.D., Burns, A.J., et al. (2020). Endotheliopathy in COVID-19-associated coagulopathy: evidence from a single-centre, cross-sectional study. Lancet Haematol. *7*, e575–e582.

Grifoni, A., Sidney, J., Zhang, Y., Scheuermann, R.H., Peters, B., and Sette, A. (2020). A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. Cell Host Microbe *27*, 671-680.e2.

Gu, Y., Cao, J., Zhang, X., Gao, H., Wang, Y., Wang, J., Zhang, J., Shen, G., Jiang, X., Yang, J., et al. (2020). Interaction network of SARS-CoV-2 with host receptome through spike protein. BioRxiv 2020.09.09.287508.

Gupta, R., Jung, E., and Brunak, S. (2004). Prediction of N-glycosylation sites in human proteins. In preparation.

Guven-Maiorov, E., Hakouz, A., Valjevac, S., Keskin, O., Tsai, C.J., Gursoy, A., and Nussinov, R. (2020). HMI-PRED: A Web Server for Structural Prediction of Host-Microbe Interactions Based on Interface Mimicry. J. Mol. Biol. *432*, 3395–3403.

Handford, P., Downing, A.K., Rao, Z., Hewett, D.R., Sykes, B.C., and Kielty, C.M. (1995). The calcium binding properties and molecular organization of epidermal growth factor-like domains in human fibrillin-1. J. Biol. Chem. *270*, 6751–6756.

Heile, J.M., Fong, Y.-L., Rosa, D., Berger, K., Saletti, G., Campagnoli, S., Bensi, G., Capo, S., Coates, S., Crawford, K., et al. (2000). Evaluation of Hepatitis C Virus Glycoprotein E2 for Vaccine Design: an Endoplasmic Reticulum-Retained Recombinant Protein Is Superior to Secreted Recombinant Protein and DNA-Based Vaccine Candidates. J. Virol. *74*, 6885 LP – 6892.

Holm, L., and Laakso, L.M. (2016). Dali server update. Nucleic Acids Res. *44*, W351–W355.

Huang, A.T., Garcia-Carreras, B., Hitchings, M.D.T., Yang, B., Katzelnick, L.C., Rattigan, S.M., Borgert, B.A., Moreno, C.A., Solomon, B.D., Trimmer-Smith, L., et al. (2020). A systematic review of antibody mediated immunity to coronaviruses: kinetics, correlates of protection, and association with severity. Nat. Commun. *11*, 4704.

Hubbard, S., and Thornton, J. (1993). NACCESS, Department of Biochemistry Molecular Biology, University College London.

Hussain, A., Hasan, A., Nejadi Babadaei, M.M., Bloukh, S.H., Chowdhury, M.E.H., Sharifi, M., Haghighat, S., and Falahati, M. (2020). Targeting SARS-CoV2 Spike Protein Receptor Binding Domain by Therapeutic Antibodies. Biomed. Pharmacother. *130*, 110559.

Hutchinson, E.G., and Thornton, J.M. (1996). PROMOTIF--a program to identify and analyze structural motifs in proteins. Protein Sci. *5*, 212–220.

Iba, T., Levy, J.H., Levi, M., Connors, J.M., and Thachil, J. (2020). Coagulopathy of Coronavirus Disease 2019. Crit. Care Med. *48*, 1358–1364.

Ichimura, T., Mori, Y., Aschauer, P., Padmanabha Das, K.M., Padera, R.F., Weins, A., Nasr, M.L., and Bonventre, J. V (2020). KIM-1/TIM-1 is a Receptor for SARS-CoV-2 in Lung and Kidney. MedRxiv 2020.09.16.20190694.

Iesa, M.A.M., Osman, M.E.M., Hassan, M.A., Dirar, A.I.A., Abuzeid, N., Mancuso, J.J., Pandey, R., Mohammed, A.A., Borad, M.J., Babiker, H.M., et al. (2020). SARS-CoV-2 and Plasmodium falciparum common immunodominant regions may explain low COVID-19 incidence in the malaria-endemic belt. New Microbes New Infect. *38*, 100817.

Igonet, S., Vulliez-Le Normand, B., Faure, G., Riottot, M.-M., Kocken, C.H.M., Thomas, A.W., and Bentley, G.A. (2007). Cross-reactivity Studies of an Anti-Plasmodium vivax Apical Membrane Antigen 1 Monoclonal Antibody: Binding and Structural Characterisation. J. Mol. Biol. *366*, 1523–1537.

Jamilloux, Y., Henry, T., Belot, A., Viel, S., Fauter, M., El, T., Walzer, T., François, B., and Sève, P. (2020). Should we suppress or stimulate immune responses for covid-19. Autoimmun Rev. *19*.

Jankowiak, Ł., Rozsa, L., Tryjanowski, P., and Møller, A.P. (2020). A negative covariation between toxoplasmosis and CoVID-19 with alternative interpretations. Sci. Rep. *10*, 12512.

