

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

Published as part of *The Journal of Physical Chemistry virtual special issue "Ruth Nussinov Festschrift"*.

Gennady M. Verkhivker* and Luisa Di Paola

Cite This: *J. Phys. Chem. B* 2021, 125, 850–873

Read Online

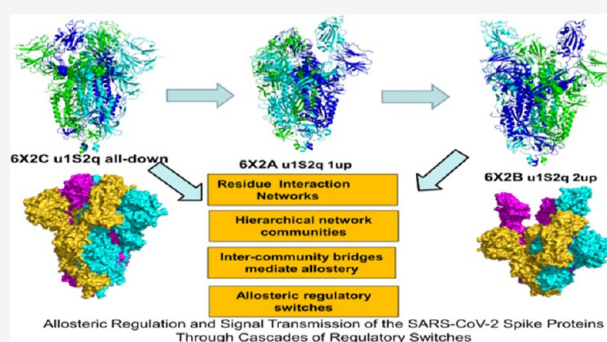
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: The rapidly growing body of structural and biochemical studies of the SARS-CoV-2 spike glycoprotein has revealed a variety of distinct functional states with radically different arrangements of the receptor-binding domain, highlighting a remarkable function-driven conformational plasticity and adaptability of the spike proteins. In this study, we examined molecular mechanisms underlying conformational and dynamic changes in the SARS-CoV-2 spike mutant trimers through the lens of dynamic analysis of allosteric interaction networks and atomistic modeling of signal transmission. Using an integrated approach that combined coarse-grained molecular simulations, protein stability analysis, and perturbation-based modeling of residue interaction networks, we examined how mutations in the regulatory regions of the SARS-CoV-2 spike protein can differentially affect dynamics and allosteric signaling in distinct functional states. The results of this study revealed key functional regions and regulatory centers that govern collective dynamics, allosteric interactions, and control signal transmission in the SARS-CoV-2 spike proteins. We found that the experimentally confirmed regulatory hotspots that dictate dynamic switching between conformational states of the SARS-CoV-2 spike protein correspond to the key hinge sites and global mediating centers of the allosteric interaction networks. The results of this study provide a novel insight into allosteric regulatory mechanisms of SARS-CoV-2 spike proteins showing that mutations at the key regulatory positions can differentially modulate distribution of states and determine topography of signal communication pathways operating through state-specific cascades of control switch points. This analysis provides a plausible strategy for allosteric probing of the conformational equilibrium and therapeutic intervention by targeting specific hotspots of allosteric interactions and communications in the SARS-CoV-2 spike proteins.



The results of this study revealed key functional regions and regulatory centers that govern collective dynamics, allosteric interactions, and control signal transmission in the SARS-CoV-2 spike proteins. We found that the experimentally confirmed regulatory hotspots that dictate dynamic switching between conformational states of the SARS-CoV-2 spike protein correspond to the key hinge sites and global mediating centers of the allosteric interaction networks. The results of this study provide a novel insight into allosteric regulatory mechanisms of SARS-CoV-2 spike proteins showing that mutations at the key regulatory positions can differentially modulate distribution of states and determine topography of signal communication pathways operating through state-specific cascades of control switch points. This analysis provides a plausible strategy for allosteric probing of the conformational equilibrium and therapeutic intervention by targeting specific hotspots of allosteric interactions and communications in the SARS-CoV-2 spike proteins.

INTRODUCTION

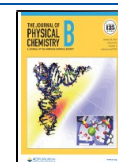
The coronavirus disease 2019 (COVID-19) pandemic associated with the severe acute respiratory syndrome (SARS) has emerged as a global international health crisis spread over the world with serious challenges for global economy, peace, and security.^{1,2} Genomic studies have shown that sequences of the novel coronavirus SARS-CoV-2 virus displayed high sequence similarity with SARS and Middle East respiratory syndrome (MERS) viral proteins involved in the replication cycle.^{3–6} SARS-CoV-2 infection is transmitted when one of the main structural proteins, the viral spike (S) glycoprotein,^{7–9} binds to the host cell receptor angiotensin-converting enzyme 2 (ACE2)^{10–12} followed by entry of the spike protein into host cells via membrane fusion. SARS spike (S) proteins are trimeric fusion proteins, with two main domains, an amino (N)-terminal S1 subunit and a carboxyl

(C)-terminal S2 subunit. The full-length SARS-CoV-2 S protein consists of two main domains, an amino (N)-terminal S1 subunit and a carboxyl (C)-terminal S2 subunit. The subunit S1 is involved in 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 of SARS coronavirus established that the key interactions between SARS-CoV/SARS-CoV-2 proteins and

