

# Coevolutionary Analysis and Perturbation-Based Network Modeling of the SARS-CoV-2 Spike Protein Complexes with Antibodies: Binding-Induced Control of Dynamics, Allosteric Interactions and Signaling

Gennady M. Verkhivker,<sup>1,2 ‡</sup> Luisa Di Paola<sup>3</sup>

<sup>1</sup>Keck Center for Science and Engineering, Schmid College of Science and Technology,  
Chapman University, One University Drive, Orange, CA 92866, USA

<sup>2</sup>Department of Biomedical and Pharmaceutical Sciences, Chapman University School of  
Pharmacy, Irvine, CA 92618, USA

<sup>3</sup>Unit of Chemical-Physics Fundamentals in Chemical Engineering, Department of Engineering,  
Università Campus Bio-Medico di Roma, via Álvaro del Portillo 21, 00128 Rome, Italy

‡corresponding author

E-mail: [verkhivk@chapman.edu](mailto:verkhivk@chapman.edu)

## Abstract

The structural and biochemical studies of the SARS-CoV-2 spike glycoproteins and complexes with highly potent antibodies have revealed multiple conformation-dependent epitopes highlighting the link between conformational plasticity of spike proteins and capacity for eliciting specific binding and broad neutralization responses. In this study, we used coevolutionary analysis, molecular simulations, and perturbation-based hierarchical network modeling of the SARS-CoV-2 S complexes with H014, S309, S2M11 and S2E12 antibodies targeting distinct epitopes to explore molecular mechanisms underlying binding-induced modulation of dynamics, stability and allosteric signaling in the spike protein trimers. The results of this study revealed key regulatory centers that can govern allosteric interactions and communications in the SARS-CoV-2 spike proteins. Through coevolutionary analysis of the SARS-CoV-2 spike proteins, we identified highly coevolving hotspots and functional clusters forming coevolutionary networks. The results revealed significant coevolutionary couplings between functional regions separated by the medium-range distances which may help to facilitate a functional cross-talk between distant allosteric regions in the SARS-CoV-2 spike complexes with antibodies. We also discovered a potential mechanism by which antibody-specific targeting of coevolutionary centers can allow for efficient modulation of allosteric interactions and signal propagation between remote functional regions. Using a hierarchical network modeling and perturbation-response scanning analysis, we demonstrated that binding of antibodies could leverage direct contacts with coevolutionary hotspots to allosterically restore and enhance couplings between spatially separated functional regions, thereby protecting the spike apparatus from membrane fusion. The results of this study also suggested that antibody binding can induce a switch from a moderately cooperative population-shift mechanism, governing structural changes of the ligand-

free SARS-CoV-2 spike protein, to antibody-induced highly cooperative mechanism that can better withstand mutations in the functional regions without significant deleterious consequences for protein function. This study provides a novel insight into allosteric regulatory mechanisms of SARS-CoV-2 S proteins, showing that antibodies can modulate allosteric interactions and signaling of spike proteins, providing a plausible strategy for therapeutic intervention by targeting specific hotspots of allosteric interactions in the SARS-CoV-2 proteins.

## Introduction

The coronavirus disease 2019 (COVID-19) pandemic associated with the severe acute respiratory syndrome (SARS)<sup>1-5</sup> has been at the focal point of biomedical research. SARS-CoV-2 infection is transmitted when the viral spike (S) glycoprotein binds to the host cell receptor leading to the entry of S protein into host cells and membrane fusion.<sup>6-8</sup> The full-length SARS-CoV-2 S protein consists of two main domains, amino (N)-terminal S1 subunit and carboxyl (C)-terminal S2 subunit. The subunit S1 is involved in the interactions with the host receptor and includes an N-terminal domain (NTD), the receptor-binding domain (RBD), and two structurally conserved subdomains (SD1 and SD2). Structural and biochemical studies have shown that the mechanism of virus infection may involve spontaneous conformational transformations of the SARS-CoV-2 S protein between a spectrum of closed and receptor-accessible open forms, where RBD continuously switches between “down” and “up” positions where the latter can promote binding with the host receptor ACE2.<sup>9-11</sup> The crystal structures of the S-RBD in the complexes with human ACE2 enzyme revealed structurally conserved binding mode shared by the SARS-CoV and SARS-CoV-2 proteins in which an extensive interaction network is formed by the receptor binding motif (RBM) of the RBD region.<sup>12-16</sup> These studies established that binding of the SARS-CoV-RBD to the ACE2 receptor can be a critical initial step for virus entry into target cells. The rapidly growing body of cryo-EM structures of the SARS-CoV-2 S proteins detailed distinct conformational arrangements of S protein trimers in the prefusion form that are manifested by a dynamic equilibrium between the closed (“RBD-down”) and the receptor-accessible open (“RBD-up”) form required for the S protein fusion to the viral membrane.<sup>17-26</sup> The cryo-EM characterization of the SARS-CoV-2 S trimer demonstrated that S protein may populate a spectrum of closed states by fluctuating between structurally rigid locked-closed form

and more dynamic, closed states preceding a transition to the fully open S conformation.<sup>26</sup> Structural and biophysical studies employed protein engineering to generate prefusion-stabilized SARS-CoV-2 S variants by introducing disulfide bonds and proline mutations to modulate stability of the S2 subunit and the inter-subunit boundaries, and consequently prevent refolding changes that accompany acquisition of the postfusion state.<sup>27</sup> By combining targeted mutagenesis and cryo-EM structure determination, recent biophysical investigations demonstrated that modifications in the contact regions between the RBD and S2 domains via S383C/D985C double mutation can lead to the thermodynamic prevalence of the closed-down conformation, while the quadruple mutant (A570L/T572I/F855Y/N856I) perturbing the inter-protomer contacts can shift the equilibrium towards the open form with the enhanced binding propensities for the ACE2 host receptor.<sup>28</sup> Protein engineering and cryo-EM studies of a prefusion-stabilized SARS-CoV-2 S ectodomain trimer using the inter-protomer disulfide bonds (S383C/D985C, G413C/P987C, T385C/T415C) between RBD and S2 regions can lock the trimer in the closed state and enhance the SARS-CoV-2 S resistance to proteolysis.<sup>29</sup> Targeted design of thermostable SARS-CoV-2 spike trimers further specified how disulfide-bonded S-protein trimer variants imposing stabilization in the strategically located inter-protomer positions (S383-D385) and (G413-V987) can promote dramatic thermodynamic shifts towards the prefusion closed states with only ~20% of the population corresponding to the open state.<sup>30</sup> Recent biochemical studies of the SARS-CoV-2 S mutants with the enhanced infectivity profile<sup>39-41</sup> discovered that a highly active D614G mutation can exert its dramatic functional effect on virus infectivity by radically shifting the population of the SARS-CoV-2 S trimer towards open states.<sup>31</sup> The cryo-EM and sophisticated tomography tools determined the high-resolution structure and distribution of S trimers in situ on the virion surface.<sup>32</sup> These studies confirmed a general mechanism of population

shifts between different functional states of the SARS-CoV-2 S trimers, suggesting that RBD epitopes can become stochastically exposed to the interactions with the host receptor ACE2. Biophysical analysis of SARS-CoV-2 S trimer on virus particles revealed four distinct conformational states for the S protein and a sequence of conformational transitions through an obligatory intermediate in which all three RBD domains in the closed conformations are oriented towards the viral particle membrane.<sup>33</sup> Cryo-EM structural studies also mapped a mechanism of conformational events associated with ACE2 binding, showing that the compact closed form of the SARS-CoV-2 S protein becomes weakened after furin cleavage between the S1 and S2 domains, leading to the increased population of partially open states and followed by ACE2 recognition that can accelerate transformation to a fully open and ACE2-bound form priming the protein for fusion activation.<sup>34</sup>

The early biochemical studies of SARS S proteins with antibodies (Abs) suggested that RBD regions of S proteins contain multiple conformation-dependent epitopes capable of inducing potent neutralizing Ab responses, thus revealing the link between conformational heterogeneity of S proteins and capacity for eliciting binding with highly potent neutralizing Abs.<sup>35</sup> Subsequently, it was shown that major neutralizing epitopes of SARS-CoV may have been preserved during cross-species transmission, and that RBD-targeted Abs have a potential for broad protection against both human and animal SARS-CoV variants.<sup>36</sup> The SARS-CoV-2 S protein-targeting monoclonal antibodies (mAbs) with potent neutralizing activity are of paramount importance and are actively pursued as therapeutic interventions for COVID-19 virus.<sup>37-40</sup> The rapidly growing structural studies of SARS-CoV and SARS-CoV-2-neutralizing Abs targeting the RBD have suggested potential mechanisms underlying inhibition of the association between the S protein and ACE2 host receptor. The early structure of SARS-CoV-

RBD complex with a neutralizing Ab 80R showed that the epitope on the S1 RBD overlapped closely with the ACE2-binding site, suggesting that a direct interference mechanism may be responsible for the neutralizing activity.<sup>41</sup> However, several SARS-CoV-specific neutralizing Abs such as m396, 80R, and F26G19 that block the RBM motif in the open S conformation did not exhibit a strong neutralizing activity against SARS-CoV-2 protein. The crystal structure of a neutralizing Ab CR3022 in the complex with the SARS-CoV-2 S-RBD revealed binding to a highly conserved epitope that is located away from the ACE2-binding site but could only be accessed when two RBDs adopt the “up” conformation.<sup>42</sup> Subsequent structural and surface plasmon resonance studies confirmed that CR3022 binds the RBD of SARS-CoV-2 displaying strong neutralization by allosterically perturbing the interactions between the RBD regions and ACE2 receptor.<sup>43</sup> The proposed neutralization mechanism of SARS-CoV-2 through destabilization of the prefusion S conformation can provide a resistance mechanism to virus escape which can be contrasted with Abs directly competing for the ACE2-binding site and often susceptible to immune evasion. Potent neutralizing Abs from COVID-19 patients examined through electron microscopy studies confirmed that the SARS-CoV-2 S protein features multiple distinct antigenic sites, including RBD-based and non-RBD epitopes.<sup>44</sup> These studies also suggested that some Abs may function by allosterically interfering with the host receptor binding and causing conformational changes in the S protein that can obstruct other epitopes and block virus infection without directly interfering with ACE2 recognition. Cryo-EM characterization of the SARS-CoV-2 S trimer in complex with the H014 Fab fragment revealed a new conformational epitope that is accessible only when the RBD is in the up conformation.<sup>45</sup> Biochemical and virological studies demonstrated that H014 prevents attachment of SARS-CoV-2 to the host cell receptors and can exhibit broad cross-neutralization activities by leveraging conserved nature of

the RBD epitope and a partial overlap with ACE2-binding region. The recently reported mAb S309 potentially neutralizes both SARS-CoV-2 and SARS-CoV through binding to a conserved RBD epitope which is distinct from the RBM region and accessible in both open and closed states, so that there is no competition between S309 and ACE2 for binding to the SARS-CoV-2 S protein.<sup>46</sup> Two ultra-potent Abs S2M11 and S2E12 targeting the overlapping RBD epitopes were recently reported, revealing Ab-specific modulation of protein responses and adaptation of different functional states for the S trimer.<sup>47</sup> Cryo-EM structures showed that S2M11 can recognize and stabilize S protein in the closed conformation by binding to a quaternary epitope spanning two RBDs of the adjacent protomers in the S trimer, while S2E12 binds to a tertiary epitope contained within one S protomer and shifts the conformational equilibrium towards a fully open S trimer conformation.<sup>47</sup> The mAbs isolated from 10 convalescent COVID-19 patients showed neutralizing activities against authentic SARS-CoV-2, with the mAb 4A8 displaying high potency by binding to the NTD of the S protein conformation with one RBD in “up” conformation and the other two RBDs in “down” conformation.<sup>48</sup> Interestingly, none of the isolated mAbs recognize the RBD and inhibit binding of SARS-CoV-2 S protein to ACE2, suggesting that allosterically regulated mechanisms may underlie the functional effects and experimentally observed efficient cross-neutralization.<sup>48</sup> Moreover, it was proposed that combining NTD-targeting 4A8 with RBD-targeting Abs may help in the design of “cocktail” therapeutics to combat the escaping mutations of the virus.

The most recent investigation reported discovery of an ultra-potent synthetic nanobody Nb6 that neutralizes SARS-CoV-2 by stabilizing the fully inactive down S conformation preventing binding with ACE2 receptor.<sup>49</sup> Affinity maturation and structure-guided design produced a trivalent nanobody, mNb6-tri that simultaneously binds to all three RBDs, yielding the



remarkably high affinity for S protein and completely blocking the S-ACE2 interactions by occupying the binding site and locking spike protein in the inactive, receptor-inaccessible state. In general, Abs tend to bind to the most easily accessible regions of the virus, where viruses can tolerate mutations and thereby escape immune challenge. The emerging body of recent studies suggested that properly designed cocktails of Abs can provide a broad and efficient cross-neutralization effects through synergistic targeting of conserved and more variable SARS-CoV-2 RBD epitopes, thereby offering a robust strategy to combat virus resistance.<sup>45-49</sup>

Computational modeling and molecular dynamics (MD) simulations have been instrumental in predicting conformational and energetic mechanisms of SARS-CoV-2 functions.<sup>50-55</sup> Microsecond, all-atom MD simulations of the full-length SARS-CoV-2 S glycoprotein embedded in the viral membrane, with a complete glycosylation profile were recently reported, providing the unprecedented level of details about open and closed structures.<sup>51</sup> MD simulations of the SARS-CoV-2 spike glycoprotein identified differences in flexibility of functional regions that may be important for modulating the equilibrium changes and binding to ACE2 host receptor.<sup>52</sup> Computational studies examined SARS-CoV-2 S trimer interactions with ACE2 enzyme using the recent crystal structures<sup>53-62</sup> providing insights into the key determinants of the binding affinity and selectivity. A comprehensive study employed MD simulations to reveal a balance of hydrophobic interactions and elaborate hydrogen-bonding network in the SARS-CoV-2-RBD interface.<sup>59</sup> Molecular mechanisms of the SARS-CoV-2 binding with ACE2 enzyme were analyzed in our recent study using coevolution and conformational dynamics.<sup>62</sup> Using protein contact networks and perturbation response scanning based on elastic network models, we recently discovered existence of allosteric sites on the SARS-CoV-2 spike protein.<sup>63</sup> By using molecular simulations and network modeling we recently presented the first evidence that the

SARS-CoV-2 spike protein can function as an allosteric regulatory engine that fluctuates between dynamically distinct functional states.<sup>64</sup>

In this study, we used a battery of computational approaches to explore and simulate molecular mechanisms underlying responses of the SARS-CoV-2 S proteins to binding of a panel of Abs (H014, S309, S2M11 and S2E12) that target distinct epitopes in the RBD regions. Using coevolutionary analysis, molecular simulations, and perturbation-based hierarchical network modeling of the SARS-CoV-2 S complexes with these Abs, we examined binding-induced modulation of dynamics, stability and allosteric interactions in the S protein trimers. The results of this study revealed structural topography of coevolutionary couplings and network connectivity that may determine mechanisms of allosteric signaling in the SARS-CoV-2 S proteins. We show that specific Ab targeting of conserved centers and coevolutionary hotspots in the S protein that are distinct from RBM region can allow not only for productive binding, but also for efficient Ab-induced modulation of long-range interactions between the S1 and S2 units of the SARS-CoV-2 S protein. Using perturbation-based network modeling, we find that targeted binding of the Abs could leverage direct contacts with coevolutionary hotspots to effectively restore allosteric potential of the S1 regions in the open states, thereby strengthening the allosteric interaction network and protecting the S protein machinery from dissociation of S1 subunit required for membrane fusion. The results of this study provide a novel insight into allosteric regulatory mechanisms of SARS-CoV-2 S proteins showing that the examined Abs can uniquely modulate signal communication providing a plausible strategy for therapeutic intervention by targeting specific hotspots of allosteric interactions in the SARS-CoV-2 proteins.

## Materials and Methods

### Sequence Conservation and Coevolutionary Analyses

Multiple sequence alignment (MSA) was obtained using the MAFFT approach<sup>65</sup> and homologues were obtained from UNIREF90.<sup>66,67</sup> We employed Kullback-Leibler (KL) sequence conservation score *KLConsScore* using MSA profiles generated by hidden Markov models in Pfam database for the SARS-CoV S glycoproteins.<sup>68,69</sup> Three Pfam domains were utilized corresponding to S1, the NTD (*bCoV\_S1\_N*, Betacoronavirus-like spike glycoprotein S1, N-terminal, Pfam:PF16451, Uniprot SPIKE\_CVHSA, pdb id 6CS0, residues 33-324), the RBD (*bCoV\_S1\_RBD*, Betacoronavirus spike glycoprotein S1, receptor binding, Pfam:PF09408, Uniprot SPIKE\_CVHSA, pdb id 6CS0, residues 335-512) and the new C-terminal domain, CTD (*CoV\_S1\_C* Coronavirus spike glycoprotein S1, C-terminal. Pfam:PF19209, Uniprot SPIKE\_CVHSA, pdb id 6CS0, residues 522-580). S2 is described in the family Pfam:PF01601 (Uniprot SPIKE\_CVHSA, pdb id 6CS0, residues 622-1120) which contains an additional S2' cleavage site, a fusion peptide, internal fusion peptide, heptad repeat 1/2 domains, and the transmembrane domain. The following Uniprot entries were used for sequence analysis : P59594: SPIKE\_SARS (previously SPIKE\_CVHSA) (pdb id 6CS0) and P0DTC2: SPIKE\_SARS2 (pdb id 6VXX, 6VYB).

The KL conservation is calculated according to the following formula:

$$KLConsScore_i = \sum_{i=1}^N \ln \frac{P(i)}{Q(i)} \quad (1)$$

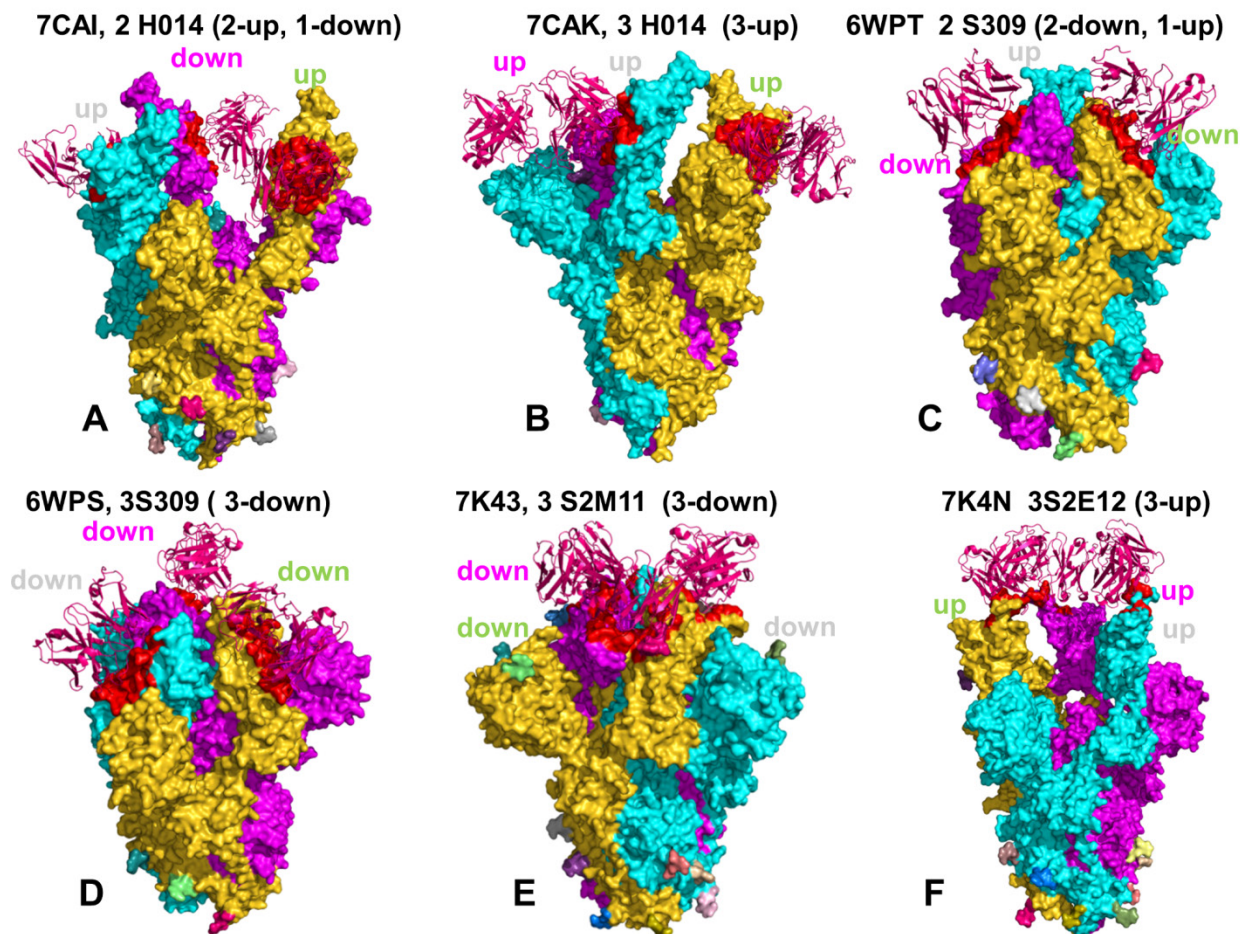
Here,  $P(i)$  is the frequency of amino acid  $i$  in that position and  $Q(i)$  is the background frequency of the amino acid in nature calculated using an amino acids background frequency distribution obtained from the UniProt database.<sup>70</sup> To evaluate coevolutionary couplings in the SARS-CoV-

2 S glycoproteins, we used MISTIC approach<sup>71-73</sup> in which computation of residue covariation was done with three different direct coupling analysis (DCA) methods: mean field DCA (mfDCA),<sup>74-76</sup> pseudo-likelihood maximization DCA (plmDCA)<sup>77,78</sup> and multivariate Gaussian modeling DCA (gaussianDCA).<sup>79,80</sup> For each residue, we computed cumulative covariation score (CScore) parameter, that evaluates to what degree a given position participates in the coevolutionary network. CScore is a derived score per position that characterizes the extent of coevolutionary couplings shared by a given residue. This score is calculated as the sum of covariation scores above a certain threshold (typically top 5% of the covariation scores) for every position pair where the particular position appears.

### **Coarse-Grained Molecular Simulations**

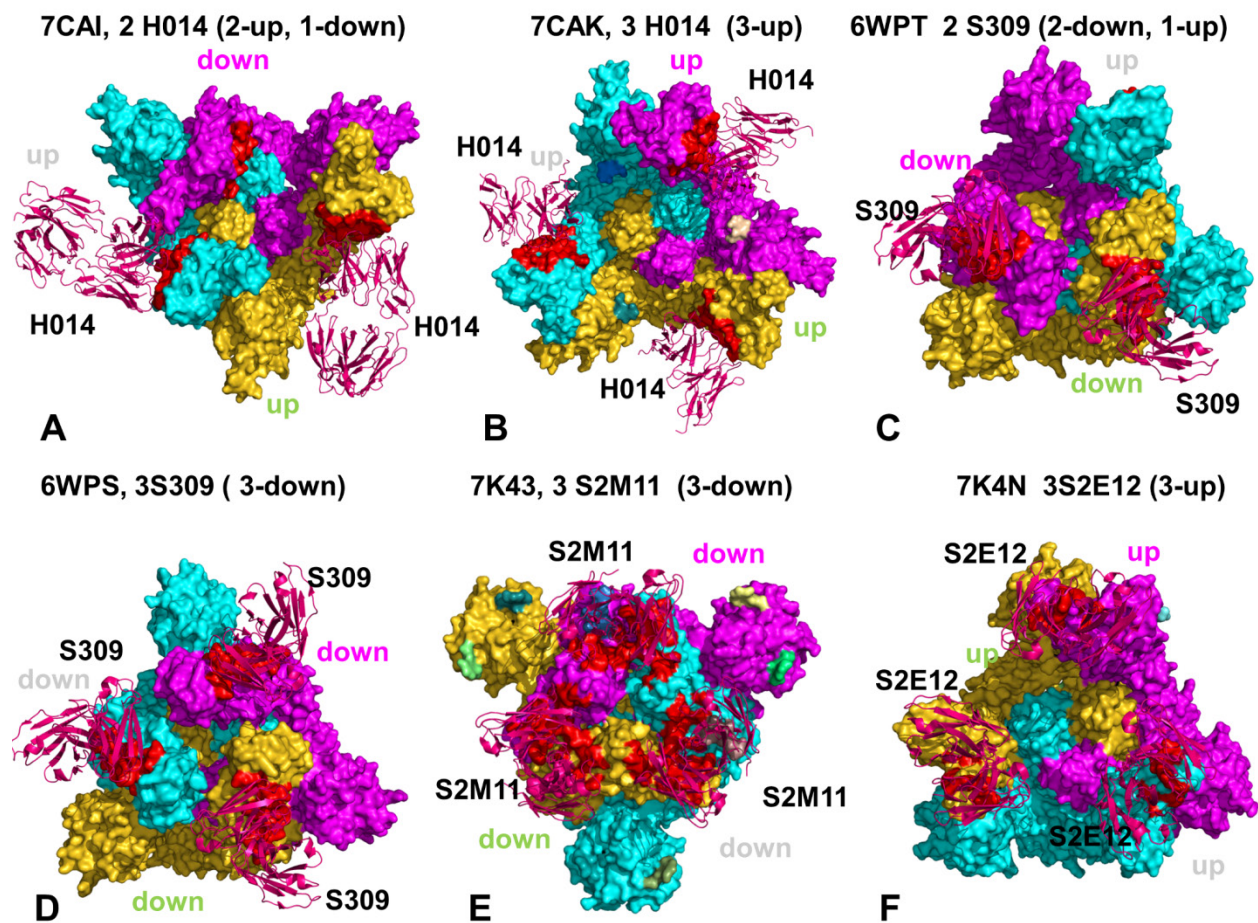
Coarse-grained (CG) models are computationally effective approaches for simulations of large systems over long timescales. In this study, CG-CABS model<sup>81-85</sup> was used for simulations of the cryo-EM structures of the SARS-CoV-2 S complexes with H014, S309, S2M11, and S2E12 Abs. In this model, the amino acid residues are represented by  $C\alpha$ ,  $C\beta$ , the center of mass of side chains and another pseudoatom placed in the center of the  $C\alpha$ - $C\alpha$  pseudo-bond.<sup>81-83</sup> We employed CABS-flex approach that efficiently combines a high-resolution coarse-grained model and efficient search protocol capable of accurately reproducing all-atom MD simulation trajectories and dynamic profiles of large biomolecules on a long time scale.<sup>81-85</sup> The sampling scheme of the CABS model used in our study is based on Monte Carlo replica-exchange dynamics and involves a sequence of local moves of individual amino acids in the protein structure as well as moves of small fragments.<sup>81-83</sup> CABS-flex standalone package dynamics implemented as a Python 2.7 object-oriented package was used for fast simulations of protein structures.<sup>85</sup> A total of 1,000 independent CG-CABS simulations were performed for each of the

studied systems. In each simulation, the total number of cycles was set to 10,000 and the number of cycles between trajectory frames was 100. The cryo-EM structures of the SARS-CoV-2 S trimer complexes with a panel of Abs including H014, S309, S2M11, and S2E12 were used in CG-CABS simulations (Figures 1,2). These structures included the partially open and fully open forms of the SARS-CoV-2 S trimer in the complex with H014 (Figure 1A,B), the partially closed and fully closed S trimer forms bound with S309 (Figure 1C,D), the fully closed S trimer form complexed with S2M11 (Figure 1E), and the fully open S trimer form in the complex with S2E12 (Figure 1F). All structures were obtained from the Protein Data Bank.<sup>86,87</sup> During structure preparation stage, protein residues in the crystal structures were inspected for missing residues and protons. Hydrogen atoms and missing residues were initially added and assigned according to the WHATIF program web interface.<sup>88,89</sup> The structures were further pre-processed through the Protein Preparation Wizard (Schrödinger, LLC, New York, NY) and included the check of bond order, assignment and adjustment of ionization states, formation of disulphide bonds, removal of crystallographic water molecules and co-factors, capping of the termini, assignment of partial charges, and addition of possible missing atoms and side chains that were not assigned in the initial processing with the WHATIF program. The missing loops in the cryo-EM structures were also reconstructed using template-based loop prediction approaches ModLoop<sup>90</sup> and ArchPRED<sup>91</sup> The conformational ensembles were also subjected to all-atom reconstruction using PULCHRA method<sup>92</sup> and CG2AA tool<sup>93</sup> to produce atomistic models of simulation trajectories. The protein structures were then optimized using atomic-level energy minimization with a composite physics and knowledge-based force fields as implemented in the 3Drefine method.<sup>94</sup>



**Figure 1.** The cryo-EM structures of the SARS-CoV-2 S protein trimer complexes with a panel of Abs used in this study. (A) The cryo-EM structure of the SARS-CoV-2 S protein trimer with two RBDs in the open state complexed with two H014 Fab (pdb id 7CAI).<sup>45</sup> (B) The cryo-EM structure of the SARS-CoV-2 S protein trimer with three RBD in the open state complexed with three H014 Fab (pdb id 7CAK).<sup>45</sup> (C) The cryo-EM structure of the SARS-CoV-2 S protein trimer with two RBDs in the closed form and one RBD in the open state bound with the two S309 neutralizing Fab fragments (pdb id 6WPT).<sup>46</sup> (D) The cryo-EM structure of the SARS-CoV-2 S protein trimer with all three RBDs in the closed form bound with the three S309 neutralizing Fab fragments (pdb id 6WPS).<sup>46</sup> (E) The cryo-EM structure of the SARS-CoV-2 S protein trimer

with all three RBDs in the closed-down form bound with the three S2M11 neutralizing Fab fragments (pdb id 7K43).<sup>47</sup> (F) The cryo-EM structure of the SARS-CoV-2 S protein trimer with all three RBDs in the open-up form bound with the three S2ME12 neutralizing Fab fragments (pdb id 7K4N).<sup>47</sup> The SARS-CoV-2 S proteins are shown in surface representation, with protomer A in green, protomer B in cyan, and protomer C in magenta. The Ab structures are shown in ribbons and colored in maroon. All structures are annotated and open/closed (up/down) conformations of S protomers are indicated.