Java, A., Apicelli, A.J., Kathryn Liszewski, M., Coler-Reilly, A., Atkinson, J.P., Kim, A.H.J., and Kulkarni, H.S. (2020). The complement system in COVID-19: Friend and foe? JCI Insight *5*, 1–12.

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers *22*, 2577–2637.

Kalungi, A., Kinyanda, E., Akena, D.H., Kaleebu, P., and Bisangwa, I.M. (2021). Less Severe Cases of COVID-19 in Sub-Saharan Africa: Could Co-infection or a Recent History of Plasmodium falciparum Infection Be Protective? Front. Immunol. *12*, 360.

Kanduc, D., and Shoenfeld, Y. (2020). Molecular mimicry between SARS-CoV-2 spike glycoprotein and mammalian proteomes: implications for the vaccine. Immunol. Res. *68*, 310–313.

Klammert, U., Mueller, T.D., Hellmann, T. V, Wuerzler, K.K., Kotzsch, A., Schliermann, A., Schmitz, W., Kuebler, A.C., Sebald, W., and Nickel, J. (2015). GDF-5 can act as a context-dependent BMP-2 antagonist. BMC Biol. *13*, 77.

Klok, F.A., Kruip, M.J.H.A., van der Meer, N.J.M., Arbous, M.S., Gommers, D.A.M.P.J., Kant, K.M., Kaptein, F.H.J., van Paassen, J., Stals, M.A.M., Huisman, M. V., et al. (2020). Incidence of thrombotic complications in critically ill ICU patients with COVID-19. Thromb. Res. *191*, 145–147.

Kmiecik, S., Gront, D., Kolinski, M., Wieteska, L., Dawid, A.E., and Kolinski, A. (2016). Coarse-Grained Protein Models and Their Applications. Chem. Rev. *116*, 7898–7936.

Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S.E., Xia, B., Hall, D.R., and Vajda, S. (2013). How Good is Automated Protein Docking? Proteins *81*, 2159–2166.

Kozakov, D., Hallc, D.R., Xiab, B., Porterb, K.A., Padhornya, D., Yuehb, C., Beglovb, D., and Vajda, S. (2017). The ClusPro web server for protein-protein docking. Nat. Protoc. *12*, 255–278.

Krishna, S.S., and Grishin, N. V (2004). Structurally Analogous Proteins Do Exist! Structure *12*, 1125–1127.

Krissinel, E. (2012). Enhanced fold recognition using efficient short fragment clustering. J Mol Biochem *1*, 76–85.

Krissinel, E., and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. Sect. D Biol. Crystallogr. *60*, 2256–2268.

Krissinel, E., and Henrick, K. (2005). Multiple Alignment of Protein Structures in Three Dimensions BT - Computational Life Sciences. M. R. Berthold, R.C. Glen, K. Diederichs, O. Kohlbacher, and I. Fischer, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 67–78.

Krissinel, E., and Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. J. Mol. Biol. *372*, 774–797.

Kuriata, A., Gierut, A.M., Oleniecki, T., Ciemny, M.P., Kolinski, A., Kurcinski, M., and Kmiecik, S. (2018). CABS-flex 2.0: a web server for fast simulations of flexibility of protein structures. Nucleic Acids Res. *46*, W338–W343.

Lasso, G., Honig, B., and Shapira, S.D. (2021). A Sweep of Earth's Virome Reveals Host-Guided Viral Protein Structural Mimicry and Points to Determinants of Human Disease. Cell Syst. *12*, 82-91.e3.

Leahy, D.J., Aukhil, I., and Erickson, H.P. (1996). Crystal Structure of a Four-Domain Segment of Human Fibronectin Encompassing the RGD Loop and Synergy Region. Cell *84*, 155–164.

Lee, S.-C., Lin, C.-C., Wang, C.-H., Wu, P.-L., Huang, H.-W., Chang, C.-I., and Wu, W. (2014). Endocytotic routes of cobra cardiotoxins depend on spatial distribution of positively charged and hydrophobic domains to target distinct types of sulfated glycoconjugates on cell surface. J. Biol. Chem. *289*, 20170–20181.

Li, F. (2015). Receptor Recognition Mechanisms of Coronaviruses: a Decade of Structural Studies. J. Virol. *89*, 1954–1964.

Li, W., Moore, M.J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M.A., Somasundaran, M., Sullivan, J.L., Luzuriaga, K., Greenough, T.C., et al. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature *426*, 450–454.

Li, Z., Jaroszewski, L., Iyer, M., Sedova, M., and Godzik, A. (2020). FATCAT 2.0: towards a better understanding of the structural diversity of proteins. Nucleic Acids Res. *48*, W60–W64.

Lin, L., Ting, S., Yufei, H., Wendong, L., Yubo, F., and Jing, Z. (2020). Epitope-based peptide vaccines predicted against novel coronavirus disease caused by SARS-CoV-2. Virus Res. *288*.

Ling, R., Dai, Y., Huang, B., Huang, W., Yu, J., Lu, X., and Jiang, Y. (2020). In silico design of antiviral peptides targeting the spike protein of SARS-CoV-2. Peptides *130*, 170328.

