THE JOURNAL OF PHYSICAL CHEMISTRY B
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCB

Article

Integrated Biophysical Modeling of the SARS-CoV-2 Spike Protein Binding and Allosteric Interactions with Antibodies

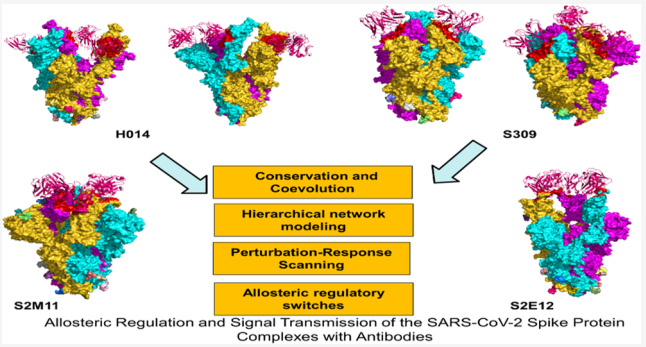
Gennady M. Verkhivker* and Luisa Di Paola

Cite This: J. Phys. Chem. B 2021, 125, 4596–4619

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Structural and biochemical studies of the severe acute respiratory syndrome (SARS)-CoV-2 spike glycoproteins and complexes with highly potent antibodies have revealed multiple conformation-dependent epitopes highlighting 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 spike protein complexes with a panel of antibodies targeting distinct epitopes to explore molecular mechanisms underlying binding-induced modulation of dynamics and allosteric signaling in the spike proteins. Through coevolutionary analysis of the SARS-CoV-2 spike proteins, we identified highly coevolving hotspots and functional clusters that enable a functional cross-talk between distant allosteric regions in the SARS-CoV-2 spike complexes with antibodies. Coarse-grained and all-atom molecular dynamics simulations combined with mutational sensitivity mapping and perturbation-based profiling of the SARS-CoV-2 receptor-binding domain (RBD) complexes with CR3022 and CB6 antibodies enabled a detailed validation of the proposed approach and an extensive quantitative comparison with the experimental structural and deep mutagenesis scanning data. By combining in silico mutational scanning, perturbation-based modeling, and network analysis of the SARS-CoV-2 spike trimer complexes with H014, S309, S2M11, and S2E12 antibodies, we demonstrated that antibodies can incur specific and functionally relevant changes by modulating allosteric propensities and collective dynamics of the SARS-CoV-2 spike proteins. The results provide a novel insight into regulatory mechanisms of SARS-CoV-2 S proteins showing that antibody-escaping mutations can preferentially target structurally adaptable energy hotspots and allosteric effector centers that control functional movements and allosteric communication in the complexes.

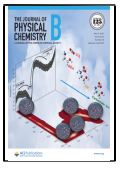


Allosteric Regulation and Signal Transmission of the SARS-CoV-2 Spike Protein Complexes with Antibodies

INTRODUCTION

Severe acute respiratory syndrome (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.[1−3] 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 revealed that spontaneous conformational transformations of the SARS-CoV-2 S protein between a spectrum of closed and receptor-accessible open forms are central to the mechanism of binding with the host receptor angiotensin-converting enzyme 2 (ACE2) and virus transmission.[4−6] The crystal structures of the S-RBD in the complexes with human ACE2 revealed a 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.[7−11] The rapidly growing body of cryo-electron microscopy (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 closed ("RBD-down") and receptor-accessible open ("RBD-up") forms required for the S protein fusion to the viral membrane.[12−21] By combining targeted mutagenesis and cryo-EM structure determination, recent biophysical investigations engineered thermostable SARS-CoV-2 S spike timers in which modifications in the contact regions between the RBD and S2 domains and the introduction of the interprotomer disulfide bonds can preferentially shift the equilibrium toward the closed or open form with enhanced binding propensities for the ACE2 host

Received: January 15, 2021
Revised: April 16, 2021
Published: April 30, 2021



https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596–4619

receptor.[22−25] The cryo-EM high-resolution structure characterization of SARS-CoV-2 S trimers in situ on the virion surface[26] and biophysical analysis of the SARS-CoV-2 S trimer on virus particles[27] revealed 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 toward the viral particle membrane. 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 ACE2 recognition that can accelerate transformation to a fully open and ACE2-bound form primed for fusion activation.[28] These investigations 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. Deep mutagenesis scanning studies have detailed SARS-CoV-2 interactions with the ACE2 host receptor and key energetic hotspots of binding and stability showing that many mutations of the RBD residues can be well tolerated with respect to both RBD folding and binding.[29,30] Functional studies characterized the key amino acid residues of the RBD for binding with human ACE2, revealing two groups of amino acid residues to modulate binding, where the SARS-CoV-2-RBD mutations to their SARS-CoV counterparts N439/R426, L452/K439, T470/N457, E484/P470, Q498/Y484, and N501/T487 can result in the enhanced binding affinity for ACE2.[31]

Computational modeling and molecular dynamics (MD) simulations have been instrumental in predicting conformational and energetic mechanisms of SARS-CoV-2 functions.[32−37] One of the first reported all-atom MD simulations of the SARS-CoV-2 S protein confirmed dynamic fluctuations between open and closed spike states by constructing the free energy landscapes and minimum energy pathways, also revealing that RBD switches to the up position through an obligatory semiopen intermediate that reduces the free energy barrier between functional forms and could serve as a prerequisite state for the host cell recognition.[32] The development of a fully glycosylated full-length SARS-CoV-2 S protein in a viral membrane provided the infrastructure for more rigorous simulations of the SARS-CoV-2 S trimer structures in a glycosylated environment.[33] 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 an unprecedented level of details about open and closed structures.[34] 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 the ACE2 host receptor.[35] 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.[36] Computational studies of the SARS-CoV-2 S trimer interactions with ACE2 using the recent crystal structures[38−42] also provided important insights into the key determinants of the binding affinity and selectivity. Molecular mechanisms of the SARS-CoV-2 binding with ACE2 were analyzed in our recent study using coevolution and conformational dynamics.[41] A series of all-atom MD simulations totaling 16.5 μs of the SARS-CoV-2 S-RBDS complex with ACE2 in the absence and presence of external force examined the effects of alanine substitutions and charge-reversal mutations of the RBD residues, showing that the hydrophobic end of RBD serves as the main energetic hotspot for ACE2 binding.[42]

The rapidly growing structural studies of SARS-CoV-2 antibodies have delineated molecular mechanisms underlying binding competition with the ACE2 host receptor, showing that the SARS-CoV-2 S protein features multiple distinct antigenic sites, where efficient cross-neutralization can be achieved through synergistic targeting of conserved and variable SARS-CoV-2-RBD epitopes.[43−47] A wide spectrum of SARS-CoV-2 antibodies can be divided into several main classes, of which class 1 and class 2 antibodies target epitopes that overlap with the ACE2-binding site.[44,45] The crystal structure of a neutralizing Ab CR3022 in the complex with the SARS-CoV-2 S-RBD revealed binding to a highly conserved cryptic epitope located away from the ACE2-binding site that can be accessed only when two RBDs adopt the "up" conformation.[48] Subsequent studies confirmed that CR3022 binds the RBD of SARS-CoV-2, displaying strong neutralization by allosterically perturbing the interactions between the RBD regions and the ACE2 receptor.[49] Although the CR3022 cryptic site is not accessible when RBD is in the "down" state, the intrinsic structural plasticity of the RBD capable of adopting a variety of conformations might transiently expose the spike surface to CR3022 recognition. Structural and biochemical studies characterized binding epitopes and binding mechanisms for a number of SARS-CoV-2 antibodies targeting RBD regions including REGN10933 and REGN10987,[50] B38 and H14,[51] P2B-2F6,[52] CA1 and CB6,[53] CC12.1 and CC12.3,[54] C105,[55] and BD-23.[56]

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.[57] 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 the conserved nature of the RBD epitope and a partial overlap with the ACE2-binding region. The recently reported S309 antibody potently 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 completion between S309 and ACE2 for binding to the SARS-CoV-2 S protein.[58] Two ultrapotent S2M11 and S2E12 antibodies 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.[59] 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 toward a fully open S trimer conformation.[59] Functional mapping of mutations in the SARS-CoV-2 S-RBD that escape antibody binding using deep mutational scanning showed that the escape mutations cluster in several RBD regions and have large effects on antibody escape and a negligible negative impact on ACE2 binding and RBD folding.[60] This illuminating study demonstrated that escape sites from antibodies can be constrained with respect to their effects on the expression of properly

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

folded RBD and ACE2 binding, suggesting that escape-resistant antibody cocktails can compete for binding to the same RBD region but have different escape mutations. There is growing evidence that properly designed cocktails of antibodies can provide 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. These studies also suggested that some SARS-CoV-2 antibodies can allosterically interfere with the host receptor binding and modulating conformational changes that can obstruct other epitopes and block virus infection without directly interfering with ACE2 recognition.

In this study, we employed an integrated biophysical strategy that encompasses several synergistic computational approaches to explore mechanisms of SARS-CoV-2 S binding and allosteric interactions with a panel of antibodies including CR3022, CB6, H014, S309, S2M11, and S2E12 that target distinct epitopes in the RBD regions. By expanding our recent studies,[61,62] we propose that the SARS-CoV-2 spike protein can function as an allosteric regulatory engine that fluctuates between dynamically distinct functional states. Using coevolutionary analysis, molecular simulations, and perturbation-based hierarchical network modeling, we examined how binding can incur antibody-specific modulation of dynamics, stability, and allosteric interactions in the SARS-CoV-2 S complexes. The results reveal structural topography of coevolutionary couplings and network connectivity that may determine mechanisms of allosteric signaling in the SARS-CoV-2 S proteins. Coarse-grained (CG) and atomistic MD simulations combined with mutational sensitivity mapping and perturbation-based profiling of the SARS-CoV-2 S-RBD complexes with CR3022 and CB6 antibodies enabled a detailed validation of the proposed approach and quantitative comparison with the experimental deep mutagenesis scanning. The results provide support to a model in which antibody-escaping mutations may preferentially target structurally adaptable energy hotspots and allosteric effector centers that control functional movements and allosteric communication in the complexes. Through in silico mutational scanning and perturbation-based network modeling of the SARS-CoV-2 S trimer complexes with H014, S309, S2M11, and S2E12 antibodies, we examined how antibodies can differentially modulate the allosteric potential of the spike residues and exert control over long-range communications. The results provide a novel insight into regulatory mechanisms of SARS-CoV-2 S proteins, showing that antibody-escaping mutations can target specific effector centers regulating global motions and allosteric interactions in the complexes. This study suggests that the SARS-CoV-2 spike protein may function as a functionally adaptable allosteric machine exploiting the plasticity of regulatory centers to fine-tune the response to antibody binding, which may be useful for therapeutic intervention by targeting specific hotspots of allosteric interactions.

MATERIALS AND METHODS

Sequence Conservation and Coevolutionary Analyses. Sequences of spike glycoproteins corresponding to NCBI GenBank accession ids MN908947 (human SARS-CoV-2), MN996532 (bat coronavirus RaTG13), AY278741 (SARS coronavirus Urbani), KY417146 (bat SARS-like coronavirus), and MK211376 (coronavirus BtRs BetaCoV/YN2018B) were previously used for the multiple sequence alignment (MSA) of protein sequences using the MAFFT approach[63] for coevolutionary detection.[64] In our study, the employed sequences were initially collected from the Pfam database.[65] MSA profiles were then derived with the aid of hidden Markov models using Pfam SARS-CoV-2 special release[66] in which a full collection of spike glycoprotein sequences was assembled and categorized in different Pfam domains. Three Pfam domains were considered for the S1 domain: the NTD (*bCoV_S1_N*, betacoronavirus-like spike glycoprotein S1, N-terminal, Pfam: PF16451, 50 sequences from 49 species, Uniprot SPIKE_CVHSA, Protein Data Bank (PDB) id 6CS0, residues 33−324), the RBD (*bCoV_S1_RBD*, betacoronavirus spike glycoprotein S1, receptor binding, Pfam: PF09408, 50 sequences from 45 species, 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, 92 sequences from 92 species, Uniprot SPIKE_CVHSA, PDB id 6CS0, residues 522−580). The S2 domain is characterized in the family Pfam: PF01601, which contains an additional S2′ cleavage site, a fusion peptide (FP), an internal fusion peptide, heptad repeat (HR) 1/2 domains, and the transmembrane (TM) domain (TD) (99 sequences from 99 species, Uniprot SPIKE_CVHSA, PDB id 6CS0, residues 622−1120). The following Uniprot entries were used for sequence−structure analysis and comparison: P59594: SPIKE_SARS (previously SPIKE_CVHSA) (PDB id 6CS0) and P0DTC2: SPIKE_SARS2 (PDB id 6VXX, 6VYB).

All sequences in the full MSA score within curated thresholds (*E*-value = 10^{-2} and a column-inclusion threshold of 80%) are included in the sequence alignment. A statistically significant and diverse number of spike glycoprotein sequences in the Pfam SARS-CoV-2 special release (∼400 sequences from Pfam domains) provided input for MSA and coevolutionary computations. Through iterations of control experiments in which the number of sequences from the Pfam SARS-CoV-2 domains used for MSA was gradually reduced, we determined that a lower bound of ∼150−170 sequences is sufficient to yield robust MSA profiles and consistent and statistically meaningful coevolutionary relationships. To discriminate coevolutionary associations driven by functional constraints from those determined by common ancestry, the covariance metric was adjusted by the average product correction (APC).[67−69] To evaluate coevolutionary couplings in the SARS-CoV-2 S glycoproteins, we used the MISTIC approach,[70−72] where sequence clustering is implemented to reduce sequence redundancy and sequence clusters are defined at a sequence identity threshold of 62%.

