

# Structural basis of SARS-CoV-2 spike protein induced by ACE2

Tomer Meirson<sup>a,b,\*</sup>, David Bomze<sup>c</sup>, Gal Markel<sup>b,d,\*</sup>

<sup>a</sup> Azrieli Faculty of Medicine, Bar-Ilan University, Israel

<sup>b</sup> Ella Lemelbaum Institute for Immuno-oncology, Sheba Medical Center, Ramat-Gan 526260, Israel

<sup>c</sup> Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 6997801, Israel

<sup>d</sup> Department of Clinical Microbiology and Immunology, Sackler Faculty of Medicine, Tel Aviv University, Tel-Aviv, Israel

\* Correspondence should be addressed to Tomer Meirson or Gal Markel at Ella Lemelbaum Institute for Immuno-Oncology and Melanoma, Sheba Medical Center, Ramat-Gan 526260, Israel. E-mail addresses: [tomermrsn@gmail.com](mailto:tomermrsn@gmail.com) or [gal.markel@sheba.health.gov.il](mailto:gal.markel@sheba.health.gov.il)

## **Abstract**

### **Motivation**

The recent emergence of the novel SARS-coronavirus 2 (SARS-CoV-2) and its international spread pose a global health emergency. The viral spike (S) glycoprotein binds the receptor (angiotensin-converting enzyme 2) ACE2 and promotes SARS-CoV-2 entry into host cells. The trimeric S protein binds the receptor using the distal receptor-binding domain (RBD) causing conformational changes in S protein that allow priming by host cell proteases. Unravelling the dynamic structural features used by SARS-CoV-2 for entry might provide insights into viral transmission and reveal novel therapeutic targets. Using structures determined by X-ray crystallography and cryo-EM, we performed structural analysis and atomic comparisons of the different conformational states adopted by the SARS-CoV-2-RBD.

### **Results**

Here, we determined the key structural components induced by the receptor and characterized their intramolecular interactions. We show that  $\kappa$ -helix (also known as polyproline II) is a predominant structure in the binding interface and in facilitating the conversion to the active form of the S protein. We demonstrate a series of conversions between switch-like  $\kappa$ -helix and  $\beta$ -strand, and conformational variations in a set of short  $\alpha$ -helices which affect the proximal hinge region. This conformational changes lead to an alternating pattern in conserved disulfide bond configurations positioned at the hinge, indicating a possible disulfide exchange, an important allosteric switch implicated in viral entry of various viruses, including HIV and murine coronavirus. The structural information presented herein enables us to inspect and understand the important dynamic features of SARS-CoV-2-RBD and propose a novel potential therapeutic strategy to block viral entry. Overall, this study provides guidance for the design and optimization of structure-based intervention strategies that target SARS-CoV-2.

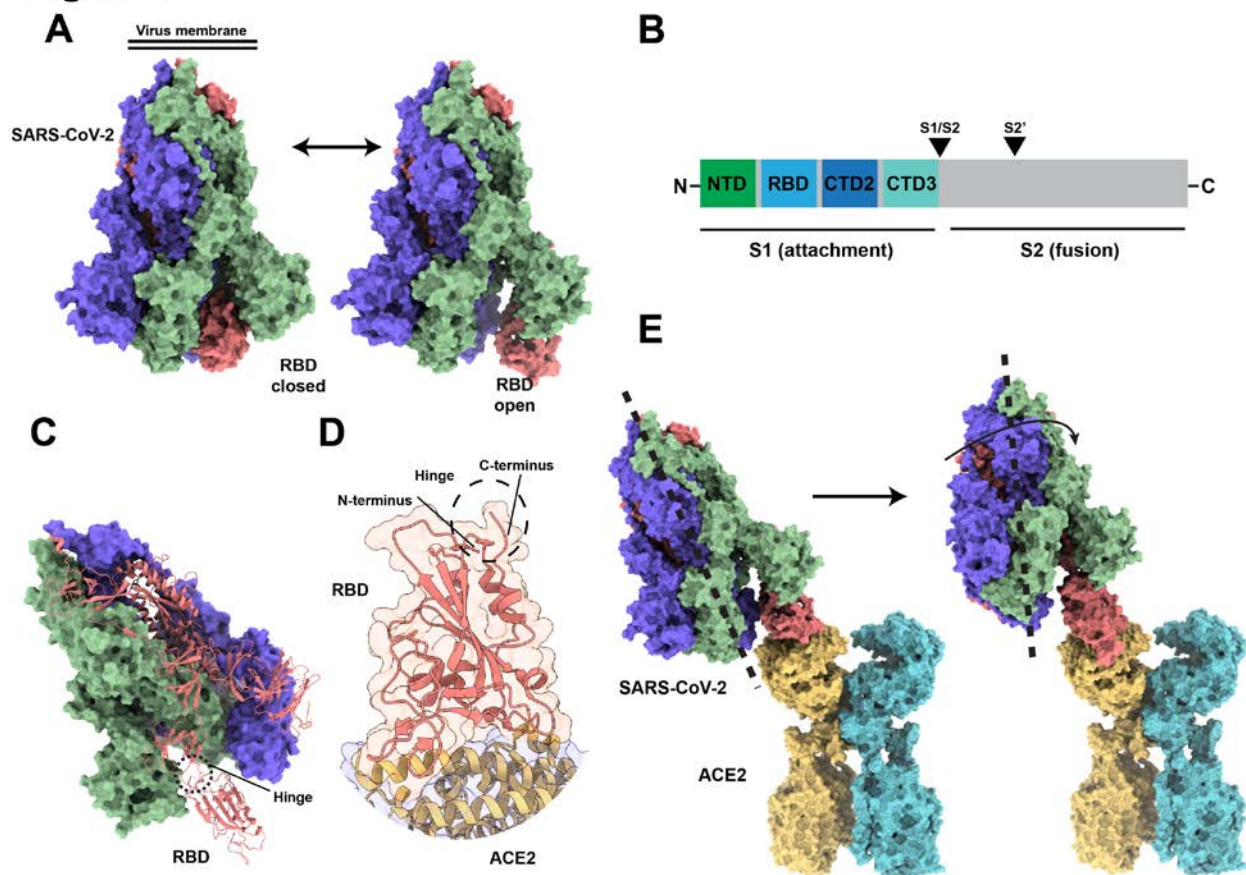
## Introduction

Several members of the coronavirus family circulate in the human population and usually manifest mild respiratory symptoms (Su, et al., 2016). However, over the past two decades, emerging coronaviruses (CoV) have raised great public health concerns worldwide. The highly pathogenic severe acute respiratory syndrome-related coronavirus (SARS-CoV) (Drosten, et al., 2003; Ksiazek, et al., 2003) and Middle-East respiratory syndrome coronavirus (MERS-CoV) (Zaki, et al., 2012) have crossed the species barrier and cause deadly pneumonia in afflicted individuals, SARS and MERS, respectively. SARS-CoV first emerged in humans in Guangdong province of China in 2002, and its global spread was associated with 774 deaths among 8,096 cases (de Wit, et al., 2016; WHO, 2004). In 2012, MERS-CoV was first identified in the Arabian Peninsula and spread to 27 countries, infecting a total of 2,494 people and claiming 858 lives (WHO, 2020). While the SARS pandemic was finally stopped by conventional control measures, including patient isolation and travel restrictions, new cases of MERS have been reported (Yoon and Kim, 2019).

In December 2019, a previously unknown CoV, named SARS-CoV-2, was discovered in Wuhan, Hubei province of China (Huang, et al., 2020; Wang, et al., 2020; Zhu, et al., 2020). The sudden emergence of the novel SARS-CoV-2 has rapidly evolved into a pandemic that posed a serious threat to global health and economy (Gates, 2020). Although SARS and MERS have a higher mortality rate, SARS-CoV-2 infection (COVID-19) spreads much more rapidly (To, et al., 2020). As of May 2020, more than 4,000,000 confirmed infections were reported in 216 countries, including over 300,000 deaths (WHO, 2020). MERS-CoV, SARS-CoV, and SARS-CoV2 were suggested to originate from bats and most likely serve as a reservoir host for these viruses (Ge, et al., 2013; Haagmans, et al., 2014; Li, et al., 2005; Memish, et al., 2013; Zhou, et al., 2020). Detailed investigations of the zoonotic origin of human CoVs indicate that SARS-CoV was transmitted from palm civets to humans and MERS-CoV from dromedary camels to humans (Guan, et al., 2003; Haagmans, et al., 2014; Kan, et al., 2005). However, the intermediate host for zoonotic transmission of SARS-CoV-2, linked to a wet animal market in Wuhan, is still under investigation (Walls, et al., 2020; Ye, et al., 2020).