Lo, W.-C., Lee, C.-Y., Lee, C.-C., and Lyu, P.-C. (2009). iSARST: an integrated SARST web server for rapid protein structural similarity searches. Nucleic Acids Res. *37*, W545–W551.

Luca, V.C., Jude, K.M., Pierce, N.W., Nachury, M. V, Fischer, S., and Garcia, K.C. (2015). Structural basis for Notch1 engagement of Delta-like 4. Science (80-.). *347*, 847 LP – 853.

Luca, V.C., Kim, B.C., Ge, C., Kakuda, S., Wu, D., Roein-Peikar, M., Haltiwanger, R.S., Zhu, C., Ha, T., and Garcia, K.C. (2017). Notch-Jagged complex structure implicates a catch bond in tuning ligand sensitivity. Science *355*, 1320–1324.

Lucchese, G., and Flöel, A. (2020). Molecular mimicry between SARS-CoV-2 and respiratory pacemaker neurons. Autoimmun. Rev. *19*, 102556.

Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. *47*, W636–W641.

Magro, G. (2020). SARS-CoV-2 and COVID-19: Is interleukin-6 (IL-6) the 'culprit lesion' of ARDS onset? What is there besides Tocilizumab? SGP130Fc. Cytokine X *2*, 100029.

Magro, C., Mulvey, J.J., Berlin, D., Nuovo, G., Salvatore, S., Harp, J., Baxter-Stoltzfus, A., and Laurence, J. (2020). Complement associated microvascular injury and thrombosis in the pathogenesis of severe COVID-19 infection: A report of five cases. Transl. Res. *220*, 1–13.

Martínez, Y.A., Guo, X., Portales-Pérez, D.P., Rivera, G., Castañeda-Delgado, J.E., García-Pérez, C.A., Enciso-Moreno, J.A., and Lara-Ramírez, E.E. (2021). The analysis on the human protein domain targets and host-like interacting motifs for the MERS-CoV and SARS-CoV/CoV-2 infers the molecular mimicry of coronavirus. PLoS One *16*, e0246901.

Mastellos, D., Morikis, D., Isaacs, S.N., Holland, M.C., Strey, C.W., and Lambris, J.D. (2003). Complement: structure, functions, evolution, and viral molecular mimicry. Immunol. Res. *27*, 367–386.

McFarland, A.J., Yousuf, M.S., Shiers, S., and Price, T.J. (2021). Neurobiology of SARS-CoV-2 interactions with the peripheral nervous system: implications for COVID-19 and pain. PAIN Reports *6*.

Mendiratta, S.S., Sekulic, N., Hernandez-Guzman, F.G., Close, B.E., Lavie, A., and Colley, K.J. (2006). A Novel α-Helix in the First Fibronectin Type III Repeat of the Neural Cell Adhesion Molecule Is Critical for N-Glycan Polysialylation. J. Biol. Chem. *281*, 36052–36059.

Mendoza, J.L., Schneider, W.M., Hoffmann, H.-H., Vercauteren, K., Jude, K.M., Xiong, A., Moraga, I., Horton, T.M., Glenn, J.S., de Jong, Y.P., et al. (2017). The IFN-λ-IFN-λR1-IL-10Rβ Complex Reveals Structural Features Underlying Type III IFN Functional Plasticity. Immunity *46*, 379–392.

Mirzaie, H., Vahidi, M., Shokoohi, M., Darvishian, M., Sharifi, H., Sharafi, H., and Karamouzian, M. (2020). COVID-19 among patients with hepatitis B or hepatitis C: A systematic review. MedRxiv 2020.10.22.20216317.

Moonens, K., Gideonsson, P., Subedi, S., Bugaytsova, J., Romaõ, E., Mendez, M., Nordén, J., Fallah, M., Rakhimova, L., Shevtsova, A., et al. (2016). Structural Insights into Polymorphic ABO Glycan Binding by Helicobacter pylori. Cell Host Microbe *19*, 55–66.

Moutal, A., Martin, L.F., Boinon, L., Gomez, K., Ran, D., Zhou, Y., Stratton, H.J., Cai, S., Luo, S., Gonzalez, K.B., et al. (2020). SARS-CoV-2 Spike protein co-opts VEGF-A/Neuropilin-1 receptor signaling to induce analgesia. Pain *162*, 243–252.

Nelson, G., Buzko, O., Spilman, P., Niazi, K., Rabizadeh, S., and Soon-Shiong, P. (2021). Molecular dynamic simulation reveals E484K mutation enhances spike RBD-ACE2 affinity and the combination of E484K, K417N and N501Y mutations (501Y.V2 variant) induces conformational change greater than N501Y mutant alone, potentially resulting in an escap. BioRxiv 2021.01.13.426558.

Nonaka, M., and Yoshizaki, F. (2004). Evolution of the complement system. Mol. Immunol. *40*, 897–902.

Nwanochie, E., and Uversky, V.N. (2019). Structure Determination by Single-Particle Cryo-Electron Microscopy: Only the Sky (and Intrinsic Disorder) is the Limit. Int. J. Mol. Sci. *20*.