By employing the MISTIC approach,[70−72] we computed the Kullback−Leibler (KL) sequence conservation score KLConsScore according to the following formula

$$\mathrm{KLConsScore}_i = \sum_{i=1}^{N} \ln \frac{P(i)}{Q(i)} \tag{1}$$

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 acid background frequency distribution obtained from the UniProt database.[73] Coevolutionary inter-residue couplings in the SARS-CoV-2 S glycoproteins were computed in MISTIC with three different direct coupling analysis (DCA) methods: mean field DCA (mfDCA),[74−76] pseudo-likelihood maximization DCA (plmDCA),[77,78] and multivariate Gaussian modeling DCA (gaussianDCA).[79,80] For each residue, we computed the

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

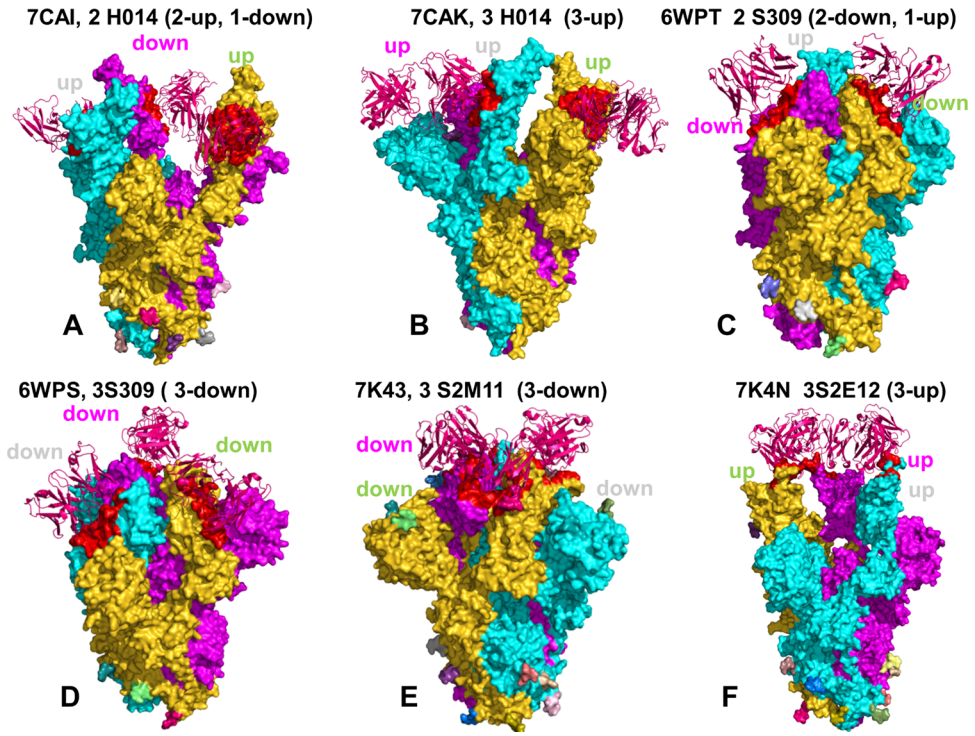


Figure 1. Cryo-EM structures of the SARS-CoV-2 S protein trimer complexes with a panel of antibodies used in this study. (A) Cryo-EM structure of the SARS-CoV-2 S protein trimer with two RBDs in the open state complexed with two H014 Fab fragments (PDB id 7CAI).[45] (B) Cryo-EM structure of the SARS-CoV-2 S protein trimer with three RBDs in the open state complexed with three H014 Fab fragments (PDB id 7CAK).[45] (C) 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).[46] (D) 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).[46] (E) 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).[47] (F) 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).[47] 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.

cumulative covariation score (CScore) parameter, which 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 time scales. In this study, the CG-CABS model[81−85] was used for simulations of the cryo-EM structures of the SARS-CoV-2 S complexes with CR3022, CB6, H014, S309, S2M11, and S2E12 antibodies. In this model, the amino acid residues are represented by C_α, C_β, the center of mass of side chains, and another pseudoatom placed in the center of the $C_\alpha-C_\alpha$ pseudo-bond.[81−83] The position of C_α atoms is confined to a cubic lattice of a grid equal to 0.61 Å. The position of the side chain is dependent on the $C_\alpha-C_\alpha-C_\alpha$ angle of the main chain and the amino acid type. We specifically opted for the CABS model as this is a high-resolution knowledge-based coarse-grained force field that is based on potentials of the mean force obtained from statistical analysis of known protein structures and structural correlations of solved protein structures.[81−83] The appealing advantage of the CG-CABS knowledge-based force field is that residue−residue interactions are considered in a context-dependent manner, where the potential depends on the mutual orientations of the contacting side chains and by implicitly taking into account multibody and solvent effects encoded in mean force statistical parameters of the model.[81−83]

We employed the CABS-flex approach that efficiently combines a high-resolution coarse-grained model and an efficient search protocol capable of accurately reproducing all-atom MD simulation trajectories and dynamic profiles of large biomolecules on a long time scale.[81−85] The sampling scheme of the CABS model used in our study is based on Monte Carlo replica-exchange dynamics and is modeled as a very long random sequence of small local moves of individual amino acids in the protein structure as well as moves of small fragments consisting of two and three residues. The default settings were used for CG-CABS simulations in which soft nativelike restraints are imposed only on pairs of residues where the distance between their C_α atoms is smaller than 8 Å and both residues are part of the same secondary structure elements. No additional custom-designed distance restraints were applied to the simulation scheme. CABS-flex standalone package dynamics implemented as a Python 2.7 object-oriented package was used for fast simulations of protein structures.[85]

A series of several 1000 independent CG-CABS replica-exchange simulations were performed for each of the studied

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

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 antibodies including H014, S309, S2M11, and S2E12 were used in CG-CABS simulations (Figures 1 and S1). 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). In addition, we simulated structures of the SARS-CoV-2 S-RBD complexes with CR3022 and CB6 antibodies using both CABS-CG and all-atom MD simulations.

Structure Preparation and Analysis. All structures were obtained from the Protein Data Bank.[86,87] Hydrogen atoms and missing residues were initially added and assigned according to the WHATIF program web interface.[88,89] The structures were further preprocessed 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 disulfide bonds, removal of crystallographic water molecules and cofactors, 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[90] and ArchPRED.[91] The conformational ensembles were also subjected to MODELLER-based all-atom reconstruction including hydrogen atoms to produce atomistic models of simulation trajectories.[92,93] The side-chain rotamers were refined and optimized by the SCWRL4 tool.[94] The protein structures were then optimized using atomic-level energy minimization with composite physics and knowledge-based force fields using the 3Drefine method.[95] The shielding of the receptor-binding sites by glycans is an important feature of viral glycoproteins, and glycosylation on SARS-CoV proteins can camouflage immunogenic protein epitopes.[96,97]

In addition to the experimentally resolved glycan residues present in the crystallographic and cryo-EM structures of studied SARS-CoV-2 S proteins, the reconstructed atomistic samples from CG-CABS simulation trajectories were elaborated by adding *N*-acetyl glycosamine (NAG) glycan residues and optimized. The structure of glycans at particular glycosites of the closed and open states of the SARS-CoV-2 S protein was previously determined, and these glycans were incorporated in atomistic modeling of the SARS-CoV-2 S protein complexes with antibodies. The glycosylated microenvironment for atomistic models of the simulation trajectories was mimicked using the structurally resolved glycan conformations for 16 out of 22 most occupied *N*-glycans in each protomer (N122, N165, N234, N282, N331, N343, N603, N616, N657, N709, N717, N801, N1074, N1098, N1134, N1158) as determined in the cryo-EM structures of the SARS-CoV-2 spike S trimer in the closed state (K986P/V987P) (PDB id 6VXX) and the open state (PDB id 6VYB) and the cryo-EM structure SARS-CoV-2 spike trimer (K986P/V987P) in the open state (PDB id 6VSB). The glycan-decorated atomistic models were subsequently used for ensemble-based structural, energetic, and network analyses of the SARS-CoV-2 S protein complexes with a panel of antibodies.

MD Simulations of the SARS-CoV-2 S-RBD Complexes with CR3022 and CB6 Antibodies. The crystal structures of the SARS-CoV-2-RBD complexes were simulated in a box size of 85 Å × 85 Å × 85 Å with a buffering distance of 12 Å. Assuming normal charge states of ionizable groups corresponding to pH = 7, sodium (Na^+) and chloride (Cl^-) counterions were added to achieve charge neutrality and a salt concentration of 0.15 M NaCl was maintained. All Na^+ and Cl^- ions were placed at least 8 Å away from any protein atoms and from each other. All-atom MD simulations were performed for an N, P, T ensemble in the explicit solvent using the NAMD 2.13 package[98] with the CHARMM36 force field.[99] Long-range nonbonded van der Waals interactions were computed using an atom-based cutoff of 12 Å with switching van der Waals potential beginning at 10 Å. Long-range electrostatic interactions were calculated using the particle mesh Ewald method[100] with a real-space cutoff of 1.0 nm and a fourth-order (cubic) interpolation. The SHAKE method was used to constrain all bonds associated with hydrogen atoms. Simulations were run using a leap-frog integrator with a 2 fs integration time step. Energy minimization after addition of solvent and ions was carried out using the steepest descent method for 100 000 steps. All atoms of the complex were first restrained at their crystal structure positions with a force constant of 10 kcal mol^{-1} $Å^{-2}$. Equilibration was done in steps by gradually increasing the system temperature in steps of 20 K starting from 10 K until 310 K, and at each step, 1 ns equilibration was done keeping a restraint of 10 kcal mol^{-1} $Å^{-2}$ on the protein C_α atoms. After the restraints on the protein atoms were removed, the system was equilibrated for an additional 10 ns. An NPT production simulation was run on the equilibrated structures for 500 ns keeping the temperature at 310 K and constant pressure (1 atm). In simulations, the Nosé−Hoover thermostat[101] and the isotropic Martyna−Tobias−Klein barostat[102] were used to maintain the temperature at 310 K and pressure at 1 atm, respectively.

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. The BeAtMuSiC approach was employed, which is based on statistical potentials describing the pairwise inter-residue distances, backbone torsion angles, and solvent accessibilities and considers the effect of mutation on the strength of the interactions at the interface and on the overall stability of the complex.[103−105] The binding free energy of the 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

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

The change in the binding energy due to a mutation was then calculated as follows

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

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 1000 equilibrium samples for each of the studied systems.

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

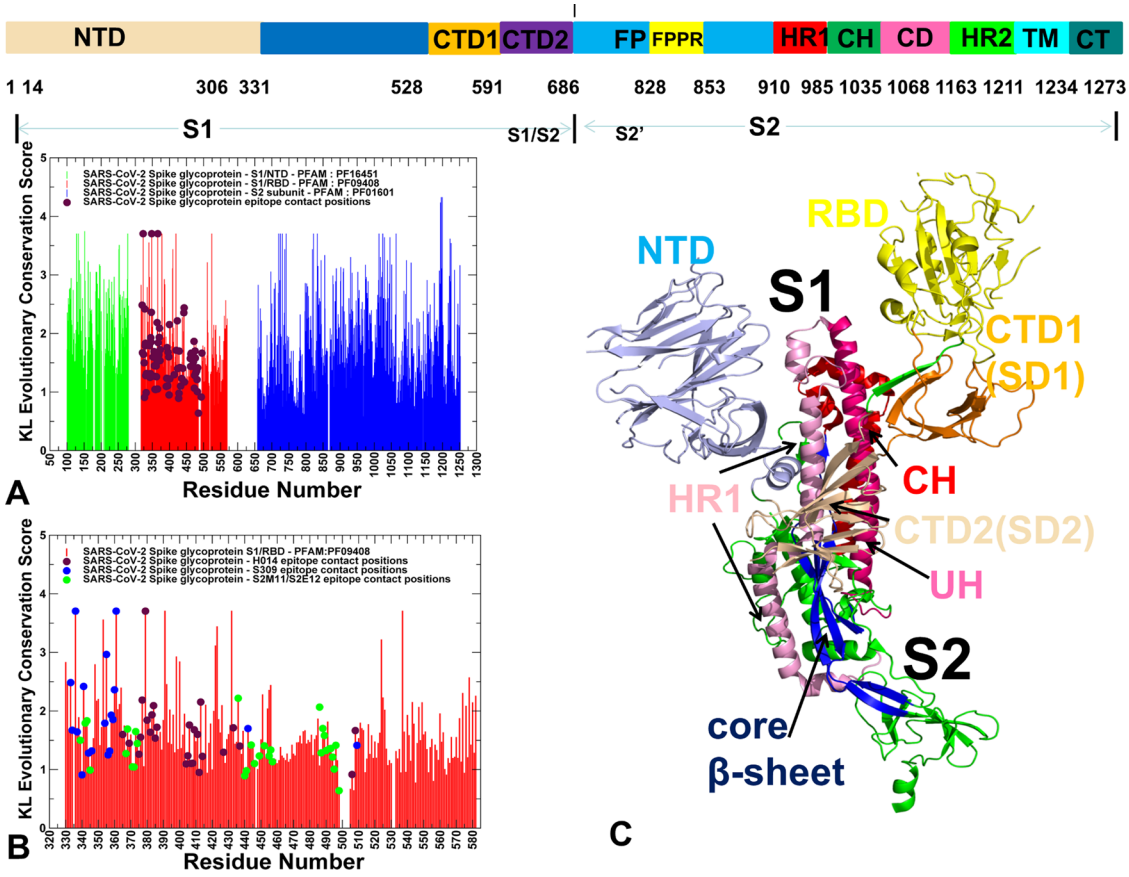


Figure 2. Sequence conservation analysis of the SARS-CoV-2 S protein. (Top panel) Schematic representation of domain organization and residue range for the full-length SARS-CoV-2 S protein. Subunits S1 and S2 include NTD RBD, C-terminal domain 1 (CTD1), C-terminal domain 2 (CTD2), the S1/S2 cleavage site (S1/S2), the S2′ cleavage site (S2′), the fusion peptide (FP), the fusion peptide proximal region (FPPR), heptad repeat 1 (HR1), the central helix (CH) region, the connector domain (CD), heptad repeat 2 (HR2), the transmembrane (TM) domain, and the cytoplasmic tail (CT). (A) KL conservation score. High KL scores indicate highly conserved sites, and low scores correspond to more variable positions. The KL scores for the S1-NTD residues are in green bars, for the S1-RBD regions in red bars, and for S2 residues in blue bars. The KL conservation scores for the epitope residues are shown in filled maroon-colored circles. (B) A close-up view of KL conservation scores for RBD regions of the SARS-CoV-2 S protein 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). (C) Subunit 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; and core β-sheet (residues 711−736, 1045−1076) in blue.

Network Modeling of Residue Interactions. The protein contact network is an undirected, unweighted graph and is built on the basis of the distance matrix $\boldsymbol{d}$, whose generic element d_{ij} records the Euclidean distance between the ith and the jth residue (measured between the corresponding α carbons). A detailed description of the network construction and significance of network descriptors was presented in our previous studies.[106−108] The active network links are defined using a range of contacts between 4 and 8 Å. Once the network is divided into a given number of clusters (powers of 2), we define the participation coefficient as

$$P_i = 1 - \left(\frac{k_i}{k_{si}}\right)^2 \tag{4}$$

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 in intercluster communication. We designate as highly active communication residues the nodes with $P > 0.75$. The proposed methodology of network clustering was implemented as Cytoscape plugin.[109]

We also employed a graph-based representation of protein structures[110−112] with residues as network nodes and the inter-residue edges as residue interactions to construct the residue interaction networks using dynamic correlations[112] and coevolutionary residue couplings,[113] as detailed in our previous studies.[113−115] The ensemble of shortest paths is determined from the matrix of communication distances by the Floyd−Warshall algorithm.[116] Network graph calculations were performed using the python package NetworkX.[117] 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

$$C_b(n_i) = \sum_{j<k}^{N} \frac{g_{jk}(i)}{g_{jk}} \tag{5}$$

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

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.

The Girvan−Newman algorithm[118,119] is used to identify local communities. To characterize global bridges from a community structure, we computed the intercommunity bridgeness parameter. The algorithmic details have been specified in our recent studies.[108,120,121] The bridgeness parameter uses community detection as the input

$$G(i) = \sum_{j \in J} l_{IJ}\delta_{iJ} \tag{6}$$

where the sum is over communities J (different from the community of node i, denoted I) and δ_{ij} is equal to 1 if there is a link between node i and community J and 0 otherwise. l_{ij} corresponds to the distance between community I and community J and is measured by the inverse of the number of links between them. The network parameters were evaluated by the Cytoscape package for network analysis.[122−124]

Perturbation Response Scanning (PRS). The perturbation response scanning (PRS) approach[125,126] has successfully identified hotspot residues driving allosteric mechanisms in single protein domains and large multidomain assemblies.[127−132] The implementation of this approach follows the protocol originally proposed by Bahar and colleagues[127,128] and was described in details in our previous study.[62] 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, the $3N \times 3N$ Hessian matrix $\boldsymbol{H}$, whose elements represent second derivatives of the potential at the local minimum, connects the perturbation forces to the residue displacements. The $3N$-dimensional vector $\Delta\boldsymbol{R}$ of node displacements in response to the $3N$-dimensional perturbation force follows Hooke's law $\boldsymbol{F} = \boldsymbol{H} \times \Delta\boldsymbol{R}$. A perturbation force is applied to one residue at a time, and the response of the protein system is measured by the displacement vector $\Delta\boldsymbol{R}(i) = \boldsymbol{H}^{-1}\boldsymbol{F}^{(i)}$, which is then translated into the $N \times N$ PRS matrix. The second derivative matrix $\boldsymbol{H}$ is obtained from simulation trajectories for each protein structure, with residues represented by C_α atoms, and the deviation of each residue from an average structure was calculated by $\Delta\boldsymbol{R}_j(t) = \boldsymbol{R}_j(t) - \langle\boldsymbol{R}_j(t)\rangle$, and corresponding covariance matrix $\boldsymbol{C}$ was then calculated by $\Delta\boldsymbol{R}\Delta\boldsymbol{R}^{\mathrm{T}}$. 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.[62] The displacement change, $\boldsymbol{\Delta R}^i$, is a $3N$-dimensional vector describing the linear response of the protein and deformation of all of the residues. Using the residue displacements upon multiple external force perturbations, we compute the magnitude of the response of residue k as $\langle\|\Delta\boldsymbol{R}_k^{(i)}\|^2\rangle$ averaged over multiple perturbation forces $\boldsymbol{F}^{(i)}$, yielding the ikth element of the $N \times N$ PRS matrix. The average effect of the perturbed effector site i on all other residues is computed by averaging over all sensor (receiver) residues j and can be expressed as $\langle(\boldsymbol{\Delta R}^i)^2\rangle_{\text{effector}}$. The effector profile determines the global influence of a given residue node on the perturbations in other protein residues and can be used as a proxy for detecting allosteric regulatory hotspots in the interaction networks. In turn, the jth 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 $\langle(\boldsymbol{\Delta R}^i)^2\rangle_{\text{sensor}}$. 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 and Coevolutionary Relationships of the SARS-CoV-2 Proteins Reveal Conserved Regulatory Centers and Functional Role of the Epitope Regions. 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 the KL sequence conservation score as implemented in the MISTIC software package.[70−72] Consistent with previous studies,[133−135] we found that S1-RBD is less conserved than domains in the S2 subunit (Figure 2A,B). The S2 subunit contains an N-terminal hydrophobic fusion peptide (FP), a fusion peptide proximal region (FPPR), a heptad repeat 1 (HR1), a central helix (CH) region, a connector domain (CD), a heptad repeat 2 (HR2), a transmembrane domain (TD), and a cytoplasmic tail (CT). 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 (residues 1211−1234) regions (Figure 2).