Received: November 26, 2020

Revised: January 8, 2021

Published: January 15, 2021



the host receptor ACE2 enzyme are mediated by the receptor binding motif (RBM) of the RBD, suggesting that these interaction mechanisms may regulate both the cross-species and human-to-human infection transmissions.^{14,15} The S1 subunit undergoes spontaneous conformational transformations between ensembles of closed and open receptor-accessible forms, switching between a standing up RBD position amenable for ACE2 binding and a closed-down position eliciting immune evasion. The S2 subunit is highly evolutionary conserved fusion machinery that contains an N-terminal hydrophobic fusion peptide (FP), fusion peptide proximal region (FPPR), heptad repeat 1 (HR1), central helix region (CH), connector domain (CD), heptad repeat 2 (HR2), transmembrane domain (TM), and cytoplasmic tail (CT). The S1 domains are situated above the S2 subunit, covering and protecting the fusion apparatus. Upon proteolytic activation at the S1/S2 boundary and transition to the fully open form of the S trimer, the S1 subunit can release the dynamic breaks, promote dissociation from S2 followed by a cascade of massive structural rearrangements of the S2 subunit to mediate the fusion of the viral and cellular membranes and host cell takeover.^{13–16}

The cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins revealed that S protein trimers can exist in a dynamic equilibrium between a closed (“RBD-down”) prefusion conformation that undergoes large structural rearrangements and hinge-based functional motions to adopt a receptor-accessible open (“RBD-up”) state in which the S protein can fuse the viral membrane.^{17–26} The early studies showed that engineered coronavirus spikes through S-2P mutations (K986P and V987P in SARS-CoV-2) are resistant to conformational changes induced by receptor recognition and can prevent the transition from the prefusion to postfusion states.²⁷ These early studies of SARS-CoV S proteins showed that the majority of the S-2P proteins were in the single RBD-up conformation (58%), with double-up and triple-up S conformations populating a smaller fraction of the equilibrium (39% and 3% respectively), indicating that S-2P mutations can destabilize the closed form but still maintain stability of the prefusion trimer. The structures of the full-length S protein in the prefusion and postfusion conformations provided an unprecedented look and insight into tectonic rearrangements during transition to the postfusion state.^{28–30} Using mass spectrometry and single-particle EM, the molecular events associated with refolding of the prefusion S glycoprotein to the postfusion form were mapped, revealing a strong similarity between postfusion organization of coronaviruses S and a conserved mechanism of mediating membrane fusion entry of these viruses.²⁸ The cryo-EM structure of the postfusion SARS-CoV trimer showed dramatic conformational changes of the SARS-CoV S2 machinery during membrane fusion, where the HR1 and the CH helical segments, which are arranged into an antiparallel orientation in the prefusion state, can form a long continuous stem helix pointing toward the target membrane in the postfusion trimer.²⁹ The rapidly growing body of cryo-EM structures of the SARS-CoV-2 S protein in distinct functional conformations detailed structural changes in the trimer that are manifested in large movements of the RBD regions, highlighting a remarkable function-driven conformational plasticity and adaptability of the spike proteins.^{31–34} The recent cryo-EM characterization of the SARS-CoV-2 S trimer demonstrated that S protein may not only fluctuate between the dynamic open and closed states but also exhibit subtle

dynamic variability by adopting structurally rigid locked-closed form that revealed the ordered FPPR segment and cross-talk between stabilized RBD and FPPR regions.³⁴