Panda, A.K., Tripathy, R., and Das, B.K. (2020). Plasmodium falciparum Infection May Protect a Population from Severe Acute Respiratory Syndrome Coronavirus 2 Infection. J. Infect. Dis. *222*, 1570–1571.

Pandey, P., Rane, J.S., Chatterjee, A., Kumar, A., Khan, R., Prakash, A., and Ray, S. (2020). Targeting SARS-CoV-2 spike protein of COVID-19 with naturally occurring phytochemicals: an in silico study for drug development. J. Biomol. Struct. Dyn. *0*, 1–11.

Patra, T., Meyer, K., Geerling, L., Isbell, T.S., Hoft, D.F., Brien, J., Pinto, A.K., Ray, R.B., and Ray, R. (2020). SARS-CoV-2 spike protein promotes IL-6 trans-signaling by activation of angiotensin II receptor signaling in epithelial cells. PLOS Pathog. *16*, e1009128.

Pišlar, A., Mitrović, A., Sabotič, J., Pečar Fonović, U., Perišić Nanut, M., Jakoš, T., Senjor, E., and Kos, J. (2020). The role of cysteine peptidases in coronavirus cell entry and replication: The therapeutic potential of cathepsin inhibitors. PLOS Pathog. *16*, e1009013.

Pizarro, J.C., Chitarra, V., Verger, D., Holm, I., Pêtres, S., Dartevelle, S., Nato, F., Longacre, S., and Bentley, G.A. (2003). Crystal Structure of a Fab Complex Formed with PfMSP1-19, the C-terminal Fragment of Merozoite Surface Protein 1 from Plasmodium falciparum: A Malaria Vaccine Candidate. J. Mol. Biol. *328*, 1091–1103.

Ploss, A., and Evans, M.J. (2012). Hepatitis C virus host cell entry. Curr. Opin. Virol. *2*, 14–19.

Raham, T.F. (2021). Influence of Malaria Endemicity and Tuberculosis Prevalence on COVID-19 Mortality. Public Health *194*, 33–35.

Ramlall, V., Thangaraj, P.M., Meydan, C., Foox, J., Butler, D., May, B., De Freitas, J.K., Glicksberg, B.S., Mason, C.E., Tatonetti, N.P., et al. (2020). Identification of Immune complement function as a determinant of adverse SARS-CoV-2 infection outcome. MedRxiv.

Ranucci, M., Sitzia, C., Baryshnikova, E., Di Dedda, U., Cardani, R., Martelli, F., and Corsi Romanelli, M. (2020). Covid-19-Associated Coagulopathy: Biomarkers of Thrombin Generation and Fibrinolysis Leading the Outcome. J. Clin. Med. *9*.

Reddy, K.R. (2020). SARS-CoV-2 and the Liver: Considerations in Hepatitis B and Hepatitis C Infections. Clin. Liver Dis. *15*, 191–194.

Reguera, J., Mudgal, G., Santiago, C., and Casasnovas, J.M. (2014). A structural view of coronavirus–receptor interactions. Virus Res. *194*, 3–15.

Renko, M., Sabotič, J., Mihelič, M., Brzin, J., Kos, J., and Turk, D. (2010). Versatile Loops in Mycocypins Inhibit Three Protease Families. J. Biol. Chem. *285*, 308–316.

Renu, K., Prasanna, P.L., and Valsala Gopalakrishnan, A. (2020). Coronaviruses pathogenesis, comorbidities and multi-organ damage – A review. Life Sci. *255*, 117839.

Rhea, E.M., Logsdon, A.F., Hansen, K.M., Williams, L.M., Reed, M.J., Baumann, K.K., Holden, S.J., Raber, J., Banks, W.A., and Erickson, M.A. (2021). The S1 protein of SARS-CoV-2 crosses the blood–brain barrier in mice. Nat. Neurosci. *24*, 368–378.

Richardson, S., Hirsch, J.S., Narasimhan, M., Crawford, J.M., McGinn, T., Davidson, K.W., and Consortium, and the N.C.-19 R. (2020). Presenting Characteristics, Comorbidities, and Outcomes

Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. JAMA *323*, 2052–2059.

Rizzo, P., Vieceli Dalla Sega, F., Fortini, F., Marracino, L., Rapezzi, C., and Ferrari, R. (2020). COVID-19 in the heart and the lungs: could we "Notch" the inflammatory storm? Basic Res. Cardiol. *115*, 1–8.

Rodrigues, J., Teixeira, J.M.C., Trellet, M., and Bonvin, A. (2018). pdb-tools: a swiss army knife for molecular structures. F1000Research *7*.

Rovina, N., Akinosoglou, K., Eugen-Olsen, J., Hayek, S., Reiser, J., and Giamarellos-Bourboulis, E.J. (2020). Soluble urokinase plasminogen activator receptor (suPAR) as an early predictor of severe respiratory failure in patients with COVID-19 pneumonia. Crit. Care *24*, 187.

Rowan, E.G. (2001). What does beta-bungarotoxin do at the neuromuscular junction? Toxicon *39*, 107–118.