**Figure 2.** The binding epitopes of the SARS-CoV-2 S protein trimer complexes with a panel of Abs used in this study. The top view highlighting the binding epitopes is shown for the cryo-EM structure of the SARS-CoV-2 S protein trimer with H014 (A,B), S309 (C,D), S2M11 (E), an S2E12 (F). Note, S2M11 recognizes a quaternary epitope comprising distinct regions of two neighboring RBDs within an S trimer (E). The SARS-CoV-2 S proteins are shown in surface representation, with protomer A in green, protomer B in cyan, and protomer C in magenta. The Ab structures are shown in ribbons and colored in maroon. All structures are annotated and open/closed (up/down) conformations of S protomers are indicated.

### Protein Stability and Mutational Scanning Analysis

To compute protein stability changes in the SARS-CoV-2 trimer mutants, we conducted a systematic alanine scanning of protein residues in the SARS-CoV-2 trimer mutants. BeAtMuSiC approach was employed that is based on statistical potentials describing the pairwise inter-residue distances, backbone torsion angles and solvent accessibilities, and considers the effect of the mutation on the strength of the interactions at the interface and on the overall stability of the complex.<sup>95-97</sup> The binding free energy of protein-protein complex can be expressed as the difference in the folding free energy of the complex and folding free energies of the two protein binding partners:

$$\Delta G_{bind} = G^{com} - G^A - G^B \quad (2)$$

The change of the binding energy due to a mutation was calculated then as the following:

$$\Delta\Delta G_{bind} = \Delta G_{bind}^{mut} - \Delta G_{bind}^{wt} \quad (3)$$



We leveraged rapid calculations based on statistical potentials to compute the ensemble-averaged binding free energy changes using equilibrium samples from MD trajectories. The binding free energy changes were computed by averaging the results over 1,000 equilibrium samples for each of the studied systems.

### Protein Contact Networks and Network Clustering

The protein contact network is a network whose nodes are the protein residues and links are active contacts between residues in the protein structure. The protein contact network is an undirected, unweighted graph; it is built on the basis of the distance matrix  $\mathbf{d}$ , whose generic element  $d_{ij}$  records the Euclidean distance between the  $i$ th and the  $j$ th residue (measured between the corresponding  $\alpha$  carbons). A detailed description of network construction and significance of network descriptors is presented in our previous studies.<sup>98-100</sup> The active network links are defined using a range of contacts between 4 Å and 8 Å. The description of the network is given by the following adjacency matrix :

$$A_{ij} = \begin{cases} 1 & \text{if } 4 \text{ \AA} < d_{ij} < 8 \text{ \AA} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where  $d_{ij}$  is the distance between the residues. The node degree describes the number of links each residue has with other residues, defined as:

$$k_i = \sum_j A_{ij} \quad (5)$$

We previously demonstrated that spectral network clustering targets functional modules in proteins.<sup>98</sup> The network clustering is based on the spectral decomposition of the network Laplacian  $L$  defined as:

$$L = D - A \quad (6)$$

where  $D$  is the degree matrix, a diagonal matrix whose diagonal is the degree vector, and  $A$  is the adjacency matrix. We used the eigenvector corresponding to the second minor eigenvalue  $\mathbf{v}_2$  of the Laplacian (Fiedler's vector) to assign nodes (residues) to different clusters. We introduced a novel feature in the hierarchical binary algorithm to compute any number of clusters  $n_{clus}$  (no more only powers of two): we parted the whole range of values of  $\mathbf{v}_2$  into  $n_{clus}$  parts, of length  $l_{clus}$ ; so residues are assigned to the first cluster if the corresponding component falls between  $\min(\mathbf{v}_2)$  and  $\min(\mathbf{v}_2) + l_{clus}$  the generic  $i$ -th cluster, thus, is that made of residues corresponding to  $\mathbf{v}_2$  components comprised in the range  $[\min(\mathbf{v}_2) + (i - 1) \cdot l_{clus}, \min(\mathbf{v}_2) + i \cdot l_{clus}]$ .

Once the network is divided into a given number of clusters (powers of two), we define the participation coefficient, defined as:

$$P_i = 1 - \left(\frac{k_i}{k_{si}}\right)^2 \quad (7)$$

where  $k_i$  is the overall node degree, while  $k_{si}$  is the node degree including only links with nodes (residues) that belong to their own cluster. The participation coefficient  $P$  describes the propensity of residue nodes to participate into inter-cluster communication. We designate as highly active communication residues the nodes with  $P > 0.75$ , based on our previous studies showing that such residues typically correspond to important regulatory centers of signal transmission between protein domains.<sup>98</sup> The proposed methodology of network clustering was implemented as Cytoscape plugin.<sup>101</sup>

In the framework of hierarchical network modeling approach, we also employed a graph-based representation of protein structures<sup>102-104</sup> with residues as network nodes and the inter-residue edges as residue interactions to construct the residue interaction networks using dynamic correlations<sup>104</sup> and coevolutionary residue couplings<sup>105</sup> as detailed in our previous studies.<sup>105-107</sup>

The ensemble of shortest paths is determined from matrix of communication distances by the Floyd-Warshall algorithm.<sup>108</sup> Network graph calculations were performed using the python package NetworkX.<sup>109</sup> Using the constructed protein structure networks, we computed the residue-based betweenness parameter. The short path betweenness centrality of residue  $i$  is defined to be the sum of the fraction of shortest paths between all pairs of residues that pass through residue  $i$ :

$$C_b(n_i) = \sum_{j < k} \frac{g_{jk}(i)}{g_{jk}} \quad (8)$$

where  $g_{jk}$  denotes the number of shortest geodesics paths connecting  $j$  and  $k$ , and  $g_{jk}(i)$  is the number of shortest paths between residues  $j$  and  $k$  passing through the node  $n_i$ .

### **Perturbation Response Scanning**

Perturbation Response Scanning (PRS) approach<sup>110,111</sup> was used to estimate residue response to external forces applied systematically to each residue in the protein system. This approach has successfully identified hotspot residues driving allosteric mechanisms in single protein domains and large multi-domain assemblies.<sup>112-117</sup> The implementation of this approach follows the protocol originally proposed by Bahar and colleagues<sup>112,113</sup> and was described in details in our previous studies.<sup>64</sup> In brief, through monitoring the response to forces on the protein residues, the PRS approach can quantify allosteric couplings and determine the protein response in functional movements. In this approach, it  $3N \times 3N$  Hessian matrix  $\mathbf{H}$  whose elements represent second derivatives of the potential at the local minimum connect the perturbation forces to the residue displacements. The  $3N$ -dimensional vector  $\Delta\mathbf{R}$  of node displacements in response to  $3N$ -

dimensional perturbation force follows Hooke's law  $\mathbf{F} = \mathbf{H} * \Delta\mathbf{R}$ . A perturbation force is applied to one residue at a time, and the response of the protein system is measured by the displacement vector  $\Delta\mathbf{R}(i) = \mathbf{H}^{-1}\mathbf{F}^{(i)}$  that is then translated into  $N \times N$  PRS matrix. The second derivatives matrix  $\mathbf{H}$  is obtained from simulation trajectories for each protein structure, with residues represented by  $C_\alpha$  atoms and the deviation of each residue from an average structure was calculated by  $\Delta\mathbf{R}_j(t) = \mathbf{R}_j(t) - \langle \mathbf{R}_j(t) \rangle$ , and corresponding covariance matrix  $\mathbf{C}$  was then calculated by  $\Delta\mathbf{R}\Delta\mathbf{R}^T$ . We sequentially perturbed each residue in the SARS-CoV-2 spike structures by applying a total of 250 random forces to each residue to mimic a sphere of randomly selected directions.<sup>64</sup> The displacement changes,  $\Delta\mathbf{R}^i$  is a  $3N$ -dimensional vector describing the linear response of the protein and deformation of all the residues. Using the residue displacements upon multiple external force perturbations, we compute the magnitude of the response of residue  $k$  as  $\langle \|\Delta\mathbf{R}_k^{(i)}\|^2 \rangle$  averaged over multiple perturbation forces  $\mathbf{F}^{(i)}$ , yielding the  $ik^{\text{th}}$  element of the  $N \times N$  PRS matrix.<sup>112,113</sup> The average effect of the perturbed effector site  $i$  on all other residues is computed by averaging over all sensors (receivers) residues  $j$  and can be expressed as  $\langle (\Delta\mathbf{R}^i)^2 \rangle_{effector}$ . The effector profile determines the global influence of a given residue node on the perturbations in other protein residues and can be used as proxy for detecting allosteric regulatory hotspots in the interaction networks. In turn, the  $j^{\text{th}}$  column of the PRS matrix describes the sensitivity profile of sensor residue  $j$  in response to perturbations of all residues and its average is denoted as  $\langle (\Delta\mathbf{R}^i)^2 \rangle_{sensor}$ . The sensor profile measures the ability of residue  $j$  to serve as a receiver (or transmitter) of dynamic changes in the system.

## Results and Discussion

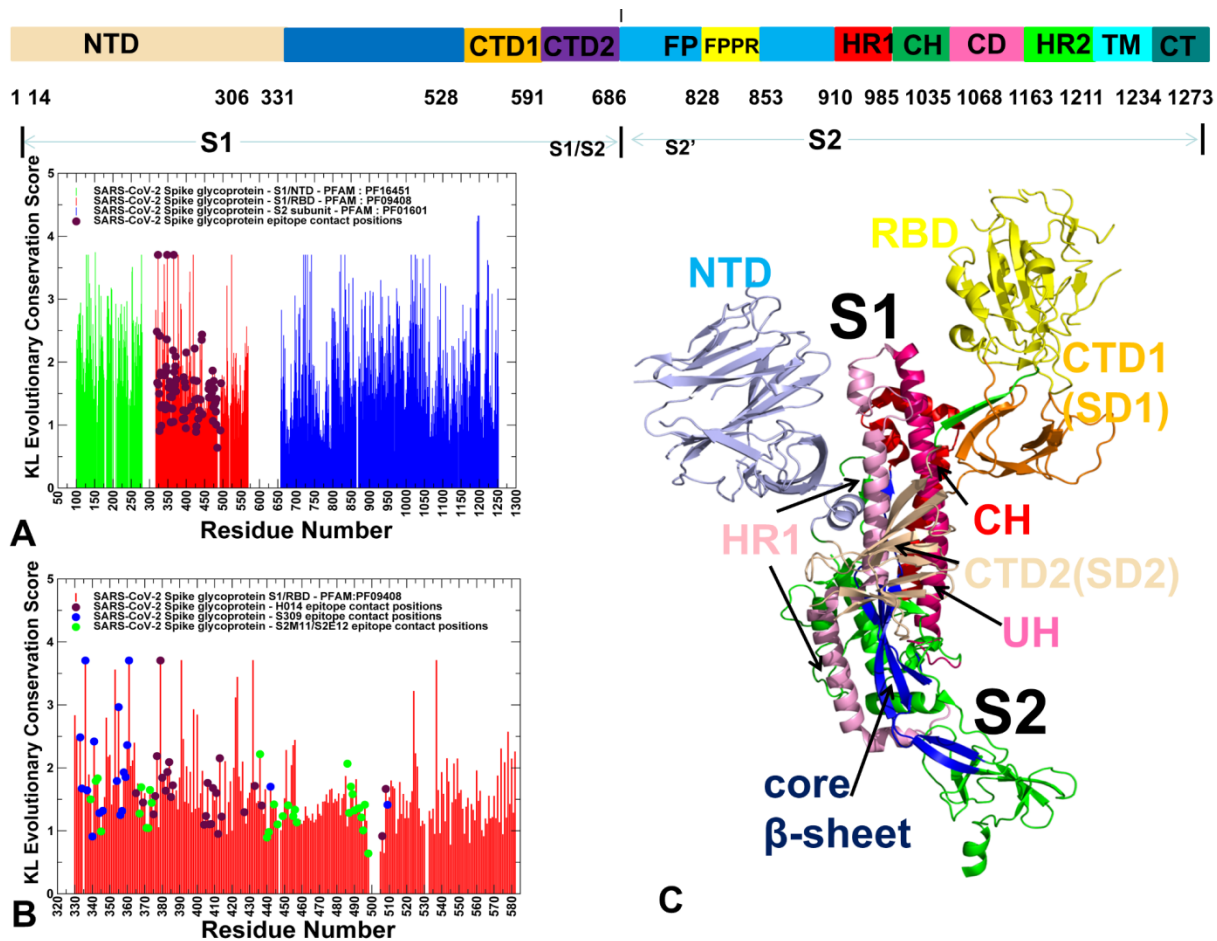
### Sequence Analysis Links Evolutionary Patterns in SARS-CoV S Proteins with Antibody Binding Preferences

To determine the evolutionary patterns in the SARS-CoV S proteins and characterize the extent of conservation and variability of the S1 and S2 subunits, we utilized KL sequence conservation score.<sup>71-73</sup> Consistent with previous studies<sup>118-120</sup> we found that S1 RBD is less conserved than domains in the S2 subunit (Figure 3A,B). The S2 subunit contains an N-terminal hydrophobic fusion peptide (FP), fusion peptide proximal region (FPPR), heptad repeat 1 (HR1), central helix region (CH), connector domain (CD), heptad repeat 2 (HR2), transmembrane domain (TD) and cytoplasmic tail (CT). The S1 domains are situated above the S2 subunit, covering and protecting the fusion apparatus. The results confirmed the higher conservation of the S2 subunit particularly highlighting conservation of the HR1 (residues 910-985), CH (residues 986-1035), CD (residues 1068-1163), HR2 (residues 1163-1211), and TD regions (residues 1211-1234) (Figure 3). The S2 subunit regions are highly conserved in the SARS-CoV and the SARS-CoV-2 variants while the S1 subunit was more diverse in the NTD and RBD regions.

Among most conserved residues in the S2 subunit are clusters of conserved cysteine residues forming disulfide bridges that are crucial for stabilization of both pre-fusion and post-fusion SARS-CoV-2 spike protein conformation. The S proteins can contain up to 40 cysteine residues, 36 of which are conserved in the S proteins of various SARS-coronaviruses.<sup>121</sup> The conserved cysteine cluster in the TD region 1220-CCMTSCCSC-1228 displayed high conservation scores, with C1121, M1222, and C1225 featuring top 1% of conservation scores (Figure 3A). Indeed, mutagenesis of the cysteine cluster I (1220-CCMTS-24), located immediately proximal to the TD showed the 55% reduction in S-mediated cell fusion as compared to the wild-type S protein.<sup>122</sup>

The proximal cysteine cluster 1225-CCSC-1228 is similarly important as alanine mutations in this cluster resulted in the 60% reduction of S-mediated cell fusion.<sup>122</sup> At the same time, the nearest cysteine-rich cluster 1230-CSCGSCCK-1237 featured only one highly conserved C1235. According to the experimental data, mutations in this region caused only a moderate 15% reduction in cell fusion<sup>122</sup>, indicating that functional role of these clusters may be closely linked with the conservation level of cysteine residues.

As cysteine residues play critical roles in structural stabilization of proteins, we focused our attention on these residues and their locations in the sequences. Interestingly, the most conserved S2 positions included cysteine residues C720, C725, C731, C742, C822, C822, C833, C1014, C1025, and C1064 (Figure 3A). A conserved region flanked by C822 and C833 is known to be important for interactions with components of the SARS-CoV S trimer to control the activation of membrane fusion.<sup>123</sup> In addition, other conserved residues included Y819, I800, L803, D830, L831, Y855, H1030, P1039, H1046 (Figure 3A,C). This is consistent with the experimental mutagenesis study based on cell-cell fusion and pseudovirion infectivity assay showing a critical role of the core-conserved residues C822, D830, L831, and C833 residues.<sup>123</sup> Some of these residues are located C-terminal to the SARS-CoV S2 cleavage site at R797 forming a highly conserved region 798-SFIEDLLFNKVTLADAGF-815 that plays an important role for membrane fusion.<sup>124</sup> Among highly conserved S protein regions are also six clusters of cysteine residues in the S2 subunit forming disulfide bridges crucial for stabilization of both pre-fusion and post-fusion SARS-CoV-2 spike protein conformations<sup>125-127</sup> (Figures 3,4).



**Figure 3.** Sequence conservation analysis of the SARS-CoV-2 S glycoprotein. (Top panel) A schematic representation of domain organization and residue range for the full-length SARS-CoV-2 spike (S) protein. The subunits S1 and S2 include NTD RBD, C-terminal domain 1(CTD1), C-terminal domain 2 (CTD2), S1/S2 cleavage site (S1/S2), S2' cleavage site (S2'), fusion peptide (FP), fusion peptide proximal region (FPPR), heptad repeat 1 (HR1), central helix region (CH), connector domain (CD), heptad repeat 2 (HR2), transmembrane domain (TM), and cytoplasmic tail (CT).

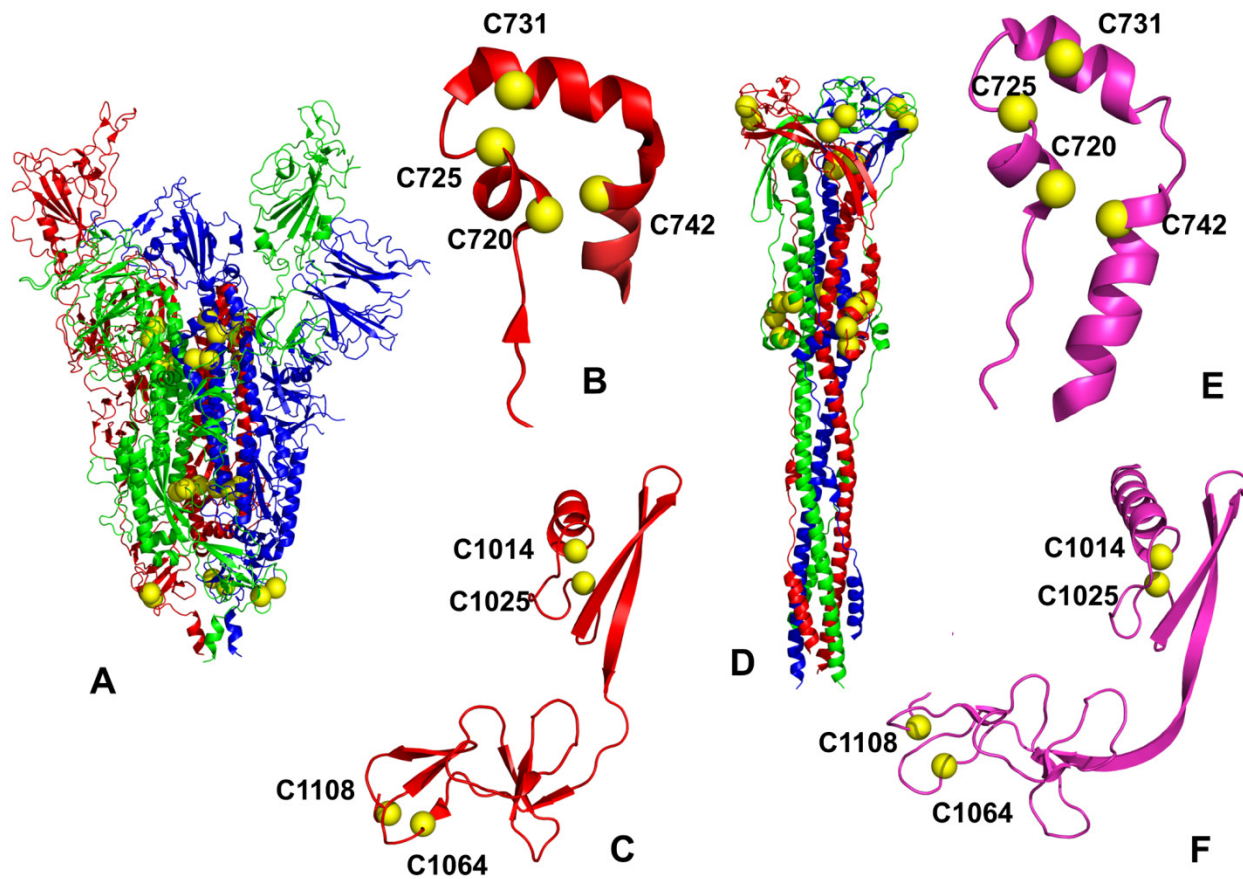
(Panel A) The KL conservation score for SARS-CoV-RBD S protein. High KL scores indicate highly conserved sites and low scores correspond to more variable positions. Three Pfam domains

were utilized corresponding to S1, the NTD (*bCoV\_S1\_N*, Betacoronavirus-like spike glycoprotein S1, N-terminal, Pfam:PF16451, Uniprot SPIKE\_CVHSA, pdb id 6CS0, residues 33-324), the RBD (*bCoV\_S1\_RBD*, Betacoronavirus spike glycoprotein S1, receptor binding, Pfam:PF09408, Uniprot SPIKE\_CVHSA, pdb id 6CS0, residues 335-512) and the new C-terminal domain, CTD (*CoV\_S1\_C* Coronavirus spike glycoprotein S1, C-terminal. Pfam:PF19209, Uniprot SPIKE\_CVHSA, pdb id 6CS0, residues 522-580). S2 is described in the family Pfam:PF01601 (Uniprot SPIKE\_CVHSA, pdb id 6CS0, residues 622-1120). The KL scores for the S1-NTD residues are shown in green bars, for the S1-RBD regions in the red bars, and for S2 residues in blue bars. The KL conservation scores for the epitope residues of all studied Abs are shown in filled maroon-colored circles.

(Panel B) A close-up view of KL conservation scores for RBD regions of the SARS-CoV-2 S protein (Pfam:PF09408, Uniprot P0DTC2: SPIKE\_SARS2 (pdb id 6VXX, 6VYB residue numbering) is shown in red bars. The KL scores are highlighted for the binding epitope residues of H014 (filled maroon-colored circles), S309 (filled blue circles), and S2M11/S2E12 (filled green circles).

(Panel C) The structural organization of the SARS-CoV-2 S protein major domains is shown for a single protomer. The subunits S1 regions are annotated as follows : NTD (residues 14-306) in light blue; RBD (residues 331-528) in yellow; CTD1 (residues 528-591) in orange; CTD2 (residues 592-686) in wheat color ; upstream helix (UH) (residues 736-781) in red; HR1 (residues 910-985) in pink; CH (residues 986-1035) in hot pink; antiparallel core  $\beta$ -sheet (residues 711-736, 1045-1076) (in blue).





**Figure 4.** Sequence and structural conservation of cysteine clusters in the SARS-CoV-2 spike protein in the prefusion and postfusion states. (A) The cryo-EM structure of the SARS-CoV-2 S protein in the prefusion form is shown in ribbons. The protomer A is in green, protomer B in red, and protomer C in blue colors. The positions of the conserved cysteine clusters are shown in yellow spheres. (B) Structural arrangement of the conserved cluster formed by C720, C725, C731, and C742 in the UH region. The UH fragment is shown in red ribbons and conserved cysteine sites are shown in yellow spheres and annotated. (C) Structural organization of the conserved cysteine cluster in the  $\beta$ -hairpin region formed by C1014 and C1025, C1064 and C1108. The protein fragment is shown in red ribbons and conserved cysteines are in yellow spheres and annotated. (D) The cryo-EM structure of the SARS-CoV-2 S protein in the postfusion form. The protomer A is in green,

protomer B in red, and protomer C in blue colors. The conserved cysteine clusters are in yellow spheres. (E) Structural arrangement of conserved cluster formed by C720, C725, C731, and C742 in the postfusion state. (F) Structural organization of a conserved cysteine cluster (C1014 and C1025, C1064 and C1108) in the post-fusion state. The conserved cysteine sites are shown in yellow spheres and annotated.

Structural analysis demonstrated that the core elements of S2 regions anchored by several cysteine clusters are highly preserved in the SARS-CoV-2 spike prefusion and postfusion states (Figure 4A,D) despite massive conformational changes of the SARS-CoV-2 S2 machinery.<sup>125,126</sup> Some of these regions include cysteine clusters formed by C720, C725, C731, and C742 in the upstream helix (UH) regions. (Figure 4B,E). Another conserved cysteine cluster of disulfide bonds is formed in the  $\beta$ -hairpin region (residues 1045-1076) located downstream of the CH region by residues C1014 and C1025 (C1032 and C1043 respectively in SPIKE\_SARS2 sequence numbering) as well as residues C1064 and C1108 (C1082 and C1126 in SPIKE\_SARS2 sequence numbering) (Figure 4C,F). This conserved segment of S2 subunit is a part of the antiparallel core  $\beta$ -sheet assembled from an N-terminal  $\beta$ -strand ( $\beta_{46}$ ) and a C-terminal  $\beta$ -hairpin ( $\beta_{49}$ – $\beta_{50}$ ). The top conserved RBD positions included C336, R355, C361, F374, F377, C379, L387, C391, D398, G413, N422, Y423, L425, F429, C432, and W436 (Figure 3B). The RBD region includes eight conserved cysteine residues, six of which form three disulfide linkages (C336–C361, C379–C432 and C391–C525), which stabilize the  $\beta$ -sheet RBD structure in the SARS-CoV-2 S protein. The crystal structures of S-proteins highlighted that two of these disulfide bonds are potentially redox-active, facilitating the primal interaction between the receptor and the spike protein.<sup>121</sup>

We particularly focused on conservation patterns of the SARS-CoV-2 RBD residues forming binding epitopes for H014, S309, S2E12 and S2M11 Abs. The H014 epitope is fairly large and broadly distributed across RBD regions (residues 368 to 386, 405 to 408 and 411 to 413, 439, and 503) forming a cavity on one side of the RBD (Tables S1 and S2, Supporting Information). Although most of contacts are formed with moderately conserved residues, H014 makes favorable interactions with two most highly conserved F377 and C379 positions in the RBD region (Figure 3B). Additionally, H014 makes strong interaction contacts with several other conserved RBD positions including Y380, S383, P384, K386, G413, and W436 (Figure 3B). Of particular importance are H014 contacts with S383 and G413 residues that are located at the inter-protomer boundaries (S383-D385) and (G413-V987) and could function as regulatory switches of S protein equilibrium.<sup>30</sup> It is worth mentioning that disulfide-bonded S-protein trimer variants S383C/D985C at the RBD to S2 boundaries can lead to a predominant population of the prefusion closed states.<sup>30</sup> Notably, the conformational epitope for H014 is only accessible when the RBD in an open conformation. According to our analysis, H014 interactions with conserved positions F377, C379, S383 located away from the RBM region could be important for binding and modulation of the enhanced cross-neutralization activities.