Among most conserved residues in the S2 subunit are clusters of conserved cysteine residues forming disulfide bridges that are crucial for stabilization of both prefusion and postfusion 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.[136] The conserved cysteine cluster in the TD region 1220-CCMTSCCSC-1228 displayed high conservation scores, with C1121, M1222, and C1225 featuring the top 1% of the conservation scores (Figure 2A). Indeed, mutagenesis of cysteine cluster I (1220-CCMTS-24), located immediately proximal to the TD, showed 55% reduction in S-mediated cell fusion as compared to the wild-type S protein.[137] The proximal cysteine cluster 1225-CCSC-1228 is similarly important as alanine mutations in this cluster resulted in 60% reduction of S-mediated cell fusion.[137] 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,[137] indicating that the functional role of these clusters may be closely linked with the conservation level of cysteine residues. Interestingly, the most conserved S2 positions included cysteine residues C720, C725, C731, C742, C822, C822, C833, C1014, C1025, and C1064 (Figure 2A). Consistent with these observations, a conserved region flanked by C822 and C833 residues is critical for interactions with components of the SARS-CoV S trimer to control the activation of membrane fusion.[138]

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 prefusion and postfusion SARS-CoV-2 spike protein conformations[139,140] (Figure S2, Supporting Information). Some of these regions include cysteine clusters formed by C720, C725, C731, and C742 in the upstream helix (UH) regions. Another conserved cysteine cluster of disulfide bonds is formed in the β-hairpin (BH) 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, respectively,

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

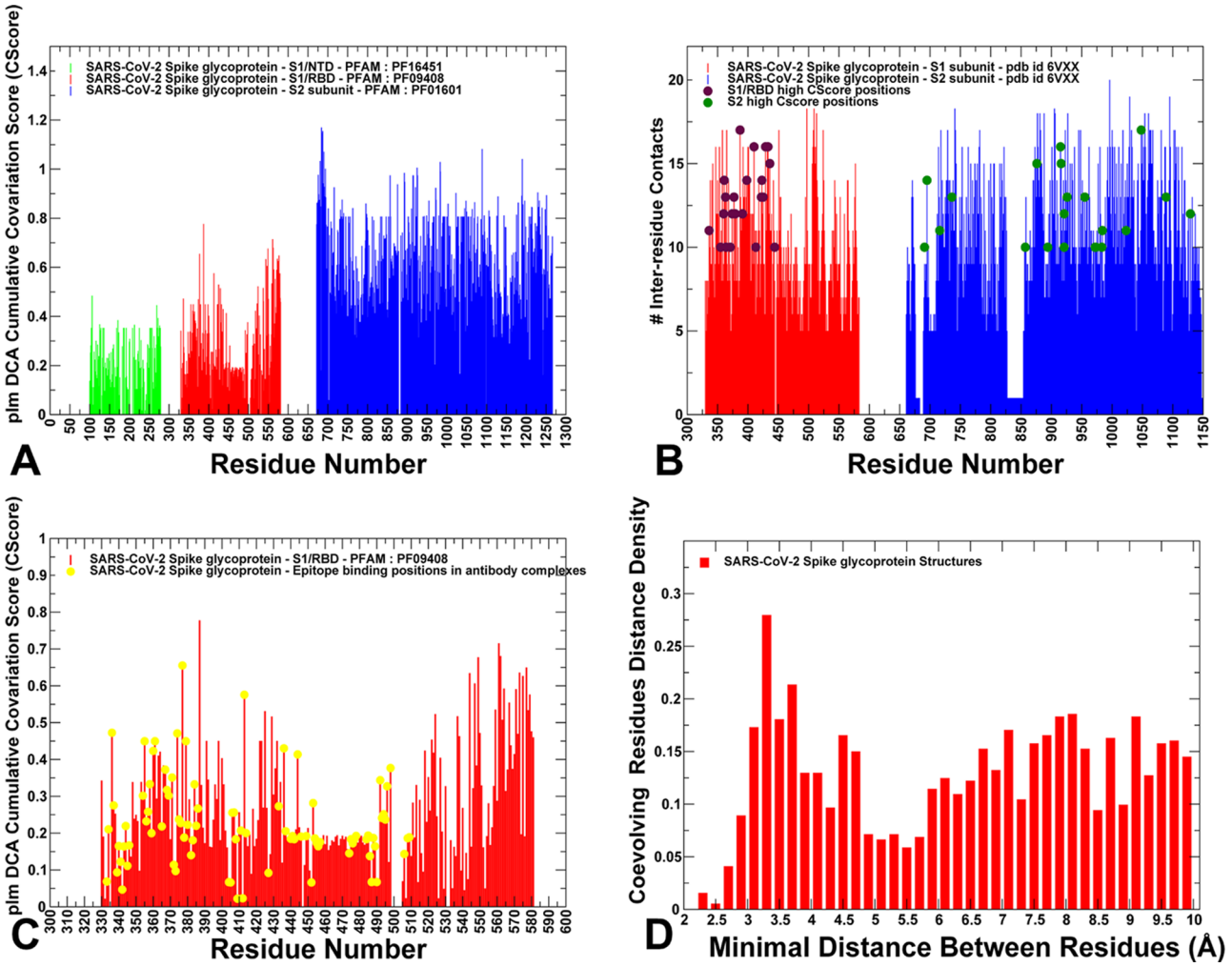


Figure 3. Coevolutionary profiles of the SARS-CoV-2 S proteins. (A) plmDCA-based coevolutionary CScore profile for the SARS-CoV-2 S proteins (P0DTC2: SPIKE_SARS2 sequence numbering). The CScore values are shown for the S1-NTD residues in green bars (Pfam: PF16451), for the RBD in red bars (Pfam: PF09408), 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 in filled green circles. (C) 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.

in SPIKE_SARS2 sequence numbering) (Figure S2, Supporting Information). 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 β-sheet RBD structure in the SARS-CoV-2 S protein.

Using the MISTIC approach,[70−72] we determined coevolutionary dependencies between S protein residues using the plmDCA model (Figure 3). To identify critical nodes of this coevolutionary network, we computed plmDCA-based CScore profiles (Figure 3A) 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 identified coevolutionary couplings for functionally important regions and mapped high CScore positions onto the binding epitopes for the studied SARS-CoV-2 complexes (Figure 3B). First, the results revealed an appreciable density of coevolving centers in the S1 subunit, in the RBD, and especially in CTD1 regions. This pattern can be further illustrated by a circular representation of the pairwise coevolutionary scores (Figure S3, Supporting Information) showing a greater concentration of coevolutionary links that are anchored by the CTD1 regions (residues 528−591). The distribution of CScores pointed to the higher density of coevolutionary couplings in the S2 subunit (Figure 3A). The residues with significant CScore values are distributed across various S2 regions, including UH (residues 736−781), CH (residues 986−1035), HR1 (residues 910−985), HR2 (residues 1163−1211), and β-hairpin (BH) (residues 1035−1071) regions. A dense network of coevolutionary coupled residues in the S2 regions is evident from a graphical annotation of the pairwise coevolutionary scores (Figure S4, Supporting Information). Interestingly, the distribution of CScore values for the epitope residues showed that many contact positions are aligned with highly coevolving residues (Figure 3B).

Structural analysis of coevolutionary hotspots corresponding to the local maxima of the distribution revealed clusters situated in the functional regions (Figure S5, Supporting

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

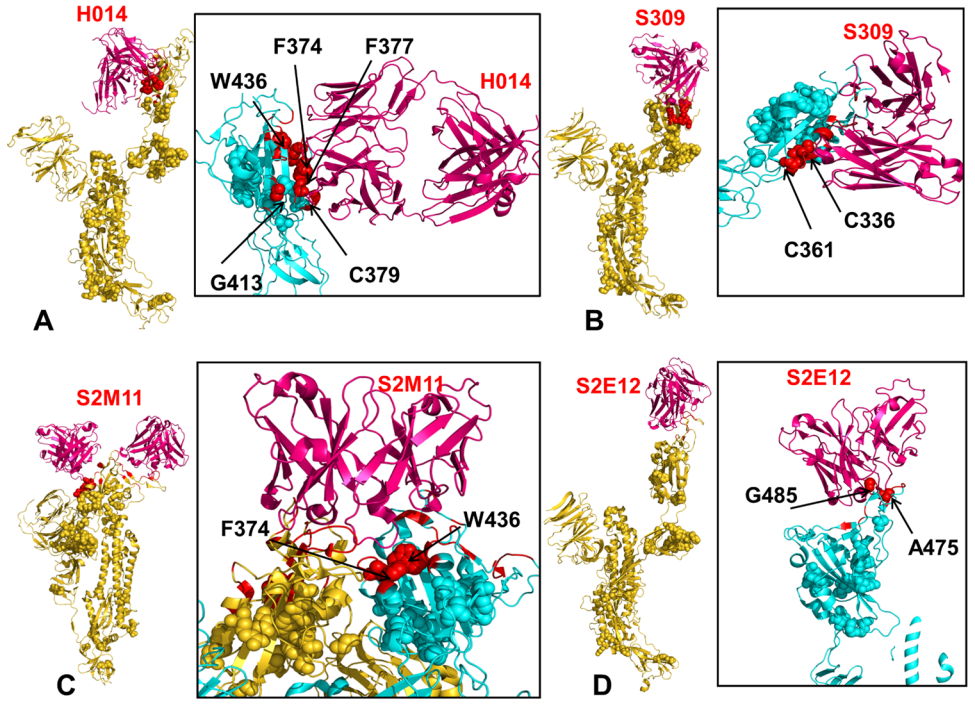


Figure 4. Structural maps of coevolutionary centers in the epitope regions of the SARS-CoV-2 complexes with antibodies. (A) Structural map of coevolutionary centers in the S complex with H014 (PDB id 7CAI/7CAK) projected onto a single "up" protomer shown in green ribbons. The coevolutionary centers are in spheres, and high CScore hotspots from the binding epitope are in red spheres. A close-up of the H014 binding epitope with the coevolving centers involved in direct contact with H014 is 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 contact with S309 are 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 contact with S2M11 and S2E12 are in red spheres and annotated.

Information). We observed that coevolutionary centers could 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 the S2 subunit as well as RBD and CTD1 regions in the S1 domain (Figure S5, Supporting Information). Interestingly, the coevolutionary signal can be significant for the S2 positions involved in multiple interactions in the UH, HR1, and CH regions. We also noticed that some of the important clusters of coevolutionary centers in the CTD1 region and near-critical junctures UH/HR1/CH are located in close proximity to the glycan attachment sites, suggesting the important role of the glycan microenvironment for modulating couplings between distant sites in the SARS-CoV-2 S trimer structures. We also computed the average number of inter-residue contacts for each residue and aligned this distribution with the top CScore positions (Figure 3C). It appeared that coevolutionary centers might have a significant number of interacting contacts. In particular, coevolutionary hotspots in the RBD regions are often aligned with the peaks of the contact distribution, suggesting that the level of coevolution may be greater in residues involved in multiple interactions. These findings are consistent with the notion that the increased structural and functional constraints for sites involved in a large number of inter-residue contacts can often imply higher coevolution values.[141]

The distance probability distribution of coevolving directly coupled residue pairs in the studied SARS-CoV-2 S structures showed several local maxima at 3.2 Å and 4.7 Å and a shallow peak near 7−8 Å (Figure 3D). The first two peaks reflect physical interactions between residues. Hence, direct coevolutionary residue couplings in the SARS-CoV-2 S structures are dominated by spatially proximal residue pairs, suggesting that coevolutionary signals are stronger for locally interacting residues than for residues involved in long-range interactions.[142] The distribution highlighted an intermediate range of coevolutionary couplings at 7−8 Å, which may reflect couplings between spatially proximal functional regions (Figure 3D). This third distribution peak can correspond to coevolutionary couplings anchored by CTD1 regions (residues 529−591) in the S1 subunit. A significant coevolutionary signal between functional regions separated by medium-range distances may 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 toward the CTD1 region and regions in the S2 subunit (Figure 4). Several high CScore residues W436, G413, F374, F377, and C379 are involved in the interactions with H014. Multiple favorable interactions are formed by F377 with Y105, T58, S59, D60, and Y50 of H014 and by C379 with N55, T58, G56, and G57 positions of H014 (Figure 4A). S309 interacts with the highly conserved C336 and C361 positions, which also correspond to coevolutionary hotspots (Figure 4B). S2M11 interacts with the conserved RBD sites F374 and W436, which also displayed high CScore values (Figure 4C). A smaller patch of coevolutionary centers is involved in contact with S2E12 (Figure 4D). 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 of the SARS-CoV-2 Spike Complexes with Antibodies Examined through CG-