Structural and biophysical studies employed protein engineering to generate prefusion-stabilized SARS-CoV-2 S variants by introducing disulfide bonds and proline mutations to modulate the stability of the S2 subunit and intersubunit boundaries and consequently prevent refolding changes that accompany acquisition of the postfusion spike state.³⁵ Cryo-EM structures of SARS-CoV-2 S variant with several proline substitutions (F817P, A892P, A899P, A942P) featured two distinct conformations—one with a single RBD in the up conformation and the other with two RBDs in the up conformation. These results suggested that modulating stability of S2 domains in the SARS-CoV-2 spike trimer may have a strong allosteric effect on dynamic couplings between S1 and S2 regions, altering the equilibrium and leading to a significant population of partially open and open states.³⁵ By combining targeted mutagenesis and cryo-EM structure determination, a recent biophysical study revealed a differential stabilization of the mutant constructs in the closed and open spike forms.³⁶ While modifications of contact regions between the RBD and S2 domains via S383C/D985C mutations leads to the thermodynamic prevalence of the closed-down conformation, the quadruple mutant (A570L/T572I/F855Y/N856I) introducing modifications at the interprotomer contacts between SD1 and S2 can shift the equilibrium away from the closed-down state, yielding a significant fraction of spike conformations in the open receptor-accessible form and showing the enhanced binding with ACE2 receptor.³⁶ This pioneering study indicated that the key interdomain contacts may act as potential global dynamic switches and constitute regulatory centers responsible for modulation of large scale changes and the equilibrium partitioning between the closed and open states. Protein engineering and cryo-EM studies of a prefusion-stabilized SARS-CoV-2 S ectodomain trimer using the interprotomer disulfide bonds (S383C/D985C, G413C/P987C, T385C/T415C) between RBD and S2 regions can lock the trimer in the closed state enhancing the SARS-CoV-2 S resistance to proteolysis.³⁷ Targeted design and structure determination of thermostable SARS-CoV-2 spike trimers unmasked the heterogeneity of the SARS-CoV-2 landscape showing that disulfide-bonded S-protein trimer variants imposing stabilization in strategically located interprotomer positions (S383-D385) and (G413-V987) can lead to a predominant population of the prefusion closed states with only ~20% population sampling the open state.³⁸ In contrast, proline substitutions K986P and V987P switch this equilibrium toward the receptor-accessible open form with only ~20% of the trimers in the closed state and ~80% of molecules adopting the open form.³⁸ SARS-CoV-2 S mutants with the enhanced infectivity profile^{39–41} have been a primary focus of the recent structural and biochemical studies, revealing that D614G mutation can radically shift the population of the SARS-CoV-2 S trimer toward open states.⁴² While in the native D614 form, SARS-CoV-2 S protein populates mostly the closed form (53% of the equilibrium) and single-up open states (47%), the introduction of D614G mutation yielded a markedly reduced 5% population of the closed conformation and widely open topology of the “up” protomers with 36% of the population adopting a single open protomer, 39% with two open protomers, and 20% with all three protomers in the open conformation.⁴²

The cryo-EM and tomography tools were employed to image the intact SARS-CoV-2 virions and determine the high-resolution structure, conformational flexibility and distribution of S trimers in situ on the virion surface.⁴³ These studies confirmed that the opening of the RBD observed in recombinant S trimers can also take place on the virus surface, and that the structure of SARS-CoV-2 S trimers on the virion surface is very similar to the closed prefusion form of the trimeric ectodomain stabilized by double proline mutations. Importantly, these studies confirmed that a general mechanism of spontaneous conformational changes and population shifts between different functional states can be maintained in different biological environments, reflecting the intrinsic properties of conformational landscapes for SARS-CoV-2 S trimers and suggesting that RBD epitopes become stochastically exposed to interactions with ACE2 and with antibodies.⁴³ Single-molecule fluorescence (Förster) resonance energy transfer (smFRET) studies of SARS-CoV-2 S trimer on virus particles revealed at least four distinct conformational states for the SARS-CoV-2 S trimer, also detailing a sequence of conformational transitions from the closed state to the receptor-bound open state through one necessary intermediate in which all three RBD domains in the closed conformations are oriented toward the viral particle membrane.⁴⁴ This study showed that the host receptor ACE2 activates spike proteins on the virus by shaping the conformational landscape toward stabilizing the “RBD-up” conformation, suggesting that mechanisms of conformational selection and receptor-induced structural adaptation can be intertwined and showing that allosteric stabilization of the closed S protein by antibodies can be effective strategy to achieve SARS-CoV-2 neutralization.⁴⁴ Cryo-EM structural studies also mapped a mechanism of conformational events associated with ACE2 binding, showing 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, while ACE2 binding to an open RBD can further accelerate transformation to a fully open and ACE2-bound form, priming the protein for fusion activation.⁴⁵