The spike (S) glycoprotein is a class I fusion protein that mediates the entry of CoVs into target cells (Tortorici, et al., 2019). The S protein forms homotrimers that protrude from the viral surface (Figure 1A) and comprises two functional subunits, which facilitate viral attachment to the surface of host cells (S1 subunit) and fusion of the viral and cellular membranes (S2 subunit) (Figure 1B). The distal part of S1 subunit, the receptor-binding domain (RBD), is linked through two anti-parallel hinge linkers which connect the domain to the N-terminal domain (NTD) and C-terminal domain 2 (CTD2) and allow the transition between closed and open conformations (Gui, et al., 2017; Song, et al., 2018; Walls, et al., 2020; Yuan, et al., 2017) (Figure 1A-C). The open conformation of the S1 subunit facilitates interaction with angiotensin-converting enzyme 2 (ACE2) (Figure 1D-E), which contributes to the stabilization of the prefusion state of the S2 subunit that contains the fusion machinery (Song, et al., 2018; Walls, et al., 2020). Activated S protein is cleaved by host proteases at the S1/S2 and S2' site resulting in a cleaved S2' subunit that drives the fusion of viral and cellular membranes (Belouzard, et al., 2009; Millet and Whittaker, 2014; Pallesen, et al., 2017). The critical step in SARS-CoV-2 infection which involves the transition between a metastable prefusion state to a stable post-fusion state is triggered by binding to ACE2 which induces conformational changes in the RBD and the hinge region (Gui, et al., 2017; Pallesen, et al., 2017; Song, et al., 2018; Walls, et al., 2020). While the interaction between the S protein and ACE2 has been extensively studied (Han, et al., 2006; Lan, et al., 2020; Letko, et al., 2020; Li, 2013; Shang, et al., 2020; Song, et al., 2018; Wan, et al., 2020; Wang, et al., 2020; Yan, et al., 2020), the key determinants in the activation of the virus upon binding to the host receptor is poorly understood. We aimed to investigate the structural basis of the SARS-CoV-2 S protein modulation induced by virus-receptor interaction.

## Figure 1



**Figure 1.** Structural comparison of SARS-CoV-2 S protein conformational states. (A) Surface diagram of SARS-CoV-2 homotrimeric structure in the unbound- closed and open conformations. (B) Structural illustration of S protein, including functional domains (NTD, N-terminal domain; RBD, receptor-binding domain; CTD2, C-terminal domain 2; CTD, C-terminal domain 3; and proteolytic cleavage sites (S1/S2, S2')). (C) S trimer with one RBD in the open conformation and (D) RBD-ACE2 complex shown as a cartoon. (E) Superposed structures depicting the conformational changes between the unbound-open (left) to the ACE2-bound state.

## Methods

### Structure collection and modeling

We searched the PDB database for high-quality structures of the SARS-CoV-2 S protein at closed, open, and bound conformations. For each conformation, we used the structure with the lowest resolution that contains the RBD and the hinge region. For the unbound-closed, unbound-open, and bound conformations, we used PDB IDs 6VXX:B, 6VYB:B, and 6M0J:E, respectively. To model the position of SARS-CoV-2 S protein compared to ACE2, we structurally aligned the models to the RBD of ACE-B<sup>o</sup>AT1 complex (Yan, et al., 2020) (6M17) and SARS-CoV in the bound conformation with the highest degree of RBD opening (Song, et al., 2018) (6ACK). The structures were preprocessed with Dock Prep and aligned using UCSF Chimera v1.13 (Pettersen, et al., 2004). To compare the secondary structure compositions of attachment proteins, we used structures with the lowest resolution from different families of enveloped viruses with class I fusion proteins (White, et al., 2008). These structures were compared to a random dataset comprising the first 1000 non-viral PDB proteins deposited in 2020 with a resolution <3 Å.

### Secondary structure assignment

The initial assignment of secondary structure was performed using DSSP (Touw, et al., 2015). To assign  $\kappa$ -helix (alternative designation for polyproline II [PPII]), we used the method which was recently introduced (Meirson, et al., 2020; Meirson, et al., 2020). Briefly, we calculated the root-mean-square dihedral deviations (RMSdD) of the peptide backbone torsional angles  $\phi$  and  $\psi$  as a measure of the average deviation from a reference  $\kappa$ -helix. To include short segments of  $\kappa$ -helix, at least two consecutive residues with mean RMSdD below the cutoff ( $\epsilon$ ) of 17 (Mansiaux, et al., 2011) were defined as the criteria for the assignment.

### Structural characterization

To analyze the conformational changes induced by ACE2 binding, we compared the bound with the unbound form. We included both the open and closed states of the unbound structures to highlight the specific effects induced by the interaction with ACE2. The backbone dihedral angles ( $\phi$ ,  $\psi$ ) were converted to generic helix parameters  $\vartheta$



(angular step per residue) and  $d$  (rise per residue) as described by Miyazawa (Miyazawa, 1961) to describe the geometrical variations intuitively. Van der Waals (VDW) intramolecular interactions were determined using a distance cutoff of 3.5 Å. Only interactions that were gained or lost by the bound compared to both unbound conformations were considered. To identify interacting residues with considerable shifts, displacements smaller than 0.5 Å were excluded. To identify important backbone hydrogen bonds (H-bonds), the mean difference in the electrostatic interaction energy between the bound and the unbound conformations was calculated using DSSP. For each acceptor-donor pair, gain, or loss of H-bond was defined if the mean energy difference is  $<-1\text{kcal/mol}$  or  $>1\text{kcal/mol}$ , respectively. The magnitude of the energy threshold represents twice the standard cutoff ( $-0.5\text{kcal/mol}$ ) for the existence of a H-bond (Zhang and Sagui, 2015). Structural analyses was performed with the Bio3D package (Grant, et al., 2006) in R version 3.6.

### **Disulfide bond analysis**

Due to inconsistencies in the reported composition of disulfide bonds in the RBD in different structures (Lavillette, et al., 2006; Song, et al., 2018; Walls, et al., 2020; Wrapp, et al., 2020; Yuan, et al., 2017) and to account for possible errors in modeling disulfide bonds (Carpentier, et al., 2010; Kleywegt and Jones, 1995; Villa and Lasker, 2014; Wlodawer, et al., 2008), we used Disulfide by Design 2.0 (DbD2) to calculate  $\chi_3$  torsion angles and bond energies (Craig and Dombkowski, 2013). The DbD2 algorithm could accurately predict the chiralities and positions of disulfide bonds based on energy function, reflecting the geometric characteristics of disulfide bonds among high-quality crystal structures (Craig and Dombkowski, 2013; Wiedemann, et al., 2020). We used the estimated disulfide energy threshold of  $<2.17\text{kcal/mol}$  that applies to most naturally occurring disulfides,  $\chi_3$  angles of  $-87$  or  $97 \pm 20$  (Craig and Dombkowski, 2013) and disulfide bond distance of  $2.03 \pm 5\%$  (Spek, 1990) to indicate a high probability of a disulfide bond that satisfies stereochemical constraints.

## Graphical visualization

The models were visualized using UCSF ChimeraX v0.94 (Goddard, et al., 2018). Models with reassigned secondary structures were visualized using the academic version of Schrodinger Maestro v11.1 (Bell, et al., 2006).  $\kappa$ -helices and  $3_{10}$ -helices were represented as ribbons and tubes, respectively. Ribbons were drawn, passing through carbon alphas. Trajectories were produced by interpolating between the bound and unbound-closed conformation and visualized using VMD (Humphrey, et al., 1996).

## Results

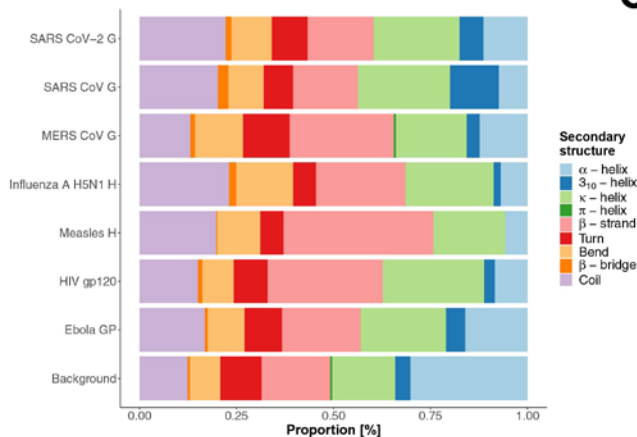
### Secondary structure determination

To characterize the structural composition of the RBD of SARS-CoV-2 and other members of the class I fusion proteins, we calculated the distribution of secondary structure assignment. Since  $\kappa$ -helix (PPII) often serves functional purposes in proteins (Adzhubei, et al., 2013; Meirson, et al., 2020) and DSSP does not assign the conformation, we performed reassignment of the secondary structures to reveal potential  $\kappa$ -helix conformations. Compared to a random dataset in the PDB, structures of the attachment proteins of CoVs, influenza, measles, HIV and Ebola viruses display higher proportions of  $\beta$ -strands and  $\kappa$ -helices, whereas  $\alpha$ -helix is under-represented (Figure 2A). Among non-regular secondary structures, the random coil is the most over-represented assignment. The reassigned SARS-CoV-2 structure reveals a diverse distribution of  $\kappa$ -helices throughout the domain, whereas other secondary structures are clustered more closely together (Figure 2B-C). The most common secondary structure in the ACE2 binding interface is  $\kappa$ -helix (Figure 2C).