Saez, N.J., Mobli, M., Bieri, M., Chassagnon, I.R., Malde, A.K., Gamsjaeger, R., Mark, A.E., Gooley, P.R., Rash, L.D., and King, G.F. (2011). A dynamic pharmacophore drives the interaction between Psalmotoxin-1 and the putative drug target acid-sensing ion channel 1a. Mol. Pharmacol. *80*, 796–808.

Šali, A. (2019). MODELLER: A Program for Protein Structure Modeling. 779–815.

Šali, A., and Blundell, T.L. (1993). Comparative Protein Modelling by Satisfaction of Spatial Restraints. J. Mol. Biol. *234*, 779–815.

Saputri, D.S., Li, S., van Eerden, F.J., Rozewicki, J., Xu, Z., Ismanto, H.S., Davila, A., Teraguchi, S., Katoh, K., and Standley, D.M. (2020). Flexible, Functional, and Familiar: Characteristics of SARS-CoV-2 Spike Protein Evolution. Front. Microbiol. *11*, 2112.

Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. (2005). The FoldX web server: an online force field. Nucleic Acids Res. *33*, W382-8.

Shang, Y., Smith, S., and Hu, X. (2016). Role of Notch signaling in regulating innate immunity and inflammation in health and disease. Protein Cell *7*, 159–174.

Sigrist, C.J., Bridge, A., and Le Mercier, P. (2020). A potential role for integrins in host cell entry by SARS-CoV-2. Antiviral Res. *177*.

Sillitoe, I., Dawson, N., Lewis, T.E., Das, S., Lees, J.G., Ashford, P., Tolulope, A., Scholes, H.M., Senatorov, I., Bujan, A., et al. (2019). CATH: Expanding the horizons of structure-based functional annotations for genome sequences. Nucleic Acids Res. *47*, D280–D284.

Smart, O.S., Womack, T.O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., and Bricogne, G. (2012). Exploiting structure similarity in refinement: automated NCS and target-structure restraints in {\it BUSTER}. Acta Crystallogr. Sect. D *68*, 368–380.

Spinello, A., Saltalamacchia, A., and Magistrato, A. (2020). Is the Rigidity of SARS-CoV-2 Spike Receptor-Binding Motif the Hallmark for Its Enhanced Infectivity? Insights from All-Atom Simulations. J. Phys. Chem. Lett. *11*, 4785–4790.

Steentoft, C., Vakhrushev, S.Y., Joshi, H.J., Kong, Y., Vester-Christensen, M.B., Schjoldager, K.T.B.G., Lavrsen, K., Dabelsteen, S., Pedersen, N.B., Marcos-Silva, L., et al. (2013). Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J. *32*, 1478–1488.

Stenflo, J., Stenberg, Y., and Muranyi, A. (2000). Calcium-binding EGF-like modules in coagulation proteinases: function of the calcium ion in module interactions. Biochim. Biophys. Acta *1477*, 51–63.

Stout, T. (2004). Show charged amino acids on PyMOL.

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. *47*, D607–D613.

Tamada, K., Shimozaki, K., Chapoval, A.I., Zhai, Y., Su, J., Chen, S.F., Hsieh, S.L., Nagata, S., Ni, J., and Chen, L. (2000). LIGHT, a TNF-like molecule, costimulates T cell proliferation and is required for dendritic cell-mediated allogeneic T cell response. J. Immunol. *164*, 4105–4110.

Tang, X., Yang, M., Duan, Z., Liao, Z., Liu, L., Cheng, R., Fang, M., Wang, G., Liu, H., Xu, J., et al. (2020). Transferrin receptor is another receptor for SARS-CoV-2 entry. BioRxiv 2020.10.23.350348.

Tombling, B.J., Wang, C.K., and Craik, D.J. (2020). EGF-like and Other Disulfide-rich Microdomains as Therapeutic Scaffolds. Angew. Chemie Int. Ed. *59*, 11218–11232.

Tortorici, M.A., and Veesler, D. (2019). Structural insights into coronavirus entry. Adv. Virus Res. *105*, 93–116.

Vajda, S., Yueh, C., Beglov, D., Bohnuud, T., Mottarella, S.E., Xia, B., Hall, D.R., and Kozakov, D. (2017). New additions to the ClusPro server motivated by CAPRI. Proteins Struct. Funct. Bioinforma. *85*, 435–444.

Versteeg, G.A., van de Nes, P.S., Bredenbeek, P.J., and Spaan, W.J.M. (2007). The Coronavirus Spike Protein Induces Endoplasmic Reticulum Stress and Upregulation of Intracellular Chemokine mRNA Concentrations. J. Virol. *81*, 10981–10990.

Volpon, L., Lamthanh, H., Barbier, J., Gilles, N., Molgó, J., Ménez, A., and Lancelin, J.-M. (2004). NMR Solution Structures of δ-Conotoxin EVIA from Conus ermineus That Selectively Acts on Vertebrate Neuronal Na+ Channels. J. Biol. Chem. *279*, 21356–21366.