Glycosylation of the SARS-CoV-2 proteins and shielding of the receptor binding sites by glycans is a common mechanism of viral glycoproteins for evading the immune system and recent studies showed that glycosylation on SARS-CoV proteins can camouflage immunogenic protein epitopes.^{46–48} The site-specific analyses of N-linked glycosylation on soluble SARS, MERS, and SARS-CoV S glycoproteins revealed an extensive heterogeneity and identified hotspot clusters suggesting that N-linked glycan modifications of SARS-CoV S proteins may not constitute an effective shield and exhibit numerous vulnerabilities in comparison to the glycan shields of other viruses.⁴⁷ This study highlighted that boundaries between buried and exposed surface regions may often define potential sites of glycan shield vulnerability.⁴⁷ Site-specific mass spectrometric mapping of the glycan-processing states and cryo-EM analysis of the SARS-CoV-2 S protein showed that SARS-CoV-2 is highly glycosylated and established the composition and functional role of 22 N-linked glycosylation sites and 4 O-linked glycosylation sites of S protein that occlude distinct regions on the protein surface.⁴⁸ This study confirmed that the glycan shield of SARS-CoV-2 S proteins may exhibit various vulnerabilities in the RBD regions, particularly highlighting the role of proximal glycosylation sites N165, N234, and N343 in shielding of the receptor

binding sites especially when the receptor binding domain is in the closed-down conformation.⁴⁸ Using a combination of antigenic screens and high-resolution cryo-EM structure determination, it was shown that an N-glycan deletion at N234 results in a dramatically reduced population of the SARS-CoV-2 S trimer with “up” state RBDs, while glycan deletion at position N165 results in the increase of the open states.⁴⁹ This study demonstrated that these glycans can act not only as elements of the protective shield to antibody-mediated immunity but are also capable of modulating the conformational landscape of the SARS-CoV-2 S protein, thereby providing another layer of regulatory switch points to control shifts between different functional states.

Computational modeling and molecular dynamics (MD) simulations have been instrumental in predicting shapes and motions of glycans for glycoproteins.^{50–52} The developed structural models of glycoforms of the SARS-CoV-2 spike protein and MD simulations probed the effect of glycan heterogeneity on the antigenicity of the S glycoprotein showing that different glycan forms can shield ~40% of the protein surface on the S glycoprotein from antibody recognition, while the most accessible antigenic surface on the S protein and regions of vulnerabilities correspond to the ACE2 binding domain, where glycan shielding or mutational changes are unable to evade host immune response.^{53,54} The development of a fully glycosylated full-length SARS-CoV-2 S protein in a viral membrane enabled atomistic simulations of the SARS-CoV-2 S trimer structures in a properly glycosylated environment.⁵⁵ MD simulations of the SARS-CoV-2 S glycoprotein embedded in the viral membrane with a complete glycosylation profile reported a great level of details about conformational dynamics of the open and closed structures, suggesting that the shape and dynamics of the glycan shield can be coupled to allosteric conformational changes of the S protein and provide an additional layer of molecular control over spike response to the host receptor.⁵⁶ The development of a “bottom-up” coarse-grained model of the SARS-CoV-2 virion was recently presented in which cryo-EM, X-ray crystallography, and computational predictions were integrated to build a multiscale model of structural SARS-CoV-2 proteins and a complete virion model.^{57,58} Another large-scale MD study of a 4.1 million atom system containing a patch of viral membrane with four full-length, fully glycosylated S proteins was reported, where structural dynamics and the analysis of steric accessibility were used for epitope mapping, revealing signatures of the flexible glycan coat that can shield a larger surface area than can be derived from static structures.⁵⁹ 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.^{60,61} Our recent study examined the molecular basis of the SARS-CoV-2 binding with the ACE2 enzyme through the lens of coevolution and conformational dynamics that conspire to drive cooperative binding interactions and signal transmission.⁶² Using protein contact networks and perturbation response scanning based on elastic network models, we recently presented an integrated approach for detecting allosteric sites on the SARS-CoV-2 spike protein, showing convergence to a single allosteric region.⁶³ Molecular simulations and network modeling approaches were also used on our most recent investigation to present evidence that the SARS-CoV-2 spike protein can function as an allosteric

regulatory engine that fluctuates between dynamically distinct functional states.⁶⁴

Allosteric molecular events can involve complex cascades of thermodynamic and dynamic changes that occur on different spatial and temporal scales. The thermodynamic-centric energy landscape concepts and conformational selection models of allosteric regulation have gained a considerable prominence based on the notion that statistical ensembles of preexisting conformational states and communication pathways are intrinsic to a given protein and provide molecular machinery for allosteric modulation of the underlying landscapes induced by external perturbations, ligand binding, and mutations.^{65–69} These models can also link the allosteric mechanisms of proteins to their functional role in regulatory processes and signaling events.⁶⁹