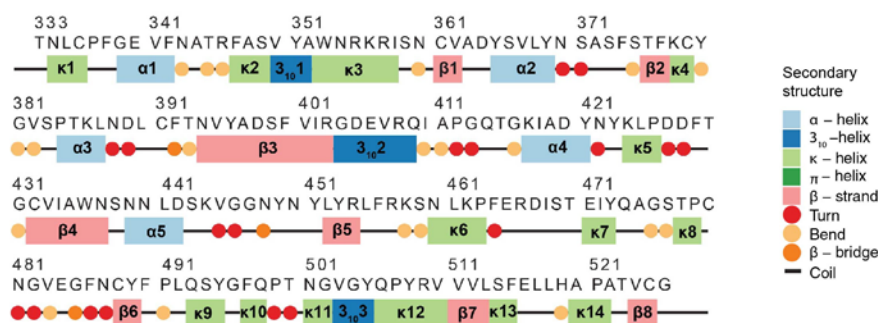


## Figure 2

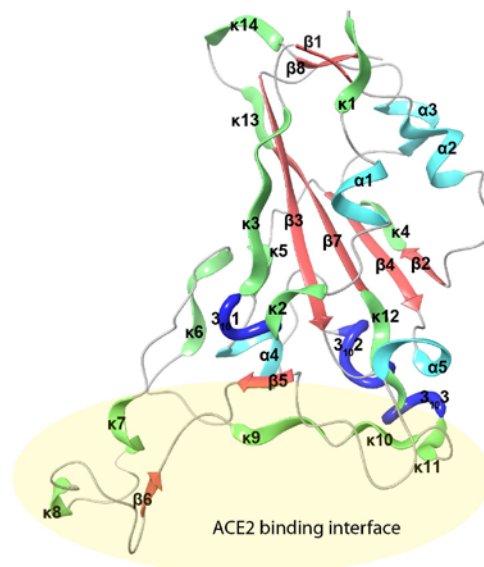
A



B



C



**Figure 2.** Secondary structure assignment of the receptor-binding domain. (A) Comparison of attachment proteins of coronaviruses and other members of the class I fusion proteins. Position of secondary structure assignment (B) and cartoon representation (C) of SARS-CoV-2-RBD. The secondary structural elements are labeled according to their occurrence in sequence.  $\alpha$ -helices (cyan),  $\beta$ -strands (red), and  $\kappa$ -helices (green) are illustrated as ribbons and  $3_{10}$ -helices (blue) as thick tubes.

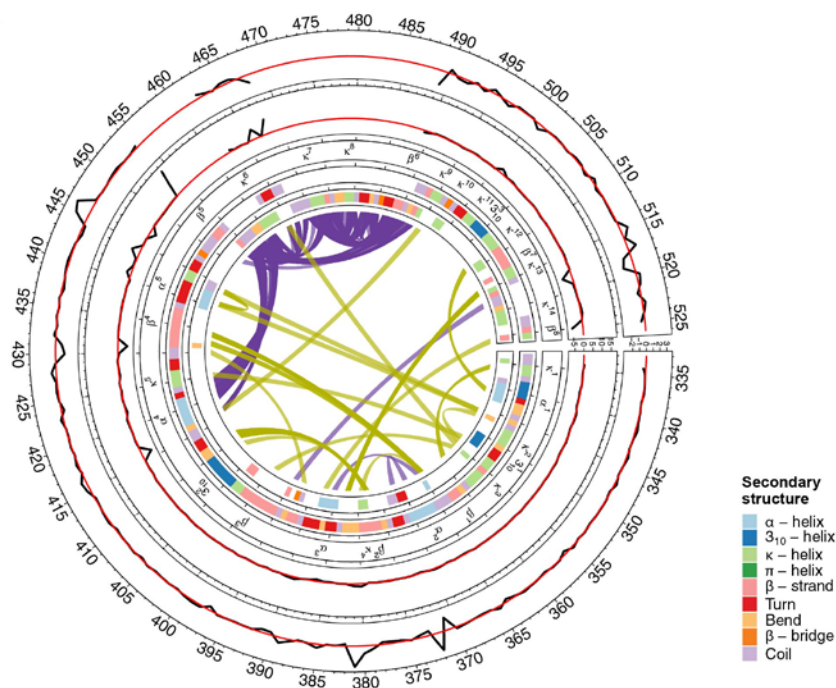
### Structural variations of SARS-CoV-2-RBD bound with ACE2 receptor

To gain insights into the effects of ACE2 interaction on the S protein of SARS-COV-2, we analyzed the intramolecular structural variations in the bound versus the unbound-closed and unbound-open conformations. Figure 3 depicts a summary of the main differences in the intramolecular interaction and H-bond profile upon receptor binding and allows to follow the interconnectivity path leading to the hinge region at the termini. While small changes in the rotational angles or rise of residues occur throughout the domain (Figure

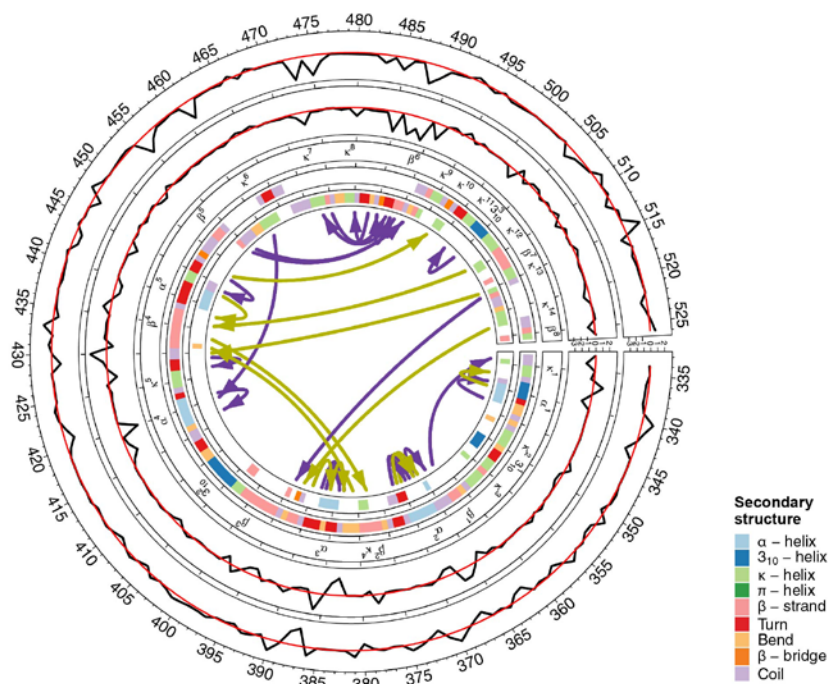
3A), the most substantial changes happen at non-regular secondary structures such as coils and turns. Compared with the unbound-open state, the bound conformation is mainly associated with the formation of  $\kappa$ -helices (20 residues), followed by  $\alpha$ -helices (13 residues),  $\beta$ -strands (10 residues), and a  $3_{10}$ -helix (3 residues). The interaction map demonstrates a rich rearrangement of intramolecular interactions (Figure 3A). Interactions are gained mostly in missing loops at the binding interface (455-491), whereas lost interactions are scattered throughout the domain. Figure 3B shows the redistribution of main-chain H-bonds energy and rearrangement of acceptor and donor H-bond networks. Consistent with the interconnectivity map, the pairing of H-bonds occurs at the ACE2 binding interface. Interestingly, local rearrangements of H-bonds occur mainly at  $\alpha$ -helices, whereas  $\kappa$ -helices facilitate distant H-bonds. This indicates that  $\kappa$ -helices mediate switch-like interactions.