Walls, A.C., Xiong, X., Park, Y.-J., Tortorici, M.A., Snijder, J., Quispe, J., Cameroni, E., Gopal, R., Dai, M., Lanzavecchia, A., et al. (2019). Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. Cell *176*, 1026-1039.e15.

Wang, L., and Xiang, Y. (2020). Spike Glycoprotein-Mediated Entry of SARS Coronaviruses. Viruses *12*.

Wang, C., Bradley, P., and Baker, D. (2007). Protein-Protein Docking with Backbone Flexibility. J. Mol. Biol. *373*, 503–519.

Wang, J., Zheng, X., Peng, Q., Zhang, X., and Qin, Z. (2020). Eph receptors: the bridge linking host and virus. Cell. Mol. Life Sci. *77*, 2355–2365.

Wang, N., Shi, X., Jiang, L., Zhang, S., Wang, D., Tong, P., Guo, D., Fu, L., Cui, Y., Liu, X., et al. (2013). Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. Cell Res. *23*, 986–993.

Watson, A.A., Lebedev, A.A., Hall, B.A., Fenton-May, A.E., Vagin, A.A., Dejnirattisai, W., Felce, J., Mongkolsapaya, J., Palma, A.S., Liu, Y., et al. (2011). Structural Flexibility of the Macrophage Dengue Virus Receptor CLEC5A: IMPLICATIONS FOR LIGAND BINDING AND SIGNALING*. J. Biol. Chem. *286*, 24208–24218.

Westall, F.C. (2006). Molecular mimicry or structural mimicry? Mol. Immunol. *43*, 1062–1064.

Wu, P.-L., Lee, S.-C., Chuang, C.-C., Mori, S., Akakura, N., Wu, W., and Takada, Y. (2006). Non-cytotoxic cobra cardiotoxin A5 binds to alpha(v)beta3 integrin and inhibits bone resorption. Identification of cardiotoxins as non-RGD integrin-binding proteins of the Ly-6 family. J. Biol. Chem. *281*, 7937–7945.

Xu, K., Tzvetkova-Robev, D., Xu, Y., Goldgur, Y., Chan, Y.-P., Himanen, J.P., and Nikolov, D.B. (2013). Insights into Eph receptor tyrosine kinase activation from crystal structures of the EphA4 ectodomain and its complex with ephrin-A5. Proc. Natl. Acad. Sci. *110*, 14634 LP – 14639.

Xue, X., Wu, J., Ricklin, D., Forneris, F., Di Crescenzio, P., Schmidt, C.Q., Granneman, J., Sharp, T.H., Lambris, J.D., and Gros, P. (2017). Regulator-dependent mechanisms of C3b processing by factor I allow differentiation of immune responses. Nat. Struct. Mol. Biol. *24*, 643–651.

Yang, C., Zhang, Y., Chen, H., Chen, Y., Yang, D., Shen, Z., Wang, X., Liu, X., Xiong, M., and Huang, K. (2020). Kidney injury molecule-1 is a potential receptor for SARS-CoV-2. BioRxiv 2020.10.09.334052.

Younes, Z., Florian, P., Isabelle, A., Abdallah, N., Mounia, O., Loubna, K., Youness, L., Nabil, Z., Khalid, S., Rafiqua, B.E.H., et al. (2020). Platelets Can Associate With SARS-CoV-2 RNA and Are Hyperactivated in COVID-19. Circ. Res. *127*, 1404–1418.

Yu, G., Boone, T., Delaney, J., Hawkins, N., Kelley, M., Ramakrishnan, M., McCabe, S., Qiu, W.R., Kornuc, M., Xia, X.Z., et al. (2000). APRIL and TALL-I and receptors BCMA and TACI: system for regulating humoral immunity. Nat. Immunol. *1*, 252–256.

Yu, J., Yuan, X., Chen, H., Chaturvedi, S., Braunstein, E.M., and Brodsky, R.A. (2020). Direct activation of the alternative complement pathway by SARS-CoV-2 spike proteins is blocked by factor D inhibition. Blood *136*, 2080–2089.

Zhang, Y., and Skolnick, J. (2004). Scoring function for automated assessment of protein structure template quality. Proteins *57*, 702–710.

Zhang, Y., and Skolnick, J. (2005). TM-align: A protein structure alignment algorithm based on the TM-score. Nucleic Acids Res. *33*, 2302–2309.

Zhang, H., Casasnovas, J.M., Jin, M., Liu, J., Gahmberg, C.G., Springer, T.A., and Wang, J. (2008). An Unusual Allosteric Mobility of the C-Terminal Helix of a High-Affinity αL Integrin I Domain Variant Bound to ICAM-5. Mol. Cell *31*, 432–437.

Zhang, X., Wu, K., Wang, D., Yue, X., Song, D., Zhu, Y., and Wu, J. (2007). Nucleocapsid protein of SARS-CoV activates interleukin-6 expression through cellular transcription factor NF-kappaB. Virology *365*, 324–335.

Zheng, B., Yuan, M., Ma, Q., Wang, S., Tan, Y., Xu, Y., Ye, J., Gao, Y., Sun, X., Yang, Z., et al. (2021). Landscape of SARS-CoV-2 spike protein-interacting cells in human tissues. Int. Immunopharmacol. 107567.