In this study, we expanded our comparative analysis of mechanisms of allosteric signal transmission for the SARS-CoV-2 S states in the closed-a down form, partially open (1-up) and open (2-up) conformations determined for a double cysteine mutant S383C/D985C (rS 2d, RBD to S2 double mutant) quadruple mutant A570L T572I F855Y N856I (u1S2q, subdomain 1 to S2 quadruple mutant).³⁶ Using an integrated approach that combined coarse-grained (CG) simulations, protein stability analysis, and multifaceted dynamic modeling of residue interaction networks, we examined how mutations can differentially affect dynamics, stability and allosteric signaling in distinct functional states if the SARS-CoV-2 S protein. A detailed atomistic network analysis and modeling of allosteric communications in the SARS-CoV-2 S trimer mutants revealed key regulatory centers and ensemble of pathways that mediate allosteric interactions and dictate conformational changes, thereby governing mechanisms of signaling and binding. The results of this study provide a novel insight into the allosteric regulatory mechanisms of SARS-CoV-2 S proteins showing that mutations at the key regulatory positions can differentially modulate distribution of states and shape up signal transmission pathways using cascades of control switch points. This analysis provides a plausible strategy for therapeutic intervention and allosterically driven modulation of the conformational equilibrium by targeting specific hotspots of allosteric interactions and communications in the SARS-CoV-2 proteins.

MATERIALS AND METHODS

Coarse-Grained Simulations and Elastic Models.

Coarse-grained (CG) models are computationally effective approaches that leverage topology-based framework for characterization of protein structure and dynamics, enabling simulations of large systems over long time scales. In this study, CG-CABS model was used for simulations of the cryo-EM structures of the SARS-CoV-2 spike trimer mutants in distinct conformational states. In this high-resolution model, the amino acid residues are represented by α , $C\beta$, the center of mass of side chains and another pseudoatom placed in the center of the α – α pseudobond. The sampling scheme of the CABS model used in our study is based on Monte Carlo replica-exchange dynamics and involves a sequence of local moves of individual amino acids in the protein structure as well as moves of small fragments.^{70–72} The solvent effect is incorporated implicitly through statistical potentials of the CABS force-field. We employed CABS-flex approach that efficiently combines a high-resolution coarse-grained model and efficient search

protocol capable of accurately reproducing all-atom MD simulation trajectories and dynamic profiles of large biomolecules on a long time scale.^{70–74}

CABS-flex standalone package dynamics implemented as a Python 2.7 object-oriented package was used for fast simulations of protein structures.^{74–76} This approach also allowed for MODELER-based reconstruction of generated models and simulation trajectories to all-atom representation. The default settings were used in which soft native-like restraints are imposed only on pairs of residues fulfilling the following conditions: the distance between their C^α atoms was smaller than 8 Å, and both residues belong to the same secondary structure elements.^{75,76} The loop regions are completely flexible in CABS-CG simulations and regions of secondary structure are restrained only locally by applying soft distance restraints to residues if their distance deviates from the distance in the crystal structure by more than 1 Å.^{75,76}

A total of 1000 independent CG-CABS simulations were performed for each of the studied systems. In each simulation, the total number of cycles was set to 10 000 and the number of cycles between trajectory frames was 100. The structures of the SARS-CoV-2 spike (S) protein trimer were used in simulations and modeling (Table 1, Figure 1): the cryo-EM structure of

Table 1. Structures of SARS-CoV2 Spike Protein Mutants Examined in This Study

pdb code	description	RBDs position
6X29	rS 2d double mutant S383C D985C	down
6X2A	u1S2q quadruple mutant A 507 L T572I F855Y N856I	1 up
6X2B	u1S2q quadruple mutant A 507 L T572I F855Y N856I	2 up
6X2C	u1S2q quadruple mutant A 507 L T572I F855Y N856I	down

the full-length S protein mutant rS 2d S383C/D985C in the closed prefusion conformation (pdb id 6X29),³⁶ the cryo-EM structure of the full-length S u1S2q quadruple mutant (A507L T572I F855Y N856I) in the all-down closed prefusion conformation (pdb id 6X29),³⁶ the cryo-EM structure of the full-length S u1S2q quadruple mutant (A507L T572I F855Y N856I) in the partially open (1-up) conformation (pdb id 6X2A),³⁶ and the cryo-EM structure of the full-length S u1S2q quadruple mutant (A507L T572I F855Y N856I) in the open (2-up) conformation (pdb id 6X2B).³⁶