## Figure 3

A



B



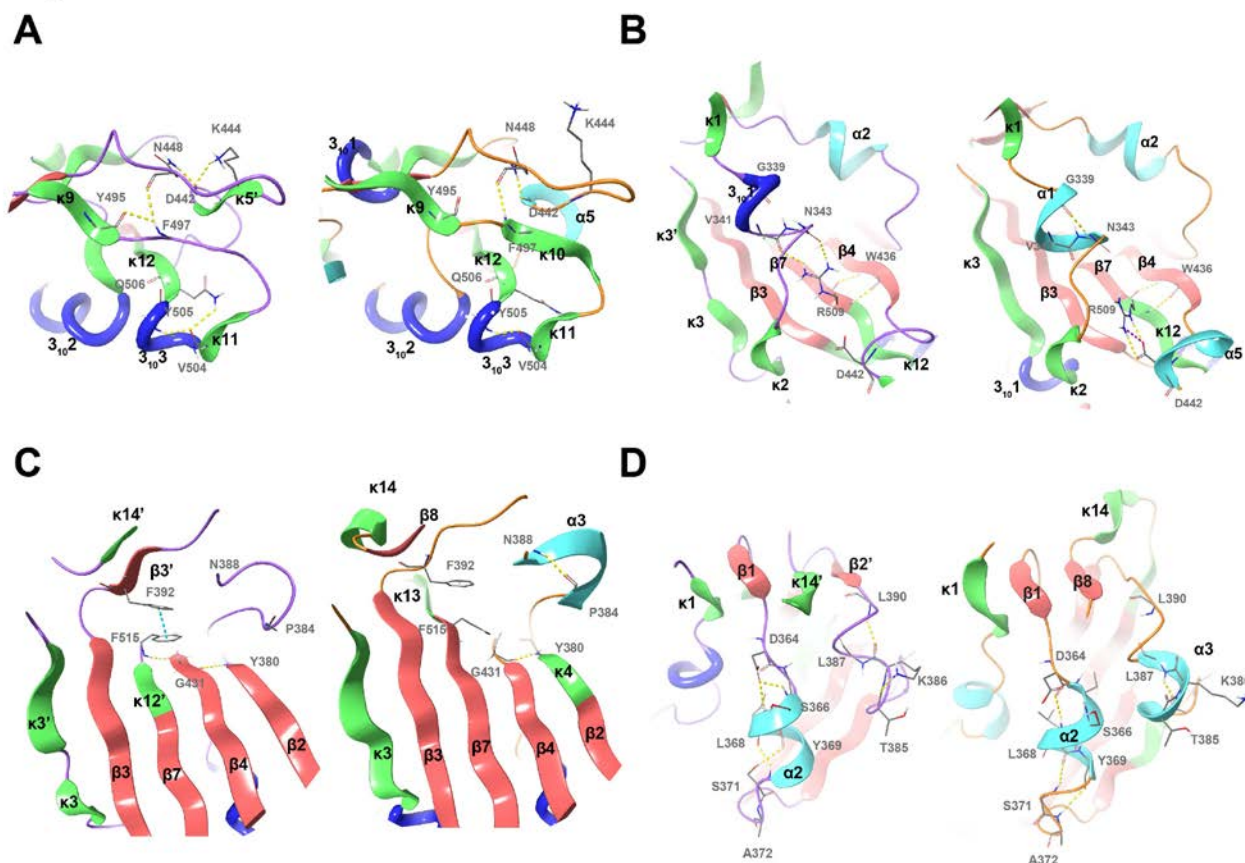
**Figure 3.** Intramolecular interactions induced by SARS-CoV-2-RBD and ACE2 complex. (A) Circos plot depicting the main differences in intramolecular interactions ( $< 3.5\text{\AA}$ ) between the unbound (6VXX:B and 6VYB:B) and bound (6M0J:E) conformations of the SARS-CoV-2-RBD. Track order: shift in the angular step ( $\vartheta$ ) per residue (i); shift in the rise ( $d$ ) per residue (ii); labels of secondary structural elements (iii); secondary structure assignment of the unbound-open conformation (iv); transformed secondary structure

of the bound conformation, representing the induced conformational changes (v). Gained and lost interactions are shown as magenta and yellow lines, respectively. (B) Circos plot depicting the main differences in main-chain H-bond profile between the unbound and bound conformations of the SARS-CoV-2-RBD. Track order: mean difference in the electrostatic interaction energy of the donor (i) and acceptor (ii); labels of secondary structural elements (iii); secondary structure assignment of the unbound-open conformation (iv); transformed secondary structure of the bound conformation, representing the induced conformational changes (v). Gained and lost H-bonds ( $|\Delta E| > 1\text{kcal/mol}$ ) illustrating the direction of donor-acceptor are shown as magenta and yellow arrows, respectively.

A closer inspection into the structural variations induced by the interaction SARS-CoV-2-RBD and ACE2 is shown in Figure 4. Upon binding, three H-bond networks comprising residues V504, Y505 with Q506, D442, Y495 with F497, and D442, N448 with K444, are destabilized (Figure 4A). The loss of H-bonds of Y495 and the side chains of K443 and Y505 is associated with conversion between  $\kappa 5'$  into a coil and the formation of a new  $\kappa$ -helix  $\kappa 10$ . This results in the advancement of D442 and the stabilization of  $\alpha$ -helix  $\alpha 5$ . The amino-acid R509 is rotated and switches between a  $\beta$ -strand ( $\beta 7$ ) to a  $\kappa$ -helix ( $\kappa 12$ ) while gaining a solid network of H-bonds and salt bridge with the side chain of D442 in the newly formed  $\alpha 5$ , at the expense of losing H-bonds with the backbone of V341 and F342 (Figure 4B). Consequently, a new H-bond is formed between N343 and G339, which converts  $3_{10}1'$  into  $\alpha 1$ . The adoption of a shorter-pitched  $\alpha$ -helix pulls  $\kappa 1$ , which constructs the N-terminus of the hinge region. The main-chain H-bond of F515 with G314 and the  $\pi$ - $\pi$  interaction with F392 are lost (Figure 4C) while a stronger pairing of H-bond forms between G431 and Y380 (Figure 3B). The rearrangement of the interactions is associated with conformational changes, including the conversion between  $\kappa 14'$  to  $\beta 8$ , which constructs the C-terminus of the hinge region. Also, the loop downstream to Y380 is displaced, and  $\alpha$ -helix  $\alpha 3$  is formed by a contribution of H-bond between the backbones of P384 and N388. The H-bond between L397 and 390 is lost, and the new  $\alpha 3$  packs closely together with  $\alpha 2$ , which includes VDW interactions between S366 and T385 (Figure 4D). Notably, this interaction is among the only gained contacts observed in the RBD, not involving the missing loops (Figure 3A). Also, a concentration

of H-bonds is re-distributed along  $\alpha 2$ , repositioning the  $\alpha$ -helix one step back (from 366-370 to 365-369). This transition is associated with a small movement of  $\beta 1$  at the medial region of the hinge.

**Figure 4**



**Figure 4.** Detailed structural comparison between the unbound and bound SARS-CoV-2-RBD and ACE2.

(A-D) Representative structures comparing the unbound-open (6VYB:B) and bound (6M0J:E)

conformations are shown as cartoon representation. Key contacts are labeled and shown as sticks. The secondary structural elements are labeled according to their occurrence in sequence.  $\alpha$ -helices (cyan),  $\beta$ -strands (red), and  $\kappa$ -helices (green) are illustrated as ribbons and  $3_{10}$ -helices (blue) as thick tubes.

### Dissociation between the RBD and CTD2

The  $\kappa 13/\kappa 14$  loop between residues 515 and 523 exhibits a relatively large conformational change between the unbound-closed and the bound states (Supplementary Figure 1).