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

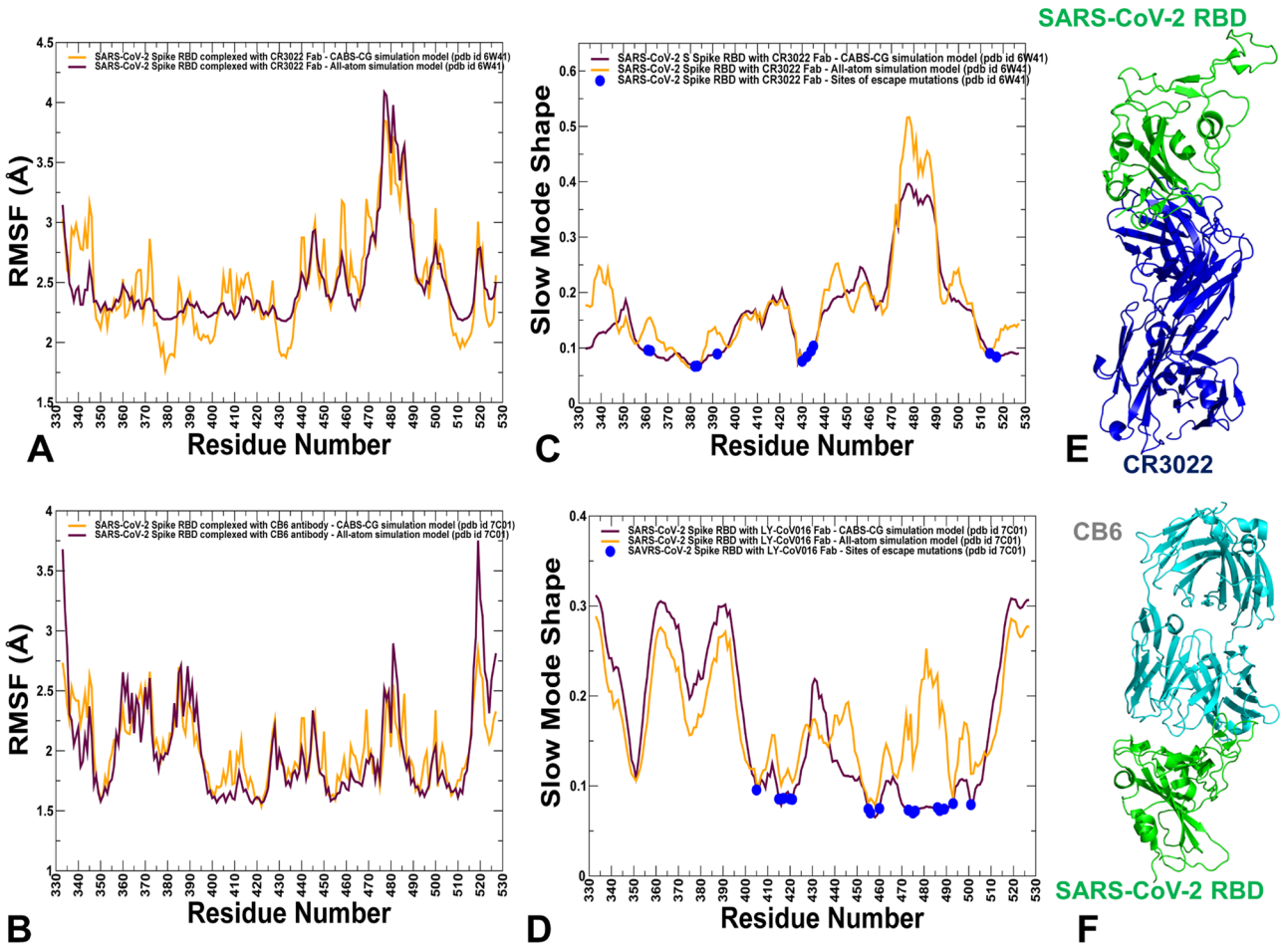


Figure 5. Conformational dynamics and essential mobility profiles of the SARS-CoV-2 S-RBD complexes with CR3022 and CB6 antibodies. (A, B) The root-mean-square fluctuations (RMSFs) obtained from CG-CABS simulations (in green lines) and all-atom MD simulations (in maroon lines) for the S-RBD complex with CR3022 and CB6 antibodies. (C, D) The essential mobility profiles for the S-RBD complexes with CR3022 and CB6 antibodies averaged over the three lowest frequency modes are obtained using principal component analysis (PCA) of trajectories from CG-CABS simulations (in green lines) and all-atom MD simulations (in maroon lines). The sites of escaping mutations for CR3022 and CB6 along the slow mode profiles are indicated in filled blue circles. (E) The structure of the S-RBD complex with CR3022 (PDB id 6W41) is shown in ribbons. The RBD is in green and CR3022 in blue. (F) The structure of the S-RBD complex with the CB6 antibody (PDB id 7C01) is shown in ribbons. The RBD is in green and CB6 in cyan.

CABS and Atomistic MD Simulations. To test the applicability of the CG-CABS approach for adequate simulation of conformational dynamics of the SARS-CoV-2 S trimers with a panel of neutralizing antibodies, we first performed a validation study in which CG-CABS and all-atom MD simulations were performed for the SARS-CoV-2 S-RBD complexes with CR3022 and CB6 antibodies (Figure 5). CABS trajectories were subjected to atomistic reconstruction and refinement, thereby allowing for a direct comparative analysis with MD simulations performed in an explicit solvent environment. Through a detailed comparison with the experimental structural and binding screening data available for the SARS-CoV-2 S complexes with CR3022 and CB6 antibodies, we tested a hypothesis according to which sites targeted by antibody escape mutations may correspond to key regulatory control points and hinge centers that coordinate functional dynamics and long-range allosteric couplings in the systems. Using a comparison of CG-CABS and MD simulations, we verified the reliability of the proposed simulation model and examined how the SARS-CoV-2 spike protein can exploit the plasticity of the RBD regions to modulate specific dynamic responses to antibody binding.

First, we compared the RMSF profiles obtained from CABS-based reconstructed trajectories and all-atom MD simulations, revealing a generally highly similar mobility distribution for the RBD residues in complexes with CR3022 (Figure 5A) and CB6 (Figure 5B) antibodies. CG-CABS trajectories featured slightly more stable profiles of the S-RBD complexes, while MD simulations highlighted a more dynamic nature of the RBD regions. Both simulation models accurately reproduced stability of the conserved core of SARS-CoV-RBD consisting of antiparallel β strands (β1 to β4 and β7) (residues 354−358, 376−380, 394−403, 431−438, 507−516) and a particularly significant stabilization of β-sheets (β5 and β6) (residues 451−454 and 492−495) that anchor the RBM region to the central core (Figure 5). The observed similarities of the dynamic profiles suggested that CG-CABS simulations accompanied by atomistic reconstruction could provide a fairly accurate and affordable simulation approach for quantifying flexibility of the larger SARS-CoV-2 S trimer complexes with the panel of antibodies.

We also characterized collective motions and determined the hinge regions in the SARS-CoV-2 S-RBD complexes using principal component analysis (PCA) of trajectories derived

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

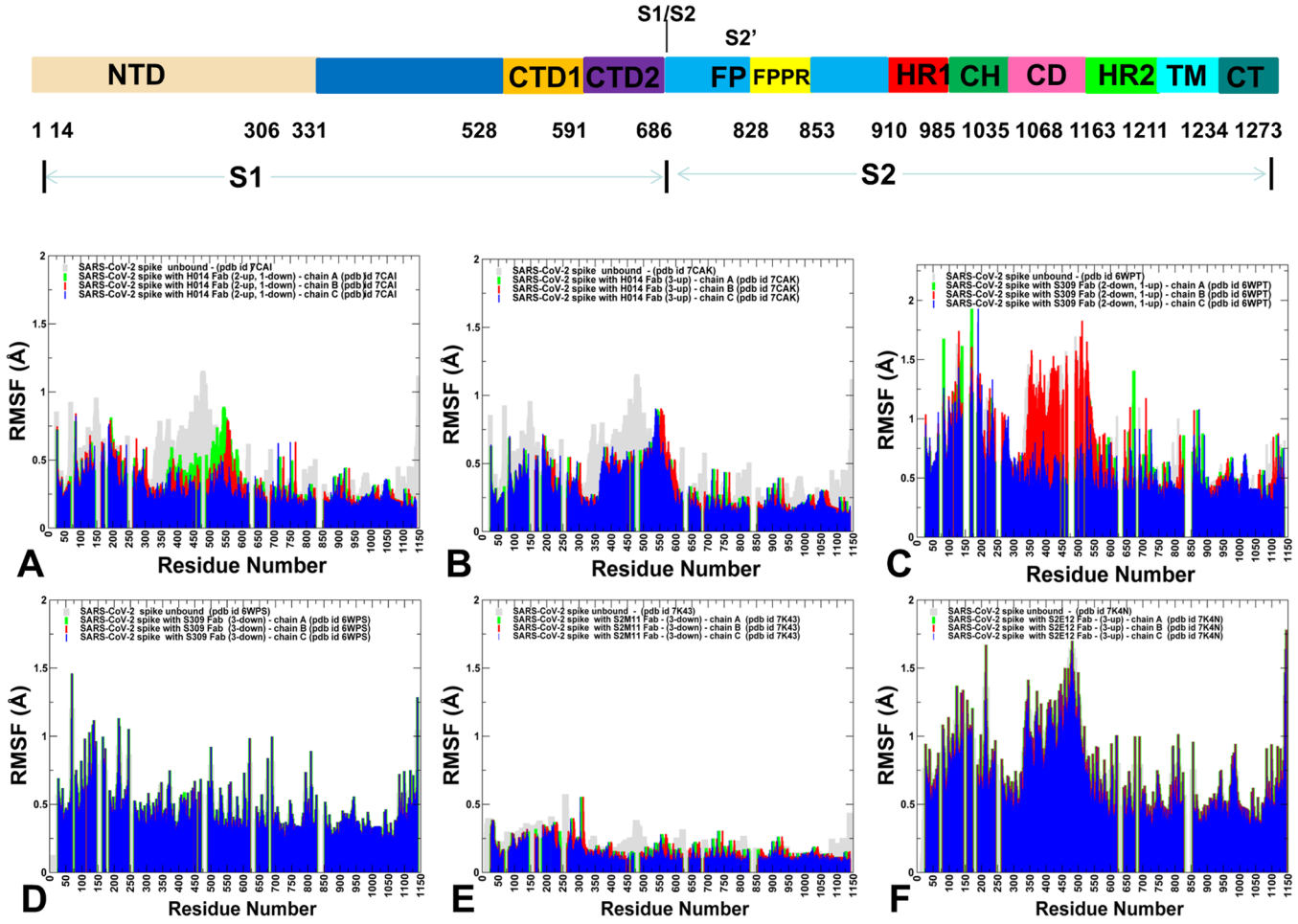


Figure 6. CABS-GG conformational dynamics of the SARS-CoV-2 S complexes. A schematic representation of domain organization and residue range for the full-length SARS-CoV-2 S protein is shown above conformational dynamic profiles. (A, B) Root-mean-square fluctuation (RMSF) profiles from simulations of the cryo-EM structures of the SARS-CoV-2 S trimer with H014. (C, D) RMSF profiles from simulations of the cryo-EM structures of the SARS-CoV-2 S trimer with S309. (E) 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) 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−C) are shown in green, red, and blue bars, respectively. The RMSF profiles for the unbound forms of the S protein trimer are shown in light-gray bars.

from CG-CABS and MD simulations using the CARMA package[143] (Figure 5C,D). The functional dynamic profiles showed a considerable overlap, indicating that both simulation models can yield a consistent prediction of hinge sites and collective motions. Using deep scanning mapping data of escape mutations for SARS-CoV-2 S-RBD binding with the CB6 antibody[50] and CR3022,[60] we projected sites of the maximum escape on the essential mobility profiles (Figure 5C,D). The results showed that key antibody-escaping mutations may often target regulatory hinge positions that coordinate collective motions and allosteric interactions.

Encouraged by these validation experiments, we then performed a series of multiple long CG-CABS simulations followed by MODELLER-based atomistic reconstruction of trajectories for the SARS-CoV-2 S trimer complexes with H014, S309, S2M11, and S2E12 antibodies (Figure 6). Molecular simulations of the SARS-CoV-2 S complexes provided a quantitative picture of the differences in flexibility of the S-protein-bound states and highlighted the effect of binding on modulation of the spike protein mobility. The inter-residue contact maps (Figure S6, Supporting Information)[144] and inter-residue distance maps (Figure S7, Supporting Information)[145] in the SARS-CoV-2 S complexes indicated that the density of the interaction contacts is significantly greater in the densely packed S2 domains.

The conformational dynamic profiles showed that H014 binding can induce dynamic changes by considerably reducing thermal fluctuations in the S1 regions of the interacting open protomers as compared to the unbound trimer form (Figure 6A,B). The H014 epitope is broadly distributed across RBD regions (residues 368−386, 405−408, 411−413, 439, and 503), forming a cavity on one side of the RBD. Although most of the contacts are formed with moderately conserved residues, H014 maintains favorable interactions with the two most highly conserved F377 and C379 positions in the RBD region (Tables S1 and S2, Supporting Information). Of particular importance are stable H014 contacts with S383 and G413 residues that are located at the interprotomer boundaries (S383−D385) and (G413−V987) and function as regulatory switches of the S protein equilibrium.[25] We also observed small thermal fluctuations with RMSF < 1.0 Å for the S1 epitope positions (residues 368−386, 405−408, 411−413, 439, and 503), which were considerably rigidified in both H014 complexes (Figure 6A,B). These findings are consistent with

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

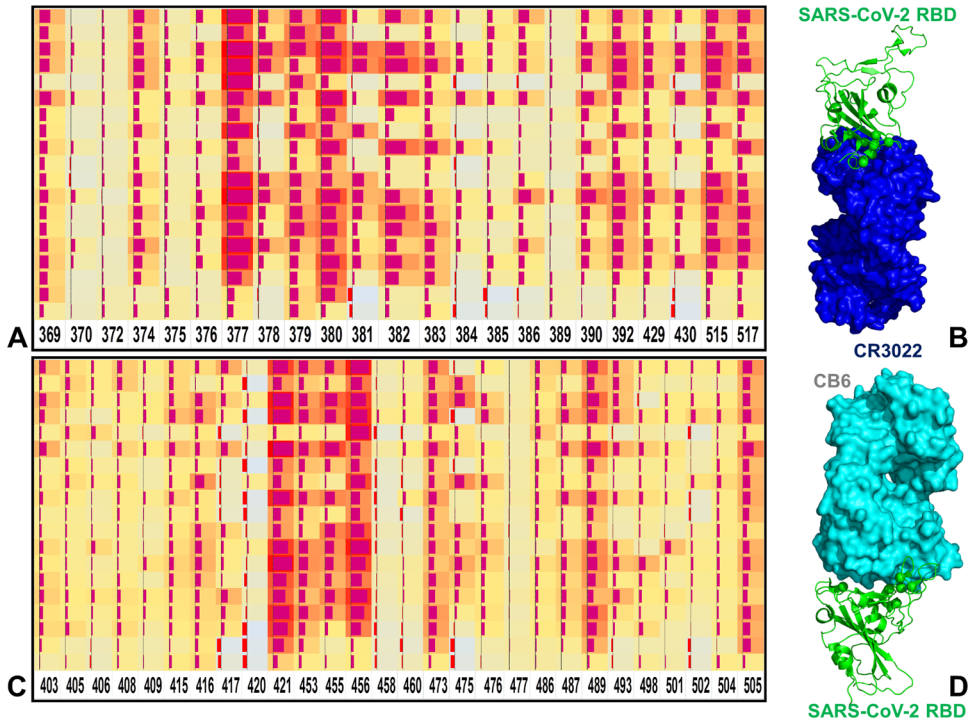


Figure 7. Mutational sensitivity heatmaps for the SARS-CoV-2 S-RBD complexes with CR3022 (A) and CB6 antibodies (C) show the effect of all single mutations at the antibody-contact sites on the computed binding free energy changes. The squares on the heatmap are colored by mutational effect in a three-colored scale from red to light blue, with red indicating the largest destabilization effect. The data bars correspond to the computed binding free energy changes, where positive values (destabilizing mutations) are shown by bars toward the right end of the cell and negative values (stabilizing mutations) are shown as bars oriented toward the left end. The length of the data bar represents the value in the square cell. (B) Structure of the S-RBD complex with CR3022 (PDB id 6W41) with the RBD in green ribbons and CR3022 in the blue surface. (D) Structure of the S-RBD complex with CR3022 (PDB id 6W41) with the RBD in green ribbons and CB6 in the cyan surface. The binding energy hotspots for CR3022 binding (F377, C379, Y380, V382, S383, F392, and F515) and for CB6 binding (Y421, L455, F456, Y473, Y489, and F505) corresponding to dark-red square cells with large data bars are shown in green spheres.

the experimental data suggesting that H014-induced changes could trigger stabilization changes in both the RBD and NTD regions.[57]

Conformational dynamics of the SARS-CoV-2 S protein complex with S309 showed relatively moderate changes in the flexibility induced by antibody binding (Figure 6C,D). In the S309 complex with the closed form of the SARS-CoV-2 S trimer, the thermal stability of the closed S protein is particularly pronounced (Figure 6D). S309 engages an epitope that is distinct from the highly mobile RBM region, making multiple stable contacts in the course of simulations with the two most conserved RBD positions C336 and C361 as well as neighboring residues L335 and P337 (Tables S3 and S4, Supporting Information). These residues form one of the disulfide linkages Cys336−Cys361 that stabilize the β-sheet RBD structure.