S309 engages an epitope distinct from the RBM making contacts with two most conserved RBD positions C336 and C361 (Figure 3B, Tables S3 and S4, Supporting Information). These residues form one of the disulfide linkages Cys336–Cys361 that stabilize the  $\beta$ -sheet RBD structure. S309 forms particularly strong contacts with neighboring residues L335 and P337 that displayed more moderate conservation. Our analysis indicated that the KL evolutionary score of the S309 contact RBD positions is considerably higher than average, with several interacting residues such

as T333, C336, V341, R355, C361 displaying particularly strong conservation (Figure 3B). We suggest that binding to these highly conserved and structurally rigid positions located away from the ACE2-binding interface may contribute to a broad neutralizing activity of S309. Indeed, viruses may have evolved to maintain the sensitive regions of their structure inaccessible to the immune system. As a result, Abs tend to bind to the most easily accessible regions of the virus, where viruses can tolerate mutations. By targeting evolutionary conserved RBD epitopes H014 and S309 Abs can potentially better combat virus resistance. Based on this analysis, we hypothesized that binding of H014 and S309 to conserved epitopes of S proteins that are distinct from ACE2-binding site may involve allosterically regulated mechanism in which Abs induce long-range alterations in the interaction networks and allosteric communications.

S2M11 recognizes a quaternary epitope through electrostatic interactions and shape complementarity, comprising distinct regions of two neighboring RBDs within an S trimer.<sup>47</sup> This Ab is believed to induce inhibition of membrane fusion through conformational trapping of SARS-CoV-2 S trimer in the closed state.<sup>47</sup> S2M11 forms contacts with highly conserved sites F374 and W436 on one RBD and makes interactions with F486 on the RBM motif of the other RBD (Table S5, Supporting Information, and Figure 3B). We also characterized the intermolecular contacts formed by S2E12 (Table S6, Supporting Information) that targets the RBM motif and can block the attachment to the ACE receptor. According to our analysis, the contact positions targeted by these Abs are only moderately conserved as the epitope overlaps with a more variable RBD region (Figure 3A,B).

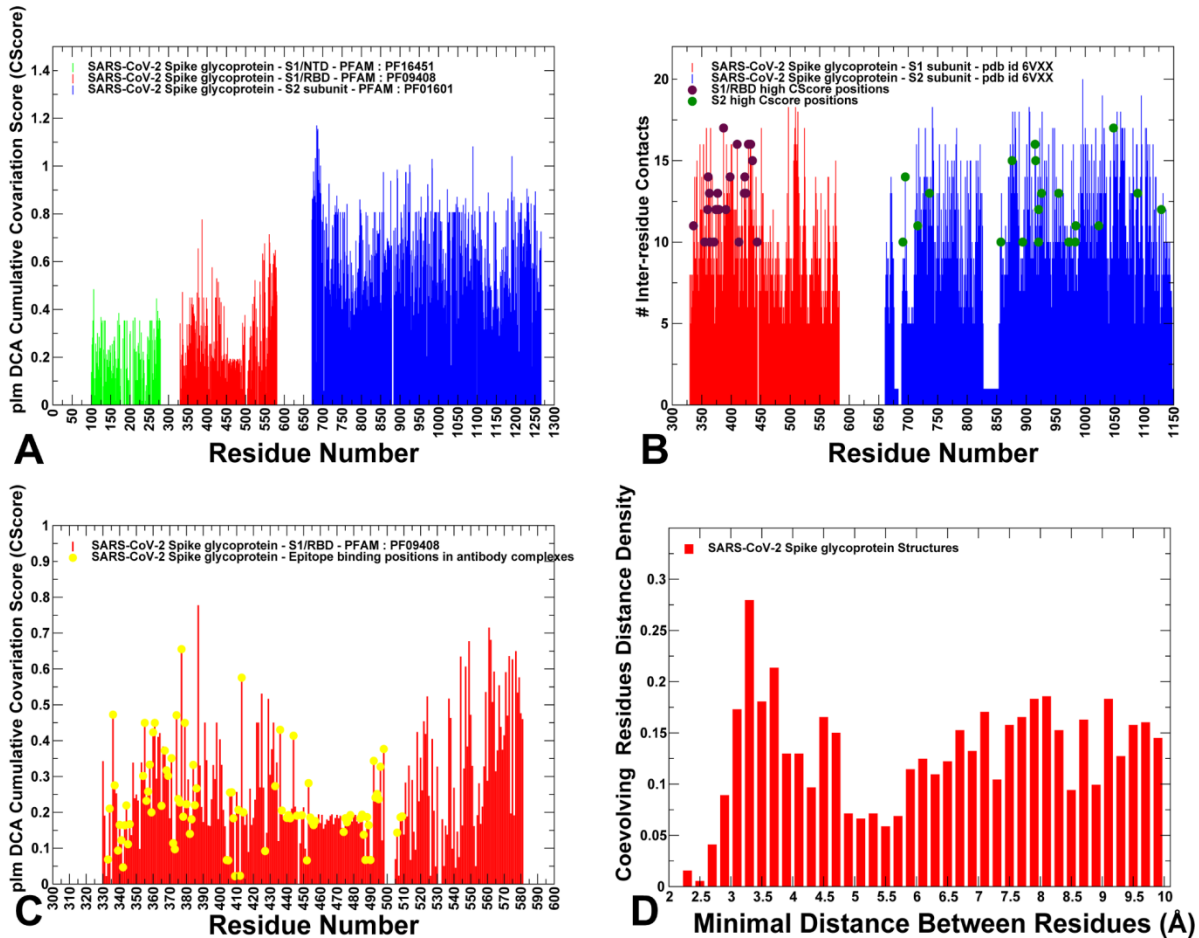
## **Coevolutionary Analysis of the SARS-CoV-2 Proteins Reveals Regulatory Centers and Functional Role of the Epitope Regions in the Network of Evolutionarily Coupled Residues**

Coevolutionary dependencies of protein residues can mediate protein recognition and are often spatially close to each other, forming clusters of interacting residues that are located near functionally important sites.<sup>129-132</sup> The coevolving residues may be clustered in mobile regions and form interaction networks of evolutionarily coupled residues that facilitate protein conformational changes.<sup>133</sup> Using MISTIC approach<sup>71-73</sup> we determined coevolutionary dependencies between S protein residues using plmDCA (Figure 5) and gaussianDCA models (Figure S1, Supporting Information). The network of coevolutionary couplings in the SARS-CoV-2 S structures was then constructed in which the nodes represented protein residues and links corresponded to coevolutionary dependencies between these residues. To identify critical nodes of this coevolutionary network that may coordinate and transmit coevolutionary signals, we computed plmDCA-based cScore profiles (Figure 5A) that measure the global influence of a given position in a coevolutionary network. This score is calculated as the sum of covariation scores above a certain threshold (top 5% of the covariation scores) for every position pair where the particular position appears. Using this approach, we quantified coevolutionary relationships between residues, identified coevolutionary couplings for functionally important regions and also mapped high CScore positions onto the binding epitopes for studied SARS-CoV-2 complexes (Figure 5B).

We focused our analysis on the computed distribution of plmDCA-based CScore profiles (Figure 5A) that revealed several important trends. First, the results revealed an appreciable density of coevolving centers in the S1 subunit, primarily in the RBD and especially CTD1 regions. This pattern can be further illustrated by a circular representation of the pairwise coevolutionary scores

(Figure S2, Supporting Information) showing the greater concentration of coevolutionary links anchored by the CTD1 regions (residues 528-591). The distribution of CScores pointed to the significantly higher density of coevolutionary couplings in the tightly packed S2 subunit (Figure 5A). The residues with significant CScore values are distributed across various S2 regions, including UH (residues 736-781), CH (residues 986-1035), HR1 region (residues 910-985), HR2 (residues 1163-1211) and  $\beta$ -hairpin (BH) region (residues 1035-1071). A dense network of coevolutionary coupled residues in the S2 regions can be therefore detected as evident from a graphical annotation of the pairwise coevolutionary scores (Figure S3, Supporting Information). Interestingly, the distribution of CScore values for the epitope residues showed that many contact positions are aligned with highly coevolving residues (Figure 5B). Among high cScore sites that establish intermolecular contacts with Abs are C336, R355, C361, F374, F377, C379, L387, C391, D398, G413, N422, Y423, L425, F429, C432, W436 (Figure 5B). Some of these residues are highly conserved (C336, C361, C379) while other sites exhibited moderate to high conservation level (L387, W436).

Structural analysis of coevolutionary hotspots corresponding to the local maxima of the distribution revealed presence of clusters situated in functional regions (Figure 6). We observed that coevolutionary centers can be localized in the key regions of the SARS-CoV-2 S protein, occupying the proximity of the SA1/S2 cleavage site, the HR1 and CH regions of S2 subunit as well as RBD and CTD1 regions in the S1 domain (Figure 6). Of particular interest are several coevolutionary hotspots located near a well-recognized cleavage site at the S1/S2 boundary. The furin cleavage site emerges as a disordered loop (residues 655-GICASYHTVSLLRST-669 in the SPIKE\_SARS sequence or residues 669-GICASYQTQT-NSPRRARSVA-688 in SPIKE\_SARS2 sequence).



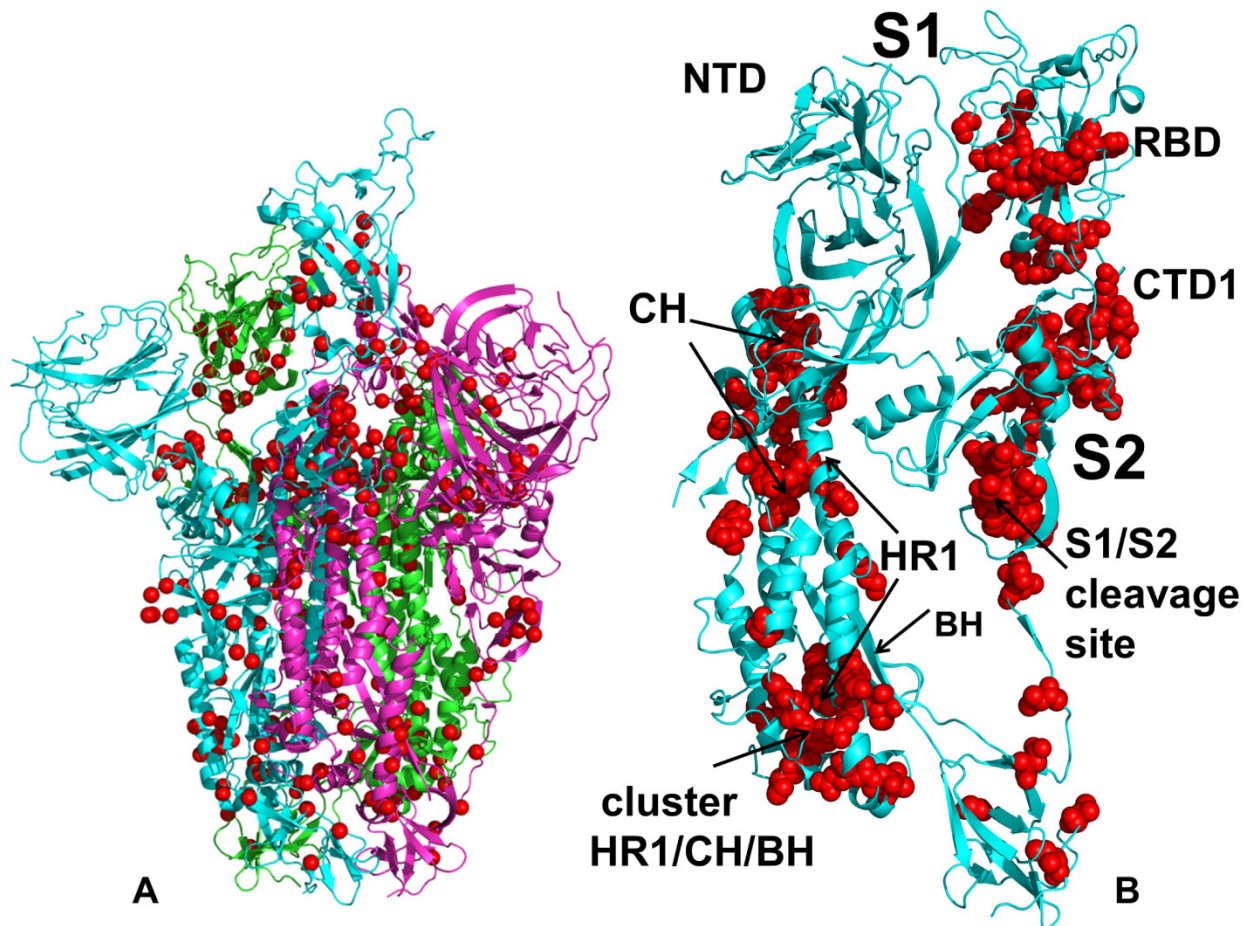
**Figure 5.** Coevolutionary profiles of the SARS CoV-2 S proteins. (A) The plmDCA-based coevolutionary Cscore profile for the SARS-CoV-2 S proteins (P0DTC2: SPIKE\_SARS2 sequence numbering). The Cscore values are shown for the S1-NTD residues in green bars (Pfam:PF16451), for the RBD in red bars (Pfam:PF09408) and for S2 regions in blue bars (Pfam:PF01601). (B) A close-up of the CSscore profile for the RBD regions is shown in red bars. The CScores for the binding epitope residues of H014, S309, S2M11, and S2E12 are shown in filled green circles. (C) The distribution of the inter-residue contacts in the S1-RBD regions (red bars) and S2 regions (blue bars). The highly coevolving centers in the RBD regions are in maroon-colored filled circles and the high CSscore residues in S2 regions are in orange-colored

filled circles. (D) The distance probability distribution of directly coupled residue pairs in the studied SARS-CoV-2 S complexes is shown in red filled bars.

Our analysis showed that residues immediately C-terminal to the S1/S2 cleavage site such as S670, K672, S673, Y676, M677, S678, S681 featured high cScore values and formed a cluster of coevolutionary centers in the S2 subunit (Figures 5,6). The multi-basic S1/S2 site in SARS-CoV-2 harbors multiple arginine residues and is involved in proteolytic cleavage of the S protein which is critical for viral entry into cells.<sup>134</sup> A particular relevance of this site stems from the fact that sequence of the S1/S2 site enables cleavage by furin in SARS-CoV-2 but not in SARS-CoV or MERS-CoV viruses.<sup>135</sup> The experimental data also showed that SARS-CoV S-mediated virus entry is based on sequential proteolytic cleavage at two distinct sites, with cleavage at the S1/S2 boundary (R667) promoting subsequent cleavage at the S2' position (R797) triggering membrane fusion.<sup>136</sup> Interestingly, our results indicated a high coevolutionary signal for R667 at the S1/S2 boundary but only a moderate Cscore for the conserved R797 position (Figure 5). These findings are generally consistent with coevolutionary patterns found in disordered protein regions showing that disordered residues whose function requires specific recognition and disorder-order transition upon binding can exhibit a high degree of coevolutionary signal.<sup>137</sup>

Although many coevolving centers in the S2 subunit are located inside the protein core and generally stable, these regions are involved in gigantic conformational rearrangements to the post-fusion state that require a nontrivial cooperation between these regions to dramatically rearrange the interaction network. Several important clusters of highly coevolving centers are localized in the HR1 regions (N925, A930, K947, N953, L959, F970, V976, L977, and L984) and CH regions (P987, E988, I993, D994, R995, I997, L1004, Y1007, T1027, and L1039) (Figures 5,6).





**Figure 6.** Structural analysis of coevolutionary hotspots in the SARS-CoV-2 S proteins.

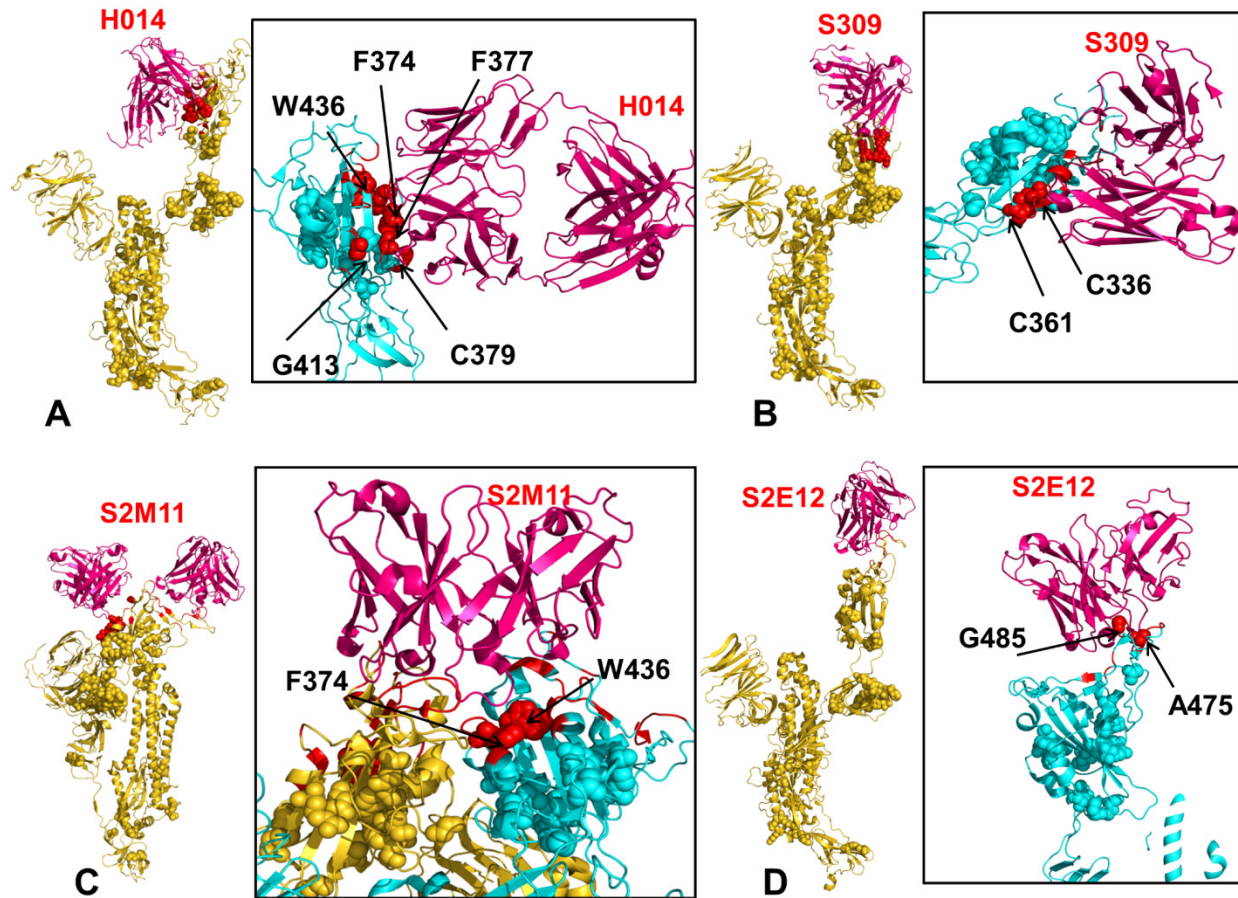
(A) Structural map of high CSscore residues shown in red spheres is projected on the cryo-EM structure of the SARS-CoV-2 S protein. (B) A close-up of coevolutionary centers mapped onto a single protomer of the S protein. The protomer is shown in cyan ribbons and high CSscore positions are depicted in red spheres. The map shows localization of coevolutionary hotspots in the key regions of the SARS-CoV-2 S protein, occupying the proximity of the S1/S2 cleavage site, the HR1 and CH regions of S2 subunit as well as RBD and CTD1 regions in the S1 domain.

An interesting cluster of coevolving centers is formed by residues from different S regions surrounding the C-terminus of the central HR1-CH helices (Figure 6). This cluster included HR1 residues Q920, N925, F927,  $\beta$ -hairpin (BH) motif residue F1052, F898 (from connecting region 841-911) and several other hydrophobic positions F800 and F802 from the region upstream of the fusion peptide FP (816-SFIEDLLFNKVTLADAGF-833) (Figure 6B). These results showed that a number of the interface core and inter-protomer centers in the S2 subunit featured a significant coevolutionary signal. Consistent with the coevolutionary studies of protein complexes,<sup>138</sup> we found that coevolutionary signal can be significant for the S2 positions involved in multiple interactions at critical junctures of UH, HR1 and CH regions. Hence, the increased structural and functional constraints for sites involved in significant number of inter-residue contacts can often imply the higher coevolution values. We found that both conservation and coevolutionary signals can increase for the S2 core residues involved in the inter-protomer interfacial contacts. However, the S2 core residues with the strongest coevolutionary signal and highest Cscore values could feature different level of conservation. In particular, the strongest coevolutionary centers in the S2 regions included fairly moderately conserved residues F1089, R983, N925 and L984 as well as strongly conserved A893, V915, L916, and A1190 positions. In general, our results indicated that S2 core residues subjected to more structural constraints and inter-residue contacts can exhibit the higher residue conservation and coevolution values.

To further probe the notion that the interface core residues can exhibit both the higher level of conservation and coevolution, we computed the average number of the inter-residue contacts for each S protein residue and aligned this distribution with the top Cscore positions in the S1/RBD and S2 regions (Figure 5C). The results indicated that coevolutionary centers tend to have a fairly significant number of the interacting contacts and can be involved in multiple interactions. In

particular, coevolutionary hotspots in the RBD regions were often aligned with the peaks of the contact distribution, supporting the notion that the level of coevolution may be greater in residues involved in multiple interactions.<sup>138</sup> It is worth noting, however, that sites with the largest number of inter-residue contacts may not necessarily correspond to the most conserved positions or residues with the highest CScore value. In fact, even though coevolutionary hotspots in the S2 subunit have a significant number of the inter-residue contacts, the distribution peaks corresponded to residues F718, V729, I742, F782 with moderate levels of conservation and coevolution (Figure 5C).

We also computed the distance probability distribution of coevolving directly coupled residue pairs in the studied SARS-CoV-2 S structures (Figure 5D). The profile showed several local maxima at 3.2 Å, 4.7 Å and a much broader area with a shallow peak near 7 Å - 8Å (Figure 5D). It is evident that the first two peaks reflect physical interactions between residues including hydrogen bonding and hydrophobic residue pairs. Hence, direct coevolutionary residue couplings in the SARS-CoV-2 S structures are dominated by spatially proximal residue pairs, that is consistent with large-scale investigations of direct coevolutionary couplings in proteins suggesting that coevolutionary signals are stronger for locally interacting residues than for residues involved in long-range interactions in allosteric networks.<sup>139</sup> Nonetheless, the distribution also highlighted another intermediate range of coevolutionary couplings at 7 Å - 8Å that is beyond direct inter-residue physical contacts and may reflect strong couplings between spatially proximal functional regions (Figure 5D). This third distribution peak can correspond to coevolutionary couplings anchored by CTD1 regions (residues 529-591) in the S1 subunit that is believed to function as allosteric connector between RBD and FPPR regions by communicating signal from and to the fusion peptide.<sup>26</sup>



**Figure 7.** Structural maps of coevolutionary centers in the epitope regions of the SARS-CoV-2 complexes with Abs. (A) Structural map of coevolutionary centers in the S complex with H014 (pdb id 7CAI/7CAK) projected onto a single “up” protomer shown in green ribbons. The coevolutionary centers are in spheres and high CSscore hotspots from the binding epitope are in red spheres. A close-up of the H014 binding epitope with the coevolving centers involved in direct contacts with H014 in red spheres and annotated. (B) Structural map and close-up of coevolutionary centers in the S complex with S309 (pdb id 6WPT). The coevolving centers involved in direct contacts with S309 in red spheres and annotated. (C,D) Structural map and close-up of coevolutionary centers in the S complex with S2M11 (pdb id 7K43) and S2E12 (pdb

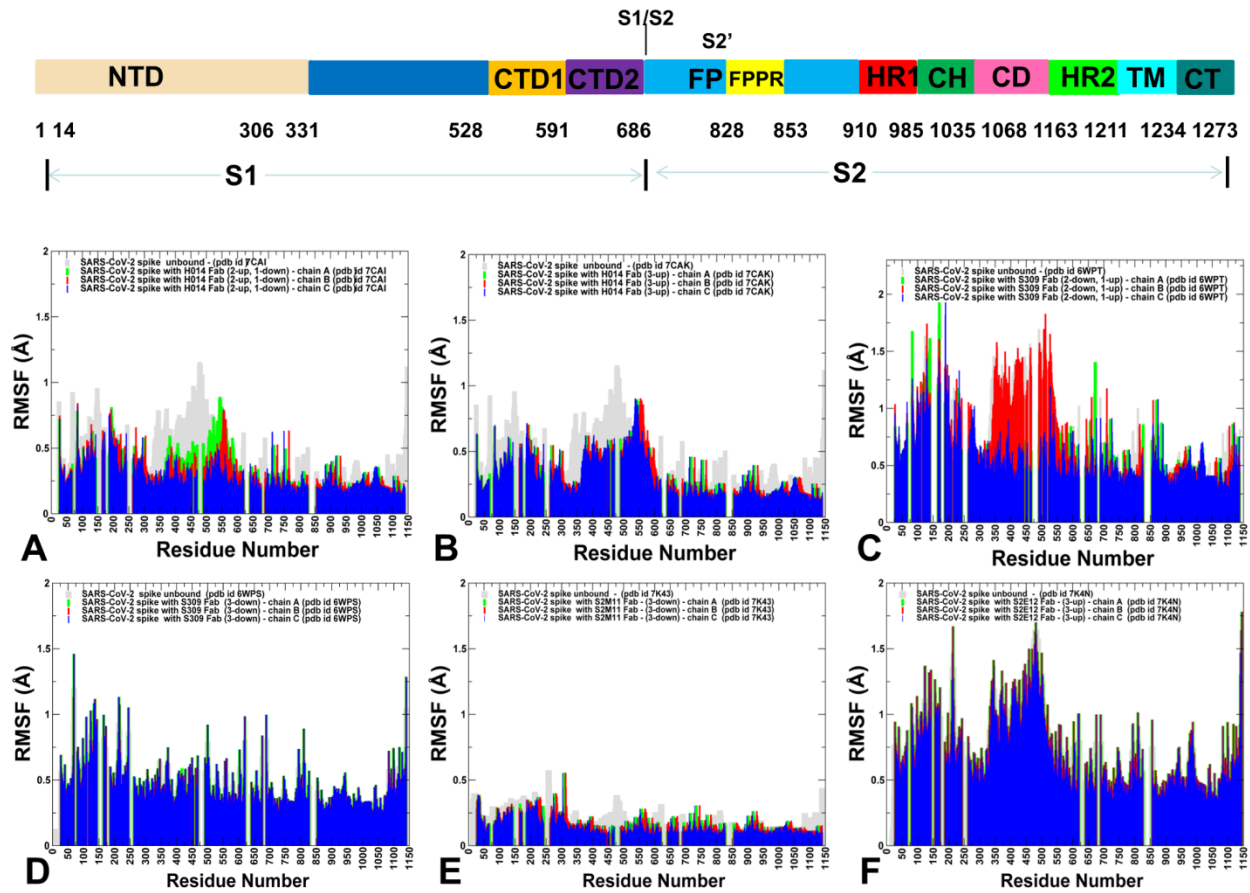
id 7K4N). The coevolving centers involved in direct contacts with S2M11 and S2E12 are in red spheres and annotated.