All structures were obtained from the Protein Data Bank.^{77,78} Structure preparation process included checking and correcting for errors, adding missing residues and atoms, filling valences with hydrogens. First, protein residues in all studied structures were systematically inspected for missing residues and protons. Hydrogen atoms and missing residues were initially added and assigned according to the WHATIF program web interface.^{79,80} The structures were then preprocessed through the Protein Preparation Wizard (Schrodinger, LLC, New York, NY) and included a systematic checking of bond order, assignment, and adjustment of ionization states, checking of disulfide bonds, removal of crystallographic water molecules and cofactors, capping of the termini, assignment of partial charges, and addition of missing atoms and side chains that were not assigned in the initial processing with the WHATIF program. The cryo-EM structures of the SARS-CoV-2 S proteins contained a number

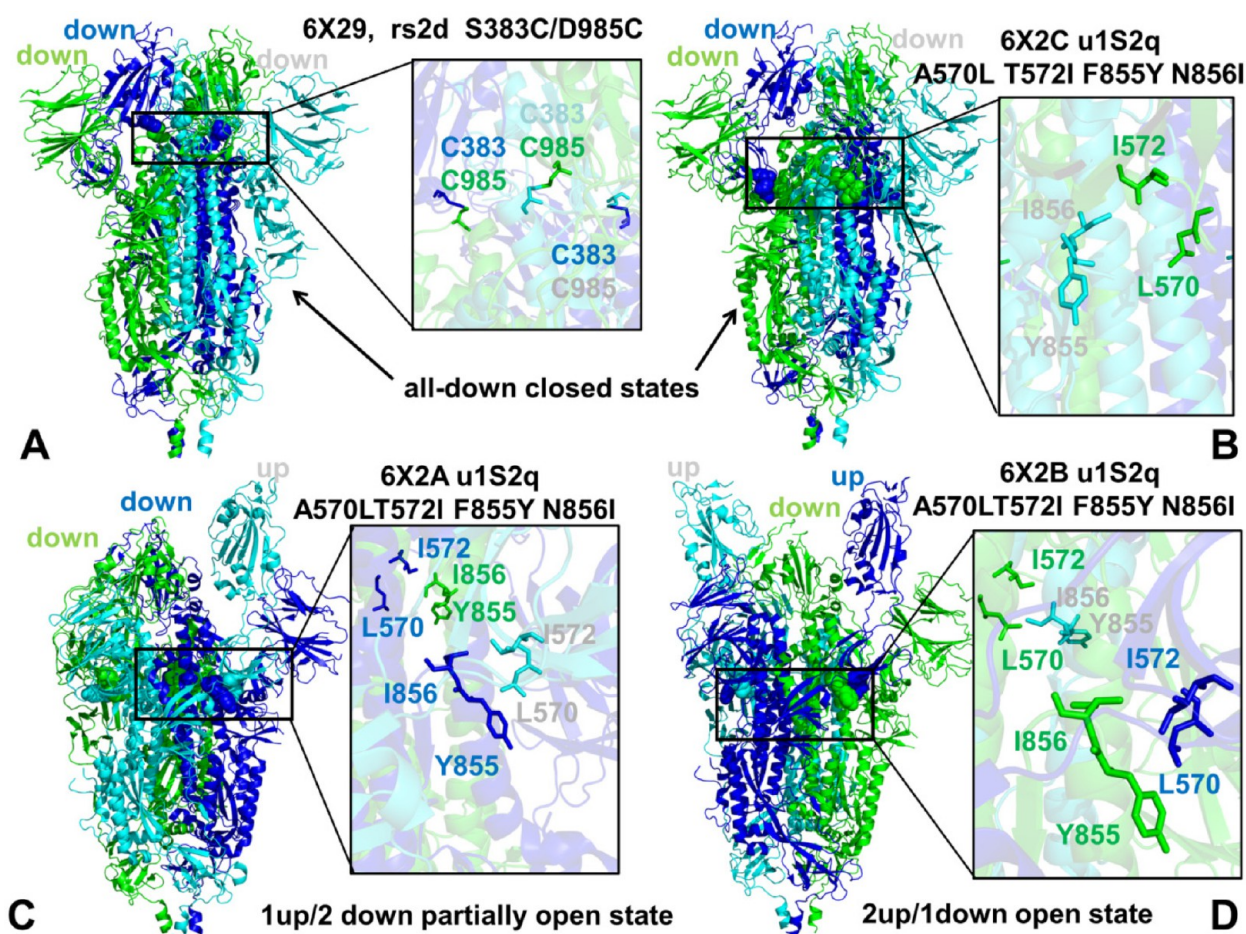


Figure 1. Cryo-EM structures of the SARS-CoV-2 spike (S) protein trimer mutants used in this study. (A) Cryo-EM structure of the full-length S mutant trimer rS 2d S383C/D985C in the closed prefusion conformation (pdb id 6X29).³⁶ The structure is shown in ribbons with protomers colored in green, cyan and blue, respectively. The down positions of the RBDs are annotated. A close-up of the mutational site S383C/D985C at the RBD to S2 boundaries. C383 and C985 are shown in sticks and colored according to their respective protomer. (B) Cryo-EM structure of the full-length S u1S2q quadruple mutant (A507L T572I F855Y N856I) in the all-down closed prefusion conformation (pdb id 6X2C).³⁶ The protomers are colored as in part A. A close-up of the mutational site at the boundary between subdomain 1 to S2 is shown with L570, I572, Y855, and I856 residues at the interprotomer interfaces. The residues are in sticks and colored according to their protomer color scheme. (C) Cryo-EM structure of the full-length S u1S2q quadruple mutant (A507L T572I F855Y N856I) in the partially open (1-up) prefusion conformation (pdb id 6X2A).³⁶ (D) The cryo-EM structure of the full-length S u1S2q quadruple mutant (A507L T572I F855Y N856I) in the open (2-up) conformation (pdb id 6X2B).³⁶ A close-up of the mutational site at the boundary between subdomain 1 to S2 is shown as in part B.