S2M11 functions by locking down the SARS-CoV-2 S trimer in the closed state through binding to a quaternary epitope comprising distinct regions of two neighboring RBDs within an S trimer. Conformational dynamic 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 displayed only very minor thermal fluctuations (Figure 6E). During simulations, S2M11 maintains a network of interaction contacts with highly conserved sites F374 and W436 from the RBD of one protomer and F486 on the RBM motif of the protomer (Table S5, Supporting Information) that collectively anchor and stabilize the binding mode of the antibody. S2E12 binds to the more dynamic RBM motif of the open protomers (Table S6, Supporting Information). A relatively small binding epitope in the S2E12 complex with the fully open form of S protein produced the dynamic profile where NTD and RBD regions showed an appreciable degree of mobility (Figure 6F). The important finding of the conformational analysis is that H014, S309, and S2M11 can induce modulation of the conformational dynamics, leading to stabilization of both S1 and S2 regions in the open protein forms, which may effectively counteract the intrinsic flexibility of the receptor-accessible, open spike conformations and thus induce potent neutralization effects. The functional dynamic analysis revealed hinge regions corresponding to sites S383, K386, D428, F518, and V539 positions (Figure S8, Supporting Information) that can regulate the interdomain movements of the NTD and RBD with respect to the S2 subunit and are preserved in S trimer structures.[108] Several hinge sites in the closed and partially open forms are located near L570, I572, Y855, I856, and S591 residues forming the interprotomer and interdomain interfaces and acting as regulatory switch centers governing the population shifts between closed and open forms.[23]

Mutational Scanning Heatmaps Reveal Binding Energy Hotspots in the SARS-CoV-2 Complexes with Antibodies. To provide further comparison between the computational and experimental data, we performed mutational sensitivity analysis and constructed mutational heatmaps for SARS-CoV-2 S-RBD binding with CR3022 and CB6

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

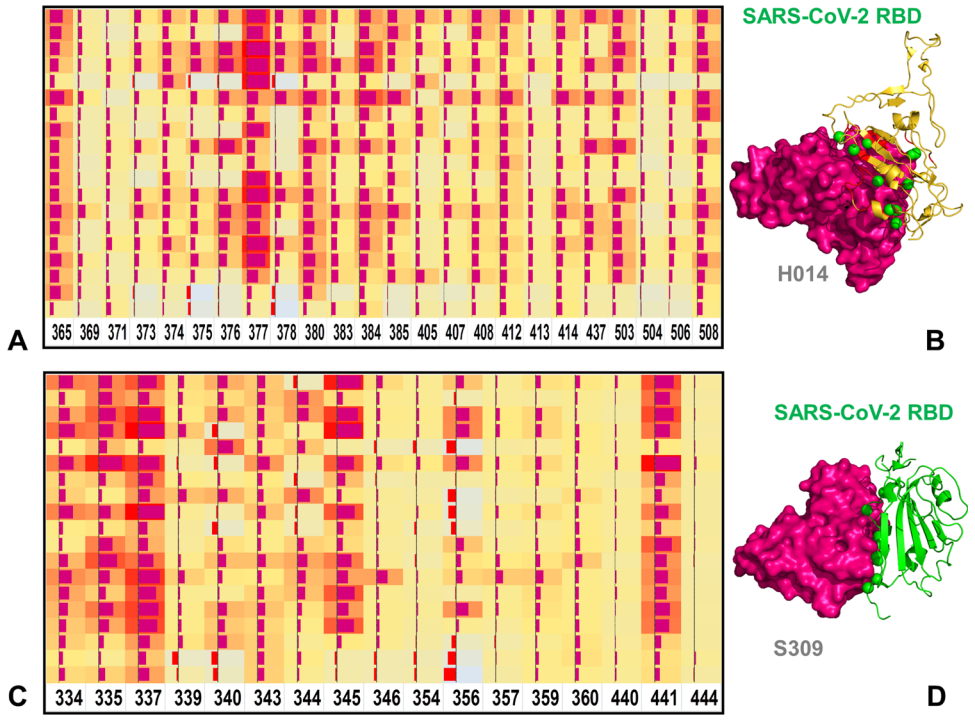


Figure 8. Mutational sensitivity heatmaps for the SARS-CoV-2 S trimer complexes with H014 (A) and S309 antibodies (C) show the effect of all single mutations at the antibody-contact sites on the computed binding free energy changes. The squares on the heatmap are colored by mutational effect in a three-colored scale from red to light blue, with red indicating the largest destabilization effect. The data bars correspond to the computed binding free energy changes. (B) Structure of the SARS-CoV-2 S trimer bound with H014 in the closed form (PDB id 7CAK). Only the S-RBD (green ribbons) is shown, and the H014 antibody is in the hot-pink-colored surface. (D) Structure of the SARS-CoV-2 S trimer with the S309 antibody in the closed form (PDB id 6WPS). The S-RBD (green ribbons) is shown, and S309 is in the hot-pink-colored surface. The binding energy hotspots for H014 binding (F377, C379, Y380, S383, P384, V503, and Y508) and for S309 binding (N334, L335, P337, N343, T345, and L441) are shown in green spheres.

antibodies (Figure 7). According to deep mutational scanning studies,[29,50,60] these antibodies target different epitopes and have a different binding affinity and stability footprint. The computed heatmap for S-RBD binding with CR3022 showed that positions F377, C379, Y380, V382, S383, F392, and F515 from the RBD epitope corresponded to strong energetic hotspots (Figure 7A,B). These results are in excellent agreement with the deep mutational scanning analysis of binding and stability.[29,60] The computed mutational maps reproduced the experimental data, highlighting the fact that important epitope residues in the cryptic binding site of CR3022 located in the RBD core can be constrained with respect to stability.[29,60] Furthermore, in silico mutational scanning is consistent with complete mapping of antibody-escaping RBD mutations, showing that CR3022 recognition can be evaded by mutations of the RBD core residues F374, G381, V382, S383, and F392.[60] For further validation and comparison with the experiment, we also computed the mutational sensitivity map of S-RBD binding with the neutralizing antibody CB6 (Figure 7C,D) that competes with ACE2 by interacting with the RBM epitope and blocking ACE2 binding.[50] The heatmap revealed the binding energy hotspots corresponding to Y421, L455, F456, Y473, Y489, and F505 residues (Figure 7C,D), which is in good agreement with deep mutational scanning data.[50] Some of these residues L455, F456, and Y489 are sites of CB6-escaping mutations, suggesting that antibody recognition can be evaded through modifications of structurally adaptable interacting sites and allosteric centers, which may not be directly involved in the binding interactions.

Equipped with this validation analysis, we used conformational ensembles generated in CG-CABS simulations to perform a systematic profiling of the S protein residues in complexes with H014, S309, S2M11, and S2E12 (Figures 8 and 9). Structural analysis indicated that H014 can sterically compete with ACE2 for RBD binding although its epitope does not overlap with the ACE2-binding site. The mutational sensitivity heatmap revealed a wide range of important binding hotspots in the SARS-CoV-2 S complexes with H014 (Figure 8A,B). In particular, in silico scanning showed a significant contribution of conserved RBD residues F377, C379, Y380, S383, P384, V503, and Y508 (Figure 8A). Among binding energy hotspots, we also detected 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 the binding signal through clusters of proximal coevolutionary pairs in the RBD and CTD1 regions. In agreement with the experimental evidence[46] and owing to the overlap in the binding epitopes of CR3022 and H014 antibodies, the mutational profiling revealed the role of conserved F377, Y380, and S383 residues as consistent and strong energetic hotspots for both CR3022 and H014 antibodies (Figure 8A,B). According to our analysis, H014 interactions with conserved positions F377, C379, and S383 located away from the RBM region could be important for binding and modulation of the enhanced cross-neutralization

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

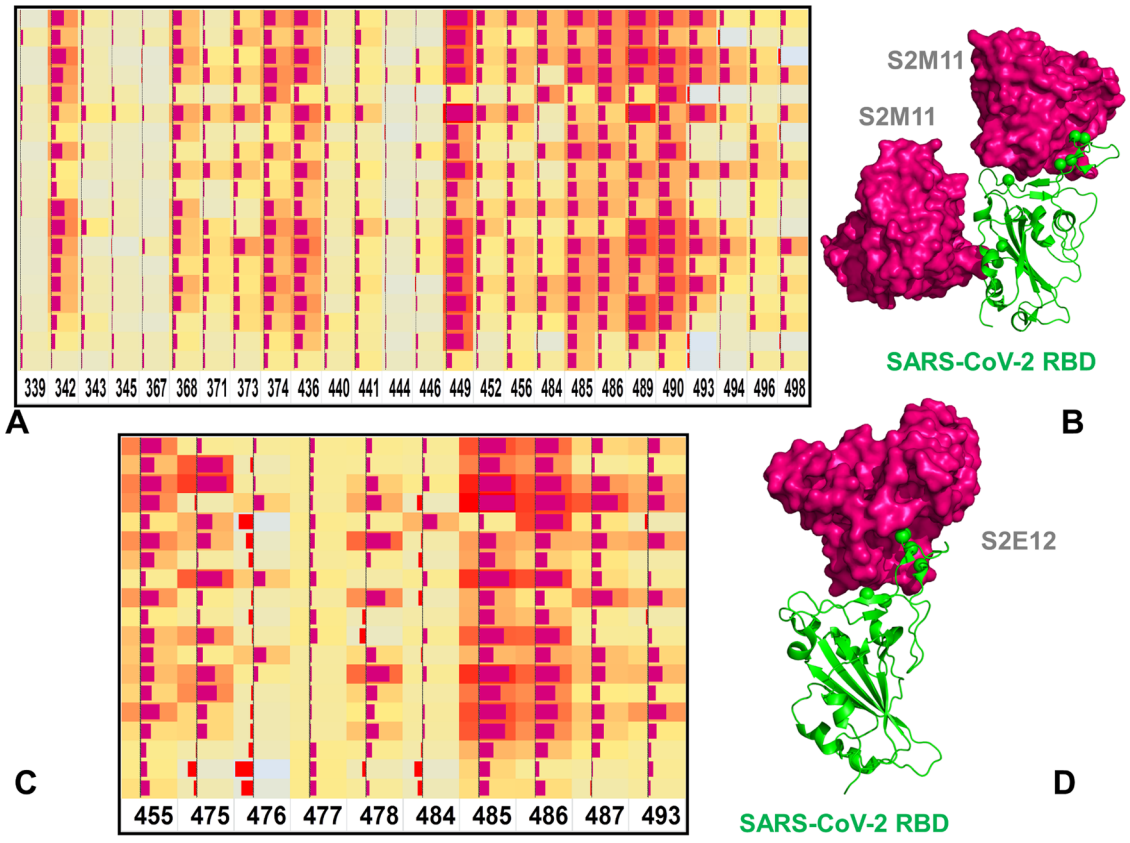


Figure 9. Mutational sensitivity heatmaps for the SARS-CoV-2 S trimer complexes with S2M11 (A) and S2E12 antibodies (C) show the effect of all single mutations at the contact sites on the computed binding free energy changes. The squares on the heatmap are colored in a three-colored scale from red to light blue, with red indicating the largest destabilization effect. (B) Structure of the SARS-CoV-2 S trimer bound with S2M11 in the closed form (PDB id 7K43). S2M11 binds to a quaternary epitope spanning two RBDs of the adjacent protomers in the S trimer. Only the S-RBDs (green ribbons) are shown, and the S2M11 antibody is in the hot-pink-colored surface. (D) Structure of the SARS-CoV-2 S trimer with S2E12 in the open form (PDB id 7K4N). The S-RBD (green ribbons) is shown, and S309 is in the hot-pink-colored surface. The binding energy hotspots for S2M11 binding (F374, W436, Y449, F456, F484, F486, Y489, and F490) and for S2E12 binding (L455, T485, and F486) are shown in green spheres.

activities. These results are also in accordance with the recent structure−functional investigation of the SARS-CoV-2 S complex with H014 and P17 forming a two-antibody cocktail with synergistic neutralization achieved by simultaneously binding to distinct RBD epitopes, where H014 strongly binds to a highly conserved patch across SARS-CoV and SARS-CoV-2, while P17 targets a SARS-CoV-2-specific patch.[146] According to this illuminating experimental study, H014 binding to conserved hotspots F377, C379, Y380, S383, and P384 is critical to promote synergistic neutralization together with the P17 antibody by acting as a shield of the S1 region and blocking the SARS-CoV-2-RBD interactions with host cellular receptors. In addition, a more detailed analysis of the H014 heatmap showed a strong destabilization effect caused by Y508 mutations, which is in line with the recent experimental evidence showing that H014 was completely incapable of neutralizing N234Q, Y508H, and D614G + A435S mutants.[147]

Another cross-neutralizing antibody S309 binds to the proteoglycan site that involves an *N*-glycan at N343 and involves 17 conserved RBD residues, which were implicated for the observed cross-reactivity and neutralization.[46] In agreement with the experimental data, we identified the binding hotspots in conserved positions N334, L335, P337, N343, T345, and L441 (Figure 8C,D). At the same time, the antibody interactions with R346, N354, R357, and K444 of SARS-CoV-2-RBD that are semiconservatively substituted in SARS-CoV RBD showed a considerable mutational tolerance for S309 binding (Figure 8C,D). The preferential S309 binding with conserved RBD sites may explain the lack of pronounced antibody-escaping mutations as modifications in the flexible region of the S309 epitope footprint can be readily tolerated and have minimal effects on serum antibody binding.[148] By targeting evolutionary conserved RBD epitopes, H014 and S309 can potentially better combat virus resistance and prevent emergence of antibody-escaping mutations.

The binding energy hotspots in the S2M11 complex with the S protein occupy two different regions, where one group includes conserved RBD sites F374 and W436, which also displayed high CScore values (Figure 9A,B). Another group of binding energy hotspot positions includes moderately conserved residues Y449, F456, F484, F486, Y489, and F490, which form a critical patch of the RBM binding interface with the host receptor. The mutational sensitivity heatmap clearly pointed to the key hotpots Y449, Y489, and F490, where all substitutions can lead to significant loss of binding affinity and stability (Figure 9A,B).