Hence, our results revealed the presence of a significant coevolutionary signal between functional regions separated by the medium-range distances which may help to facilitate a long-range cross-talk between distant allosteric regions in the S1 and S2 subunits.

Structural mapping of coevolutionary centers highlighted global connectivity of the coevolutionary network spanning from the epitope binding site towards the CTD1 region and regions in the S2 subunit (Figure 7). Collectively, these clusters could form modules of a coevolutionary network that may allow for efficient allosteric interactions and communications in the SARS-CoV-2 S proteins. In the SARS-CoV-2 complexes with H014, the contact interface is fairly large and involves a significant stretch of the RBD residues. Several high CScore residues W436, G413, F374, F377, and C379 are involved in the interactions with H014. In particular, multiple favorable interactions are formed by F377 with Y105, T58, S59, D60, Y50 of H014 and by C379 with N55, T58, G56, G57 positions of H014 (Figure 7A). S309 binding involves interactions with the highly conserved C336 and C361 positions that also correspond to coevolutionary hotspots and could anchor a network of evolutionary coupled residues in the S protein (Figure 7B). S2M11 interacts with the evolutionary conserved RBD sites F374 and W436 that also displayed high CScore values (Figure 7C). A smaller patch of coevolutionary centers is involved in contacts with S2E12 that connects the epitope with the S2 subunit via CTD1 region that serves as a mediating hub of coevolutionary clusters in the S1 (Figure 7D). Collectively, these clusters could form modules of a coevolutionary network that may allow for efficient allosteric interactions and communications in the SARS-CoV-2 S proteins.

## **Conformational Dynamics and Mutational Scanning Reveal Modulation of Protein Stability and Binding Energy Hotspots of the SARS-CoV-2 Spike Complexes**

We employed multiple CABS-CG simulations followed by atomistic reconstruction and refinement to provide a detailed comparative analysis of dynamic landscapes that are characteristic of the SARS-CoV-2 S trimer complexes with H014, S309, S2M11, and S2E12 Abs. Using these simulations, we examined how Ab binding could affect the global dynamic profiles of the closed, partially open, an open states revealing the important regions of flexibility (Figure 8). The analysis of the inter-residue contact maps (Figure S4, Supporting Information)<sup>140</sup> and inter-residue distance maps (Figure S5, Supporting Information)<sup>141</sup> in the SARS-CoV-2 S complexes with Abs indicated that the density of the interaction contacts is significantly greater in the densely packed S2 domains. The overall packing density of the closed S protein conformations complexed with S309 and S2M11 is also markedly higher as compared to the partially open and open states. Molecular simulations of the SARS-CoV-2 S complexes provided a quantitative picture of the differences in flexibility of the S protein states and the effect of Ab binding on modulation protein stability. A comparative analysis of the dynamics profiles showed that H014 binding can induce the significant dynamic changes by considerably reducing thermal fluctuations in the S1 regions of the Ab-interacting open protomers as compared to the unbound trimer form (Figure 8A,B). We observed small thermal fluctuations with RMSF < 1.0 Å for the S1 epitope positions (residues 368 to 386, 405 to 408 and 411 to 413, 439, and 503) that were considerably rigidified in both H014 complexes (Figure 8A,B). These findings are consistent with the experimental structural data suggesting that Ab-induced structural changes could trigger stabilization changes in both the RBD and NTD regions.<sup>45</sup>



**Figure 8.** CABS-GG conformational dynamics of the SARS-CoV-2 S complexes. A schematic representation of domain organization and residue range for the full-length SARS-CoV-2 S protein is shown above conformational dynamics profiles. (A,B) The spike root mean square fluctuations (RMSF) profiles from simulations of the cryo-EM structures of the SARS-CoV-2 S trimer with H014. (C,D) The RMSF profiles from simulations of the cryo-EM structures of the SARS-CoV-2 S trimer with S309. (E) The RMSF profiles from simulations of the cryo-EM structures of the SARS-CoV-2 S trimer with all three RBDs in the closed form bound with S2M11. (F) The RMSF profiles from simulations of the cryo-EM structure of the SARS-CoV-2 S protein trimer with all three RBDs in the open-up form bound with S2E12. The profiles

for protomer chains A,B and C are shown in green, red and blue bars respectively. The RMSF profiles for the unbound forms of S protein trimer are shown in light grey bars.

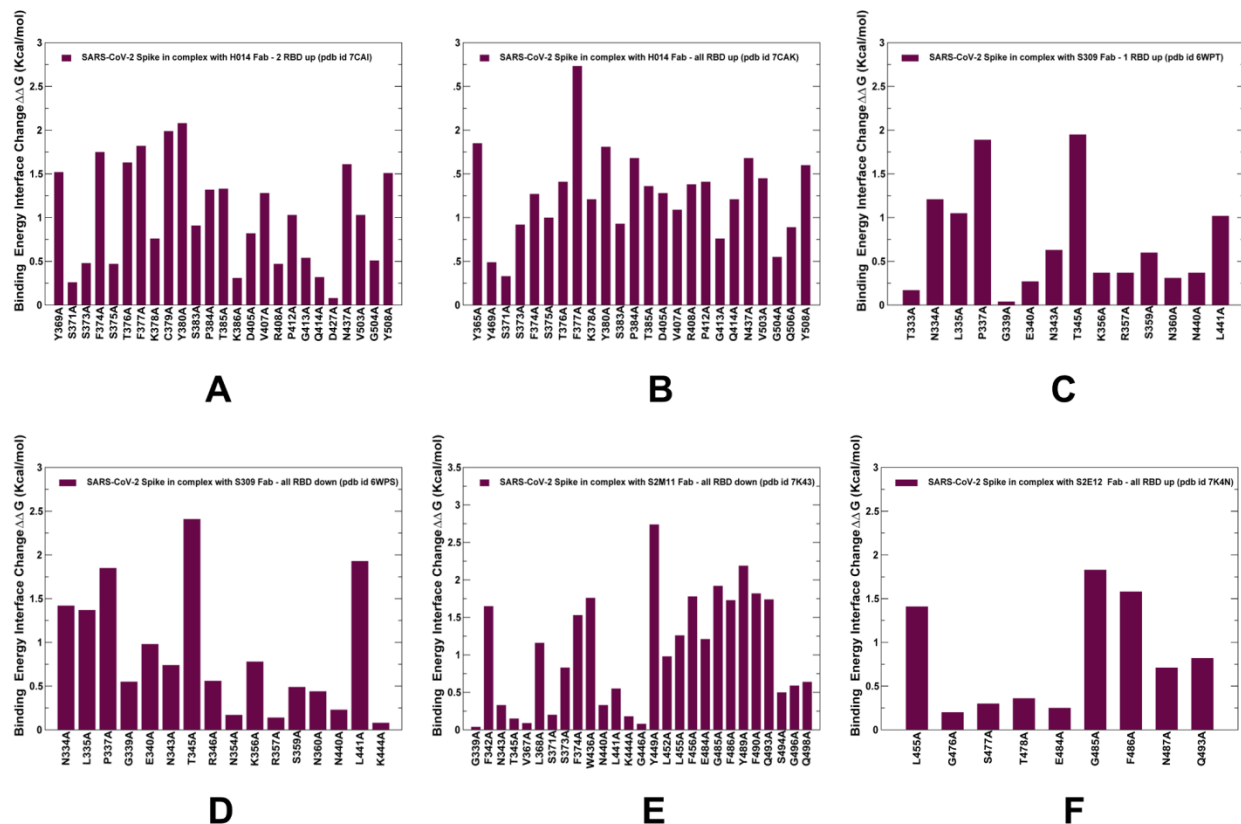
Conformational dynamics of the SARS-CoV-2 S protein complex with S309 showed only minor changes in the flexibility upon binding, particularly in the complex with S309 bound to 2 closed protomers (Figure 8C). In this case, the unbound open protomer displayed an appreciable flexibility, while the NTD regions of S309-bound closed protomers also showed some degree of mobility. In the S309 complex with 3 Abs bound to closed protomers, we found that stability of the closed S protein is protected, with only minor changes in the local dynamics between unbound and S309-bound S forms (Figure 8D). S2M11 functions by locking down the SARS-CoV-2 S trimer in the closed state through binding to a quaternary epitope. Conformational dynamics profile of the S protein complex with S2M11 in the closed form reflected this mechanism by featuring an extremely stable SARS-CoV-2 S conformation in which both S1 and S2 regions were virtually immobilized and displayed only very minor thermal fluctuations (Figure 8E). Interestingly, according to our analysis, this is the most stable bound form of the SARS-CoV-2 S protein among studied systems, suggesting that ultra-potent neutralization effect may be partly determined by the extreme thermodynamic stabilization of the closed-down S protein form. The mechanism of S2E12 neutralization of SARS-CoV-2 S protein is based on direct targeting of the RBM regions and interfering with ACE2 binding. A relatively small binding epitope in the S2E12 complex with the fully open form of S protein produced the dynamics profile where NTD and RBD regions showed an appreciable degree of mobility, while the S2 regions were mostly immobilized (Figure 8F). The important finding of this analysis was that H014, S309 and S2M11 Abs can exert modulation of the conformational dynamics leading to a significant stabilization of both S1 and



S2 regions in the open protein forms, which may effectively counteract the intrinsic flexibility of the receptor-accessible, open S conformations and thus induce potent neutralization effects.

To establish connection between dynamics and energetics of the SARS-CoV-2 binding, we employed the conformational ensembles generated in simulations and performed a systematic alanine scanning of the protein residues (Figure 9). The results revealed a wide range of important binding hotspots in the S protein complexes with H014 (Figure 9A,B). This is consistent with the dynamics profile showing a broad stabilization of the RBD regions, including the epitope residues and RBM positions. In particular, alanine scanning showed a significant contribution of conserved RBD residues F374, F377, K378, C379, Y380, P384, T385 as well as N437, V503, Y508 (Figure 9B). Among binding energy hotspots we detected some of the highly conserved positions and several coevolutionary centers such as F374, F377, and C379 residues. We argue that through interactions with major coevolutionary centers in the conserved RBD epitope, H014 may exert its long-range effect by propagating binding signal through clusters of proximal coevolutionary pairs in the RBD and CTD1 regions. The noticeably fewer number of binding hotspots were seen in the S309 complexes with partially closed (2-down) and fully closed forms of the S protein (Figure 9C,D). The determined binding hotspots L335, P337, T345, and L441 are characterized by only moderate conservation and CScore values. S309 also makes weaker contacts with conserved and coevolutionary important RBD centers C336 and C361. However, the binding free energy changes caused by alanine mutations in these positions are fairly moderate (~0.7 - 0.8 kcal/mol). The binding energy hotspots in the S2M11 complex with S protein occupy two different regions, where one group includes conserved RBD sites F374 and W436 that also displayed high CScore values (Figure 9E). Another group of binding energy hotspot positions

includes moderately conserved residues Y449, F456, F484, F486, Y489 that form a critical patch of the RBM binding interface with the host receptor.



**Figure 9.** Alanine scanning of the binding epitope residues in the SARS-CoV-2 S complexes with a panel of Abs. The binding free energy changes upon alanine mutations for the epitope residues in the SARS-CoV-2 S complex with H014 - two RBDs in the open state, pdb id 7CAI (panel A), SARS-CoV-2 S complex with H014 - three RBD in the open state, pdb id 7CAK (panel B), SARS-CoV-2 S complex with S309 - two RBDs in the closed form, pdb id 6WPT (panel C), SARS-CoV-2 S complex with S3090 - three RBDs in the closed form, pdb id 6WPS (panel D), SARS-CoV-2 S complex with S2M11 - three RBDs in the closed form, pdb id 7K43 (panel E), SARS-CoV-2 S complex with S2E12 - three RBDs in the closed form, pdb id 7K4N (panel F).

The computed binding free energy changes values are shown in bars. The binding interface residues are determined for each complex based on the average interaction contacts that persist during simulation of a given complex.

We previously showed that a conserved segment 486-FNCYFPL-492 in the RBD region emerged as a central binding energy hotspot in the SARS-CoV-2 complex with ACE2 receptor.<sup>62</sup> Hence, through binding to two adjacent protomers S2M11 can simultaneously block interface with ACE2 and influence long-range couplings using a network of coevolutionary coupled residues. In general, the results indicated that the interactions S014 and S2M11 Abs can lead to stabilization of the S conformations and emergence of multiple binding energy hotspots. By targeting these centers H014 and S309 can exert their neutralizing effect by achieving a strong binding affinity with the SARS-CoV-2 S protein and also by strengthening long-range couplings of S1 and S2 regions (Figure 9).

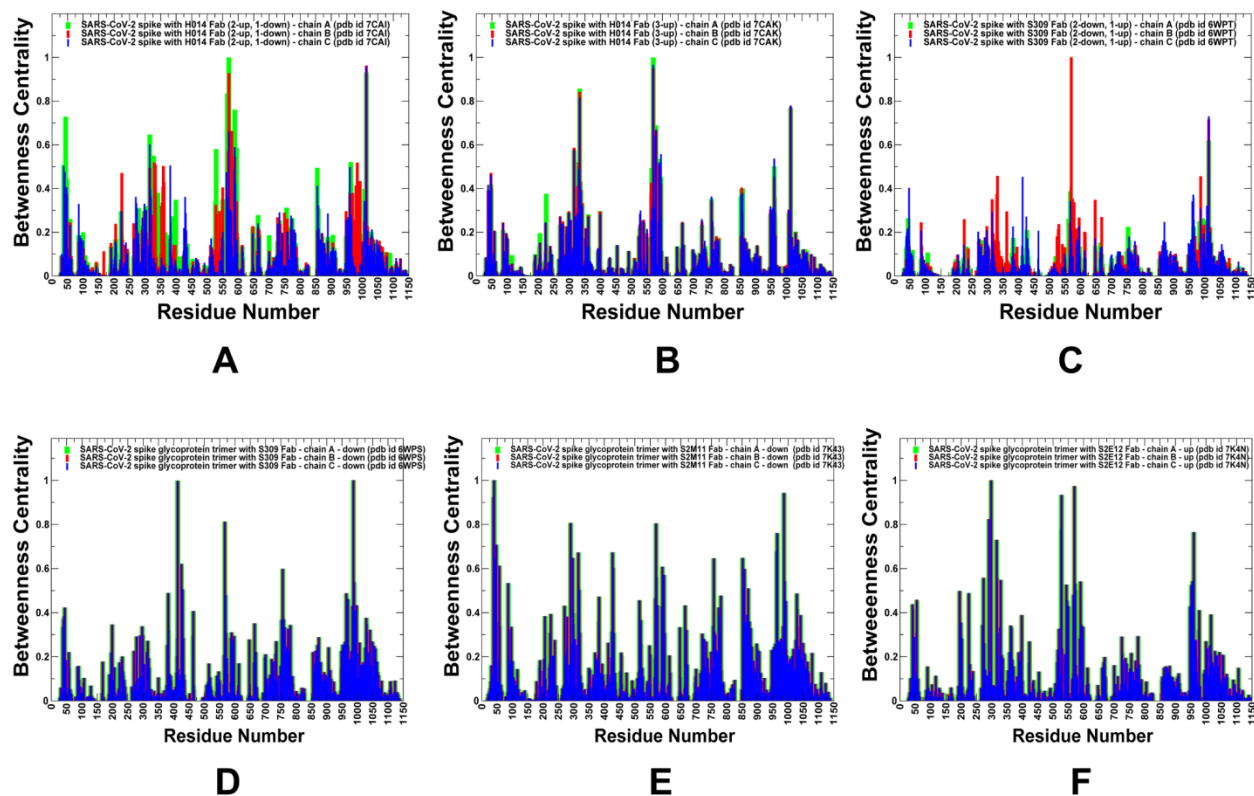
### **Hierarchical Network Modeling Reveals Mediating Centers of Allosteric Interactions in the SARS-CoV-2 Spike Complexes**

We applied a hierarchical-based network modeling approach in which the residue interactions and network couplings are described with the increasing level of atomistic details and complexity. First, a protein contact network was implemented to highlight the topological role of residues in protein structure activity and identify residues mostly responsible for signal transmission throughout the protein structure. In this simplified model, the protein residues correspond to network nodes and inter-residue contacts are considered as active links based on distance criteria as described in our previous studies.<sup>98-100</sup> Based on the hierarchical clustering algorithm, we computed the average participation coefficient  $P$  values that measure the contribution of residue

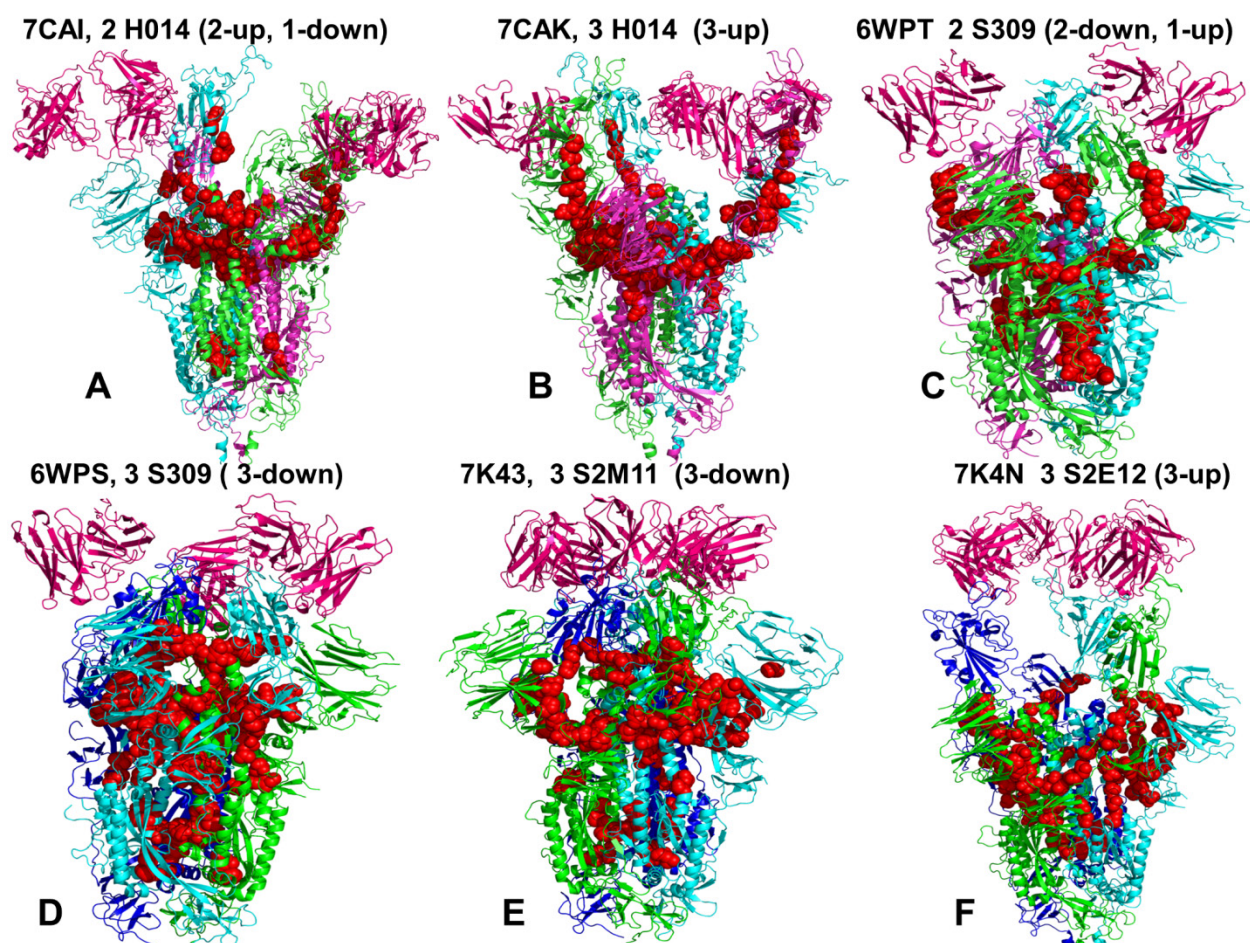
nodes in communication between different clusters (functional domains). To focus analysis on several prominent cases, we reported the communicating residues in the SARS-CoV-2 structures bound with H014 (Tables S7,S8, Supporting Information) and S309 (Tables S9,S10, Supporting Information). The results indicated that the majority of the inter-cluster communicating sites are localized in the RBD and especially CTD1 regions for SARS-CoV-2 S complexes with H014 (Tables S7,S8, Supporting Information). In this case, by attenuating mobility of the interacting RBD regions H014 binding may activate allosteric interaction networks and communications between S1 and S2 regions with CTD1 residues acting as global mediating centers of long-range interactions. The distribution of communicating positions in the SARS-CoV-2 S complexes with S309 (Tables S9,S10, Supporting Information) revealed an appreciably larger number of potential mediating centers with significant communication propensities. Moreover, these positions corresponded to different regions, including a significant number of mediating hubs in the UH, CH and HR1 regions of S2 subunit as well as residues in the CTD1 regions of S1. These preliminary findings suggested that allosteric interaction networks in the SARS-CoV-2 S complexes with S309 could be broadly distributed, which can arguably reflect strengthening of allosteric couplings between S1 and S2 subunits as S309 locks the down-regulated form of the S protein. In the framework of the hierarchical approach, we also explored a more detailed model of the residue interaction networks by using a graph-based representation with residues as network nodes and the inter-residue edges defined by both dynamic correlations<sup>104</sup> and coevolutionary residue couplings<sup>105</sup> as detailed in our previous studies.<sup>105-107</sup> Using the results of simulations, the ensemble-averaged distributions of the betweenness centrality were computed for the SARS-CoV-2 S complexes with Abs (Figure 10). We found that the high centrality residues can be assembled in tight interaction clusters localized in the key functional regions of the S protein. In

the SARS-CoV-2 S protein complexes with H014, the centrality profiles featured strong and dense peaks in the RBD and CTD1 regions of S1 as well as another peak in the CH region of S2 (residues 986-1035) (Figure 10A,B). The centrality peaks also aligned well with the hinge centers of S1 (residues 315-320, 569-572), indicating that these dynamically important control points could also mediate communication in the residue interaction networks. The network centrality analysis also revealed clusters of distribution peaks in the SARS-CoV-2 S complexes featuring the fully closed conformation (Figure 10D,E). In these structures S309 and S2M11 induce a strong stabilization effect and lock the S protein in the closed state. According to our results, these structurally stable states can also feature a broadly distributed allosteric network mediated by functional sites in both S1 and S2 subunits, primarily CTD1 regions (residues 529-591) UH (residues 736-781), CH (residues 986-1035), and  $\beta$ -hairpin (BH) region (residues 1035-1071). The dominant clusters of centrality peaks located in the RBD and CTD1 regions of S1 and CH regions of S2 can be seen in the S complex with S2E12 (Figure 10F). This showed that S2E12 binding may activate the increased mediating capacity of CTD1 regions and strengthen allosteric interactions between S1 and S2 regions.

Structural mapping of high centrality sites highlighted differences between network organizations in the SARS-CoV-2 complexes (Figure 11). In the complexes with H014 the high centrality sites are concentrated near CTD1 regions that could strengthen couplings at the inter-domain boundaries between S1 and S2 (Figure 11A,B). We argue that H014 binding may increase the allosteric potential of the RBD and CTD1 regions and activate communication between the RBD and S2 via CTD1 regions. Of particular interest is a dense network of mediating centers in the complexes with S309 and S2M11 (Figure 11C-E), showing that these Abs may facilitate a broad allosteric interaction network between S1 and S2 functional regions.



**Figure 10.** The residue-based betweenness centrality profiles in the SARS-CoV-2 S complexes with a panel of Abs. The centrality values are computed by averaging the results over 1,000 representative samples from CABS-CG simulations and atomistic reconstruction of trajectories. (A) The centrality profile is shown for the SARS-CoV-2 S complex with H014 - two RBDs in the open state (A), SARS-CoV-2 S complex with H014 - three RBD in the open state (B), SARS-CoV-2 S complex with S309 - two RBDs in the closed form (C), SARS-CoV-2 S complex with S309 - three RBDs in the closed form (D), SARS-CoV-2 S complex with S2M11 - three RBDs in the closed form (E), SARS-CoV-2 S complex with S2E12 - three RBDs in the closed form (F). The profiles for protomer chains A,B and C are shown in green, red and blue bars respectively.



**Figure 11.** Structural maps of high centrality clusters in the SARS-CoV-2 S complexes. (A) Structural map for the SARS-CoV-2 S complex with H014 - two RBDs in the open state (A), SARS-CoV-2 S complex with H014 - three RBD in the open state (B), SARS-CoV-2 S complex with S309 - two RBDs in the closed form (C), SARS-CoV-2 S complex with S3090 - three RBDs in the closed form (D), SARS-CoV-2 S complex with S2M11 - three RBDs in the closed form (E), SARS-CoV-2 S complex with S2E12 - three RBDs in the closed form (F). The protomer A is shown in green ribbons, protomer B in cyan ribbons, and protomer C in blue ribbons. The bound Abs are depicted in dark pink-colored ribbons. The high centrality residue clusters are shown in red spheres.