of missing loops of various lengths that were dispersed throughout the structure and required a rigorous template-based loop modeling and reconstruction. The modeled missing loops were located in the NTD regions (residues 16–26, 70–79, 144–164, 173–185, 246–262), RBD regions (residues 445–446, 469–488) CTD2 regions (residues 621–640), the cleavage site at the S1/S2 boundary (residues 677-QTNSPGSVSSVA-688 or 677-QTNSPRRARSVA-688), and FPPR region (residues 828–853). The furin cleavage site emerges as a disordered loop (residues 655-GICASYHTVSLLRST-669 in the SPIKE_SARS sequence or residues 669-GICASYQTQT-NSPRRARSVA-688 in SPIKE_SARS2 sequence). Some of the disordered residues are also located C-terminal to the S2 cleavage site at R815 forming a conserved region 816-SFIEDLLFNKVTLADAGF-831 that plays an important role for membrane fusion. The missing loops in the RBD regions (residues 445–446, 469–488) were reconstructed by template-based modeling using the cryo-EM structure of human ACE2 in the presence of the neutral amino acid transporter B⁰AT1 complexed with the SARS-CoV-2 RBD

protein (pdb id 6M17) that includes the complete RBD region as a template. The missing loops in the studied cryo-EM structures of the SARS-CoV-2 S protein were reconstructed and optimized using template-based loop prediction approaches ModLoop⁸¹ and ArchPRED server⁸² and further confirmed by the FALC (fragment assembly and loop closure) program.⁸³ The side chain rotamers were refined and optimized by the SCWRL4 tool.⁸⁴

The CG conformational ensembles were also subjected to all-atom reconstruction using CABS-flex standalone package dynamics and checked using PULCHRA method⁸⁵ and CG2AA tool⁸⁶ to produce atomistic models of simulation trajectories. The protein structures were then optimized using atomic-level energy minimization with a composite physics and knowledge-based force fields as implemented in the 3Drefine method.⁸⁷

In all-atom MD simulations, the glycans are typically added to the protein structure and a complete model of the glycoprotein is then subjected to simulations.^{56–59} In our study, we employed CABS-CG model for dynamics

simulations followed by all-atom reconstruction of trajectories. In addition, the reconstructed atomistic structures from simulation trajectories were decorated by *N*-acetyl glycosamine (NAG) glycan residues and optimized. The structure of glycans at particular glycosites of the closed and open states of SARS-CoV-2 S protein was previously determined and these glycans were incorporated in atomistic modeling of the SARS-CoV-2 mutant structures. The glycosylated microenvironment for atomistic models of the simulation trajectories was mimicked by 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, and 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 open state (pdb id 6VYB),³¹ and the cryo-EM structure SARS-CoV-2 spike trimer (K986P/V987P) in the open state (pdb id 6VSB).³² The glycan-decorated atomistic models were subsequently used for the ensemble-based structural, energetic and network analyses of the SARS-CoV-2 mutants examined in this work.