These results are in excellent agreement with the binding assay experiments, revealing that Y449N, E484K/Q, F490L, and S494P RBD variants led to decreased S2M11 binding to RBD.[59] Notably, the strongest experimentally observed effect

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

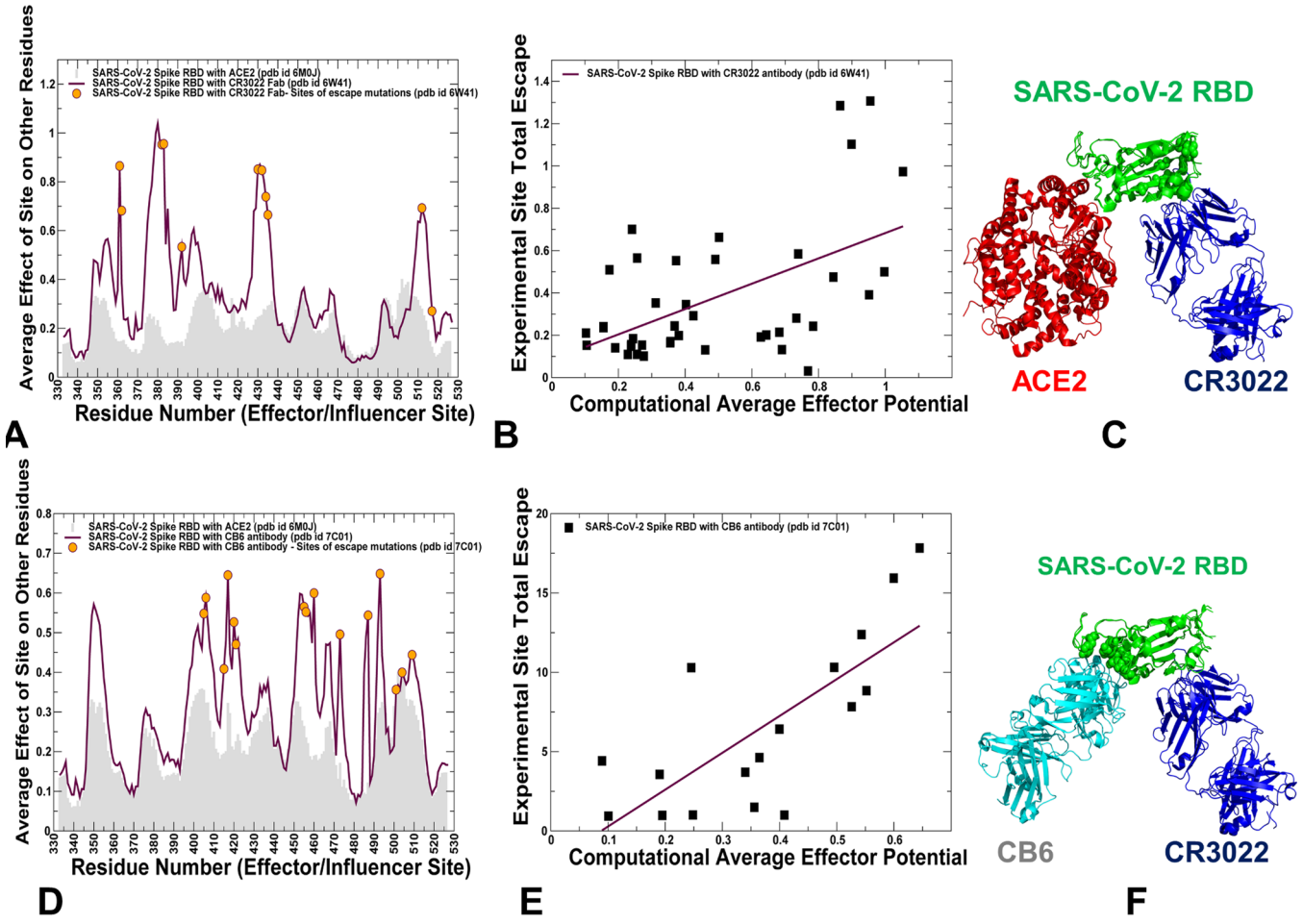


Figure 10. Analysis of the PRS effector profiles for the SARS-CoV-2 S-RBD complexes with CR3022 and CB6 antibodies. (A, D) The PRS effector profiles for the S-RBD complex with CR3022 (PDB id 6W41) and CB6 (PDB id 7C01) are shown in maroon-colored lines. The sites featuring significant total antibody escape are mapped onto the profiles and shown as orange-colored filled circles. The PRS effector profiles are superimposed with the PRS effector distribution for the S-RBD complex with ACE2 shown in light gray (PDB id 6M0J). (B, E) Correlation between the computed effector potential and the experimentally determined total escape for the S-RBD residues features escaping mutations for CR3022 and CB6 antibodies, respectively. (C) Superposition of the S-RBD (in green) bound with ACE2 (in red) and CR3022 (in blue). The CR3022 binding epitope has no overlap with the ACE2-binding site. (F) Superposition of the S-RBD (in green) bound with CB6 (in cyan) and CR3022 (in blue).

was seen for Y449N and F490L mutations.[59] We examined in more details the sensitivity profiles for these RBD positions in the S protein complex with S2M11 and found that these mutations induce the largest destabilization change on S2M11 binding with SARS-CoV-2 S-RBD binding (Figure S9, Supporting Information). S2E12 recognizes an RBD epitope overlapping with the RBM and can only interact with open RBDs through electrostatic and van der Waals interactions with residues L455, A475, G476, S477, T478, F484, T485, F486, N487, and Q493. Consistent with the experimental binding data,[59] the computed heatmap clearly pointed to T485 and F486 positions as key binding energy hotspots (Figure 9C,D).

To summarize, the mutational sensitivity profiling of the SARS-CoV-2 S protein binding with CR3022, CB6, H014, S309, S2M11, and S2E12 demonstrated an excellent agreement with the deep mutational scanning and binding assay experiments. The diversity of binding energy hotspots in the SARS-CoV-2 S protein complexes showed a significant heterogeneity of mutational constraints and highlighted plasticity of the RBD regions since many RBD interacting sites can be tolerant of mutations, while only several hotspot centers are strongly constrained to the wild-type amino acid. These findings are particularly interesting in light of recent functional studies[60] showing that antibodies targeting the same RBD region may often have distinct escape mutations and that escape mutations target RBD positions that have significant deleterious effects on antibody binding but little negative impact on ACE2 binding and RBD folding. We propose that escape mutations constrained by the requirements for productive ACE2 binding and preservation of RBD stability may often select structurally adaptable regulatory sites to compromise antibody recognition allosterically through antibody-specific modulation of global motions and long-range interactions.

Perturbation Response Scanning Identifies Regulatory Hotspots of Allosteric Interactions in the SARS-CoV-2 S Complexes as Preferential Sites of Antibody-Escaping Mutations. Using the PRS method,[125−132] we quantified the allosteric effect of each residue in the SARS-CoV-2 complexes with a panel of studied antibodies in response to external perturbations. The effector profiles

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

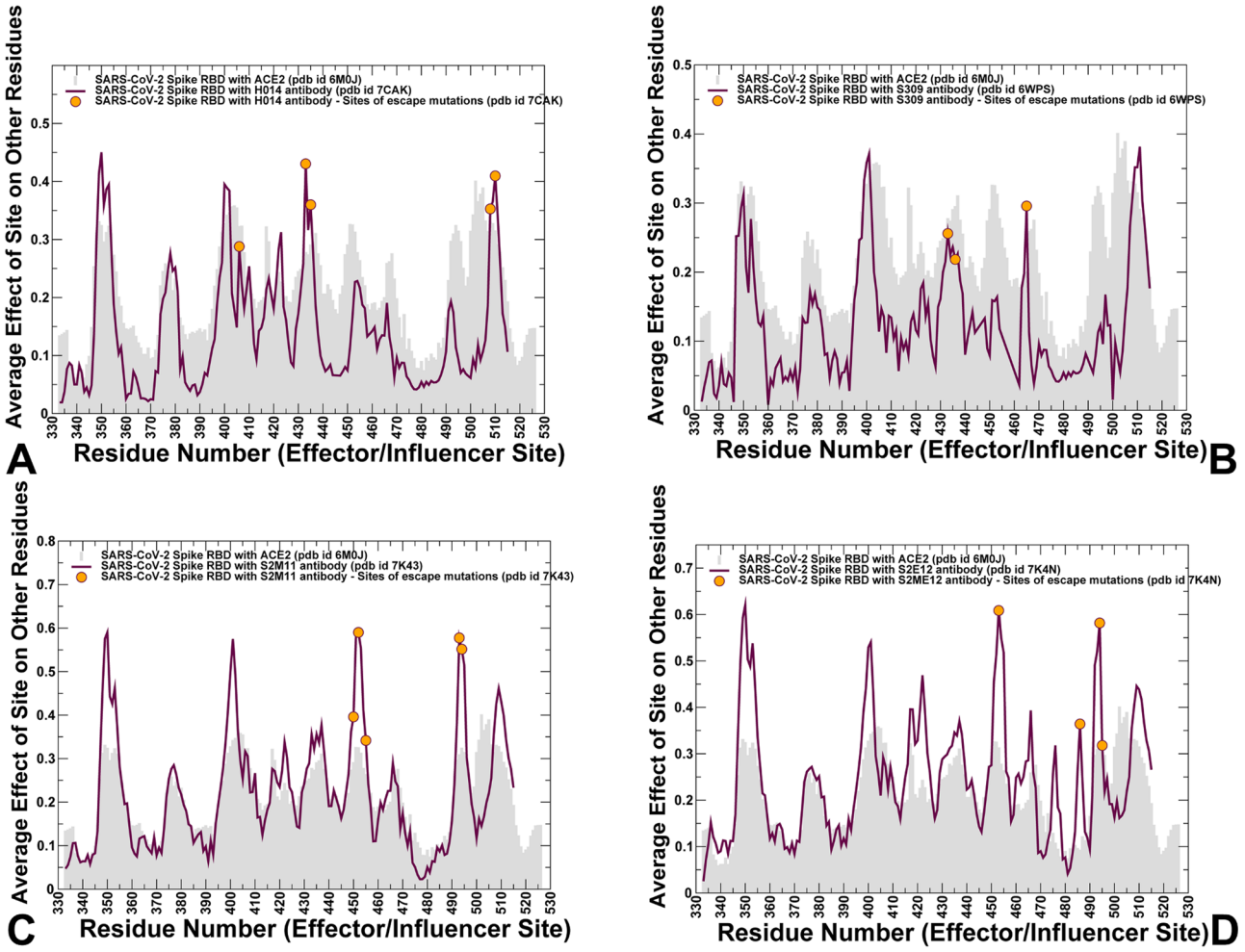


Figure 11. PRS effector profiles for the SARS-CoV-2 S trimer complexes with a panel of antibodies. The effector distributions are shown for the SARS-CoV-2 S complex with H014—three RBD in the open state (A), SARS-CoV-2 S complex with S309—three RBDs in the closed form (B), SARS-CoV-2 S complex with S2M11—three RBDs in the closed form (C), and SARS-CoV-2 S complex with S2E12—three RBDs in the closed form (D). The profiles are shown in maroon-colored lines. The sites of experimentally validated escape mutations are mapped onto the profiles and shown as orange-colored filled circles. The PRS effector profiles are superimposed with the PRS effector distribution for the S-RBD complex with ACE2 shown in light gray (PDB id 6M0J).

estimate the propensities of a given residue to influence allosteric 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 antibodies can incur specific and functionally relevant changes by modulating the effector profiles of the SARS-CoV-2 S protein complexes. We also propose an allosteric mechanism according to which antibody-escaping variations could preferentially target structurally adaptable regulatory centers of collective movements and allosteric communications in the SARS-CoV-2 S complexes. To validate this hypothesis, we first performed PRS analysis of the SARS-CoV-2 S-RBD complexes with CR3022 and CB6 antibodies (Figure 10) accompanied by a detailed comparison with the experimental data. In this analysis, we leveraged the results of deep mutational scanning of all-amino-acid mutations in the RBD that affect CR3022 and CB6 antibody binding,[50,60] providing a direct comparison between computational perturbation-based scanning and complete mapping of the RBD mutations that escape recognition of these two antibodies. To highlight the specific antibody effect, we aligned the PRS effector profiles for the SARS-CoV-2-RBD complexes with antibodies and the ACE2 host receptor. By mapping sites with the experimentally observed maximum escape onto the PRS effector profiles (Figure 10A,B), we found that these residues can be predominantly aligned with the effector peaks and correspond to the regulatory centers of long-range allosteric interactions in the complexes. According to the experimental data,[60] CR3022 binding can be escaped by mutations in the core RBD with the maximum escape site C361, V362, K378, V382, S383, F392, T430, I434, A435, and V512 residues corresponding to the effector centers of the PRS distribution (Figure 10A−C). Several of these escaping sites (C361, I434, A435, and V512) are not involved in direct intermolecular contacts with CR3022, and therefore mutational changes in these positions may be linked with long-range allosteric effects that are captured in the PRS effector profiles. The Pearson correlation coefficient ρ = 0.6 between the computed effector potential and the total escape parameters experimentally determined for sites evading CR3022 binding (Figure 10B) indicated a good agreement with the experimental data.

Our results are also consistent with the notion that the cryptic epitope targeted by CR3022 is important for stabilizing

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

the closed conformation of the prefusion trimer, reducing the likelihood that resistance mutations would occur in this region. The results also highlighted the antibody-specific effector response as compared to the S-RBD complex with ACE2 (Figure 10A−C).

The effector profile for the S-RBD complex with the CB6 antibody targeting ACE2-binding site similarly revealed a strong alignment between the distribution peaks and the experimentally determined sites of maximum escape (Figure 10D−F). Among these functionally important positions are D405, E406, T415, K417, D420, Y421, L455, F456, A475, Q493, and N501 residues. Only a small fraction of these sites corresponded to the binding energy hotspots (Y421, L455, and F456). It should be pointed out that mutational sensitivity maps for CB6 binding yielded only minor binding energy changes upon substitutions in D405, E406, K417, D420, A475, and N501 residues. Of particular interest are sites K417 and N501 implicated in global circulating mutations and the E406 residue, which is the unique position where mutations can escape the Food and Drug Administration (FDA)-approved combination of REGN10933 and REGN10987 antibodies.[50] The Pearson correlation coefficient ρ = 0.75 between the residue effector potential and the total escape for sites evading CB6 binding (Figure 10E) was fairly significant. This suggested that functionally important regulatory RBD centers revealed through the perturbation-based scanning approach can be often targeted by antibody-escaping variants. Our results support the hypothesis that the emergence of antibody-escaping mutations in the key effector positions may be attributed to their global regulatory role as control points of global stability and functional dynamics. In this model, modifications in these positions may not only weaken local interactions and binding but also compromise long-range allosteric couplings and coordinated movements of RBD to scan for optimal intermolecular arrangements with the antibody. Based on our findings, we suggest that escape mutations constrained by the requirements for ACE2 binding and preservation of RBD stability may often select structurally adaptable allosteric centers to compromise antibody recognition through modulation of allosteric interactions in the complex.

Using PRS analysis of the SARS-CoV-2 S trimer structures in the unbound forms and complexes with H014, S309, S2M11, and S2E12 antibodies, we examined how binding at different RBD epitopes may affect the stability and long-range couplings by modulating the allosteric effector potential of SARS-CoV-2 S residues (Figure S10, Supporting Information). Through a comparative analysis of the unbound and bound forms of the SARS-CoV-2 S trimer, we found that antibody binding could typically induce the increased effector potential of the RBD (residues 331−528) and CTD1 (residues 528−591) regions while modulating and reducing the allosteric potential of S2 regions (Figure S10, Supporting Information). The effector profiles also indicated the density and clustering of effector peaks distributed in UH (residues 736−781) and CH (residues 986−1035) regions.

To compare the results of perturbation-based scanning with the available experimental data, we focused on the effector distributions for the SARS-CoV-2-RBD regions (Figure 11). Although a general shape of the effector profile is fairly conserved across all complexes, the distributions reflected antibody-specific modulation of allosteric propensities for the RBD residues (Figure 11). In the SARS-CoV-2 complex with H014, the effector peak centers corresponded to RBD residues S383, F377, K378, Y380, GV382, S383, and P384 involved in direct interactions with the antibody (Figure 11A). The distribution peaks are aligned with sites of H014-escaping mutations including A435, N439, and Y508 residues.[147] Notably, A435 and N439 residues are not involved in interaction contacts with the antibody but nonetheless may develop antibody-resistant mutations.

A similar alignment between effector peaks and sites of escaping mutations was seen in the SARS-CoV-2 S complex with S309 (Figure 11B). 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. Indeed, S2M11 can induce the increased effector potential in the RBD regions, particularly residues F374, F377, K378 C379, and Y380 (Figure 11C). In addition, the effector peaks are closely aligned with sites of antibody-escaping mutations Y449, L452, L455, and S494, where only the Y449 position corresponds to the binding energy hotspot. 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 the allosteric effector potential in the S1 regions and activate the effector capacity of the conserved stretch of residues in the β-sheet of the RBD regions (Figure 11D).