## **Perturbation Response Scanning Identifies Regulatory Hotspots of Allosteric Interactions in Different Conformational States of the SARS-CoV-2 Spike Trimer**

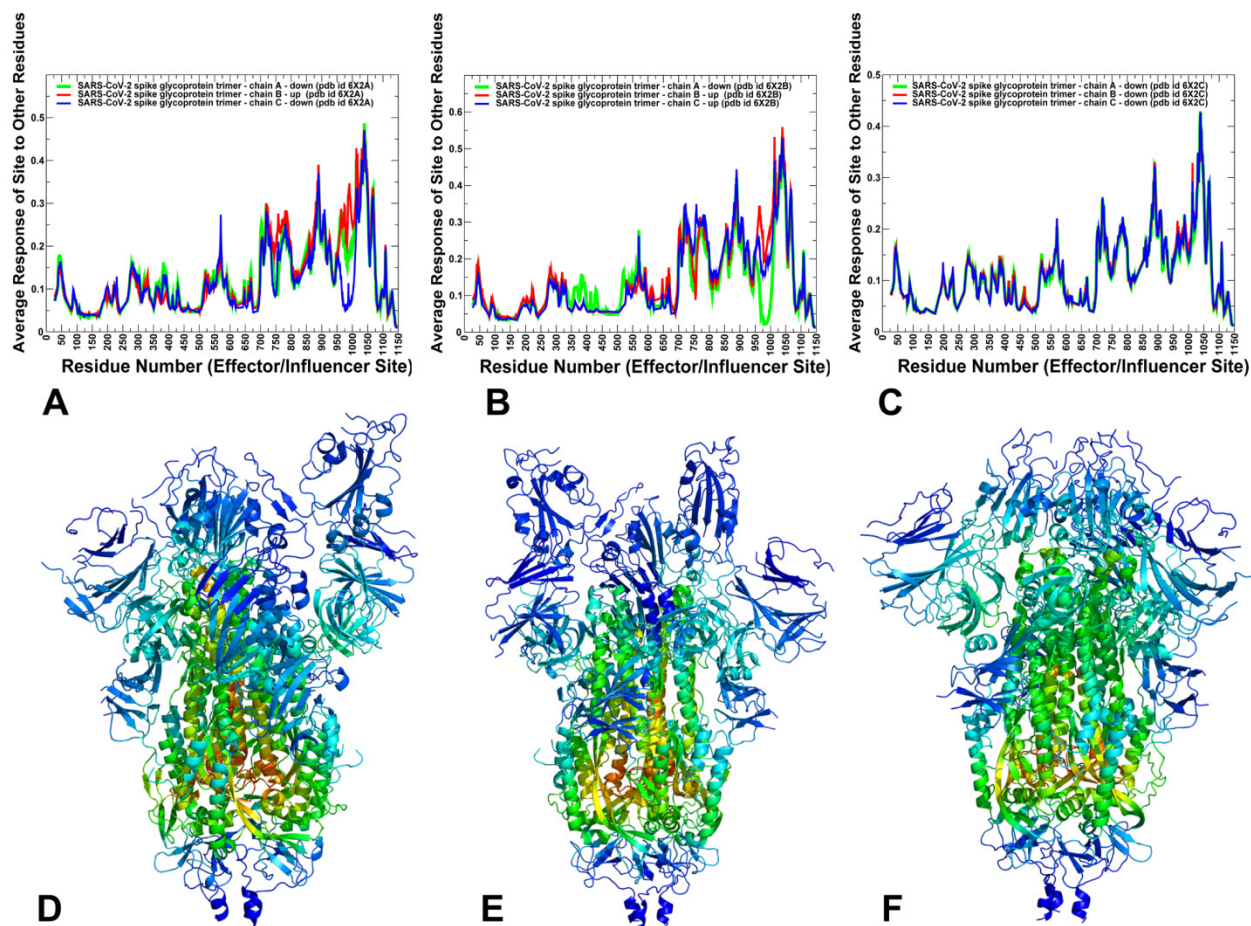
Using the PRS method<sup>112,113</sup> we quantified the allosteric effect of each residue in the SARS-CoV-2 complexes in response to external perturbations. PRS analysis produced the residue-based effector response profiles in different functional states of the unbound SARS-CoV-2 S trimer (Figure 12) and SARS-CoV-2 S complexes with H014, S309, S2M11, and S2E12 Abs (Figure 13). The effector profiles estimate the propensities of a given residue to influence dynamic changes in other residues and are applied to identify regulatory hotspots of allosteric interactions as the local maxima along the profile. The central hypothesis tested in the PRS analysis is that Ab binding can incur measurable and functionally relevant changes by modulating the effector profiles of the unbound SARS-CoV-2 S protein trimer. Moreover, we conjectured that binding can differentially affect the effector response profiles and allosteric interaction networks in distinct functional forms of the SARS-CoV-2 S protein. By systematically comparing the PRS profiles in the unbound and bound S protein forms, we determined the distribution of regulatory allosteric centers and clarified the role of specific functional regions in controlling allosteric conformational changes.

To establish the baseline for comparison of allosteric profiles in the SARS-CoV-2 complexes we first computed PRS effector profiles for the unbound S protein in the closed (3 RBDs-down), partially open (2RBDs-down, 1RBD-up) and open states (1RBD-down, 2 RBDs-up) (Figure 12). In this analysis of the unbound forms of the SARS-CoV-2 S trimer, we used the cryo-EM structure of the full-length quadruple mutant (A507L T572I F855Y N856I) in the all-down closed prefusion conformation (pdb id 6X2C), the partially open (1-up) conformation (pdb id 6X2A), and the open (2-up) conformation (pdb id 6X2B).<sup>28</sup> These structures revealed thermodynamic



and dynamic differences between different states in which dynamic switch centers are responsible for modulation of allosteric changes between the closed and open S states.<sup>28</sup> By using these structures as reference states for the PRS analysis, we also examined how Ab binding can alter the localization and effect of these regulatory switch points on allosteric interactions.

The results showed that the effector profile in the unbound S structures can remain largely conserved, displaying the largest peaks in the functional regions of the S2 subunit (Figure 12A-C). The major effector peaks corresponded to residues 756-758 in the UH region, HR1 region (residues 910-985), residues 886-890 and BH region (residues 1035-1071). In the closed and partially open forms, we also observed a secondary peak corresponding to CTD1 regions (residues 529-591) in the S1 subunit. Only a minor effector density was seen in the RBD regions (Figure 12A-C). At the same time, these regions can serve as primary sensors of allosteric signaling in the S trimers for the partially open and open states (Figure S6, Supporting Information). These results are consistent with our latest studies of the SARS-CoV-2 S structures demonstrating that the broadly distributed effector density can be seen only in the fully locked closed state, while allosteric couplings in more dynamic closed and open forms may be largely governed by the regulatory centers in core of the S2 subunit.<sup>64</sup> The reduced density of the effector centers in the RBD regions indicated that the allosteric signaling in the dynamic closed and open forms may be primarily one-directional, in which allosteric centers in the S2 core regions could dictate allosteric changes in the RBD regions. Based on this preliminary evidence, we suggested that efficient allosteric signaling between S1 and S2 subunits and broad allosteric networks may be salient features of the thermodynamically locked closed S form.



**Figure 12.** The PRS effector profiles in the closed, partially open and open states of the SARS-CoV-2 spike trimers. (A) The PRS effector profile is shown for the ligand-free SARS-CoV-2 S trimer in the partially open (1 RBD up) conformation (pdb id 6X2A). (B) The PRS effector profile for the ligand-free SARS-CoV-2 S trimer in the open (2 RBDs up) conformation (pdb id 6X2B). (C) The PRS effector profile for the ligand-free SARS-CoV-2 S trimer in the closed (3 RBDs-down) prefusion conformation (pdb id 6X2C). The profiles are shown for protomer in green lines, protomer B in red lines, and protomer C in blue lines. Structural maps of the PRS effector profiles are shown for the partially open state of the SARS-CoV-2 S prefusion trimer (D), open state (E), and closed state (F). The color gradient from blue to red indicates the increasing effector propensities.

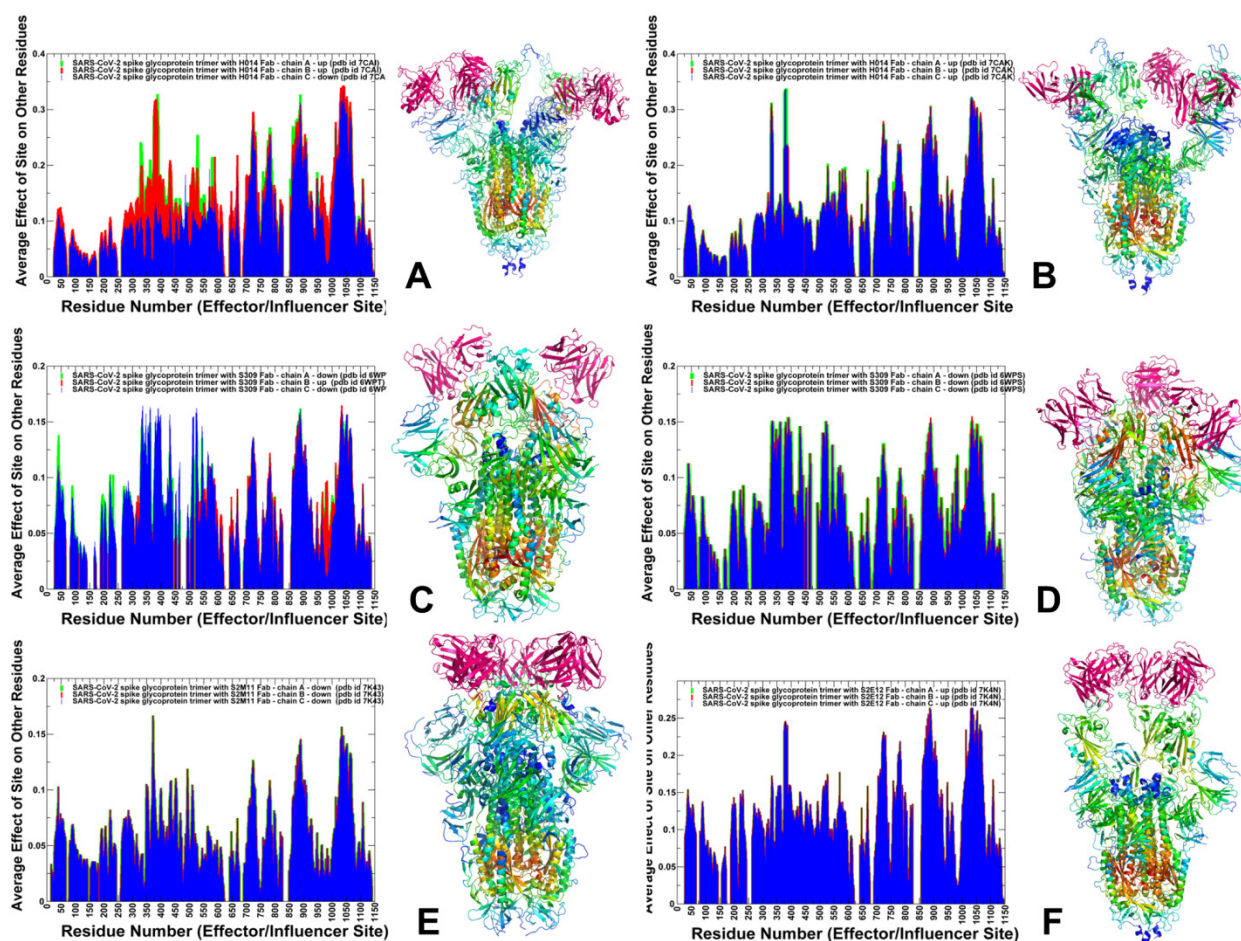
We conjectured that Ab binding at different RBD epitopes may affect not only local interactions and stability near the binding sites but also have long-range effect by modulating allosteric effector potential of SARS-CoV-2 S regions and altering allosteric interaction networks. This can allow for highly cooperative motions in which many spatially distributed effector residues are in the allosteric network that links distant S1 and S2 functional regions. In this model, allostery requires an effector ligand to stabilize the interactions in the closed S state over those in the open S states.

The central result of this analysis is that the studied neutralizing Abs could effectively restore allosteric potential of the RBD and CTD1 regions in the closed and open states without compromising the effector potential of S2 regions, thereby introducing ligand-induced cooperativity and strengthening the broad allosteric interaction network (Figure 13). Indeed, the results showed that H014 binding induced significant changes in the effector profile of the RBD and CTD1 regions for “up” protomers while preserving and enhancing the effector capacity of the CH and CD residues (Figure 13A,B). In the partially open form of the SARS-CoV-2 complex, the effector peak center corresponded to RBD residues S383, F377, K378, C379, Y380, G381, V382, S383, P384 involved in direct productive interactions with H014 (Figure 13A). Several of these effector centers (F377, C379, and S383) also corresponded to conserved positions implicated in mediating coevolutionary couplings. In addition, we noticed a significant increase of the effector potential for the CTD1 regions in the open protomers. Similar findings were observed in the open form (3RBDs - up) of the SARS-CoV-2 S complex with H014 where the allosteric potential of RBD regions was markedly enhanced (Figure 13B).

Structural maps of the effector profiles further illustrated these findings by showing the increased effector density penetrating into the RBDs of protomers that interact with H014 (Figure 13A,B).

These results suggested that H014 binding could modulate the effector propensities of the RBD residues and restore the allosteric potential of the S1 regions, leading to strengthening of the network of allosteric interactions in the complex. The effector profiles also indicated the increased density and clustering of effector peaks distributed across RBD, CTD1, UH and CH regions (Figures 13A,B).

The effector profiles of SARS-CoV-2 complexes with S309 showed the significantly increased effector potential of the RBD and CTD1 regions, which is manifested in the emergence of dominant and broad peaks in these S1 regions (Figure 13C,D). Among emerging effector peaks in the RBD regions are T333, N334, C361, V362, A363, L390 and C391. These residues correspond to highly conserved and structurally stable positions involved in stabilizing disulfide linkages, C336–C361 and C391–C525 that anchor the  $\beta$ -sheet structure in the RBD regions. The S309-induced modulation of allosteric effector propensities could be also amplified by the fact that this Ab appears to thermodynamically stabilize partially closed and closed forms of the S protein where two or all three RBDs assume the down-regulated conformation (Figure 13C,D). Structural maps highlighted a significant expansion of the effector density towards S1 regions and boundaries between S1 and S2 subunits (Figure 13C,D). By strengthening allosteric couplings between S1 and S2 subunits, S309 could arguably lock the down-regulated form to ensure S1-based protection of the fusion machinery.



**Figure 13.** The PRS effector profiles for the SARS-CoV-2 S complexes with Abs. (A) The PRS effector distributions and structural maps of the effector profile are shown for the SARS-CoV-2 S complex with H014 - two RBDs in the open state (A), SARS-CoV-2 S complex with H014 - three RBD in the open state (B), SARS-CoV-2 S complex with S309 - two RBDs in the closed form (C), SARS-CoV-2 S complex with S3090 - three RBDs in the closed form (D), SARS-CoV-2 S complex with S2M11 - three RBDs in the closed form (E), SARS-CoV-2 S complex with S2E12 - three RBDs in the closed form (F). The profiles are shown for protomer in green bars, protomer B in red bars, and protomer C in blue bars. Structural maps of the PRS effector profiles are shown with the color gradient from blue to red indicating the increasing effector propensities.

S2M11 binds a quaternary epitope comprising distinct regions of two neighboring RBDs within an S trimer and induced stabilization of SARS-CoV-2 S in the closed conformational state. We found that S2M11 binding can promote the increased effector potential of conserved and structurally stable residues that are not directly involved in binding contacts (Figure 13E). Indeed, S2M11 can induce the increased effector potential in the RBD regions, particularly residues F374, F377, K378 C379, Y380 (Figure 13E). S2E12 recognizes an RBD epitope overlapping with the RBM that is partially buried at the interface between protomers in the closed S trimer and therefore S2E12 can only interact with open RBDs. According to our results, S2E12 binding can cause a similar redistribution of allosteric effector potential in the S1 regions and activate effector capacity of conserved stretch of residues in the  $\beta$ -sheet of the RBD regions (Figure 13F). Strikingly, S2E12-induced modulation of the effector propensities in the fully open S conformation can effectively restore allosteric potential for the RBD and CTD1 S1 regions, while strengthening peaks in the UH and CH regions of S2 subunit. As a result, as highlighted by structural projection of the effector profiles, the potential effector centers become spatially distributed across the S trimer structure (Figure 13F).

The PRS sensor profiles describe the propensity of flexible residues to serve as carriers of large allosteric conformational changes. We found that Ab binding can induce changes in the shape and peaks of sensor profiles in the partially open and closed S forms (Figure S7, Supporting Information). The sensor peaks in the RBD regions that are prevalent in the unbound S protein become partially suppressed, featuring now a broader distribution of small peaks localized across S1 and S2 regions. This suggested that the sensor propensities of the RBD regions can be significantly reduced in the SARS-CoV-2 S complexes, which reflects conformational and dynamic constraints imposed by Abs on large structural changes.

The results of the PRS analysis also suggested that allosteric mechanisms underlying Ab binding to S proteins may bear signs of ligand-induced cooperativity in which the effector can shift the distribution of local interactions and energies for many residues.<sup>142</sup> Based on our results, we argue that Abs can induce a switch from a moderately cooperative population-shift mechanism of the unbound S protein to a highly cooperative ligand-induced allosteric mechanism. While the allosteric interaction network of the unbound S protein in the population-shift mechanism tends to be more dispersed and smaller, binding can induce a large and dense allosteric network that efficiently couples local changes in the distant S1 and S2 regions.

In this context, it is worth noting that cooperative allosteric mechanisms with a broad allosteric network tend to better withstand mutations in the functional regions without significant deleterious consequences for protein function.<sup>142</sup> Accordingly, it may be suggested that the ligand-induced cooperative allosteric effect produced by Ab binding may enhance resistance against mutations so that mutational changes would not easily alter conformational preferences and expose the RBD regions to interactions with the host receptor.<sup>143-147</sup> In some contrast, a less cooperative population-shift mechanism in the unbound S protein may be more susceptible and vulnerable to mutations of residues in the communication network, which may allow individual mutations at the regulatory switch centers to alter conformational equilibrium and potentially increase population of the receptor-accessible open S conformations.

## Conclusions

This study examined molecular mechanisms underlying SARS-CoV-2 S protein binding with a panel of highly potent Abs through the lens of coevolutionary relationships and ligand-induced modulation of allosteric interaction networks. The results revealed key functional regions and regulatory centers that govern coevolutionary couplings and allosteric interactions in the SARS-CoV-2S protein complexes. We found that Ab-specific targeting of coevolutionary hotspots in the S protein can allow for efficient modulation of long-range interactions between S1 and S2 units by propagating signal through clusters of spatially proximal coevolutionary coupled residues. The results revealed strong coevolutionary signal between functional regions separated by the medium-range distances which may help to facilitate a long-range cross-talk between distant allosteric regions. Conformational dynamics and binding energetics analyses showed that binding of Abs can lead to significant stabilization of both S1 and S2 regions which may be relevant in rationalization of potent neutralization effects. The PRS analysis of the unbound and bound SARS-CoV-2 S proteins showed that Abs can promote formation of highly cooperative and broad allosteric networks that restore and enhance couplings between S1 and S2 regions, thereby inhibiting dissociation of S1 subunit from the spike apparatus required for membrane fusion. By systematically comparing the PRS profiles, we clarified the role of specific functional regions in regulating allosteric interactions. The results of this study provide a novel insight into allosteric regulatory mechanisms of SARS-CoV-2 S proteins showing that Abs can uniquely modulate signal communication providing a plausible strategy for therapeutic intervention by targeting specific hotspots of allosteric interactions in the SARS-CoV-2 proteins.



## **AUTHOR INFORMATION**

### **\* Corresponding Author**

Phone: 714-516-4586; Fax: 714-532-6048; E-mail: [verkhivk@chapman.edu](mailto:verkhivk@chapman.edu)

The authors declare no competing financial interest.

**Acknowledgment** This work was partly supported by institutional funding from Chapman University. The author acknowledges support by the Kay Family Foundation Grant A20-0032.

## **ABBREVIATIONS**

SARS, Severe Acute Respiratory Syndrome; RBD, Receptor Binding Domain; ACE2, Angiotensin-Converting Enzyme 2 (ACE2); NTD, N-terminal domain; RBD, receptor-binding domain; CTD1, C-terminal domain 1; CTD2, C-terminal domain 2; FP, fusion peptide; FPPR, fusion peptide proximal region; HR1, heptad repeat 1; CH, central helix region; CD, connector domain; HR2, heptad repeat 2; TM, transmembrane anchor; CT, cytoplasmic tail.

## References

- (1) Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K. S. M.; Lau, E. H. Y.; Wong, J. Y.; Xing, X.; Xiang, N.; Wu, Y.; Li, C.; Chen, Q.; Li, D.; Liu, T.; Zhao, J.; Liu, M.; Tu, W.; Chen, C.; Jin, L.; Yang, R.; Wang, Q.; Zhou, S.; Wang, R.; Liu, H.; Luo, Y.; Liu, Y.; Shao, G.; Li, H.; Tao, Z.; Yang, Y.; Deng, Z.; Liu, B.; Ma, Z.; Zhang, Y.; Shi, G.; Lam, T. T. Y.; Wu, J. T.; Gao, G. F.; Cowling, B. J.; Yang, B.; Leung, G. M.; Feng, Z. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **2020**, 382, 1199-1207.
- (2) Wang, C.; Horby, P. W.; Hayden, F. G.; Gao, G. F. A novel coronavirus outbreak of global health concern. *Lancet* **2020**, 395, 470-473.
- (3) Yi, Y.; Lagniton, P. N. P.; Ye, S.; Li, E.; Xu, R. H. COVID-19: what has been learned and to be learned about the novel coronavirus disease. *Int. J. Biol. Sci.* **2020**, 16, 1753-1766.
- (4) Wu, A.; Peng, Y.; Huang, B.; Ding, X.; Wang, X.; Niu, P.; Meng, J.; Zhu, Z.; Zhang, Z.; Wang, J.; Sheng, J.; Quan, L.; Xia, Z.; Tan, W.; Cheng, G.; Jiang, T. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* **2020**, 27, 325-328.
- (5) Tai, W.; He, L.; Zhang, X.; Pu, J.; Voronin, D.; Jiang, S.; Zhou, Y.; Du, L. Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell. Mol. Immunol.* **2020**, 17, 613-620.
- (6) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; Müller, M. A.; Drosten, C.; Pöhlmann, S. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **2020**, 181, 271-280.e8.
- (7) Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; Bi, Y.; Ma, X.; Zhan, F.; Wang, L.; Hu, T.; Zhou, H.; Hu, Z.; Zhou, W.; Zhao, L.; Chen, J.; Meng,

- Y.; Wang, J.; Lin, Y.; Yuan, J.; Xie, Z.; Ma, J.; Liu, W. J.; Wang, D.; Xu, W.; Holmes, E. C.; Gao, G. F.; Wu, G.; Chen, W.; Shi, W.; Tan, W. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **2020**, 395, 565-574.
- (8) Duan, L.; Zheng, Q.; Zhang, H.; Niu, Y.; Lou, Y.; Wang, H. The SARS-CoV-2 spike glycoprotein biosynthesis, structure, function, and antigenicity: Implications for the design of spike-based vaccine immunogens. *Front. Immunol.* **2020**, 11, 576622.
- (9) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K. Y.; Zhou, H.; Yan, J.; Qi, J. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell* **2020**, 181, 894-904.e9.
- (10) Wan, Y.; Shang, J.; Graham, R.; Baric, R. S.; Li, F. Receptor recognition by the novel coronavirus from Wuhan: An analysis based on decade-long structural studies of SARS coronavirus. *J. Virol.* **2020**, 94, e00127-20.
- (11) Shang, J.; Wan, Y.; Luo, C.; Ye, G.; Geng, Q.; Auerbach, A.; Li, F. Cell entry mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, 117, 11727-11734.
- (12) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K. Y.; Zhou, H.; Yan, J.; Qi, J. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **2020**, 181, 894-904.e9.
- (13) Wan, Y.; Shang, J.; Graham, R.; Baric, R. S.; Li, F. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J. Virol.* **2020**, 94, e00127-20.
- (14) Li, F.; Li, W.; Farzan, M.; Harrison, S. C. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* **2005**, 309, 1864-1868.

- (15) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020**, 581, 215-220.
- (16) Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020**, 581, 221-224.
- (17) Kirchdoerfer, R. N.; Cottrell, C. A.; Wang, N.; Pallesen, J.; Yassine, H. M.; Turner, H. L.; Corbett, K. S.; Graham, B. S.; McLellan, J. S.; Ward, A. B. Pre-fusion structure of a human coronavirus spike protein. *Nature* **2016**, 531, 118-121.
- (18) Walls, A. C.; Tortorici, M. A.; Bosch, B. J.; Frenz, B.; Rottier, P. J. M.; DiMaio, F.; Rey, F. A.; Velesler, D. Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature* **2016**, 531, 114-117.
- (19) Gui, M.; Song, W.; Zhou, H.; Xu, J.; Chen, S.; Xiang, Y.; Wang, X. Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Res.* **2017**, 27, 119-129.
- (20) Walls, A. C.; Xiong, X.; Park, Y. J.; Tortorici, M. A.; Snijder, J.; Quispe, J.; Cameroni, E.; Gopal, R.; Dai, M.; Lanzavecchia, A.; Zambon, M.; Rey, F. A.; Corti, D.; Velesler, D. Unexpected receptor functional mimicry elucidates activation of coronavirus fusion. *Cell* **2019**, 176, 1026-1039.e15.
- (21) Yuan, Y.; Cao, D.; Zhang, Y.; Ma, J.; Qi, J.; Wang, Q.; Lu, G.; Wu, Y.; Yan, J.; Shi, Y.; Zhang, X.; Gao, G. F. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat. Commun.* **2017**, 8, 15092.
- (22) Song, W.; Gui, M.; Wang, X.; Xiang, Y. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLoS Pathog.* **2018**, 14, e1007236.

(23) Kirchdoerfer, R. N.; Wang, N.; Pallesen, J.; Wrapp, D.; Turner, H. L.; Cottrell, C. A.; Corbett, K. S.; Graham, B. S.; McLellan, J. S.; Ward, A. B. Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. *Sci. Rep.* **2018**, *8*, 15701.

(24) Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Velesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181*, 281-292.e6.

(25) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C. L.; Abiona, O.; Graham, B. S.; McLellan, J. S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367*, 1260-1263.

(26) Cai, Y.; Zhang, J.; Xiao, T.; Peng, H.; Sterling, S. M.; Walsh, R. M., Jr.; Rawson, S.; Rits-Volloch, S.; Chen, B. Distinct conformational states of SARS-CoV-2 spike protein. *Science* **2020**, *369*, 1586-1592.

(27) Hsieh, C. L.; Goldsmith, J. A.; Schaub, J. M.; DiVenere, A. M.; Kuo, H. C.; Javanmardi, K.; Le, K. C.; Wrapp, D.; Lee, A. G.; Liu, Y.; Chou, C. W.; Byrne, P. O.; Hjorth, C. K.; Johnson, N. V.; Ludes-Meyers, J.; Nguyen, A. W.; Park, J.; Wang, N.; Amengor, D.; Lavinder, J. J.; Ippolito, G. C.; Maynard, J. A.; Finkelstein, I. J.; McLellan, J. S. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **2020**, *369*, 1501-1505.

(28) Henderson, R.; Edwards, R. J.; Mansouri, K.; Janowska, K.; Stalls, V.; Gobeil, S. M. C.; Kopp, M.; Li, D.; Parks, R.; Hsu, A. L.; Borgnia, M. J.; Haynes, B. F.; Acharya, P. Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nat. Struct. Mol. Biol.* **2020**, *27*, 925-933.

- (29) McCallum, M.; Walls, A. C.; Bowen, J. E.; Corti, D.; Veessler, D. Structure-guided covalent stabilization of coronavirus spike glycoprotein trimers in the closed conformation. *Nat. Struct. Mol. Biol.* **2020**, *27*, 942-949.
- (30) Xiong, X.; Qu, K.; Ciazynska, K. A.; Hosmillo, M.; Carter, A. P.; Ebrahimi, S.; Ke, Z.; Scheres, S. H. W.; Bergamaschi, L.; Grice, G. L.; Zhang, Y.; Nathan, J. A.; Baker, S.; James, L. C.; Baxendale, H. E.; Goodfellow, I.; Doffinger, R.; Briggs, J. A. G. A thermostable, closed SARS-CoV-2 spike protein trimer. *Nat. Struct. Mol. Biol.* **2020**, *27*, 934-941.
- (31) Yurkovetskiy, L.; Wang, X.; Pascal, K. E.; Tomkins-Tinch, C.; Nyalile, T. P.; Wang, Y.; Baum, A.; Diehl, W. E.; Dauphin, A.; Carbone, C.; Veinotte, K.; Egri, S. B.; Schaffner, S. F.; Lemieux, J. E.; Munro, J. B.; Rafique, A.; Barve, A.; Sabeti, P. C.; Kyratsous, C. A.; Dudkina, N. V.; Shen, K.; Luban, J. Structural and functional analysis of the D614G SARS-CoV-2 spike protein variant. *Cell* **2020**, *183*, 739-751.e8.
- (32) Turoňová, B.; Sikora, M.; Schürmann, C.; Hagen, W. J. H.; Welsch, S.; Blanc, F. E. C.; von Bülow, S.; Gecht, M.; Bagola, K.; Hörner, C.; van Zandbergen, G.; Landry, J.; de Azevedo, N. T. D.; Mosalaganti, S.; Schwarz, A.; Covino, R.; Mühlebach, M. D.; Hummer, G.; Krijnse Locker, J.; Beck, M. In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science* **2020**, *370*, 203-208.
- (33) Lu, M.; Uchil, P. D.; Li, W.; Zheng, D.; Terry, D. S.; Gorman, J.; Shi, W.; Zhang, B.; Zhou, T.; Ding, S.; Gasser, R.; Prevost, J.; Beaudoin-Bussieres, G.; Anand, S. P.; Laumaea, A.; Grover, J. R.; Lihong, L.; Ho, D. D.; Mascola, J.; Finzi, A.; Kwong, P. D.; Blanchard, S. C.; Mothes, W. Real-time conformational dynamics of SARS-CoV-2 spikes on virus particles. *bioRxiv* **2020**, doi: 10.1101/2020.09.10.286948.