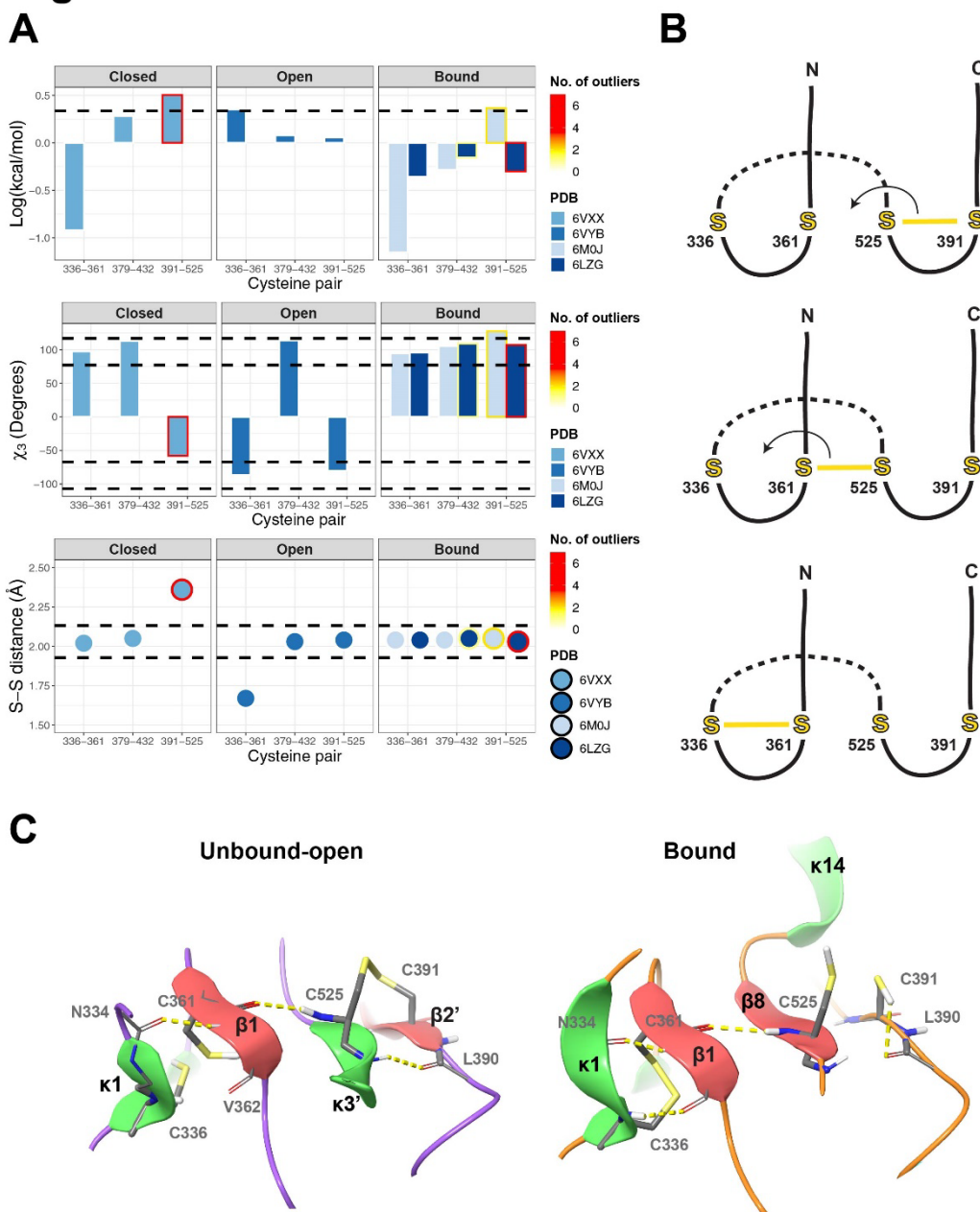


This loop is involved in the interaction between the RBD and CTD2 in the closed conformation. However, in the open conformation where RBD dissociates from CTD2, the loop is missing in both SARS-CoV and SARS-CoV-2 (Song, et al., 2018; Walls, et al., 2020; Wrapp, et al., 2020). The dissociation between the domains is required for the hinge and the RBD to move freely. Together, this indicates that the RBD-ACE2 complex stabilizes the  $\kappa 13/\kappa 14$  loop in a conformation that disfavors the interaction between RBD and CTD2.

### **Disulfide bond analysis**

The RBD-ACE2 complex inducible interactions culminate in the hinge region. The hinge, which we show to exhibit various conformational changes (Figure 3 and 4), contains two pairs of cysteines. Since cysteine pairs have the potential to act as allosteric switches (Bekendam, et al., 2016; Butera, et al., 2018; Chiu and Hogg, 2019), we hypothesized they are altered during the interaction between the S protein and ACE2 receptor. Therefore, we calculated the energy and geometrical features of the potential disulfide bonds in all high-quality structures of the RBD in SARS-CoV-2. Due to gross inconsistencies between disulfide assignments (Lavillette, et al., 2006; Song, et al., 2018; Walls, et al., 2020; Wrapp, et al., 2020; Yuan, et al., 2017) and quality issues of these pairs, we also characterized quality metrics of the cysteine pairs. Figure 5A shows an alternating pattern in the quality metrics of the N- (Cys<sup>336</sup>-Cys<sup>361</sup>) and C-terminal (Cys<sup>391</sup>-Cys<sup>525</sup>) cysteine pairs, except one structure (PDB 6VYB:B) in which the pair Cys<sup>336</sup>-Cys<sup>361</sup> in the open conformation is in the reduced form. This indicates that only one disulfide exists at a given state. Among cysteine pairs without assignment issues, the disulfide bond energy also shows an alternating pattern (Figure 5B), where the unbound-open and bound conformation have low energies in the N- and C-terminal cysteine pairs, respectively. This indicates that the cysteine pair switch between disulfide classes upon binding of the RBD to the receptor. Since the two cysteine pairs are aligned parallelly and closely together, the results suggest they are involved in disulfide shuffling (Figure 5B,C).

## Figure 5



**Figure 5.** Characterization of the conserved SARS-CoV-2-RBD cysteine pairs. (A) Shown are the disulfide bond energy (top),  $\chi_3$  torsional angle (middle), and S-S distance (bottom) in different conformational states of SARS-CoV-2-RBD (closed – 6VXX:B; open – 6VYB:B; bound – 6MOJ:E and 6LZG:B). Dashed lines show the threshold for ideal disulfide bonds (top panel – below the line; middle and bottom panel – between the lines). The number of outliers, including RSRZ (root-mean-square of Z score) outliers, non-rotameric sidechains, clashes, and bond length and angle outliers, were summarized

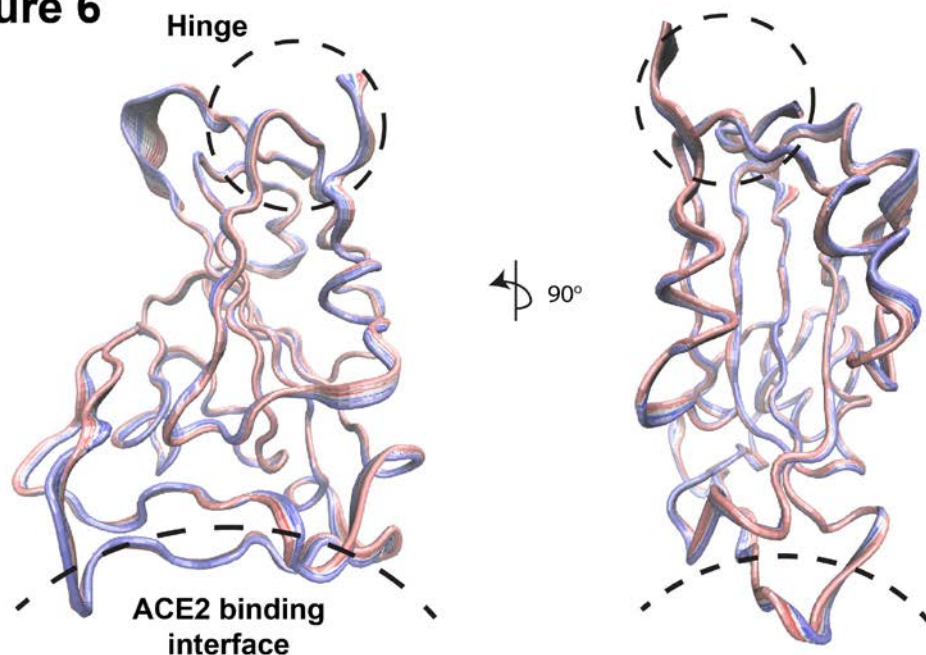


for each pair. Of note, the disulfide outliers were among the highest-ranked outliers of all the corresponding structure, indicating a poor fit to the density map. (B) Illustration of the predicted disulfide shuffling between the conserved four cysteine residues. (C) A cartoon representation of the hinge region with the two cysteine pairs depicting alternating disulfide bond configurations. The unbound-open conformation (left) demonstrates the original disulfide bond configuration. In the bound conformation (right), the Cys<sup>391-525</sup>, which poorly fit the density map, are shown as unpaired cysteines.

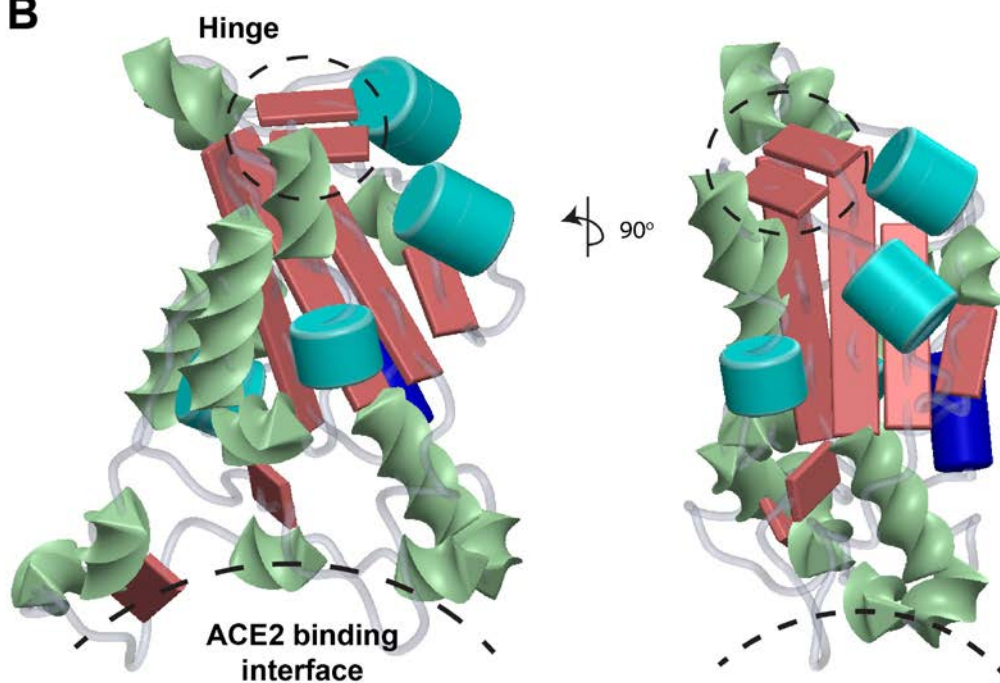
A summary of the conformational changes between the unbound and the bound RBD states are shown as an interpolated trajectory in Figure 6A. Also, a simplified model linking between the distal part and the hinge is depicted in Figure 6B.