The central result of the computational perturbation-based screening is that antibody-escaping variations could emerge in structurally adaptable regulatory centers of collective movements and allosteric communications in the SARS-CoV-2 S complexes. Our analysis suggested that these sites could act as effector hubs in which dynamic perturbations could have a measurable long-range effect on many residues. As a result, mutations in these positions could modulate allosteric communications rendering modifications that could potentially compromise antibody recognition through long-range effects. For instance, mutations in allosteric centers could modulate functional motions and curtail the ability of RBD to adopt an array of specific conformations, which may be required for productive antibody recognition. Hence, the SARS-CoV-2 S protein may function as a functionally adaptable regulatory machine that exploits allosteric effector centers to generate escape mutants in positions located outside of the antibody footprint. In support of this mechanism, the recent landscape analysis of the SARS-CoV-2 S escape variants identified spike mutations that attenuate antibody neutralization and revealed scenarios for both direct and indirect allosteric mechanisms of resistance where escape mutations emerge in positions distant from the antibody epitope through modulation of global conformational changes that shield or restrict antibody access to the spike protein.[149]

Network Modeling Links Sites of Antibody-Escaping Mutations with the Intercommunity Bridges Mediating Allosteric Communications in the SARS-CoV-2 S Complexes. We applied a network modeling approach in which 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 the protein

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

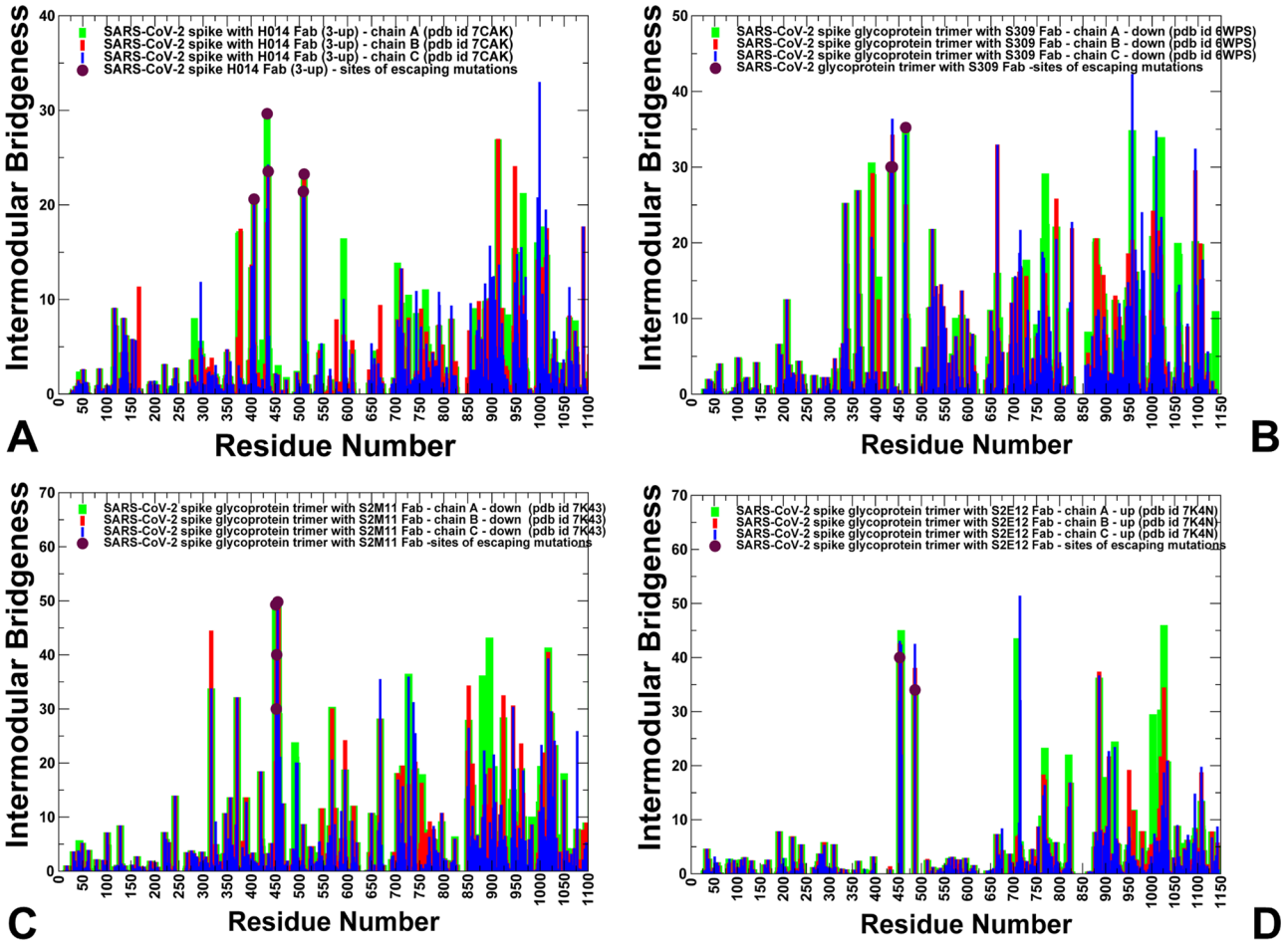


Figure 12. Intercommunity bridgeness profiles in the SARS-CoV-2 S complexes with a panel of antibodies. The profile is shown for the SARS-CoV-2 S complex with H014—three RBD in the fully open state (A), SARS-CoV-2 S complex with S309—three RBDs in the closed form (B), SARS-CoV-2 S complex with S2M11—three RBDs in the closed form (C), and SARS-CoV-2 S complex with S2E12—three RBDs in the closed form (D). The profiles for protomer chains (A−C) are shown in green, red, and blue bars, respectively. The spike protein sites targeted by antibody-escaping mutations are shown in maroon-colored filled circles.

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.[106−108] 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 and S8, Supporting Information) and S309 (Tables S9 and S10, Supporting Information). The results indicated that the majority of the intercluster communicating sites are localized in the RBD and especially CTD1 regions for SARS-CoV-2 S complexes with H014 (Tables S7 and S8, Supporting Information). The distribution of communicating positions in the SARS-CoV-2 S complexes with S309 (Tables S9 and 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 the 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 downregulated form of the S protein.

In the framework of the hierarchical approach, we also explored a more detailed model of the residue interaction networks using a graph-based representation with residues as network nodes and the inter-residue edges defined by both dynamic correlations[112] and coevolutionary residue couplings.[113] The residue interaction networks were divided into local interaction communities in which residues are densely interconnected, whereas residues from different communities may be weakly connected through the intermodular links. Based on the community decomposition of the SARS-CoV-2 S trimer complexes with antibodies, we computed the intermodular bridgeness profiles (Figure 12). Our previous network studies of protein systems suggested that this parameter could be helpful in identifying mediating centers of allosteric communications anchoring multiple communities.[150] Moreover, it was suggested that the intercommunity bridging sites may function as regulatory switch points that

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

determine the fidelity of allosteric interactions and signaling in diverse protein systems.[150]

We hypothesized that allosteric regulatory sites revealed by the perturbation-based residue scanning could be aligned with strategic positions in the residue interaction networks. The network analysis revealed a number of discrete sharp peaks of the intercommunity bridgeness that span in both S1 and S2 regions (Figure 12). By mapping positions corresponding to sites of antibody-escaping mutations, we noticed that these residues were often aligned with the distribution peaks in the RBD regions. In the SARS-CoV-2 S trimer complex with H014, the distribution peaks matched almost precisely with H014-escaping mutation site A435, N439, and Y508 residues (Figure 12A). Similarly, for other SARS-CoV-2 S complexes with S309, S2M11, and S2E12, the residues with high values of the intercommunity bridgeness metric corresponded to sites of antibody-escaping mutations (Figure 12B−D). Consistent with the PRS analysis, these results indicated that the regulatory control points could link together different local communities serving as "stepping stones" for allosteric communication paths in the system. This analysis provided an additional support to the notion that escaping mutations could select vulnerable allosteric hubs for targeting through amino acid modifications. In network terms, even moderate perturbations in these bridging positions could have a global effect on the intercommunity communications and fidelity of the interaction network, thereby compromising recognition with antibodies through long-range changes.

The ensemble-averaged distributions of the betweenness centrality were also computed for the SARS-CoV-2 S complexes with H014, S309, S2M11, and S2E12 (Figure S11, Supporting Information). 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).

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 S11, Supporting Information). 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 (residues 529−591), UH (residues 736−781), CH (residues 986−1035), and β-hairpin (BH) (residues 1035−1071) regions. The dominant clusters of centrality peaks located in the RBD and CTD1 regions of S1 and the CH region of S2 can be seen in the S complex with S2E12 (Figure S11, Supporting Information). 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 S12, Supporting Information). In the complexes with H014, the high-centrality sites are concentrated near CTD1 regions, which could strengthen couplings at the interdomain boundaries between S1 and S2 (Figure S12, Supporting Information). 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 showing that these antibodies may facilitate a broad allosteric interaction network between S1 and S2 functional regions.

CONCLUSIONS

This study examined molecular mechanisms underlying SARS-CoV-2 S protein binding with a panel of highly potent antibodies through the lens of coevolutionary relationships and ligand-induced modulation of allosteric interaction networks. Through coevolutionary analysis of the SARS-CoV-2 spike proteins, we identified highly coevolving hotspots and functional clusters forming coevolutionary networks that enable a functional cross-talk between distant allosteric regions in the SARS-CoV-2 spike complexes with antibodies. Coarse-grained and atomistic MD simulations combined with mutational sensitivity mapping and perturbation-based profiling of the SARS-CoV-2 S-RBD complexes with CR3022 and CB6 antibodies enabled a detailed validation of the proposed approach and quantitative comparison with the experimental deep mutagenesis scanning. Mutational sensitivity profiling of the SARS-CoV-2 S protein binding demonstrated a robust agreement with the deep mutational scanning and binding assay experiments. The diversity of binding energy hotspots in the SARS-CoV-2 S protein complexes highlighted the plasticity of the RBD regions since many RBD interacting sites can be tolerant of mutations, while only several hotspot centers are strongly constrained to the wild-type amino acid. By integrating insights from the perturbation-based allosteric profiling and network modeling of the SARS-CoV-2 S structures, we found that antibody-specific escape mutations may often select structurally adaptable regulatory sites to compromise antibody recognition allosterically through antibody-specific modulation of global motions and long-range interactions. The results of this study provide a novel insight into allosteric regulatory mechanisms of SARS-CoV-2 S proteins showing that antibodies 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.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.1c00395.

Top view of the binding epitopes of the SARS-CoV-2 S protein trimer complexes with a panel of antibodies used in this study (Figure S1); sequence and structural conservation of cysteine clusters in the SARS-CoV-2 spike prefusion and postfusion states (Figure S2); circular representation of the pairwise coevolutionary couplings between S protein residues in the S1 and S2 regions (Figures S3 and S4); structural analysis and mapping of coevolutionary hotspots in the SARS-CoV-2 S proteins (Figure S5); contact map for the cryo-EM structures of the SARS-CoV-2 S complexes with antibodies (Figure S6); distance maps for the cryo-EM structures of the SARS-CoV-2 S complexes (Figure S7);

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

collective dynamics and essential mobility profiles of the SARS-CoV-2 S trimer complexes with a panel of antibodies (Figure S8); mutational sensitivity analysis of the SARS-CoV-2 S trimer complex with S2M11 for key functional residues in the RBD region (Figure S9); PRS effector profiles for the unbound and antibody-bound SARS-CoV-2 S trimer complexes (Figure S10); residue-based betweenness centrality profiles in the SARS-CoV-2 S complexes with a panel of antibodies (Figure S11); structural maps of high-centrality clusters in the SARS-CoV-2 S complexes with H014, S309, S2M11, and S2E12 antibodies (Figure S12); participation coefficients for communicating residues in the SARS-CoV-2 S complexes (Tables S1−S6); intercluster communicating residues in the SARS-CoV-2 complexes with H014 (Tables S7 and S8); and intercluster communicating residues in the SARS-CoV-2 complexes with S309 (Tables S9 and S10) (PDF)

AUTHOR INFORMATION

Corresponding Author

Gennady M. Verkhivker − *Keck Center for Science and Engineering, Schmid College of Science and Technology, Chapman University, Orange, California 92866, United States; Department of Biomedical and Pharmaceutical Sciences, Chapman University School of Pharmacy, Irvine, California 92618, United States;* orcid.org/0000-0002-4507-4471; Phone: 714-516-4586; Email: verkhivk@chapman.edu; Fax: 714-532-6048

Author

Luisa Di Paola − *Unit of Chemical-Physics Fundamentals in Chemical Engineering, Department of Engineering, Università Campus Bio-Medico di Roma, 00128 Rome, Italy;* orcid.org/0000-0001-5329-8689

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.1c00395

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was partly supported by institutional funding from the 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; NTD, N-terminal 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) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **2020**, *181*, 271.e8−280.e8.

(2) Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **2020**, *395*, 565−574.

(3) 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*, No. 576622.

(4) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K. Y.; et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell* **2020**, *181*, 894.e9−904.e9.

(5) 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*, No. e00127-20.

(6) 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.

(7) 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.

(8) 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*, No. e00127-20.

(9) 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.

(10) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020**, *581*, 215−220.

(11) 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.

(12) Walls, A. C.; Tortorici, M. A.; Bosch, B. J.; Frenz, B.; Rottier, P. J. M.; DiMaio, F.; Rey, F. A.; Veesler, D. Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature* **2016**, *531*, 114−117.

(13) 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.

(14) 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.

(15) Walls, A. C.; Xiong, X.; Park, Y. J.; Tortorici, M. A.; Snijder, J.; Quispe, J.; Cameroni, E.; Gopal, R.; Dai, M.; Lanzavecchia, A.; et al. Unexpected receptor functional mimicry elucidates activation of coronavirus fusion. *Cell* **2019**, *176*, 1026.e15−1039.e15.

(16) Yuan, Y.; Cao, D.; Zhang, Y.; Ma, J.; Qi, J.; Wang, Q.; Lu, G.; Wu, Y.; Yan, J.; Shi, Y.; et al. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat. Commun.* **2017**, *8*, No. 15092.

(17) 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*, No. e1007236.

(18) 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*, No. 15701.