(34) Benton, D. J.; Wrobel, A. G.; Xu, P.; Roustan, C.; Martin, S. R.; Rosenthal, P. B.; Skehel, J. J.; Gamblin, S. J. Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. *Nature* **2020**, doi: 10.1038/s41586-020-2772-0.

(35) He, Y.; Lu, H.; Siddiqui, P.; Zhou, Y.; Jiang, S. Receptor-binding domain of severe acute respiratory syndrome coronavirus spike protein contains multiple conformation-dependent epitopes that induce highly potent neutralizing antibodies. *J. Immunol.* **2005**, 174, 4908-4915.

(36) He, Y.; Li, J.; Li, W.; Lustigman, S.; Farzan, M.; Jiang, S. Cross-neutralization of human and palm civet severe acute respiratory syndrome coronaviruses by antibodies targeting the receptor-binding domain of spike protein. *J. Immunol.* **2006**, 176, 6085-92.

(37) Chen, X.; Li, R.; Pan, Z.; Qian, C.; Yang, Y.; You, R.; Zhao, J.; Liu, P.; Gao, L.; Li, Z.; Huang, Q.; Xu, L.; Tang, J.; Tian, Q.; Yao, W.; Hu, L.; Yan, X.; Zhou, X.; Wu, Y.; Deng, K.; Zhang, Z.; Qian, Z.; Chen, Y.; Ye, L. Human monoclonal antibodies block the binding of SARS-CoV-2 spike protein to angiotensin converting enzyme 2 receptor. *Cell Mol. Immunol.* **2020**, 17, 647-649.

(38) Wrapp, D.; De Vlieger, D.; Corbett, K. S.; Torres, G. M.; Wang, N.; Van Breedam, W.; Roose, K.; van Schie, L.; Hoffmann, M.; Pöhlmann, S.; Graham, B. S.; Callewaert, N.; Schepens, B.; Saelens, X.; McLellan, J. S. Structural Basis for Potent Neutralization of Betacoronaviruses by Single-Domain Camelid Antibodies. *Cell* **2020**, 181, 1004-1015.e15.

(39) Yu, F.; Xiang, R.; Deng, X.; Wang, L.; Yu, Z.; Tian, S.; Liang, R.; Li, Y.; Ying, T.; Jiang, S. Receptor-binding domain-specific human neutralizing monoclonal antibodies against SARS-CoV and SARS-CoV-2. *Signal Transduct. Target. Ther.* **2020**, 5, 212.

(40) Wang, C.; Li, W.; Drabek, D.; Okba, N. M. A.; van Haperen, R.; Osterhaus, A.; van Kuppeveld, F. J. M.; Haagmans, B. L.; Grosveld, F.; Bosch, B. J. A human monoclonal antibody blocking SARS-CoV-2 infection. *Nat. Commun.* **2020**, 11, 2251.

(41) Hwang, W. C.; Lin, Y.; Santelli, E.; Sui, J.; Jaroszewski, L.; Stec, B.; Farzan, M.; Marasco, W. A.; Liddington, R. C. Structural basis of neutralization by a human anti-severe acute respiratory syndrome spike protein antibody, 80R. *J. Biol. Chem.* **2006**, 281, 34610-34616.

(42) Yuan, M.; Wu, N. C.; Zhu, X.; Lee, C. D.; So, R. T. Y.; Lv, H.; Mok, C. K. P.; Wilson, I. A. A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* **2020**, 368, 630-633.

(43) Huo, J.; Zhao, Y.; Ren, J.; Zhou, D.; Duyvesteyn, H. M. E.; Ginn, H. M.; Carrique, L.; Malinauskas, T.; Ruza, R. R.; Shah, P. N. M.; Tan, T. K.; Rijal, P.; Coombes, N.; Bewley, K. R.; Tree, J. A.; Radecke, J.; Paterson, N. G.; Supasa, P.; Mongkolsapaya, J.; Sreaton, G. R.; Carroll, M.; Townsend, A.; Fry, E. E.; Owens, R. J.; Stuart, D. I. Neutralization of SARS-CoV-2 by destruction of the prefusion spike. *Cell Host Microbe* **2020**, 28, 445-454.e6.

(44) Brouwer, P. J. M.; Caniels, T. G.; van der Straten, K.; Snitselaar, J. L.; Aldon, Y.; Bangaru, S.; Torres, J. L.; Okba, N. M. A.; Claireaux, M.; Kerster, G.; Bentlage, A. E. H.; van Haaren, M. M.; Guerra, D.; Burger, J. A.; Schermer, E. E.; Verheul, K. D.; van der Velde, N.; van der Kooi, A.; van Schooten, J.; van Breemen, M. J.; Bijl, T. P. L.; Slieden, K.; Aartse, A.; Derking, R.; Bontjer, I.; Kootstra, N. A.; Wiersinga, W. J.; Vidarsson, G.; Haagmans, B. L.; Ward, A. B.; de Bree, G. J.; Sanders, R. W.; van Gils, M. J. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* **2020**, 369, 643-650.

(45) Lv, Z.; Deng, Y. Q.; Ye, Q.; Cao, L.; Sun, C. Y.; Fan, C.; Huang, W.; Sun, S.; Sun, Y.; Zhu, L.; Chen, Q.; Wang, N.; Nie, J.; Cui, Z.; Zhu, D.; Shaw, N.; Li, X. F.; Li, Q.; Xie, L.;



Wang, Y.; Rao, Z.; Qin, C. F.; Wang, X. Structural basis for neutralization of SARS-CoV-2 and SARS-CoV by a potent therapeutic antibody. *Science* **2020**, 369, 1505-1509.

(46) Pinto, D.; Park, Y. J.; Beltramello, M.; Walls, A. C.; Tortorici, M. A.; Bianchi, S.; Jaconi, S.; Culap, K.; Zatta, F.; De Marco, A.; Peter, A.; Guarino, B.; Spreafico, R.; Cameroni, E.; Case, J. B.; Chen, R. E.; Havenar-Daughton, C.; Snell, G.; Telenti, A.; Virgin, H. W.; Lanzavecchia, A.; Diamond, M. S.; Fink, K.; Veessler, D.; Corti, D. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **2020**, 583, 290-295.

(47) Tortorici, M. A.; Beltramello, M.; Lempp, F. A.; Pinto, D.; Dang, H. V.; Rosen, L. E.; McCallum, M.; Bowen, J.; Minola, A.; Jaconi, S.; Zatta, F.; De Marco, A.; Guarino, B.; Bianchi, S.; Lauron, E. J.; Tucker, H.; Zhou, J.; Peter, A.; Havenar-Daughton, C.; Wojcechowskyj, J. A.; Case, J. B.; Chen, R. E.; Kaiser, H.; Montiel-Ruiz, M.; Meury, M.; Czudnochowski, N.; Spreafico, R.; Dillen, J.; Ng, C.; Sprugasci, N.; Culap, K.; Benigni, F.; Abdelnabi, R.; Foo, S. C.; Schmid, M. A.; Cameroni, E.; Riva, A.; Gabrieli, A.; Galli, M.; Pizzuto, M. S.; Neyts, J.; Diamond, M. S.; Virgin, H. W.; Snell, G.; Corti, D.; Fink, K.; Veessler, D. Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms. *Science* **2020**, 370, 950-957.

(48) Chi, X.; Yan, R.; Zhang, J.; Zhang, G.; Zhang, Y.; Hao, M.; Zhang, Z.; Fan, P.; Dong, Y.; Yang, Y.; Chen, Z.; Guo, Y.; Li, Y.; Song, X.; Chen, Y.; Xia, L.; Fu, L.; Hou, L.; Xu, J.; Yu, C.; Li, J.; Zhou, Q.; Chen, W., A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* **2020**, 369, 650-655.

(49) Schoof, M.; Faust, B.; Saunders, R. A.; Sangwan, S.; Rezelj, V.; Hoppe, N.; Boone, M.; Billesbølle, C. B.; Puchades, C.; Azumaya, C. M.; Kratochvil, H. T.; Zimanyi, M.; Deshpande, I.; Liang, J.; Dickinson, S.; Nguyen, H. C.; Chio, C. M.; Merz, G. E.; Thompson, M. C.;

Diwanji, D.; Schaefer, K.; Anand, A. A.; Dobzinski, N.; Zha, B. S.; Simoneau, C. R.; Leon, K.; White, K. M.; Chio, U. S.; Gupta, M.; Jin, M.; Li, F.; Liu, Y.; Zhang, K.; Bulkley, D.; Sun, M.; Smith, A. M.; Rizo, A. N.; Moss, F.; Brilot, A. F.; Pourmal, S.; Trenker, R.; Pospiech, T.; Gupta, S.; Barsi-Rhyne, B.; Belyy, V.; Barile-Hill, A. W.; Nock, S.; Krogan, N. J.; Ralston, C. Y.; Swaney, D. L.; García-Sastre, A.; Ott, M.; Vignuzzi, M.; Walter, P.; Manglik, A. An ultrapotent synthetic nanobody neutralizes SARS-CoV-2 by stabilizing inactive Spike. *Science* **2020**, 370, 1473-1479.

(50) Woo, H.; Park, S. J.; Choi, Y. K.; Park, T.; Tanveer, M.; Cao, Y.; Kern, N. R.; Lee, J.; Yeom, M. S.; Croll, T. I.; Seok, C.; Im, W. Developing a fully glycosylated full-length SARS-CoV-2 spike protein model in a viral membrane. *J. Phys. Chem. B* **2020**, 124, 7128-7137.

(51) Casalino, L.; Gaieb, Z.; Goldsmith, J.A.; Hjorth, C.K.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; McLellan, J.S.; Fadda, E.; Amaro, R. E. Beyond shielding: The roles of glycans in the SARS-CoV-2 spike protein. *ACS Cent. Sci.* **2020**, 6, 10, 1722-1734.

(52) Kalathiya, U.; Padariya, M.; Mayordomo, M.; Lisowska, M.; Nicholson, J.; Singh, A.; Baginski, M.; Fahraeus, R.; Carragher, N.; Ball, K.; Haas, J.; Daniels, A.; Hupp, T. R.; Alfaro, J. A. Highly conserved homotrimer cavity formed by the SARS-CoV-2 spike glycoprotein: A novel binding site. *J. Clin. Med.* **2020**, 9, 1473.

(53) Brielle, E. S.; Schneidman-Duhovny, D.; Linial, M. The SARS-CoV-2 exerts a distinctive strategy for interacting with the ACE2 human receptor. *Viruses* **2020**, 12,497.

(54) Spinello, A.; Saltalamacchia, A.; Magistrato, A. 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.* **2020**, 11, 4785-4790.

- (55) Chen, Y.; Guo, Y.; Pan, Y.; Zhao, Z. J. Structure analysis of the receptor binding of 2019-nCoV. *Biochem. Biophys. Res. Commun.* **2020**, 525, 135-40.
- (56) Othman, H.; Bouslama, Z.; Brandenburg, J. T.; da Rocha, J.; Hamdi, Y.; Ghedira, K.; Srairi-Abid, N.; Hazelhurst, S. Interaction of the spike protein RBD from SARS-CoV-2 with ACE2: Similarity with SARS-CoV, hot-spot analysis and effect of the receptor polymorphism. *Biochem. Biophys. Res. Commun.* **2020**, 527, 702-708.
- (57) Ghorbani, M.; Brooks, B. R.; Klauda, J. B. Critical Sequence Hot-spots for Binding of nCoV-2019 to ACE2 as Evaluated by Molecular Simulations. *bioRxiv* **2020**, doi: 10.1101/2020.06.27.175448.
- (58) Veeramachaneni, G. K.; Thunuguntla, V.; Bobbillapati, J.; Bondili, J. S. Structural and simulation analysis of hotspot residues interactions of SARS-CoV 2 with human ACE2 receptor. *J. Biomol. Struct. Dyn.* **2020**, 1-11.
- (59) Wang, Y.; Liu, M.; Gao, J. Enhanced receptor binding of SARS-CoV-2 through networks of hydrogen-bonding and hydrophobic interactions. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, 117, 13967-13974.
- (60) Laurini, E.; Marson, D.; Aulic, S.; Fermeglia, M.; Pricl, S. Computational alanine scanning and structural analysis of the SARS-CoV-2 Spike protein/angiotensin-converting enzyme 2 complex. *ACS Nano* **2020**, 14, 11821-11830.
- (61) Ali, A.; Vijayan, R. Dynamics of the ACE2-SARS-CoV-2/SARS-CoV spike protein interface reveal unique mechanisms. *Sci. Rep.* **2020**, 10, 14214.
- (62) Verkhivker, G.M. Coevolution, dynamics and allostery conspire in shaping cooperative binding and signal transmission of the SARS-CoV-2 spike protein with human angiotensin-converting enzyme 2. *Int. J. Mol. Sci.* **2020**, 21, 8268.

- (63) Di Paola, L.; Hadi-Alijanvand, H.; Song, X.; Hu, G.; Giuliani, A. The discovery of a putative allosteric site in the SARS-CoV-2 spike protein using an integrated structural/dynamic approach. *J. Proteome Res.* **2020**, *19*, 4576-4586.
- (64) Verkhivker, G.M. Molecular simulations and network modeling reveal an allosteric signaling in the SARS-CoV-2 spike proteins. *J. Proteome Res.* **2020**, *19*, 4587-4608.
- (65) Rozewicki, J.; Li, S.; Amada, K. M.; Standley, D. M.; Katoh, K. MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res.* **2019**, *47*, W5-W10.
- (66) Suzek, B. E.; Huang, H.; McGarvey, P.; Mazumder, R.; Wu, C. H. UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* **2007**, *23*, 1282-1288.
- (67) Suzek, B. E.; Wang, Y.; Huang, H.; McGarvey, P. B.; Wu, C. H. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* **2015**, *31*, 926-932.
- (68) El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S. R.; Luciani, A.; Potter, S. C.; Qureshi, M.; Richardson, L. J.; Salazar, G. A.; Smart, A.; Sonnhammer, E. L. L.; Hirsh, L.; Paladin, L.; Piovesan, D.; Tosatto, S. C. E.; Finn, R. D. The Pfam protein families database in 2019. *Nucleic Acids Res.* **2019**, *47*, D427-D432.
- (69) Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G. A.; Sonnhammer, E. L. L.; Tosatto, S. C. E.; Paladin, L.; Raj, S.; Richardson, L. J.; Finn, R. D.; Bateman, A., Pfam: The protein families database in 2021. *Nucleic Acids Res* **2020**. doi: 10.1093/nar/gkaa913.
- (70) Wu, C. H.; Apweiler, R.; Bairoch, A.; Natale, D. A.; Barker, W. C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; Magrane, M.; Martin, M. J.; Mazumder, R.; O'Donovan, C.; Redaschi, N.; Suzek, B. The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res.* **2006**, *34*, D187-D191.

- (71) Marino Buslje, C.; Teppa, E.; Di Domenico, T.; Delfino, J. M.; Nielsen, M. Networks of high mutual information define the structural proximity of catalytic sites: implications for catalytic residue identification. *PLoS Comput. Biol.* **2010**, *6*, e1000978.
- (72) Simonetti, F. L.; Teppa, E.; Chernomoretz, A.; Nielsen, M.; Marino Buslje, C. MISTIC: Mutual information server to infer coevolution. *Nucleic Acids Res* **2013**, *41*, W8-W14.
- (73) Colell, E. A.; Iserte, J. A.; Simonetti, F. L.; Marino-Buslje, C., MISTIC2: comprehensive server to study coevolution in protein families. *Nucleic Acids Res* **2018**, *46*, W323-W328.
- (74) Morcos, F.; Pagnani, A.; Lunt, B.; Bertolino, A.; Marks, D. S.; Sander, C.; Zecchina, R.; Onuchic, J. N.; Hwa, T.; Weigt, M. Direct-coupling analysis of residue coevolution captures native contacts across many protein families. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, E1293-E1301.
- (75) Morcos, F.; Hwa, T.; Onuchic, J. N.; Weigt, M. Direct coupling analysis for protein contact prediction. *Methods Mol. Biol.* **2014**, *1137*, 55-70.
- (76) Kaján, L.; Hopf, T. A.; Kalaš, M.; Marks, D. S.; Rost, B. FreeContact: fast and free software for protein contact prediction from residue co-evolution. *BMC Bioinformatics* **2014**, *15*, 85.
- (77) Seemayer, S.; Gruber, M.; Söding, J. CCMpred--fast and precise prediction of protein residue-residue contacts from correlated mutations. *Bioinformatics* **2014**, *30*, 3128-3130.
- (78) Ekeberg, M.; Lövkvist, C.; Lan, Y.; Weigt, M.; Aurell, E. Improved contact prediction in proteins: using pseudolikelihoods to infer Potts models. *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.* **2013**, *87*, 012707.
- (79) Baldassi, C.; Zamparo, M.; Feinauer, C.; Procaccini, A.; Zecchina, R.; Weigt, M.; Pagnani, A. Fast and accurate multivariate Gaussian modeling of protein families: predicting residue contacts and protein-interaction partners. *PLoS One* **2014**, *9*, e92721.

(80) Gueudre, T.; Baldassi, C.; Zamparo, M.; Weigt, M.; Pagnani, A., Simultaneous identification of specifically interacting paralogs and interprotein contacts by direct coupling analysis. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, 113, 12186-12191.

(81) Kolinski, A. Protein modeling and structure prediction with a reduced representation. *Acta Biochim. Pol.* **2004**, 51, 349-371.

(82) Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A.E.; Kolinski, A. Coarse-grained protein models and their applications. *Chem. Rev.* **2016**, 116, 7898-7936.

(83) Kmiecik, S.; Kouza, M.; Badaczewska-Dawid, A.E.; Kloczkowski, A.; Kolinski, A. Modeling of protein structural flexibility and large-scale dynamics: Coarse-grained simulations and elastic network models. *Int. J. Mol. Sci.* **2018**, 19, e3496.

(84) Ciemny, M.P.; Badaczewska-Dawid, A.E.; Pikuzinska, M.; Kolinski, A.; Kmiecik, S. Modeling of disordered protein structures using monte carlo simulations and knowledge-based statistical force fields. *Int. J. Mol. Sci.* **2019**, 20, e606.

(85) Kurcinski, M.; Oleniecki, T.; Ciemny, M.P.; Kuriata, A.; Kolinski, A.; Kmiecik, S. CABS-flex standalone: A simulation environment for fast modeling of protein flexibility. *Bioinformatics* **2019**, 35, 694-695.

(86) Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, 28, 235-242.

(87) Rose, P. W.; Prlic, A.; Altunkaya, A.; Bi, C.; Bradley, A. R.; Christie, C. H.; Costanzo, L. D.; Duarte, J. M.; Dutta, S.; Feng, Z.; Green, R. K.; Goodsell, D. S.; Hudson, B.; Kalro, T.; Lowe, R.; Peisach, E.; Randle, C.; Rose, A. S.; Shao, C.; Tao, Y. P.; Valasatava, Y.; Voigt, M.; Westbrook, J. D.; Woo, J.; Yang, H.; Young, J. Y.; Zardecki, C.; Berman, H. M.; Burley, S. K. The RCSB

protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Res.* **2017**, 45, D271-D281.

(88) Hooft, R. W.; Sander, C.; Vriend, G. Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures. *Proteins* **1996**, 26, 363-376.

(89) Hekkelman, M. L.; Te Beek, T. A.; Pettifer, S. R.; Thorne, D.; Attwood, T. K.; Vriend, G. WIWS: A protein structure bioinformatics web service collection. *Nucleic Acids Res.* **2010**, 38, W719-W723.

(90) Fiser, A.; Sali, A. ModLoop: Automated modeling of loops in protein structures. *Bioinformatics* **2003**, 19, 2500-25001.

(91) Fernandez-Fuentes, N.; Zhai, J.; Fiser, A. ArchPRED: A template based loop structure prediction server. *Nucleic Acids Res.* **2006**, 34, W173-W176.

(92) Rotkiewicz, P.; Skolnick, J. Fast procedure for reconstruction of full-atom protein models from reduced representations. *J. Comput. Chem.* **2008**, 29, 1460-1465.

(93) Lombardi, L. E.; Marti, M. A.; Capece, L. CG2AA: backmapping protein coarse-grained structures. *Bioinformatics* **2016**, 32, 1235-1237.

(94) Bhattacharya, D.; Nowotny, J.; Cao, R.; Cheng, J. 3Drefine: an interactive web server for efficient protein structure refinement. *Nucleic Acids Res.* **2016**, 44, W406-W409.

(95) Dehouck, Y.; Kwasigroch, J. M.; Rooman, M.; Gilis, D. BeAtMuSiC: Prediction of changes in protein-protein binding affinity on mutations. *Nucleic Acids Res.* **2013**, 41, W333-W339.

(96) Dehouck, Y.; Gilis, D.; Rooman, M. A new generation of statistical potentials for proteins. *Biophys. J.* **2006**, 90, 4010-4017.

- (97) Dehouck, Y.; Grosfils, A.; Folch, B.; Gilis, D.; Bogaerts, P.; Rooman, M. Fast and accurate predictions of protein stability changes upon mutations using statistical potentials and neural networks: PoPMuSiC-2.0. *Bioinformatics* **2009**, *25*, 2537-2543.
- (98) Tasdighian, S.; Di Paola, L.; De Ruvo, M.; Paci, P.; Santoni, D.; Palumbo, P.; Mei, G.; Di Venere, A.; Giuliani, A. Modules identification in protein Structures: The topological and geometrical solutions. *J. Chem. Inf. Model.* **2014**, *54*, 159-168.
- (99) De Ruvo, M.; Giuliani, A.; Paci, P.; Santoni, D.; Di Paola, L. Shedding light on protein-ligand binding by graph theory: The topological nature of allostery. *Biophys. Chem.* **2012**, 165-166, 21–29.
- (100) Di Paola, L.; Paci, P.; Santoni, D.; De Ruvo, M.; Giuliani, A. Proteins as sponges: A statistical journey along protein structure organization principles. *J. Chem. Inf. Model.* **2012**, *52*, 474-482.
- (101) Cumbo, F.; Paci, P.; Santoni, D.; Di Paola, L.; Giuliani, A. GIANT: A Cytoscape plugin for modular networks. *PLoS One* **2014**, *9*, e105001.
- (102) Brinda, K.V.; Vishveshwara, S. A network representation of protein structures: Implications for protein stability. *Biophys. J.* **2005**, *89*, 4159-4170.
- (103) Vijayabaskar, M.S.; Vishveshwara, S. Interaction energy based protein structure networks. *Biophys. J.* **2010**, *99*, 3704-3715.
- (104) Sethi, A.; Eargle, J.; Black, A.A.; Luthey-Schulten, Z. Dynamical networks in tRNA:protein complexes. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6620-6625.
- (105) Stetz, G.; Verkhivker, G. M. Computational analysis of residue interaction networks and coevolutionary relationships in the Hsp70 chaperones: A community-hopping model of allosteric regulation and communication. *Plos Comput. Biol.* **2017**, *13*, e1005299.



- (106) Czemerer, J.; Buse, K.; Verkhivker, G. M. Atomistic simulations and network-based modeling of the Hsp90-Cdc37 chaperone binding with Cdk4 client protein: A mechanism of chaperoning kinase clients by exploiting weak spots of intrinsically dynamic kinase domains. *PLoS One* **2017**, 12, e0190267.
- (107) Stetz, G.; Verkhivker, G. M. Dancing through life: Molecular dynamics simulations and network-centric modeling of allosteric mechanisms in Hsp70 and Hsp110 chaperone proteins. *PLoS One* **2015**, 10, e0143752.61.
- (108) Floyd, R.W. Algorithm 97: Shortest path. *Commun. A.C.M.* **1962**, 5, 345.
- (109) Hagberg, A.A.; Schult, D.A.; Swart, P.J. Exploring network structure, dynamics, and function using NetworkX, in : G. Varoquaux, T. Vaught, J. Millman (Eds.), Proceedings of the 7th Python in Science Conference (SciPy2008), Pasadena, **2008**, pp. 11-15.
- (110) Atilgan, C.; Atilgan, A. R., Perturbation-response scanning reveals ligand entry-exit mechanisms of ferric binding protein. *PLoS Comput. Biol.* **2009**, 5, e1000544.
- (111) Atilgan, C.; Gerek, Z. N.; Ozkan, S. B.; Atilgan, A. R., Manipulation of conformational change in proteins by single-residue perturbations. *Biophys. J.* **2010**, 99, 933-943.
- (112) General, I. J.; Liu, Y.; Blackburn, M. E.; Mao, W.; Gierasch, L. M.; Bahar, I., ATPase subdomain IA is a mediator of interdomain allostery in Hsp70 molecular chaperones. *PLoS Comput. Biol.* **2014**, 10, e1003624.
- (113) Dutta, A.; Krieger, J.; Lee, J. Y.; Garcia-Nafria, J.; Greger, I. H.; Bahar, I., Cooperative Dynamics of Intact AMPA and NMDA Glutamate Receptors: Similarities and Subfamily-Specific Differences. *Structure* **2015**, 23, 1692-1704.
- (114) Penkler, D.; Sensoy, O.; Atilgan, C.; Tastan Bishop, O., Perturbation-Response Scanning Reveals Key Residues for Allosteric Control in Hsp70. *J. Chem. Inf. Model.* **2017**, 57, 1359-1374.

- (115) Penkler, D. L.; Atilgan, C.; Tastan Bishop, O., Allosteric Modulation of Human Hsp90alpha Conformational Dynamics. *J. Chem. Inf. Model.* **2018**, 58, 383-404.
- (116) Stetz, G.; Tse, A.; Verkhivker, G. M., Dissecting Structure-Encoded Determinants of Allosteric Cross-Talk between Post-Translational Modification Sites in the Hsp90 Chaperones. *Sci. Rep.* **2018**, 8, 19.
- (117) Jalalypour, F.; Sensoy, O.; Atilgan, C., Perturb-Scan-Pull: A Novel Method Facilitating Conformational Transitions in Proteins. *J. Chem. Theory Comput.* **2020**, 16, 3825-3841.
- (118) Jaimes, J. A.; André, N. M.; Chappie, J. S.; Millet, J. K.; Whittaker, G. R. Phylogenetic Analysis and Structural Modeling of SARS-CoV-2 Spike Protein Reveals an Evolutionary Distinct and Proteolytically Sensitive Activation Loop. *J. Mol. Biol.* **2020**, 432, 3309-3325.
- (119) Shah, M.; Ahmad, B.; Choi, S.; Woo, H.G. Sequence variation of SARS-CoV-2 spike protein may facilitate stronger interaction with ACE2 promoting high infectivity. *Research Square* **2020**. DOI: 10.21203/rs.3.rs-16932/v1.
- (120) Armijos-Jaramillo, V.; Yeager, J.; Muslin, C.; Perez-Castillo, Y. SARS-CoV-2, an evolutionary perspective of interaction with human ACE2 reveals undiscovered amino acids necessary for complex stability. *Evol. Appl.* **2020**, 00: 1-11.
- (121) Singh, J.; Dhindsa, R. S.; Misra, V.; Singh, B. SARS-CoV2 infectivity is potentially modulated by host redox status. *Comput. Struct. Biotechnol. J.* **2020**, 18, 3705-3711.
- (122) Petit, C. M.; Chouljenko, V. N.; Iyer, A.; Colgrove, R.; Farzan, M.; Knipe, D. M.; Kousoulas, K. G. Palmitoylation of the cysteine-rich endodomain of the SARS-coronavirus spike glycoprotein is important for spike-mediated cell fusion. *Virology* **2007**, 360, 264-274.