## Figure 6

**A**



**B**



**Figure 6.** Dynamical features of the SARS-CoV-2-RBD. A summary of the conformational changes between the unbound-closed (6VXX:B in blue) and the bound (6M0J:E in red) RBD states are shown as an interpolated trajectory. (B) A simplified model, demonstrating the association between the secondary structures and the relationship between the distal part (ACE2 binding interface) and the hinge region.  $\alpha$ -helices (cyan),  $\beta$ -strands (red), and  $\kappa$ -helices (green) are illustrated as shown as 3D shapes.

## Discussion

The recent emergence of SARS-CoV-2 pandemic represents a major epidemiological challenge. ACE2 has been reported to be the receptor that initiates the activation of this novel CoV (Hoffmann, et al., 2020; Yan, et al., 2020). In this study, we determined the key structural components induced by the receptor and characterized their intramolecular interactions.

Numerous structures of the prefusion human CoV S proteins were determined at different states, and the key regions responsible for the interaction with the receptor were previously reported (Lan, et al., 2020; Shang, et al., 2020; Walls, et al., 2020; Wan, et al., 2020; Wrapp, et al., 2020; Yan, et al., 2020). However, to our knowledge, no previous study has extensively investigated the mechanism of transduction through the RBD. The structural transduction mechanism on a molecular level remains a difficult question to address experimentally. Therefore, we used current state-of-the-art structures of the SARS-CoV-2-RBD and focused on their structural organization. These structures represent snapshots of the dynamic S protein and allow to track and model the conformational transitions between the different states.

Characterizing the secondary structures of proteins is fundamental for gaining knowledge and simplifying the complicated 3D structures. We show that  $\kappa$ -helix is a predominant structure in the binding interface and in facilitating the conversion to the active form of the S protein. This conformation which is commonly known as PPII was recently designated as  $\kappa$ -helix, following the widespread criticism of the misleading name PPII (Adzhubei, et al., 2013; Hollingsworth, et al., 2009; Mansiaux, et al., 2011; Martin, et al., 2014; Meirson, et al., 2020; Meirson, et al., 2020). As many structures contain few prolines or none, the name 'polyproline' is considered inappropriate, and a more general term which abides the tradition of Latin letters to secondary structures was proposed (Meirson, et al., 2020). The role of  $\kappa$ -helix in propagating interactions and facilitating switch-like components coincides with the assessment that they represent 'functional blocks' as compared with other conformations such as  $\alpha$ -helices that often represent structural building blocks (Adzhubei, et al., 2013; Meirson, et al., 2020). The

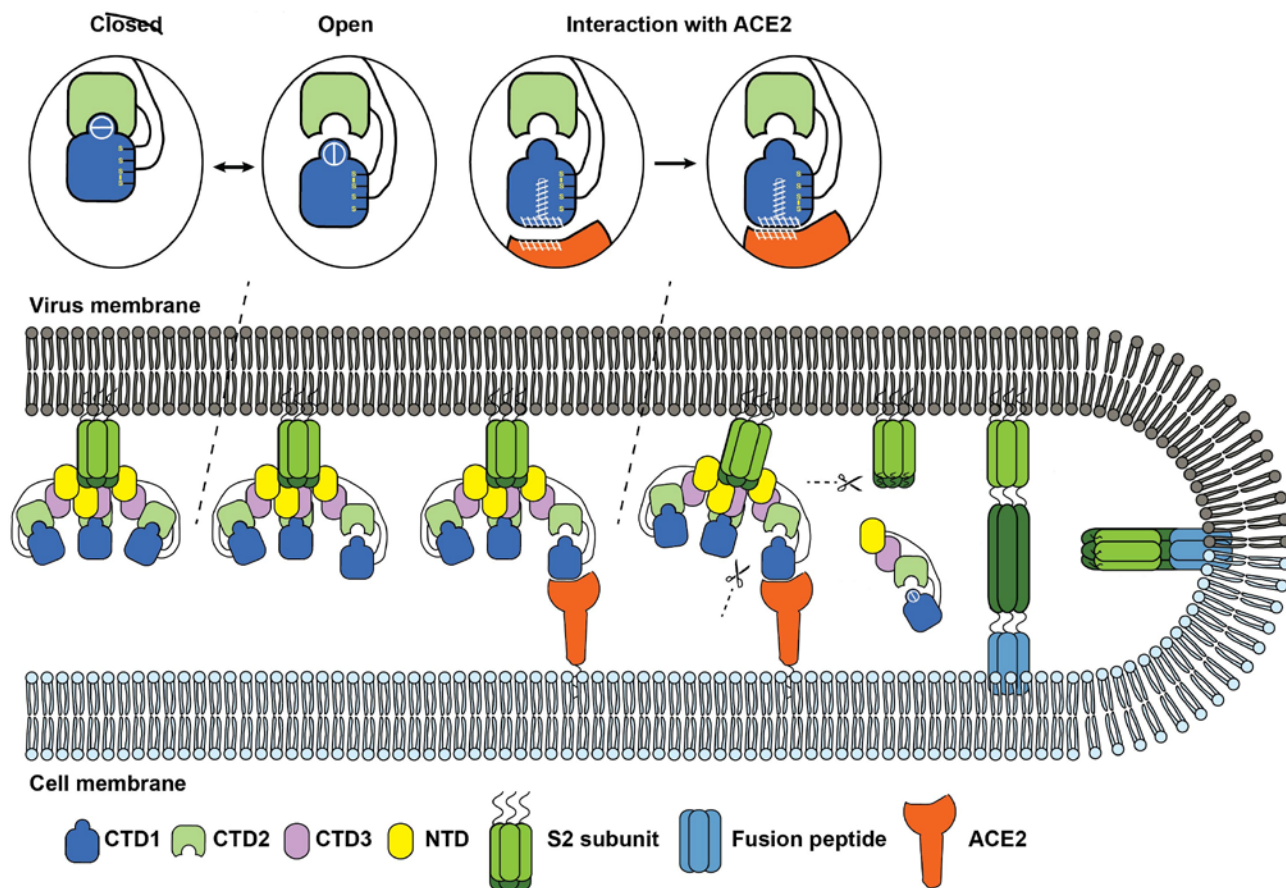
flexible and extended conformation of  $\kappa$ -helices, as well as non-regular H-bonds and preferred location on the surface of proteins, making them ideal elements for a wide range of molecular interactions (Cubellis, et al., 2005; Stapley and Creamer, 1999; Zagrovic, et al., 2005). However, despite being more common than most secondary structures, this conformation is often overlooked, apart from proline-rich regions. This is explained in part because it is not defined by H-bonds and is not assigned by the secondary structure assignment program employed in the PDB. Other reasons include a lack of graphical representation and its misleading historical name (Meirson, et al., 2020).

Our findings demonstrate that the high prevalence of  $\kappa$ -helix, as well as  $\beta$ -strands, are not unique to SARS-CoV-2 and appear to characterize other viruses. The conformational changes between different states of the SARS-CoV-2-RBD are associated with a typical transition between  $\beta$ -strand and  $\kappa$ -helix, as they are closely related in the torsional space (Hollingsworth and Karplus, 2010; Mansiaux, et al., 2011; Oh, et al., 2010). Therefore, we hypothesize that  $\kappa$ -helix could serve as an efficient evolutionary tool due to its flexible nature, which could adapt more quickly in the dynamic environment compared to more restricted secondary structures. In line with this suggestion, Austin et al. showed evolutionary conservation of  $\kappa$ -helix bias in intrinsically disordered regions that could be tuned by changing the distribution of  $\kappa$ -helix for multiple functions, including molecular recognition or allosteric regulation (Austin Elam, et al., 2013).