Zizzo, G., and Cohen, P.L. (2020). Imperfect storm: is interleukin-33 the Achilles heel of COVID-19? Lancet Rheumatol. *2*, e779–e790.
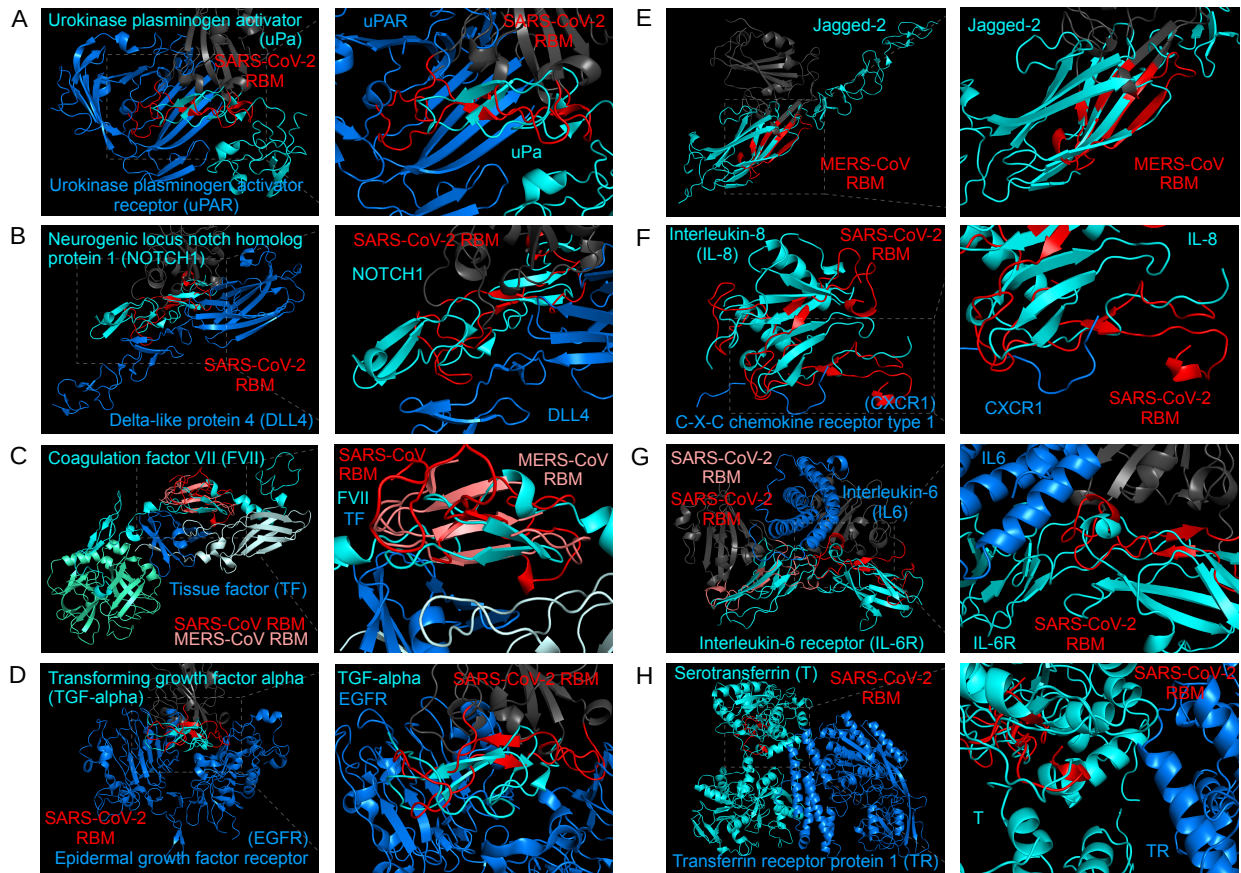
Supplementary figures and tables

RBM	Receptor-binding domain residues	Receptor-binding motif residues
SARS-CoV-2	333-523	443-457,469-506
SARS-CoV	332-511	426-443,456-492
MERS-CoV	380-586	501-515,529-558