- (123) Madu, I. G.; Belouzard, S.; Whittaker, G. R., SARS-coronavirus spike S2 domain flanked by cysteine residues C822 and C833 is important for activation of membrane fusion. *Virology* **2009**, 393, 265-71.
- (124) Madu, I. G.; Roth, S. L.; Belouzard, S.; Whittaker, G. R., Characterization of a highly conserved domain within the severe acute respiratory syndrome coronavirus spike protein S2 domain with characteristics of a viral fusion peptide. *J. Virol.* **2009**, 83, 7411-7421.
- (125) Walls, A. C.; Tortorici, M. A.; Snijder, J.; Xiong, X.; Bosch, B. J.; Rey, F. A.; Veesler, D. Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, 114, 11157-11162.
- (126) Fan, X.; Cao, D.; Kong, L.; Zhang, X., Cryo-EM analysis of the post-fusion structure of the SARS-CoV spike glycoprotein. *Nat. Commun.* **2020**, 11, 3618.
- (127) Mercurio, I.; Tragni, V.; Busto, F.; De Grassi, A.; Pierri, C. L. Protein structure analysis of the interactions between SARS-CoV-2 spike protein and the human ACE2 receptor: from conformational changes to novel neutralizing antibodies. *Cell. Mol. Life Sci.* **2020**, 1-22.
- (128) Hati, S.; Bhattacharyya, S. Impact of thiol-disulfide balance on the binding of Covid-19 spike protein with angiotensin-converting enzyme 2 receptor. *ACS Omega* **2020**, 5, 16292-16298.
- (129) Lee, B. C.; Park, K.; Kim, D. Analysis of the residue-residue coevolution network and the functionally important residues in proteins. *Proteins* **2008**, 72, 863-872.
- (130) Chakrabarti, S.; Panchenko, A. R. Coevolution in defining the functional specificity. *Proteins* **2009**, 75, 231-240.
- (131) Chakrabarti, S.; Panchenko, A. R. Structural and functional roles of coevolved sites in proteins. *PLoS One* **2010**, 5, e8591.

- (132) Teppa, E.; Wilkins, A. D.; Nielsen, M.; Buslje, C. M. Disentangling evolutionary signals: conservation, specificity determining positions and coevolution. Implication for catalytic residue prediction. *BMC Bioinformatics* **2012**, 13, 235.
- (133) Jeon, J.; Nam, H. J.; Choi, Y. S.; Yang, J. S.; Hwang, J.; Kim, S. Molecular evolution of protein conformational changes revealed by a network of evolutionarily coupled residues. *Mol. Biol. Evol.* **2011**, 28, 2675-85.
- (134) Hoffmann, M.; Kleine-Weber, H.; Pöhlmann, S. A multibasic cleavage site in the spike Pprotein of SARS-CoV-2 is essential for infection of human lung cells. *Mol. Cell* **2020**, 78 779-784.e5.
- (135) Örd, M.; Faustova, I.; Loog, M., The sequence at Spike S1/S2 site enables cleavage by furin and phospho-regulation in SARS-CoV2 but not in SARS-CoV1 or MERS-Co. *Sci. Rep.* **2020**, 10, 16944.
- (136) Belouzard, S.; Chu, V. C.; Whittaker, G. R., Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, 106 , 5871-5876.
- (137) Jeong, C. S.; Kim, D. Coevolved residues and the functional association for intrinsically disordered proteins. *Pac. Symp. Biocomput.* **2012**, 140-51.
- (138) Teppa, E.; Zea, D. J.; Marino-Buslje, C. Protein-protein interactions leave evolutionary footprints: High molecular coevolution at the core of interfaces. *Protein Sci.* **2017**, 26, 2438-2444.
- (139) Anishchenko, I.; Ovchinnikov, S.; Kamisetty, H.; Baker, D. Origins of coevolution between residues distant in protein 3D structures. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, 114, 9122-9127.
- (140) Pietal, M. J.; Tuszynska, I.; Bujnicki, J. M., PROTMAP2D: visualization, comparison and analysis of 2D maps of protein structure. *Bioinformatics* **2007**, 23, 1429-1430.

(141) Chen, J. E.; Huang, C. C.; Ferrin, T. E. RRDistMaps: a UCSF Chimera tool for viewing and comparing protein distance maps. *Bioinformatics* **2015**, 31, 1484-1486.

(142) Weinkam, P.; Pons, J.; Sali, A. Structure-based model of allostery predicts coupling between distant sites. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, 109, 4875-4880.

(143) Li, Q.; Wu, J.; Nie, J.; Zhang, L.; Hao, H.; Liu, S.; Zhao, C.; Zhang, Q.; Liu, H.; Nie, L.; Qin, H.; Wang, M.; Lu, Q.; Li, X.; Sun, Q.; Liu, J.; Huang, W.; Wang, Y. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. *Cell* **2020**, 182, 1284-1294.e9.

(144) Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E. E.; Bhattacharya, T.; Foley, B.; Hastie, K. M.; Parker, M. D.; Partridge, D. G.; Evans, C. M.; Freeman, T. M.; de Silva, T. I.; McDanal, C.; Perez, L. G.; Tang, H.; Moon-Walker, A.; Whelan, S. P.; LaBranche, C. C.; Saphire, E. O.; Montefiori, D. C. Tracking changes in SARS-CoV-2 Spike: Evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **2020**, 182, 812-827.e19.

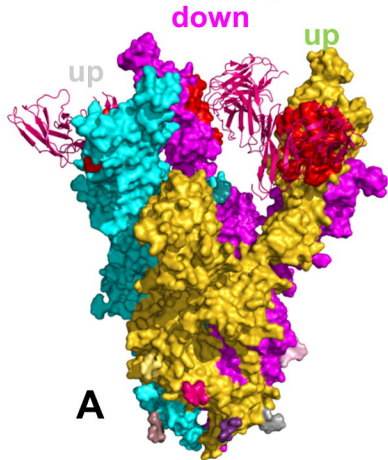
(145) Laha, S.; Chakraborty, J.; Das, S.; Manna, S. K.; Biswas, S.; Chatterjee, R. Characterizations of SARS-CoV-2 mutational profile, spike protein stability and viral transmission. *Infect. Genet. Evol.* **2020**, 85, 104445.

(146) Mansbach, R. A.; Chakraborty, S.; Nguyen, K.; Montefiori, D.; Korber, B.; Gnanakaran, S. The SARS-CoV-2 spike variant D614G favors an open conformational state. *bioRxiv* **2020**, doi: 10.1101/2020.07.26.219741. 2007.2026.219741.

(147) Gobeil, S.; Janowska, K.; McDowell, S.; Mansouri, K.; Parks, R.; Manne, K.; Stalls, V.; Kopp, M.; Henderson, R.; Edwards, R. J.; Haynes, B. F.; Acharya, P. D614G mutation

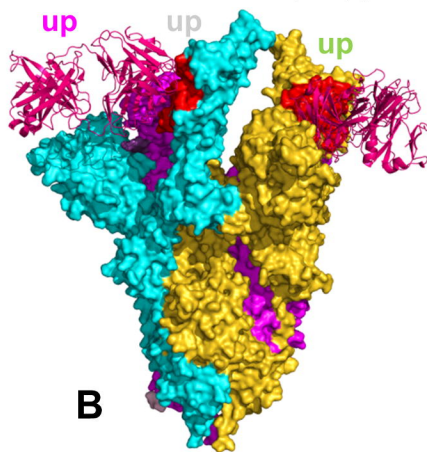
alters SARS-CoV-2 spike conformational dynamics and protease cleavage susceptibility at the S1/S2 junction. *bioRxiv* **2020**, doi: 10.1101/2020.10.11.335299.

7CAI, 2 H014 (2-up, 1-down)



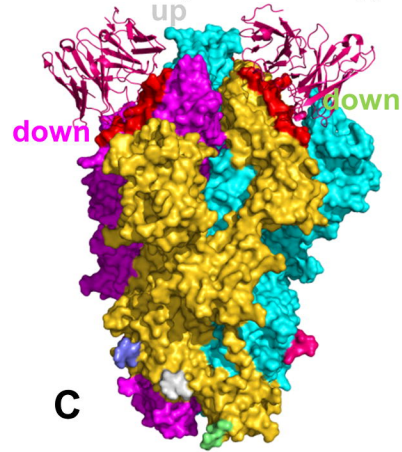
A

7CAK, 3 H014 (3-up)



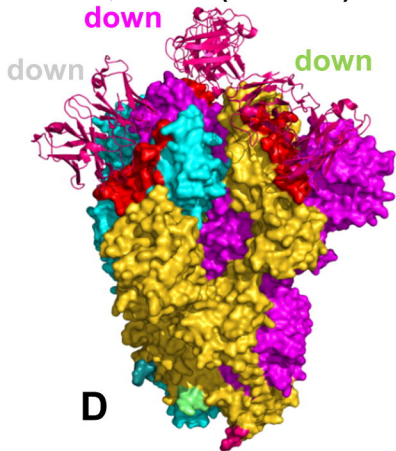
B

6WPT 2 S309 (2-down, 1-up)



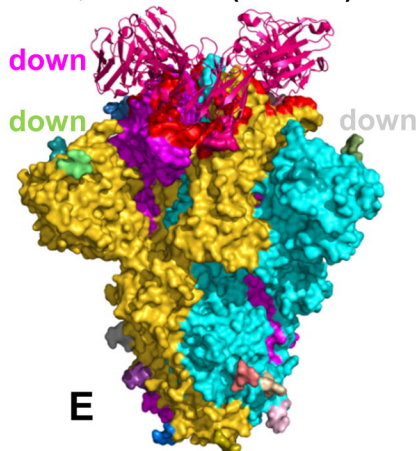
C

6WPS, 3S309 ( 3-down)



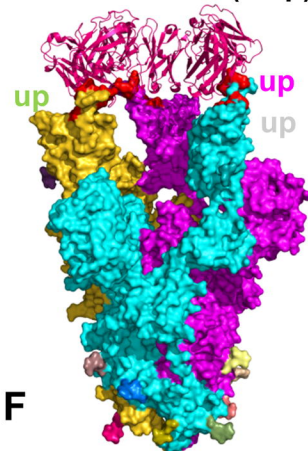
D

7K43, 3 S2M11 (3-down)



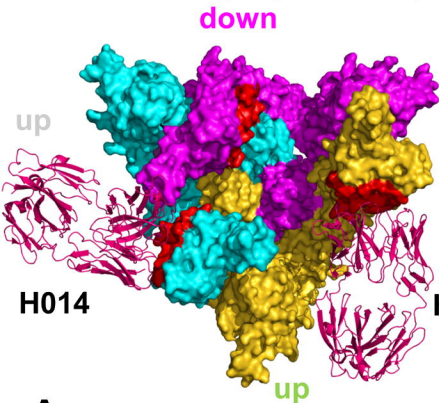
E

7K4N 3S2E12 (3-up)

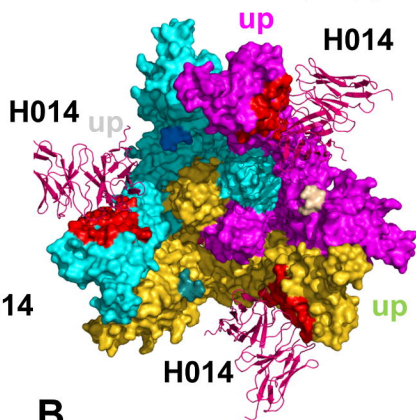


F

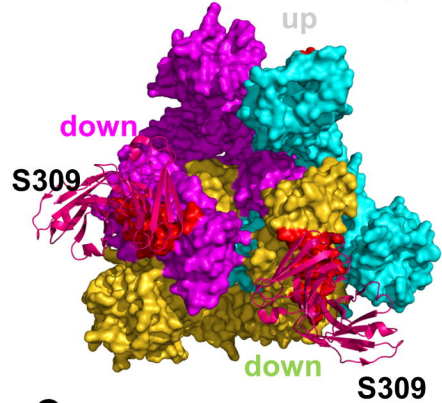
7CAI, 2 H014 (2-up, 1-down)



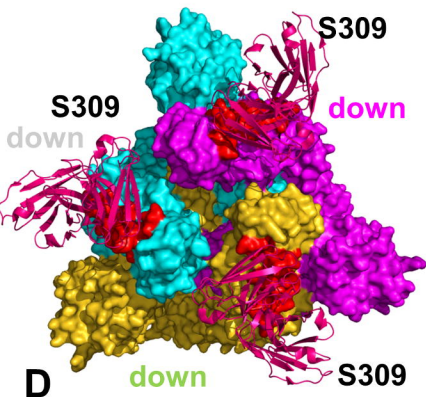
7CAK, 3 H014 (3-up)



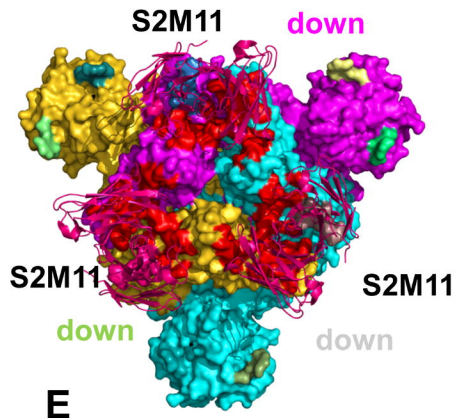
6WPT 2 S309 (2-down, 1-up)



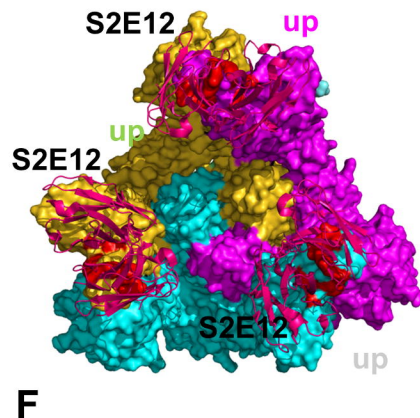
6WPS, 3S309 (3-down)



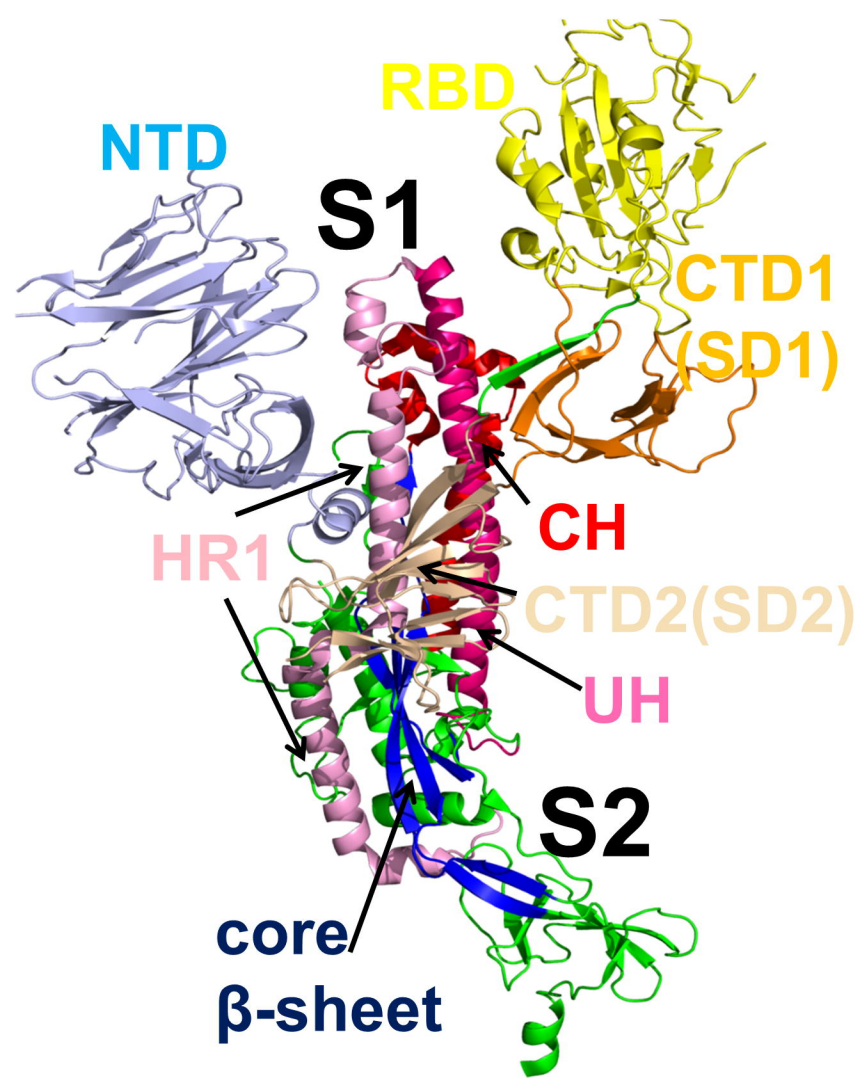
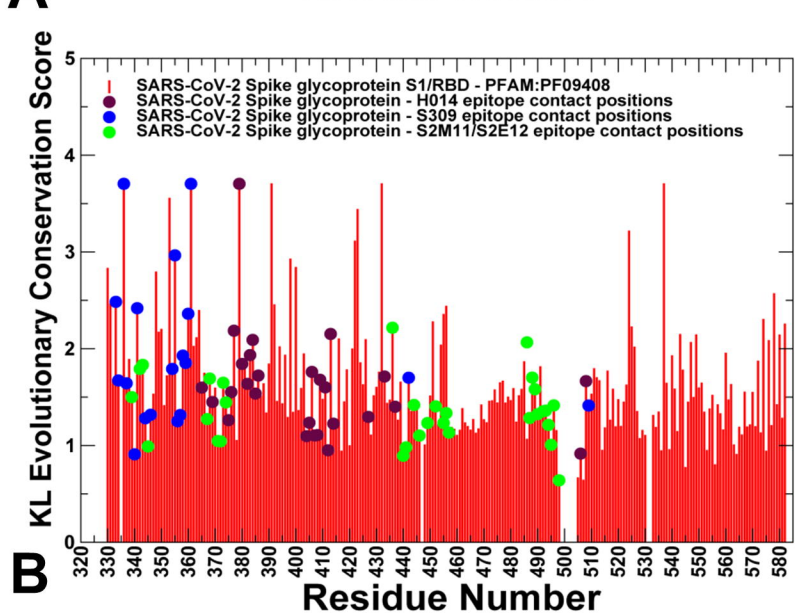
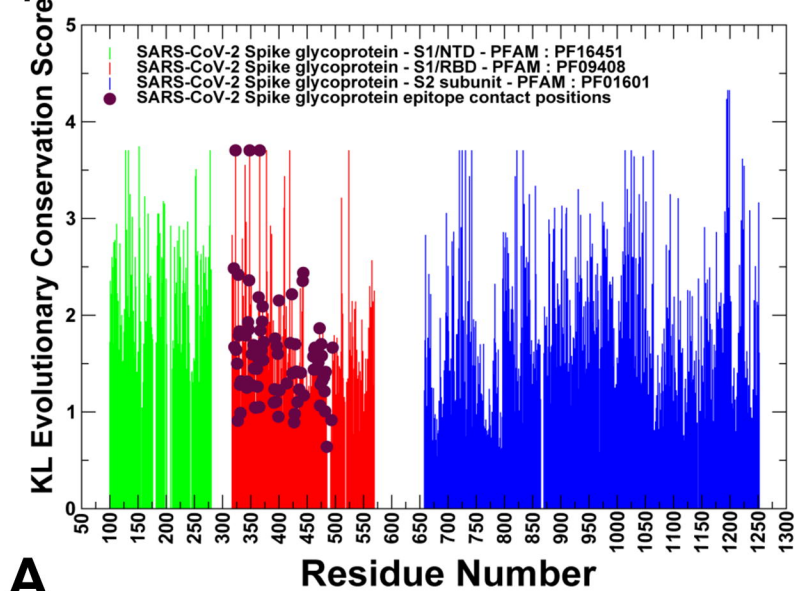
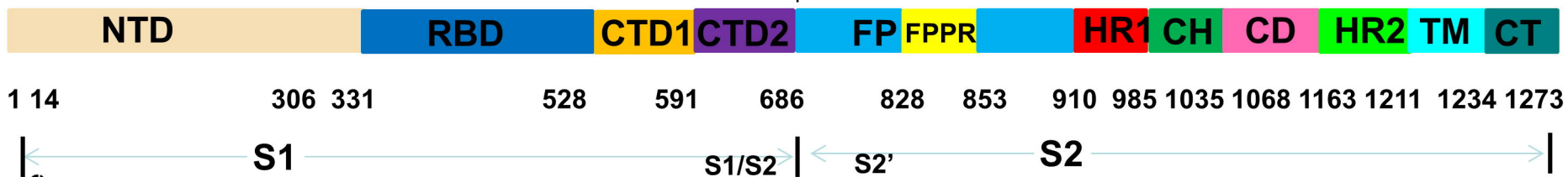
7K43, 3 S2M11 (3-down)



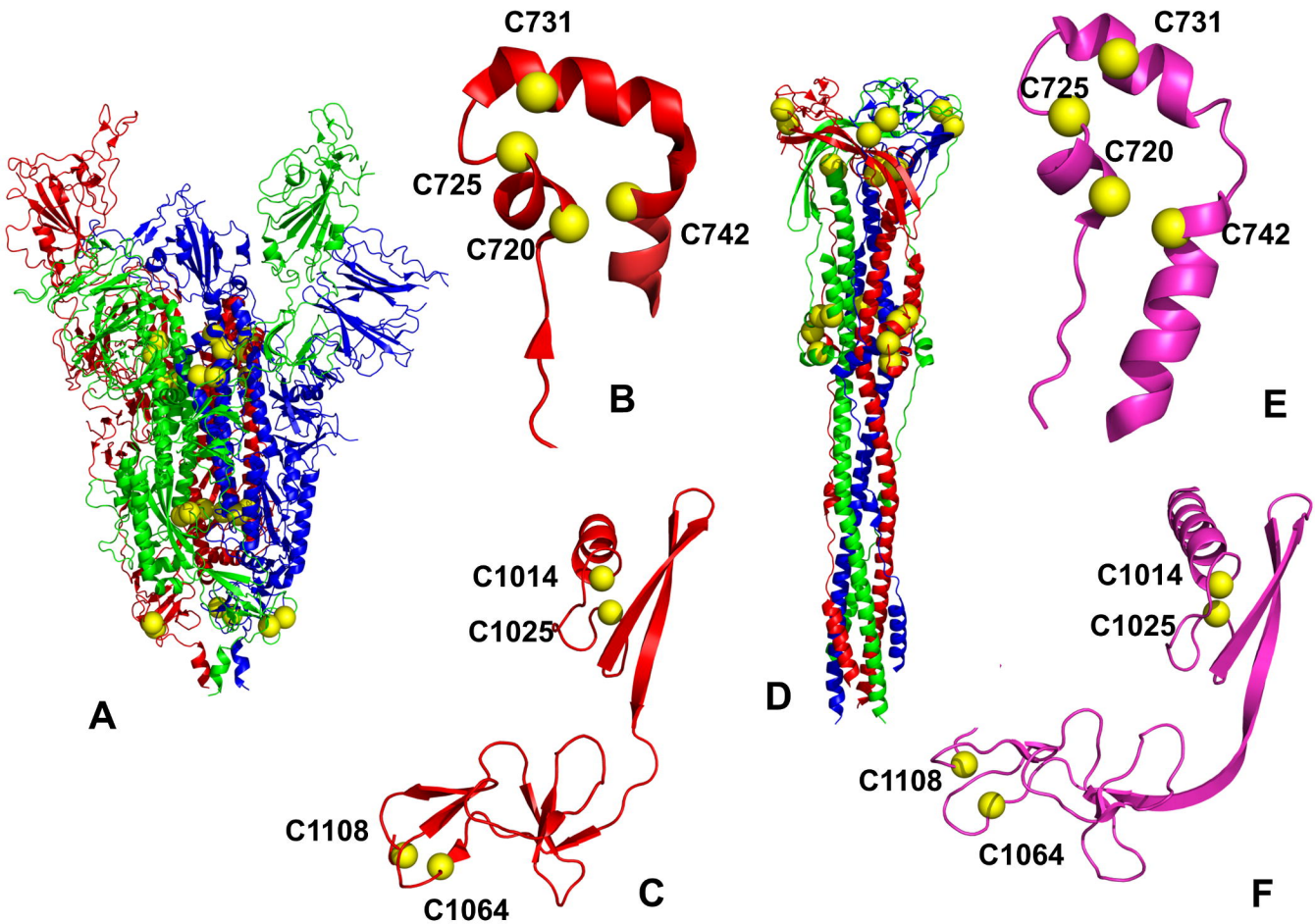
7K4N 3S2E12 (3-up)

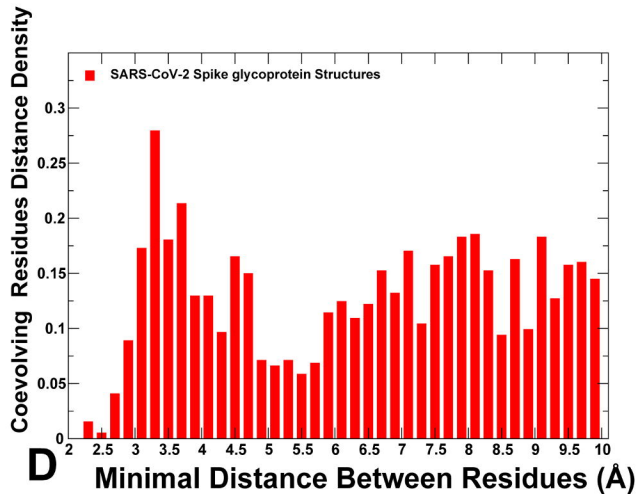
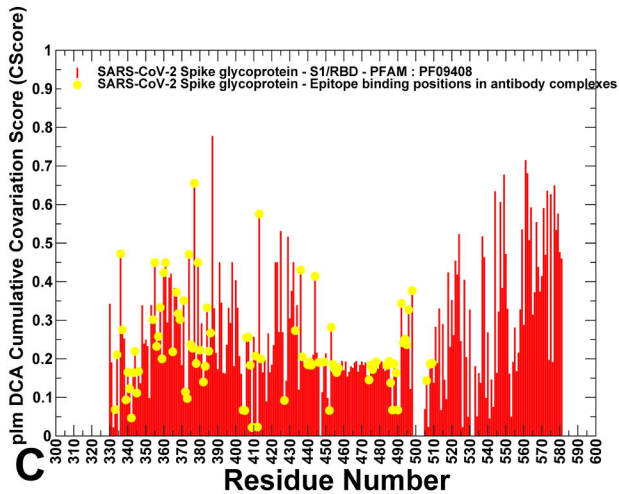
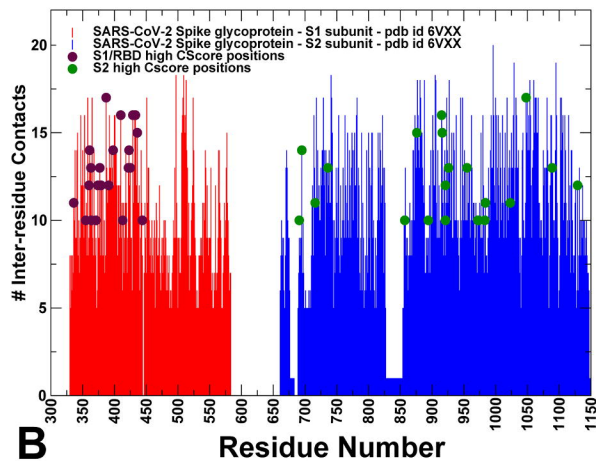
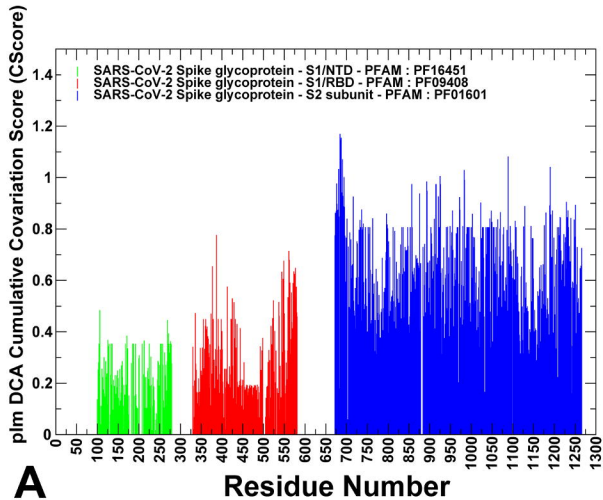