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

(20) 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.
(21) 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.
(22) 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.; et al. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **2020**, *369*, 1501−1505.
(23) Henderson, R.; Edwards, R. J.; Mansouri, K.; Janowska, K.; Stalls, V.; Gobeil, S. M. C.; Kopp, M.; Li, D.; Parks, R.; Hsu, A. L.; et al. Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nat. Struct. Mol. Biol.* **2020**, *27*, 925−933.
(24) McCallum, M.; Walls, A. C.; Bowen, J. E.; Corti, D.; Veesler, D. Structure-guided covalent stabilization of coronavirus spike glycoprotein trimers in the closed conformation. *Nat. Struct. Mol. Biol.* **2020**, *27*, 942−949.
(25) 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.; et al. A thermostable, closed SARS-CoV-2 spike protein trimer. *Nat. Struct. Mol. Biol.* **2020**, *27*, 934−941.
(26) 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.; et al. In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science* **2020**, *370*, 203−208.
(27) Lu, M.; Uchil, P. D.; Li, W.; Zheng, D.; Terry, D. S.; Gorman, J.; Shi, W.; Zhang, B.; Zhou, T.; Ding, S.; et al. Real-time conformational dynamics of SARS-CoV-2 spikes on virus particles. *Cell Host Microbe* **2020**, *28*, 880.e8−891.e8.
(28) 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**, *588*, 327−330.
(29) Starr, T. N.; Greaney, A. J.; Hilton, S. K.; Ellis, D.; Crawford, K. H. D.; Dingens, A. S.; Navarro, M. J.; Bowen, J. E.; Tortorici, M. A.; Walls, A. C.; et al. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **2020**, *182*, 1295.e20−1310.e20.
(30) Chan, K. K.; Dorosky, D.; Sharma, P.; Abbasi, S. A.; Dye, J. M.; Kranz, D. M.; Herbert, A. S.; Procko, E. Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. *Science* **2020**, *369*, 1261−1265.
(31) Yi, C.; Sun, X.; Ye, J.; Ding, L.; Liu, M.; Yang, Z.; Lu, X.; Zhang, Y.; Ma, L.; Gu, W.; et al. Key residues of the receptor binding motif in the spike protein of SARS-CoV-2 that interact with ACE2 and neutralizing antibodies. *Cell. Mol. Immunol.* **2020**, *17*, 621−630.
(32) Gur, M.; Taka, E.; Yilmaz, S. Z.; Kilinc, C.; Aktas, U.; Golcuk, M. Conformational transition of SARS-CoV-2 spike glycoprotein between its closed and open states. *J. Chem. Phys.* **2020**, *153*, No. 075101.
(33) 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.; et al. Developing a fully glycosylated full-length SARS-CoV-2 spike protein model in a viral membrane. *J. Phys. Chem. B* **2020**, *124*, 7128−7137.
(34) 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.; et al. Beyond shielding: The roles of glycans in the SARS-CoV-2 spike protein. *ACS Cent. Sci.* **2020**, *6*, 1722−1734.
(35) 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.
(36) 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.
(37) Ali, A.; Vijayan, R. Dynamics of the ACE2-SARS-CoV-2/SARS-CoV spike protein interface reveal unique mechanisms. *Sci. Rep.* **2020**, *10*, No. 14214.
(38) 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*, No. 497.
(39) Ghorbani, M.; Brooks, B. R.; Klauda, J. B. Critical Sequence Hotspots for Binding of Novel Coronavirus to Angiotensin Converter Enzyme as Evaluated by Molecular Simulations. *J. Phys. Chem. B* **2020**, *124*, 10034−10047.
(40) 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.
(41) 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*, No. 8268.
(42) Taka, E.; Yilmaz, S. Z.; Golcuk, M.; Kilinc, C.; Aktas, U.; Yildiz, A.; Gur, M. Critical Interactions between the SARS-CoV-2 Spike Glycoprotein and the Human ACE2 Receptor. *bioRxiv* **2020**, DOI: 10.1101/2020.09.21.305490.
(43) 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 Transduction Targeted Ther.* **2020**, *5*, No. 212.
(44) Gavor, E.; Choong, Y. K.; Er, S. Y.; Sivaraman, H.; Sivaraman, J. Structural Basis of SARS-CoV-2 and SARS-CoV Antibody Interactions. *Trends Immunol.* **2020**, *41*, 1006−1022.
(45) Barnes, C. O.; Jette, C. A.; Abernathy, M. E.; Dam, K. A.; Esswein, S. R.; Gristick, H. B.; Malyutin, A. G.; Sharaf, N. G.; Huey-Tubman, K. E.; Lee, Y. E.; et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **2020**, *588*, 682−687.
(46) Yuan, M.; Liu, H.; Wu, N. C.; Wilson, I. A. Recognition of the SARS-CoV-2 receptor binding domain by neutralizing antibodies. *Biochem. Biophys. Res. Commun.* **2021**, *538*, 192−203.
(47) 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.; et al. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* **2020**, *369*, 643−650.
(48) 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.
(49) 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.; et al. Neutralization of SARS-CoV-2 by destruction of the prefusion spike. *Cell Host Microbe* **2020**, *28*, 445.e6−454.e6.
(50) Starr, T. N.; Greaney, A. J.; Addetia, A.; Hannon, W. W.; Choudhary, M. C.; Dingens, A. S.; Li, J. Z.; Bloom, J. D. Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science* **2021**, *371*, 850−854.
(51) Wu, Y.; Wang, F.; Shen, C.; Peng, W.; Li, D.; Zhao, C.; Li, Z.; Li, S.; Bi, Y.; Yang, Y.; et al. A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2. *Science* **2020**, *368*, 1274−1278.
(52) Ju, B.; Zhang, Q.; Ge, J.; Wang, R.; Sun, J.; Ge, X.; Yu, J.; Shan, S.; Zhou, B.; Song, S.; et al. Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* **2020**, *584*, 115−119.
(53) Shi, R.; Shan, C.; Duan, X.; Chen, Z.; Liu, P.; Song, J.; Song, T.; Bi, X.; Han, C.; Wu, L.; et al. A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. *Nature* **2020**, *584*, 120−124.
(54) Yuan, M.; Liu, H.; Wu, N. C.; Lee, C. D.; Zhu, X.; Zhao, F.; Huang, D.; Yu, W.; Hua, Y.; Tien, H.; et al. Structural basis of a shared antibody response to SARS-CoV-2. *Science* **2020**, *369*, 1119−1123.
(55) Barnes, C. O.; West, A. P., Jr.; Huey-Tubman, K. E.; Hoffmann, M. A. G.; Sharaf, N. G.; Hoffman, P. R.; Koranda, N.; Gristick, H. B.; Gaebler, C.; Muecksch, F.; et al. Structures of Human Antibodies

Bound to SARS-CoV-2 Spike Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell* **2020**, *182*, 828.e16−842.e16.
(56) Cao, Y.; Su, B.; Guo, X.; Sun, W.; Deng, Y.; Bao, L.; Zhu, Q.; Zhang, X.; Zheng, Y.; Geng, C.; et al. Potent Neutralizing Antibodies against SARS-CoV-2 Identified by High-Throughput Single-Cell Sequencing of Convalescent Patients' B Cells. *Cell* **2020**, *182*, 73.e16−84.e16.
(57) Lv, Z.; Deng, Y. Q.; Ye, Q.; Cao, L.; Sun, C. Y.; Fan, C.; Huang, W.; Sun, S.; Sun, Y.; Zhu, L.; et al. Structural basis for neutralization of SARS-CoV-2 and SARS-CoV by a potent therapeutic antibody. *Science* **2020**, *369*, 1505−1509.
(58) Pinto, D.; Park, Y. J.; Beltramello, M.; Walls, A. C.; Tortorici, M. A.; Bianchi, S.; Jaconi, S.; Culap, K.; Zatta, F.; De Marco, A.; et al. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **2020**, *583*, 290−295.
(59) Tortorici, M. A.; Beltramello, M.; Lempp, F. A.; Pinto, D.; Dang, H. V.; Rosen, L. E.; McCallum, M.; Bowen, J.; Minola, A.; Jaconi, S.; et al. Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms. *Science* **2020**, *370*, 950−957.
(60) Greaney, A. J.; Starr, T. N.; Gilchuk, P.; Zost, S. J.; Binshtein, E.; Loes, A. N.; Hilton, S. K.; Huddleston, J.; Eguia, R.; Crawford, K. H. D.; et al. Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host Microbe* **2021**, *29*, 44.e9−57.e9.
(61) 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.
(62) 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.
(63) 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.
(64) Priya, P.; Shanker, A. Coevolutionary forces shaping the fitness of SARS-CoV-2 spike glycoprotein against human receptor ACE2. *Infect., Genet. Evol.* **2021**, *87*, No. 104646.
(65) 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.; et al. The Pfam protein families database in 2019. *Nucleic Acids Res.* **2019**, *47*, D427−D432.
(66) 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.; et al. Pfam: The protein families database in 2021. *Nucleic Acids Res.* **2021**, *49*, D412−D419.
(67) Tillier, E. R.; Lui, T. W. Using multiple interdependency to separate functional from phylogenetic correlations in protein alignments. *Bioinformatics* **2003**, *19*, 750−755.
(68) Dunn, S. D.; Wahl, L. M.; Gloor, G. B. Mutual information without the influence of phylogeny or entropy dramatically improves residue contact prediction. *Bioinformatics* **2008**, *24*, 333−340.
(69) Baker, F. N.; Porollo, A. CoeViz: a web-based tool for coevolution analysis of protein residues. *BMC Bioinf.* **2016**, *17*, No. 119.
(70) 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*, No. e1000978.
(71) 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.
(72) 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.
(73) Wu, C. H.; Apweiler, R.; Bairoch, A.; Natale, D. A.; Barker, W. C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; et al. The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res.* **2006**, *34*, D187−D191.
(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 Bioinf.* **2014**, *15*, No. 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., Nonlinear, Soft Matter Phys.* **2013**, *87*, No. 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*, No. e92721.
(80) Gueudré, 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*, No. 3496.
(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*, No. 606.
(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.; et al. 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) Martí-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sánchez, R.; Melo, F.; Šaliali, A. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291−325.
(93) Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinf.* **2016**, *54*, 5.6.1−5.6.37.

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

(94) Krivov, G. G.; Shapovalov, M. V.; Dunbrack, R. L., Jr. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* **2009**, *77*, 778−795.
(95) 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.
(96) Watanabe, Y.; Berndsen, Z. T.; Raghwani, J.; Seabright, G. E.; Allen, J. D.; Pybus, O. G.; McLellan, J. S.; Wilson, I. A.; Bowden, T. A.; Ward, A. B.; et al. Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nat. Commun.* **2020**, *11*, No. 2688.
(97) Watanabe, Y.; Allen, J. D.; Wrapp, D.; McLellan, J. S.; Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **2020**, *369*, 330−333.
(98) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(99) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; Mackerell, A. D., Jr. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain $\chi(1)$ and $\chi(2)$ dihedral angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257−3273.
(100) Di Pierro, M.; Elber, R.; Leimkuhler, B. A Stochastic Algorithm for the Isobaric-Isothermal Ensemble with Ewald Summations for All Long Range Forces. *J. Chem. Theory Comput.* **2015**, *11*, 5624−5637.
(101) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé−Hoover chains: the canonical ensemble via continuous dynamics. *J. Chem. Phys.* **1992**, *97*, 2635−2643.
(102) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **1994**, *101*, 4177−4189.
(103) 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.
(104) Dehouck, Y.; Gilis, D.; Rooman, M. A new generation of statistical potentials for proteins. *Biophys. J.* **2006**, *90*, 4010−4017.
(105) 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.
(106) 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.
(107) 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.
(108) Verkhivker, G. M.; Di Paola, L. Dynamic Network Modeling of Allosteric Interactions and Communication Pathways in the SARS-CoV-2 Spike Trimer Mutants: Differential Modulation of Conformational Landscapes and Signal Transmission via Cascades of Regulatory Switches. *J. Phys. Chem. B* **2021**, *125*, 850−873.
(109) Cumbo, F.; Paci, P.; Santoni, D.; Di Paola, L.; Giuliani, A. GIANT: A Cytoscape plugin for modular tetworks. *PLoS One* **2014**, *9*, No. e105001.
(110) Brinda, K. V.; Vishveshwara, S. A network representation of protein structures: Implications for protein stability. *Biophys. J.* **2005**, *89*, 4159−4170.
(111) Vijayabaskar, M. S.; Vishveshwara, S. Interaction energy based protein structure networks. *Biophys. J.* **2010**, *99*, 3704−3715.
(112) 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.
(113) 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*, No. e1005299.
(114) Czemeres, 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*, No. e0190267.
(115) 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*, No. e0143752.
(116) Floyd, R. W. Algorithm 97: Shortest path. *Commun. ACM* **1962**, *5*, 345.
(117) Hagberg, A. A.; Schult, D. A.; Swart, P. J. In *Exploring Network Structure, Dynamics, and Function Using NetworkX*, Proceedings of the 7th Python in Science Conference (SciPy2008), Pasadena, Varoquaux, G.; Vaught, T.; Millman, J., Eds., 2008; pp 11−15.
(118) Girvan, M.; Newman, M. E. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 7821−7826.
(119) Newman, M. E. J. Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2006**, *74*, No. 036104.
(120) Astl, L.; Verkhivker, G. M. Atomistic modeling of the ABL kinase regulation by allosteric modulators using structural perturbation analysis and community-based network reconstruction of allosteric communications. *J. Chem. Theory Comput.* **2019**, *15*, 3362−3380.
(121) Astl, L.; Verkhivker, G. M. Dynamic view of allosteric regulation in the Hsp70 chaperones by J-Domain cochaperone and post-translational modifications: Computational analysis of Hsp70 mechanisms by exploring conformational landscapes and residue interaction networks. *J. Chem. Inf. Model.* **2020**, *60*, 1614−1631.
(122) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498−2504.
(123) Kovács, I. A.; Palotai, R.; Szalay, M. S.; Csermely, P. Community landscapes: an integrative approach to determine overlapping network module hierarchy, identify key nodes and predict network dynamics. *PLoS One* **2010**, *5*, No. e12528.
(124) Szalay-Bekő, M.; Palotai, R.; Szappanos, B.; Kovacs, I. A.; Papp, B.; Csermely, P. ModuLand plug-in for Cytoscape: determination of hierarchical layers of overlapping network modules and community centrality. *Bioinformatics* **2012**, *28*, 2202−2204.
(125) Atilgan, C.; Atilgan, A. R. Perturbation-response scanning reveals ligand entry-exit mechanisms of ferric binding protein. *PLoS Comput. Biol.* **2009**, *5*, No. e1000544.
(126) 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.
(127) 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*, No. e1003624.
(128) 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.
(129) 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.
(130) Penkler, D. L.; Atilgan, C.; Tastan Bishop, O. Allosteric Modulation of Human Hsp90alpha Conformational Dynamics. *J. Chem. Inf. Model.* **2018**, *58*, 383−404.
(131) 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*, No. 6899.
(132) 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.

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619

(133) 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.
(134) Gussow, A. B.; Auslander, N.; Faure, G.; Wolf, Y. I.; Zhang, F.; Koonin, E. V. Genomic determinants of pathogenicity in SARS-CoV-2 and other human coronaviruses. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 15193−15199.
(135) Guruprasad, L. Evolutionary relationships and sequence-structure determinants in human SARS coronavirus-2 spike proteins for host receptor recognition. *Proteins* **2020**, *88*, 1387−1393.
(136) 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.
(137) 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.
(138) 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*, 265−271.
(139) 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.
(140) 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*, No. 3618.
(141) 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.
(142) 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.
(143) Koukos, P. I.; Glykos, N. M. Grcarma: A fully automated task-oriented interface for the analysis of molecular dynamics trajectories. *J. Comput. Chem.* **2013**, *34*, 2310−2312.
(144) Pietal, M. J.; Tuszynska, I.; Bujnicki, J. M. PROTMAP2D: visualization, comparison and analysis of 2D maps of protein structure. *Bioinformatics* **2007**, *23*, 1429−1430.
(145) 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.
(146) Yao, H.; Sun, Y.; Deng, Y. Q.; Wang, N.; Tan, Y.; Zhang, N. N.; Li, X. F.; Kong, C.; Xu, Y. P.; Chen, Q.; et al. Rational development of a human antibody cocktail that deploys multiple functions to confer Pan-SARS-CoVs protection. *Cell Res.* **2021**, *31*, 25−36.
(147) Li, Q.; Wu, J.; Nie, J.; Zhang, L.; Hao, H.; Liu, S.; Zhao, C.; Zhang, Q.; Liu, H.; Nie, L.; et al. The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell* **2020**, *182*, 1284.e9−1294.e9.
(148) Greaney, A. J.; Loes, A. N.; Crawford, K. H. D.; Starr, T. N.; Malone, K. D.; Chu, H. Y.; Bloom, J. D. Comprehensive mapping of mutations to the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human serum antibodies. *Cell Host Microbe* **2021**, *29*, 463−476.
(149) Liu, Z.; VanBlargan, L. A.; Bloyet, L. M.; Rothlauf, P. W.; Chen, R. E.; Stumpf, S.; Zhao, H.; Errico, J. M.; Theel, E. S.; Liebeskind, M. J.; et al. Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell Host Microbe* **2021**, *29*, 477.e4−488.e4.
(150) Astl, L.; Stetz, G.; Verkhivker, G. M. Allosteric Mechanism of the Hsp90 Chaperone Interactions with Cochaperones and Client Proteins by Modulating Communication Spines of Coupled Regulatory Switches: Integrative Atomistic Modeling of Hsp90 Signaling in Dynamic Interaction Networks. *J. Chem. Inf. Model.* **2020**, *60*, 3616−3631.

https://doi.org/10.1021/acs.jpcb.1c00395
J. Phys. Chem. B 2021, 125, 4596−4619