The hinge region was reported to facilitate the RBD motion and participate in the activation process (Gui, et al., 2017; Pallesen, et al., 2017; Song, et al., 2018; Walls, et al., 2020), and our structural analysis suggests that the conformational changes culminate at the hinge which contains four highly conserved cysteines (Shang, et al., 2020; Wang, et al., 2020). To explore a possible allosteric switching mechanism, we performed atomic comparisons of the cysteine pairs at different states of the S protein. Since inconsistencies in disulfide assignment exist and errors in structure determination are not uncommon (Carpentier, et al., 2010; Kleywegt and Jones, 1995; Villa and Lasker, 2014; Wlodawer, et al., 2008), we also assessed their geometric quality. Atomic details of the

structures at a pH range of 6.5-8.0 reveal alternating patterns in bond energy, geometric characteristics, and quality, between the pairs Cys<sup>336</sup>-Cys<sup>361</sup> and Cys<sup>391</sup>-Cys<sup>525</sup>, but not in Cys<sup>379-432</sup>. Nonetheless, Cys<sup>379-432</sup> displays a switch in the chirality of the disulfide bond in the bound state. These findings demonstrate a switching mechanism between disulfide bonds of Cys<sup>336</sup>-Cys<sup>361</sup> and Cys<sup>391</sup>-Cys<sup>525</sup>, where at each state (open, closed, or bound), only one pair of cysteine satisfies favorable disulfide configuration and quality criteria. The unfavorable configuration is associated with unphysical disulfide bond characteristics, high energy, poor quality, or their combination. This indicates that the distorted disulfide bonds entail substantial stress or that bond assignments were inaccurate, and these pairs are reduced. Both are possible as the cysteine pairs, located at the hinge, undergo significant conformational changes, and forced stretching of disulfide bonds is known to accelerate their cleavage (Zhou, et al., 2014). The four cysteine residues are adjacent and aligned suitably for disulfide exchange reactions. Such an arrangement makes it possible for a concerted series of disulfide exchange reactions to occur (Zhou, et al., 2008) and is supported by the alternating pattern of the Cys<sup>336</sup>-Cys<sup>361</sup> and Cys<sup>391</sup>-Cys<sup>525</sup> disulfide bond configurations. A proposed model of disulfide shuffling and SARS-CoV-2 viral entry is depicted in Figure 7.

## Figure 7



**Figure 7.** Proposed model for the pre- to post-fusion transition of SARS-CoV-2 S protein. The closed and open transitions are associated with the dissociation of RBD and CTD2 and conformational changes in the hinge region, including disulfide bond rearrangement. The complex of SARS-CoV-2-RBD with ACE2 induces conformational changes and rearrangement of the disulfide bonds that stabilize the S protein in the active form. The Activated S protein is cleaved by host proteases and induces the pre- to post-fusion transition of the S2 subunit, and initiates the fusion of viral and cellular membranes.

Beyond serving purely structural role, disulfide bonds can participate in redox reactions and act as allosteric switches controlling protein functions (Bekendam, et al., 2016; Butera, et al., 2018; Chiu and Hogg, 2019; Zhou, et al., 2014). Specific disulfide exchange reactions depend on a reducing agent such as thioredoxin or protein disulfide isomerase (PDI) (Zhou, et al., 2014). Rearrangement of disulfides (disulfide shuffling) can



also occur via intra-protein thiol-disulfide exchange reactions without additional agents, which depends on conformational changes (Chiu and Hogg, 2019; Zhang, et al., 2018). An increasing number of studies support an essential role for disulfide exchange in the entry of multiple viruses in susceptible cells (Stantchev, et al., 2012). In HIV, the attachment of gp120 subunit of the viral envelope (Env) to its primary receptor CD4, induces conformational changes that cause disulfide exchange in a PDI-dependent manner and is obligatory for triggering membrane-fusion process (Fenouillet, et al., 2007; Owen, et al., 2016; Stantchev, et al., 2012). More specifically, structural rearrangements of the S protein of CoV murine hepatitis virus (MHV) during cell interaction has been reported to affect cell entry using disulfide shuffling in the RBD (Gallagher, 1996; Weismiller, et al., 1990) as our result suggest for SARS-CoV-2. Surprisingly, Lavillette et al. showed that SARS-CoV S1 subunit is redox insensitive using chemical manipulation of the redox state, in contrast to various viruses including HIV and the CoV MHV (Fenouillet, et al., 2007; Lavillette, et al., 2006). However, the study utilized murine leukemia retrovirus (MLV) pseudotyped with S1 subunit, lacking the S2 subunit, a system with limited biological relevance as the subunits cooperate and form a tightly packed trimeric structure. Furthermore, the subunits remain non-covalently bound after proteolytic S1/S2 cleavage (Tortorici, et al., 2019; Walls, et al., 2016).

We propose that targeted redox exchange between conserved cysteine pairs in the S protein could conceptualize a new strategy in the development of high-affinity ligands against SARS-CoV-2, with important therapeutic implications. Recently, Hati et al. showed, using molecular dynamic simulations, that reducing all disulfide bonds in both ACE2 and SARS-CoV2 impairs their binding affinity (Bhattacharyay and Hati, 2020). However, more evidence is required to establish the role of redox potential and paired and unpaired cysteines in the S protein during viral entry. Also, our study is limited to a computational assessment of structures reconstructed using X-ray and cryoEM, and the implications of the observed structural rearrangements remain to be determined.

Currently, no efficient antivirals against SARS-CoV-2 or other CoVs are available, and numerous clinical trials are underway (Lythgoe and Middleton, 2020). In parallel, efforts



continue to develop antivirals and vaccines (Chen, et al., 2020; Liu, et al., 2020). Structure-based design of antivirals that efficiently recognize the target relies on understanding the main structural features, including local structural dynamics. Furthermore, developing selective therapies and efficient vaccines against adaptive evolutionary patterns of the virus poses a significant challenge. This challenge is amplified due to the persistency of the pandemic and the estimation that SARS-CoV-2 might continue to circulate in the population with renewed outbreaks (Kissler, et al., 2020; Tse, et al., 2020; Ye, et al., 2020). Our analysis has laid the major inducible structural features of the SARS-CoV-2-RBD and propose a new potential therapeutic strategy to block viral entry. Overall, this study may be helpful in guiding the development and optimization of structure-based intervention strategies that target SARS-CoV-2.

## **Conflict of interests**

The authors declare no conflict of interest.

## **Funding information**

Gal Markel is supported by the Samulei Foundation Grant for Integrative Immuno-Oncology and by the Israel Science Foundation IPMP Grant. Tomer Meirson is supported by the Foulkes Foundation fellowship for MD/PhD students.

## **Acknowledgements**

The authors would like to thank Haya and Nehemia Lemelbaum for their continuous generous support.

## References

- Adzhubei, A.A., Sternberg, M.J. and Makarov, A.A. Polyproline-II helix in proteins: structure and function. *Journal of molecular biology* 2013;425(12):2100-2132.
- Austin Elam, W., *et al.* Evolutionary conservation of the polyproline II conformation surrounding intrinsically disordered phosphorylation sites. *Protein Science* 2013;22(4):405-417.
- Bekendam, R.H., *et al.* A substrate-driven allosteric switch that enhances PDI catalytic activity. *Nature communications* 2016;7(1):1-11.
- Bell, J., *et al.* PrimeX and the Schrödinger computational chemistry suite of programs. *International Tables for Crystallography* 2006:534-538.
- Belouzard, S., Chu, V.C. and Whittaker, G.R. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proceedings of the National Academy of Sciences* 2009;106(14):5871-5876.
- Bhattacharyay, S. and Hati, S. Impact of Thiol-Disulfide Balance on the Binding of Covid-19 Spike Protein with Angiotensin Converting Enzyme 2 Receptor. *bioRxiv* 2020.
- Butera, D., *et al.* Autoregulation of von Willebrand factor function by a disulfide bond switch. *Science advances* 2018;4(2):eaq1477.
- Carpentier, P., *et al.* Raman-assisted crystallography suggests a mechanism of X-ray-induced disulfide radical formation and reparation. *Structure* 2010;18(11):1410-1419.
- Chen, W.-H., *et al.* The SARS-CoV-2 vaccine pipeline: an overview. *Current tropical medicine reports* 2020:1-4.
- Chiu, J. and Hogg, P.J. Allosteric disulfides: Sophisticated molecular structures enabling flexible protein regulation. *Journal of Biological Chemistry* 2019;294(8):2949-2960.
- Craig, D.B. and Dombkowski, A.A. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC bioinformatics* 2013;14(1):346.
- Cubellis, M., *et al.* Properties of polyproline II, a secondary structure element implicated in protein-protein interactions. *Proteins: Structure, Function, and Bioinformatics* 2005;58(4):880-892.
- de Wit, E., *et al.* SARS and MERS: recent insights into emerging coronaviruses. *Nature Reviews Microbiology* 2016;14(8):523.
- Drosten, C., *et al.* Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *New England journal of medicine* 2003;348(20):1967-1976.
- Fenouillet, E., Barbouche, R. and Jones, I.M. Cell entry by enveloped viruses: redox considerations for HIV and SARS-coronavirus. *Antioxidants & redox signaling* 2007;9(8):1009-1034.
- Gallagher, T.M. Murine coronavirus membrane fusion is blocked by modification of thiols buried within the spike protein. *Journal of virology* 1996;70(7):4683-4690.
- Gates, B. Responding to Covid-19—a once-in-a-century pandemic? *New England Journal of Medicine* 2020;382(18):1677-1679.
- Ge, X.-Y., *et al.* Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 2013;503(7477):535-538.
- Goddard, T.D., *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science* 2018;27(1):14-25.