Supplementary Table 1. Spike protein receptor-binding domain and motif residues selected in this study
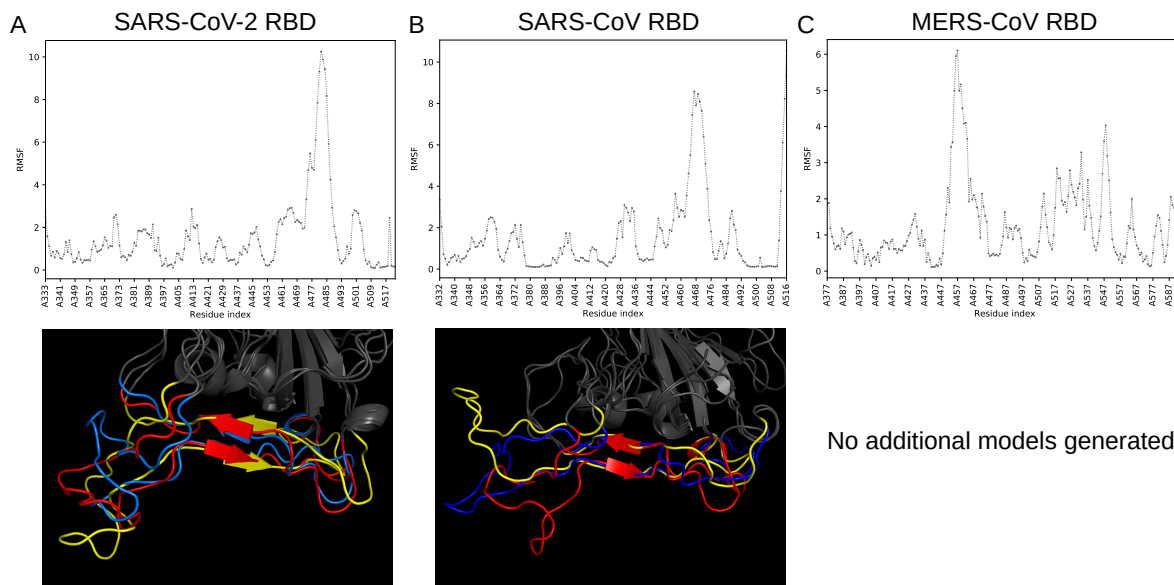

Supplementary Figure 1. CABS-Flex 2.0 analysis and models for coronavirus RBDs
The top row displays the root mean square flexibility for each residue in the SARS-CoV-2 (A), SARS-CoV (B), and MERS-CoV (C) RBDs in the CABS-Flex 2.0 analyses. The bottom row shows selected models from CABS-Flex 2.0, included for the SARS-CoV-2 and SARS-CoV structural similarity screen, in alignment with the reference RBM.

Supplementary Figure 2. Structural alignments of potential endogenous mimics

Structural alignments between the following coronavirus RBMs and potential molecular mimics are shown: SARS-CoV-2 and urokinase plasminogen activator (A), SARS-CoV-2 and NOTCH1 (B), MERS-CoV and SARS-CoV mimicking coagulation factor VIIa (C), SARS-CoV-2 and TGF-alpha (D), MERS-CoV and jagged-2 (E), SARS-CoV-2 and IL-8 (F), SARS-CoV-2 and IL-6 receptor alpha and beta chains (G), SARS-CoV-2 and serotransferrin (H). RBMs are labelled red, mimicked proteins are cyan, and potential interaction partners are marine blue.
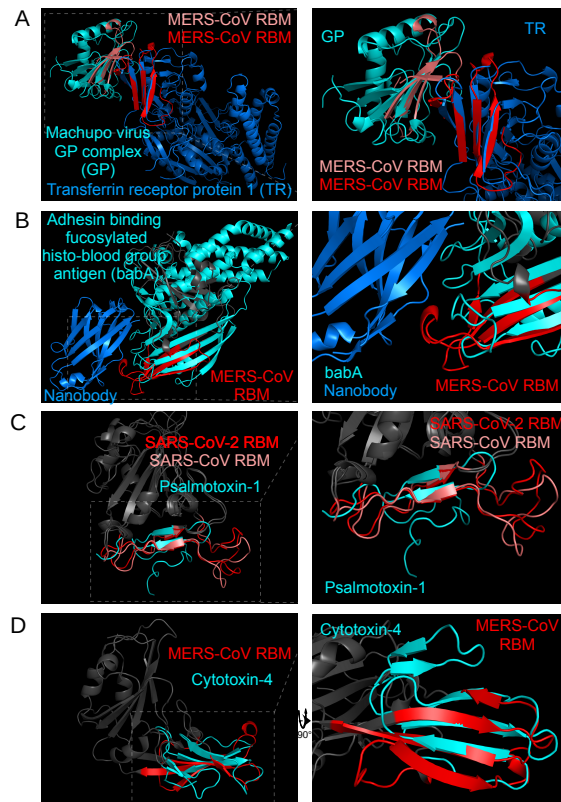
Supplementary Figure 3. Structural alignments with exogenous proteins
The MERS-CoV RBM aligned to the transferrin receptor protein 1 and Machupo virus glycoprotein polyprotein GP complex (A) and the adhesin-binding fucosylated histo-blood group antigen (B). The SARS-CoV and SARS-CoV-2 RBMs aligned to psalmotoxin-1 (C), and the MERS-CoV RBM aligned to cytotoxin 4 (D). RBMs are labelled red, mimicked proteins are cyan, and potential interaction partners are marine blue.

MERS-CoV	SARS-CoV	SARS-CoV-2	HCV PDB
x			5kzp
x	x	x	4g6a
	x	x	5nph
		x	5fgc
	x		4xvj
		x	4uoi
		x	4gag
		x	5fgb
		x	6bzu
		x	4hs6
	x	x	6bzv
	x	x	6bzw
	x		4wht

Supplementary Table 2. Hepatitis C virus (HCV) PDB codes returned in structural similarity screen