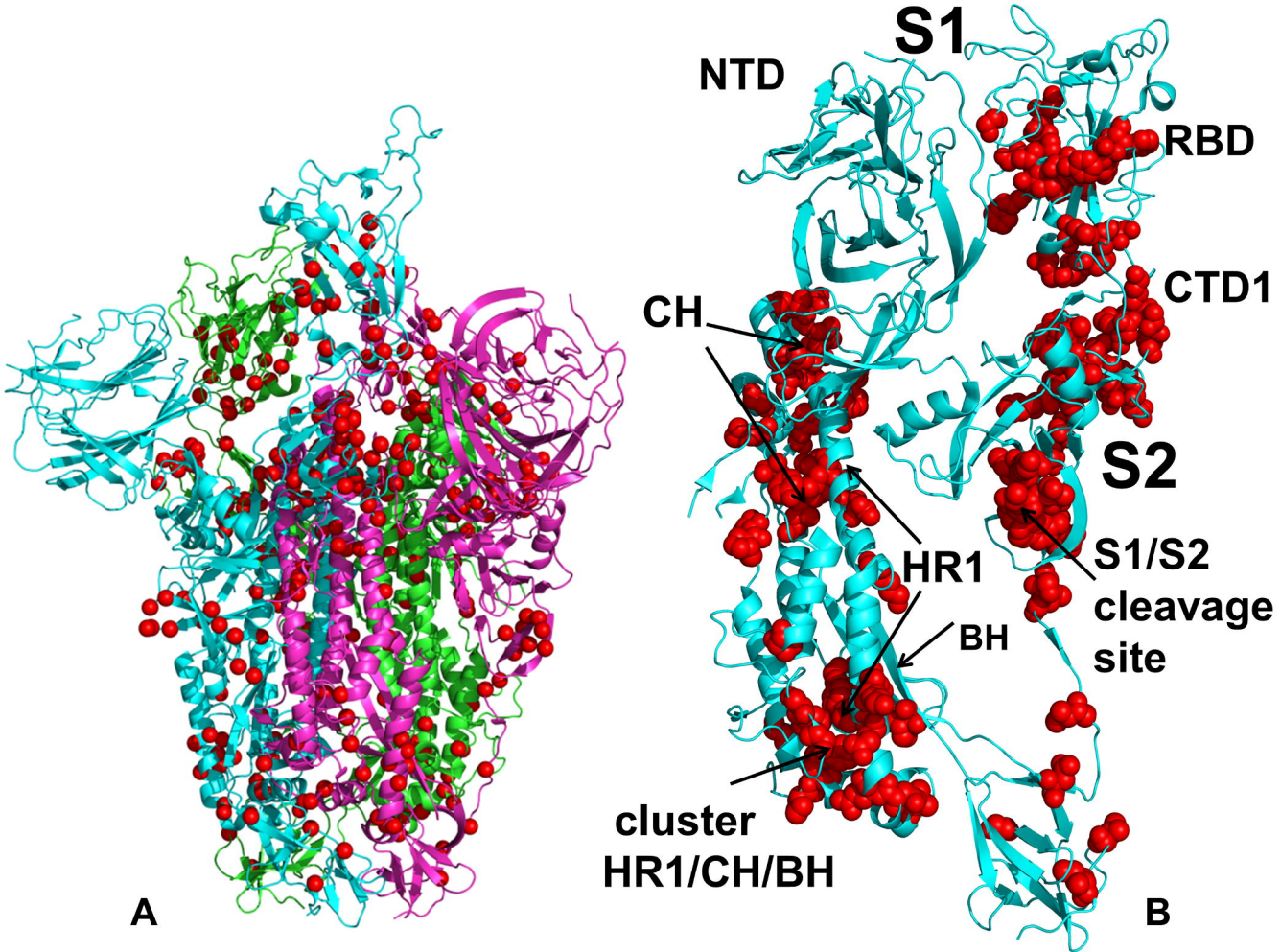


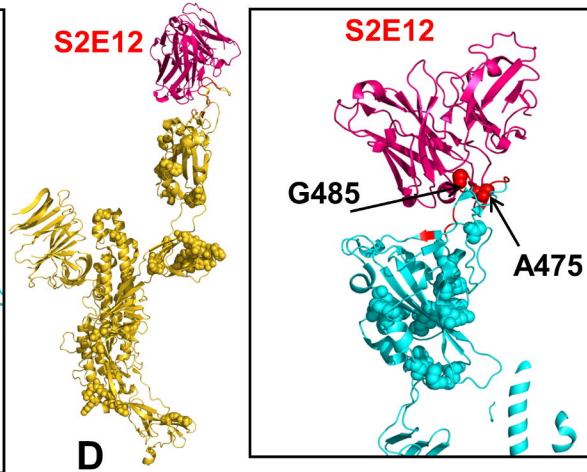
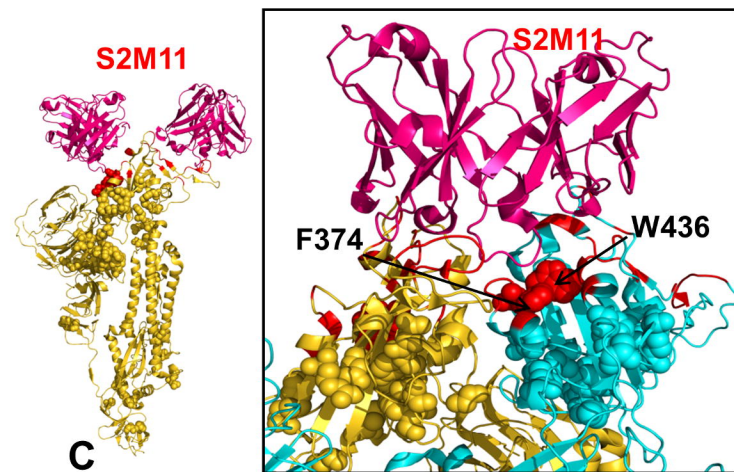
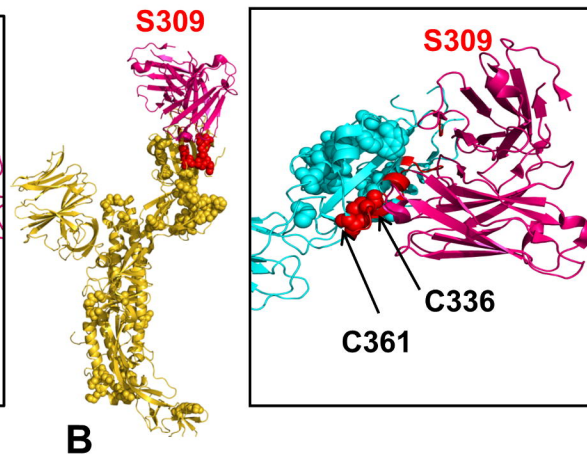
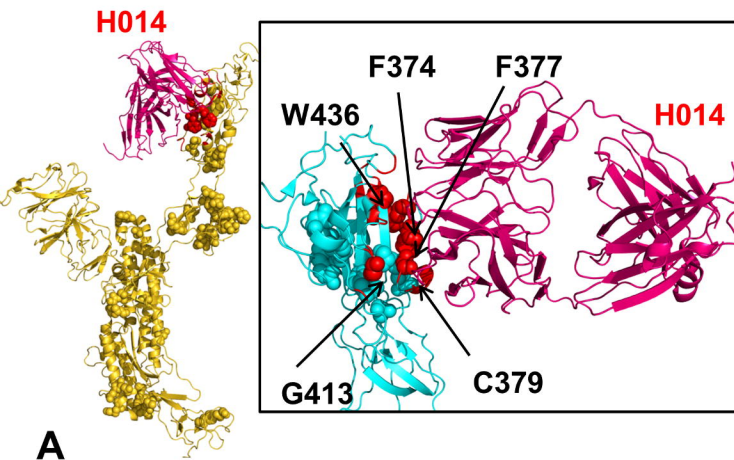


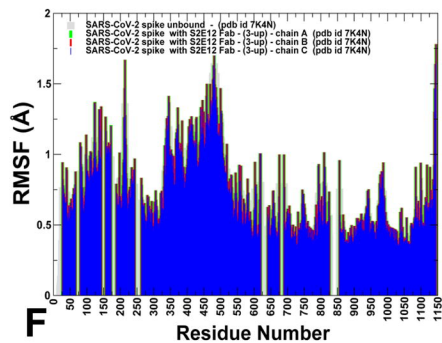
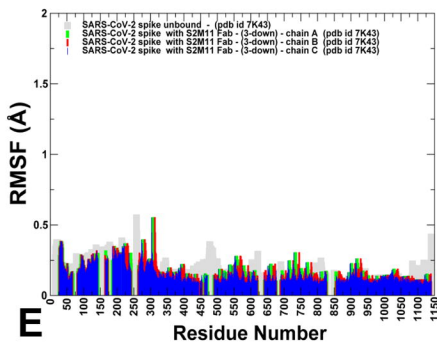
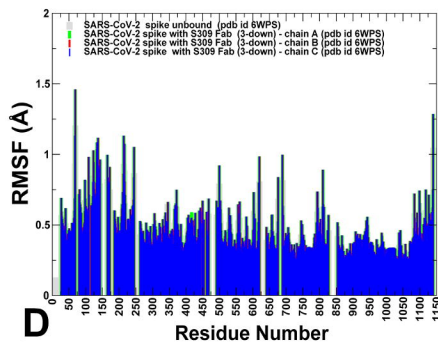
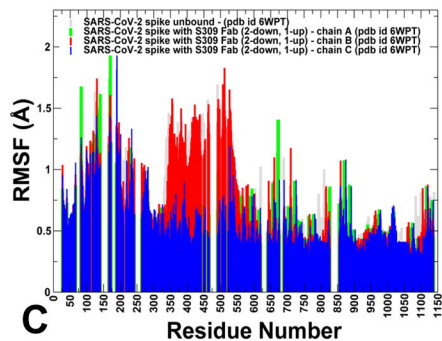
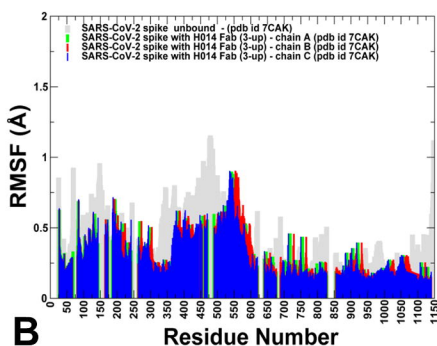
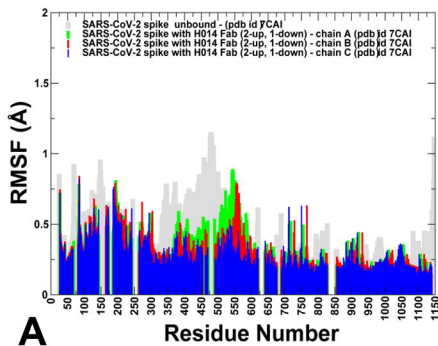
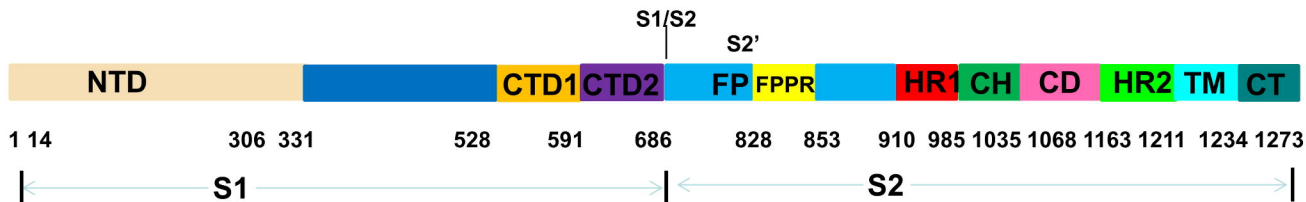
**C**

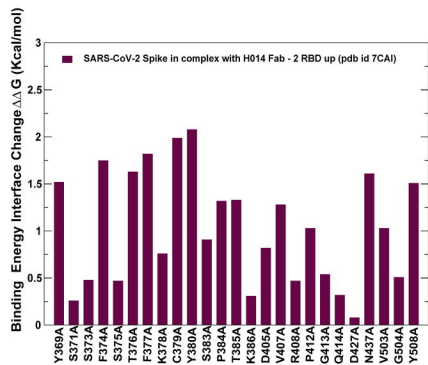




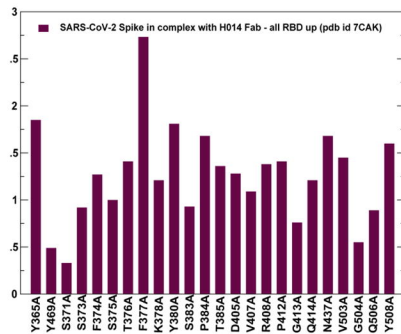




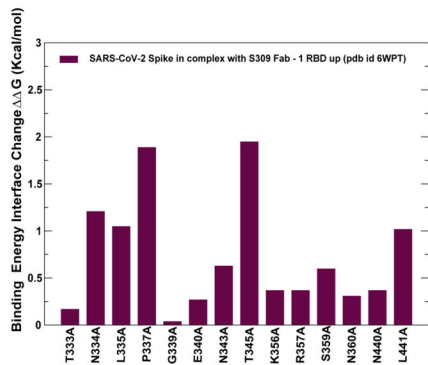




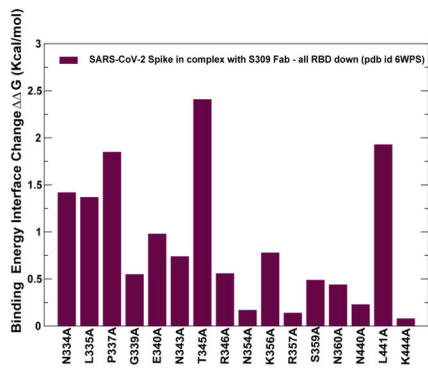
**A**



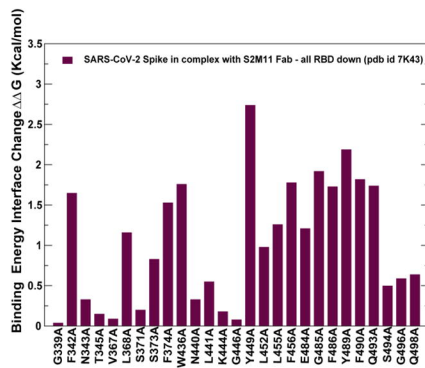
**B**



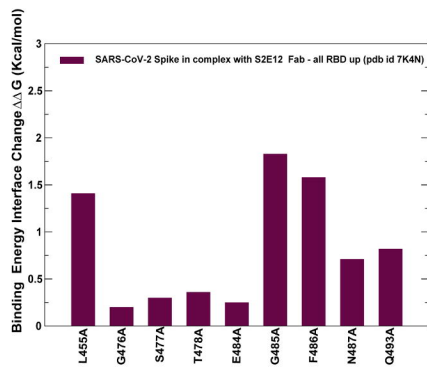
**C**



**D**

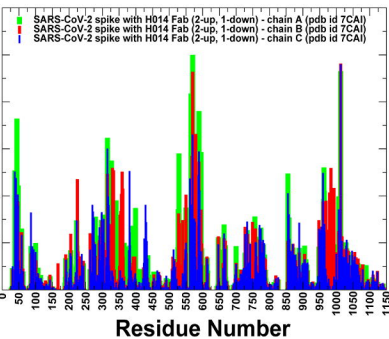


**E**

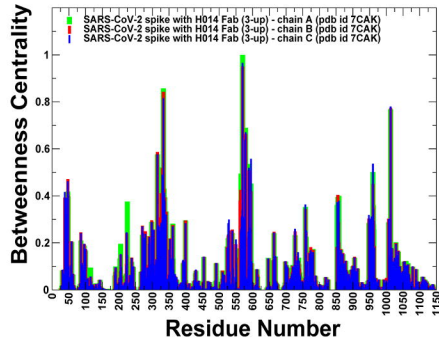


**F**

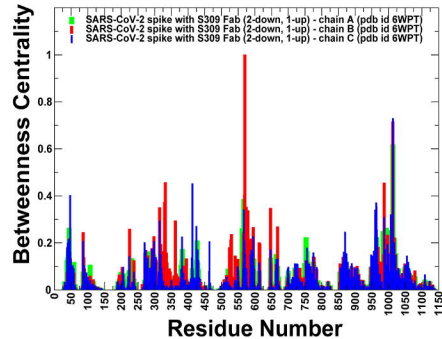
Betweenness Centrality



**A**

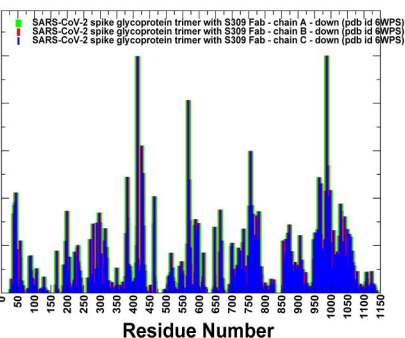


**B**

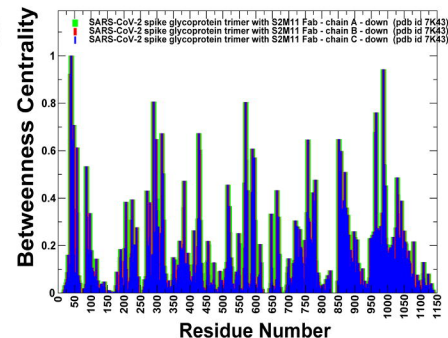


**C**

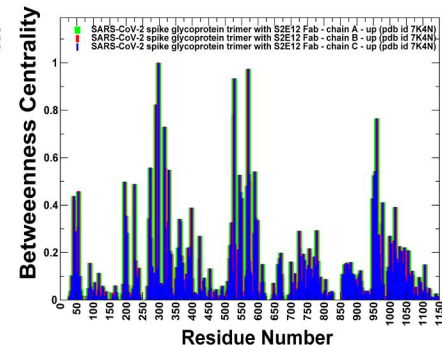
Betweenness Centrality



**D**



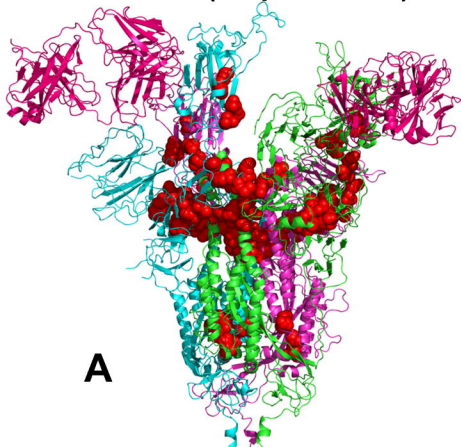
**E**



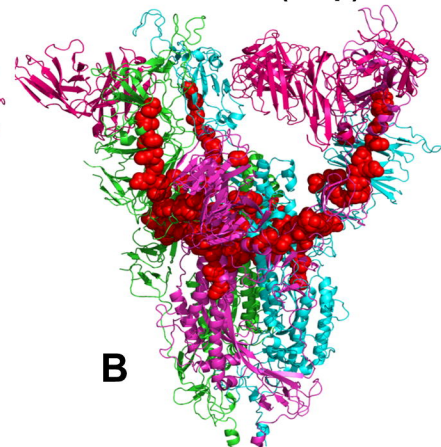
**F**



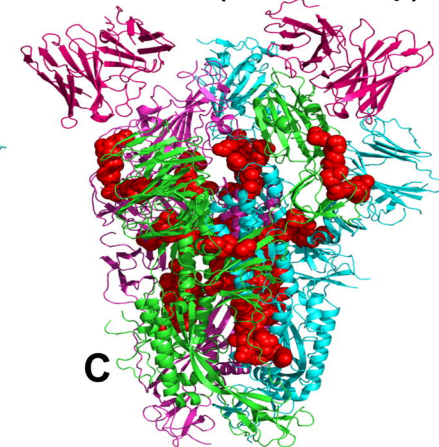
7CAI, 2 H014 (2-up, 1-down)



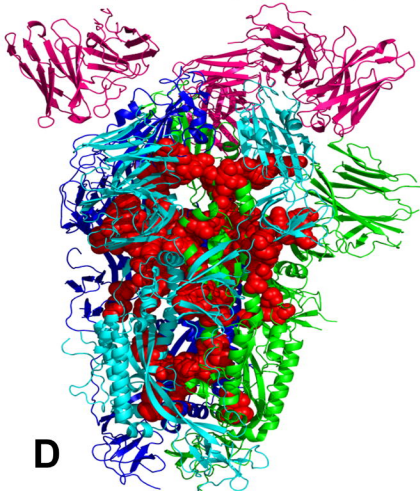
7CAK, 3 H014 (3-up)



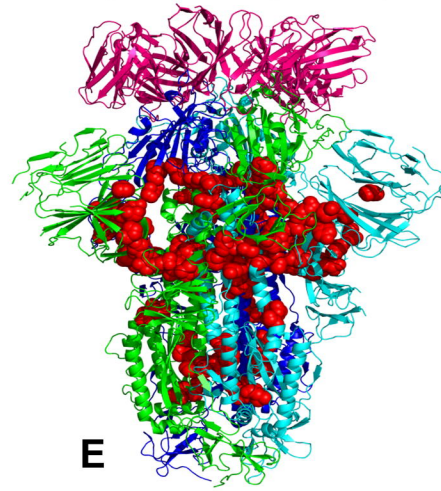
6WPT 2 S309 (2-down, 1-up)



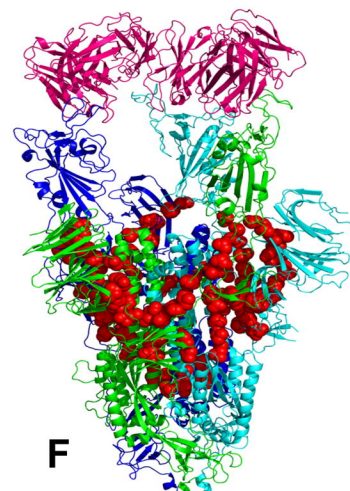
6WPS, 3 S309 (3-down)

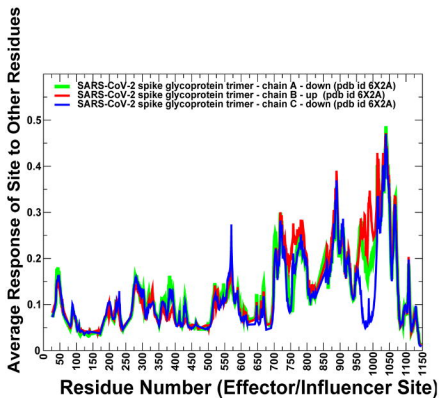


7K43, 3 S2M11 (3-down)

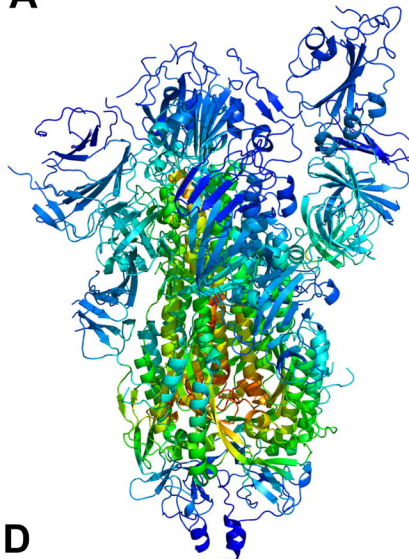


7K4N 3 S2E12 (3-up)

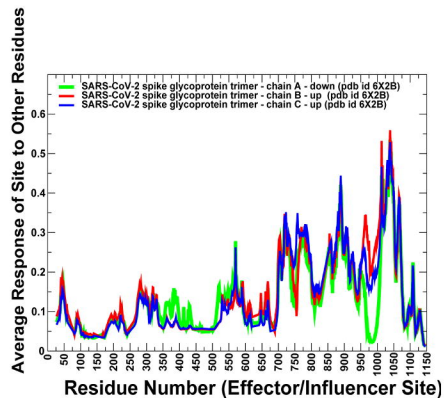




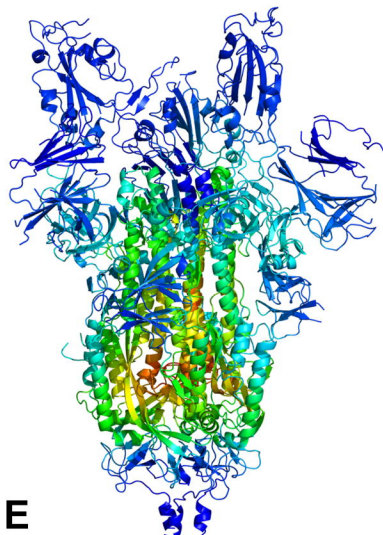
**A**



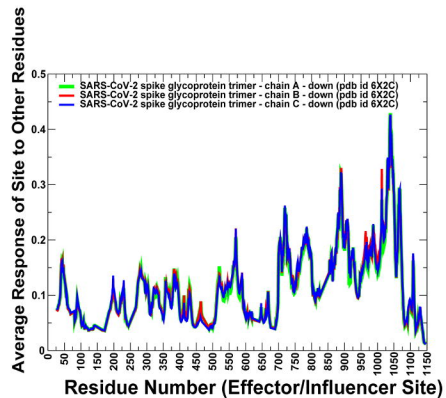
**D**



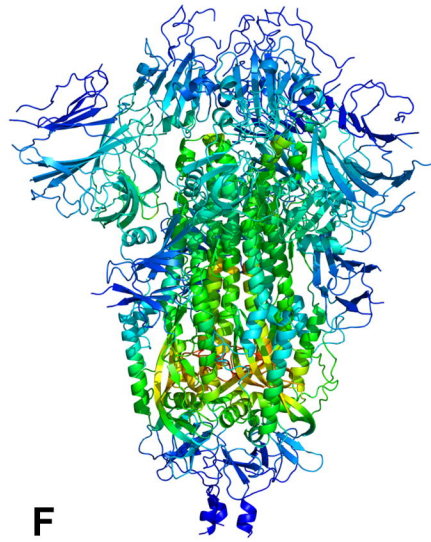
**B**



**E**



**C**



**F**