- Grant, B.J., *et al.* Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* 2006;22(21):2695-2696.
- Guan, Y., *et al.* Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science* 2003;302(5643):276-278.
- Gui, M., *et al.* Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell research* 2017;27(1):119-129.
- Haagmans, B.L., *et al.* Middle East respiratory syndrome coronavirus in dromedary camels: an outbreak investigation. *The Lancet infectious diseases* 2014;14(2):140-145.
- Han, D.P., Penn-Nicholson, A. and Cho, M.W. Identification of critical determinants on ACE2 for SARS-CoV entry and development of a potent entry inhibitor. *Virology* 2006;350(1):15-25.
- Hoffmann, M., *et al.* SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 2020.
- Hollingsworth, S.A., Berkholz, D.S. and Karplus, P.A. On the occurrence of linear groups in proteins. *Protein Science* 2009;18(6):1321-1325.
- Hollingsworth, S.A. and Karplus, P.A. A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. *Biomolecular concepts* 2010;1(3-4):271-283.
- Huang, C., *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The lancet* 2020;395(10223):497-506.
- Humphrey, W., Dalke, A. and Schulten, K. VMD: visual molecular dynamics. *Journal of molecular graphics* 1996;14(1):33-38.
- Kan, B., *et al.* Molecular evolution analysis and geographic investigation of severe acute respiratory syndrome coronavirus-like virus in palm civets at an animal market and on farms. *Journal of virology* 2005;79(18):11892-11900.
- Kissler, S.M., *et al.* Social distancing strategies for curbing the COVID-19 epidemic. *medRxiv* 2020.
- Kleywegt, G.J. and Jones, T.A. Where freedom is given, liberties are taken. *Structure* 1995;3(6):535-540.
- Ksiazek, T.G., *et al.* A novel coronavirus associated with severe acute respiratory syndrome. *New England journal of medicine* 2003;348(20):1953-1966.
- Lan, J., *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* 2020:1-6.
- Lavillette, D., *et al.* Significant redox insensitivity of the functions of the SARS-CoV spike glycoprotein comparison with HIV envelope. *Journal of Biological Chemistry* 2006;281(14):9200-9204.
- Letko, M., Marzi, A. and Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nature microbiology* 2020;5(4):562-569.
- Li, F. Receptor recognition and cross-species infections of SARS coronavirus. *Antiviral research* 2013;100(1):246-254.
- Li, W., *et al.* Bats are natural reservoirs of SARS-like coronaviruses. *Science* 2005;310(5748):676-679.
- Liu, W., *et al.* Learning from the past: possible urgent prevention and treatment options for severe acute respiratory infections caused by 2019-nCoV. *Chembiochem* 2020.
- Lythgoe, M.P. and Middleton, P. Ongoing clinical trials for the management of the COVID-19 pandemic. *Trends in Pharmacological Sciences* 2020.

- Mansiaux, Y., *et al.* Assignment of PolyProline II conformation and analysis of sequence–structure relationship. *PLoS one* 2011;6(3).
- Martin, C., *et al.* Silaproline helical mimetics selectively form an all-trans PPII helix. *Chemistry–A European Journal* 2014;20(44):14240-14244.
- Meirson, T., *et al.* A helical lock and key model of polyproline II conformation with SH3. *Bioinformatics* 2020;36(1):154-159.
- Meirson, T., *et al.*  $\kappa$ -helix and the helical lock and key model: A pivotal way of looking at polyproline II. *Bioinformatics* 2020.
- Memish, Z.A., *et al.* Middle East respiratory syndrome coronavirus in bats, Saudi Arabia. *Emerging infectious diseases* 2013;19(11):1819.
- Millet, J.K. and Whittaker, G.R. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *Proceedings of the National Academy of Sciences* 2014;111(42):15214-15219.
- Miyazawa, T. Molecular vibrations and structure of high polymers. II. Helical parameters of infinite polymer chains as functions of bond lengths, bond angles, and internal rotation angles. *Journal of Polymer Science* 1961;55(161):215-231.
- Oh, K.I., *et al.* Circular dichroism eigenspectra of polyproline II and  $\beta$ -strand conformers of trialanine in water: Singular value decomposition analysis. *Chirality* 2010;22(1E):E186-E201.
- Owen, G.R., *et al.* Human CD4 metastability is a function of the allosteric disulfide bond in domain 2. *Biochemistry* 2016;55(15):2227-2237.
- Pallesen, J., *et al.* Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proceedings of the National Academy of Sciences* 2017;114(35):E7348-E7357.
- Pettersen, E.F., *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry* 2004;25(13):1605-1612.
- Shang, J., *et al.* Structural basis of receptor recognition by SARS-CoV-2. *Nature* 2020:1-4.
- Song, W., *et al.* Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLoS pathogens* 2018;14(8):e1007236.
- Spek, A. *Acta Crystallogr., Sect. A.* 1990;46:34.
- Stantchev, T.S., *et al.* Cell-type specific requirements for thiol/disulfide exchange during HIV-1 entry and infection. *Retrovirology* 2012;9(1):97.
- Stapley, B.J. and Creamer, T.P. A survey of left-handed polyproline II helices. *Protein Science* 1999;8(3):587-595.
- Su, S., *et al.* Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends in microbiology* 2016;24(6):490-502.
- To, K.K.-W., *et al.* Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *The Lancet Infectious Diseases* 2020.
- Tortorici, M.A., *et al.* Structural basis for human coronavirus attachment to sialic acid receptors. *Nature structural & molecular biology* 2019;26(6):481-489.
- Touw, W.G., *et al.* A series of PDB-related databanks for everyday needs. *Nucleic acids research* 2015;43(D1):D364-D368.
- Tse, L.V., *et al.* The Current and Future State of Vaccines, Antivirals and Gene Therapies Against Emerging Coronaviruses. *Frontiers in Microbiology* 2020;11:658.

- Villa, E. and Lasker, K. Finding the right fit: chiseling structures out of cryo-electron microscopy maps. *Current opinion in structural biology* 2014;25:118-125.
- Walls, A.C., *et al.* Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* 2020.
- Walls, A.C., *et al.* Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature* 2016;531(7592):114-117.
- Wan, Y., *et al.* Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus. *Journal of virology* 2020;94(7).
- Wang, C., *et al.* A novel coronavirus outbreak of global health concern. *The Lancet* 2020;395(10223):470-473.
- Wang, Q., *et al.* Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell* 2020.
- Weismiller, D., *et al.* Monoclonal antibodies to the peplomer glycoprotein of coronavirus mouse hepatitis virus identify two subunits and detect a conformational change in the subunit released under mild alkaline conditions. *Journal of virology* 1990;64(6):3051-3055.
- White, J.M., *et al.* Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Critical reviews in biochemistry and molecular biology* 2008;43(3):189-219.
- WHO. Summary of probably SARS cases with onset of illness from 1 November 2002 to 31 July 2003. In.; 2004.
- WHO. Middle East respiratory syndrome coronavirus (MERS-CoV). In.; 2020.
- WHO. WHO Coronavirus Disease (COVID-19) Dashboard. In.; 2020.
- Wiedemann, C., *et al.* Cysteines and Disulfide Bonds as Structure-Forming Units: Insights From Different Domains of Life and the Potential for Characterization by NMR. *Frontiers in Chemistry* 2020;8:280.
- Wlodawer, A., *et al.* Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *The FEBS journal* 2008;275(1):1-21.
- Wrapp, D., *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 2020;367(6483):1260-1263.
- Yan, R., *et al.* Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* 2020;367(6485):1444-1448.
- Ye, Z.-W., *et al.* Zoonotic origins of human coronaviruses. *International journal of biological sciences* 2020;16(10):1686.
- Yoon, I.-K. and Kim, J.H. First clinical trial of a MERS coronavirus DNA vaccine. *The Lancet. Infectious Diseases* 2019;19(9):924.
- Yuan, Y., *et al.* Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nature communications* 2017;8:15092.
- Zagrovic, B., *et al.* Unusual compactness of a polyproline type II structure. *Proceedings of the National Academy of Sciences* 2005;102(33):11698-11703.
- Zaki, A.M., *et al.* Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *New England Journal of Medicine* 2012;367(19):1814-1820.
- Zhang, W., *et al.* Intra- and inter-protein couplings of backbone motions underlie protein thiol-disulfide exchange cascade. *Scientific reports* 2018;8(1):1-14.

Zhang, Y. and Sagui, C. Secondary structure assignment for conformationally irregular peptides: Comparison between DSSP, STRIDE and KAKSI. *Journal of Molecular Graphics and Modelling* 2015;55:72-84.

Zhou, B., *et al.* Identification of allosteric disulfides from prestress analysis. *Biophysical journal* 2014;107(3):672-681.

Zhou, P., *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *nature* 2020;579(7798):270-273.

Zhou, Y., *et al.* NMR solution structure of the integral membrane enzyme DsbB: functional insights into DsbB-catalyzed disulfide bond formation. *Molecular cell* 2008;31(6):896-908.

Zhu, N., *et al.* A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine* 2020.