We performed principal component analysis (PCA) of reconstructed trajectories derived from CABS-CG simulations using the CARMA package⁸⁸ and also determined the essential dynamics profiles in slow modes using elastic network models (ENM) analysis.⁸⁹ Previous studies indicated that CABS-CG MC simulations and ENM computations can provide similar dynamic profiles that closely reproduce the experimental dynamic data for a wide range of protein structures.^{71,72} Two elastic network models: Gaussian network model (GNM)^{90,91} and anisotropic network model (ANM) approaches⁹² were adopted to compute the amplitudes of isotropic thermal motions and directionality of anisotropic motions. The functional dynamics analysis was conducted using the GNM in which protein structure is reduced to a network of *N* residue nodes identified by C_α atoms, and the fluctuations of each node are assumed to be isotropic and Gaussian. The topology of the protein structure is described by $N \times N$ Kirchhoff matrix of inter-residue contacts Γ , where the off-diagonal elements are -1 , if the nodes are within a cutoff distance r_c , and zero otherwise. In the GNM approach^{90,91} the interaction potential for a protein of *N* residues was:

$$V_{\text{GNM}} = -\frac{\gamma}{2} \left[\sum_{i=1}^{N-1} \sum_{j=i+1}^N (R_{ij} - R_{ij}^0) \cdot (R_{ij} - R_{ij}^0) \Gamma_{ij} \right] \quad (1)$$

where R_{ij}^0 and R_{ij} were the equilibrium and instantaneous distance between residues *i* and *j*, and Γ was the $N \times N$ Kirchhoff matrix:

$$\Gamma_{ij} = \begin{cases} -1 & i \neq j, R_{ij} \leq r_c \\ 0 & i \neq j, R_{ij} > r_c \\ -\sum_{i,i \neq j} \Gamma_{ij} & i = j \end{cases} \quad (2)$$

Thus, the square fluctuations were:

$$\langle \Delta R_i \cdot \Delta R_j \rangle = (3kT/\gamma) \cdot [\Gamma^{-1}]_{ij} \quad (3)$$

The normal modes were extracted by eigenvalue decomposition of the matrix $\Gamma = U\Lambda U^T$, *U* being the orthogonal matrix

whose k^{th} column u_k was the k^{th} mode eigenvector. Λ was the diagonal matrix of eigenvalues λ_k .

In ANM approach⁹² the interaction potential for a protein of *N* residues was:

$$V_{\text{ANM}} = \frac{\gamma}{2} \sum_{i,j}^N (|R_{ij}| - |R_{ij}^0|)^2 \quad (4)$$

The $3N \times 3N$ Hessian matrix, *H*, determines the systems dynamics.

$$H_{ij} = \begin{bmatrix} \frac{\partial^2 V}{\partial X_i \partial X_j} & \frac{\partial^2 V}{\partial X_i \partial Y_j} & \frac{\partial^2 V}{\partial X_i \partial Z_j} \\ \frac{\partial^2 V}{\partial Y_i \partial X_j} & \frac{\partial^2 V}{\partial Y_i \partial Y_j} & \frac{\partial^2 V}{\partial Y_i \partial Z_j} \\ \frac{\partial^2 V}{\partial Z_i \partial X_j} & \frac{\partial^2 V}{\partial Z_i \partial Y_j} & \frac{\partial^2 V}{\partial Z_i \partial Z_j} \end{bmatrix} \quad (5)$$

where X_i , Y_i , and Z_i represented the Cartesian components of the *i*th residue and *V* was the potential energy of the system. We selected $r_c = 13$ Å. ANMs provide information about the amplitudes, as well as the direction of residue fluctuations. Conformational mobility profiles in the essential space of low frequency modes were obtained using DynOmics server⁹¹ and ANM server.⁹²

Protein Stability 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. A BeAtMuSiC approach was employed that is based on statistical potentials describing the pairwise inter-residue distances, backbone torsion angles and solvent accessibilities, and considers the effect of the mutation on the strength of the interactions at the interface and on the overall stability of the complex.^{93–95} 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.

Protein Contact Networks and Network Clustering. The protein contact network is a network whose nodes are the protein residues and links are active contacts between residues in the protein structure. The protein contact network is an undirected, unweighted graph; it is built on the basis of the distance matrix **d**, whose generic element d_{ij} records the Euclidean distance between the *i*th and the *j*th residue (measured between the corresponding α carbons). A detailed description of network construction and significance of network descriptors is presented in our previous studies.^{96–98} The active network links are defined using a range of contacts between 4 and 8 Å. The description of the network is given by the following adjacency matrix:

$$A_{ij} = \begin{cases} 1 & \text{if } 4 \text{ \AA} < d_{ij} < 8 \text{ \AA} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

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

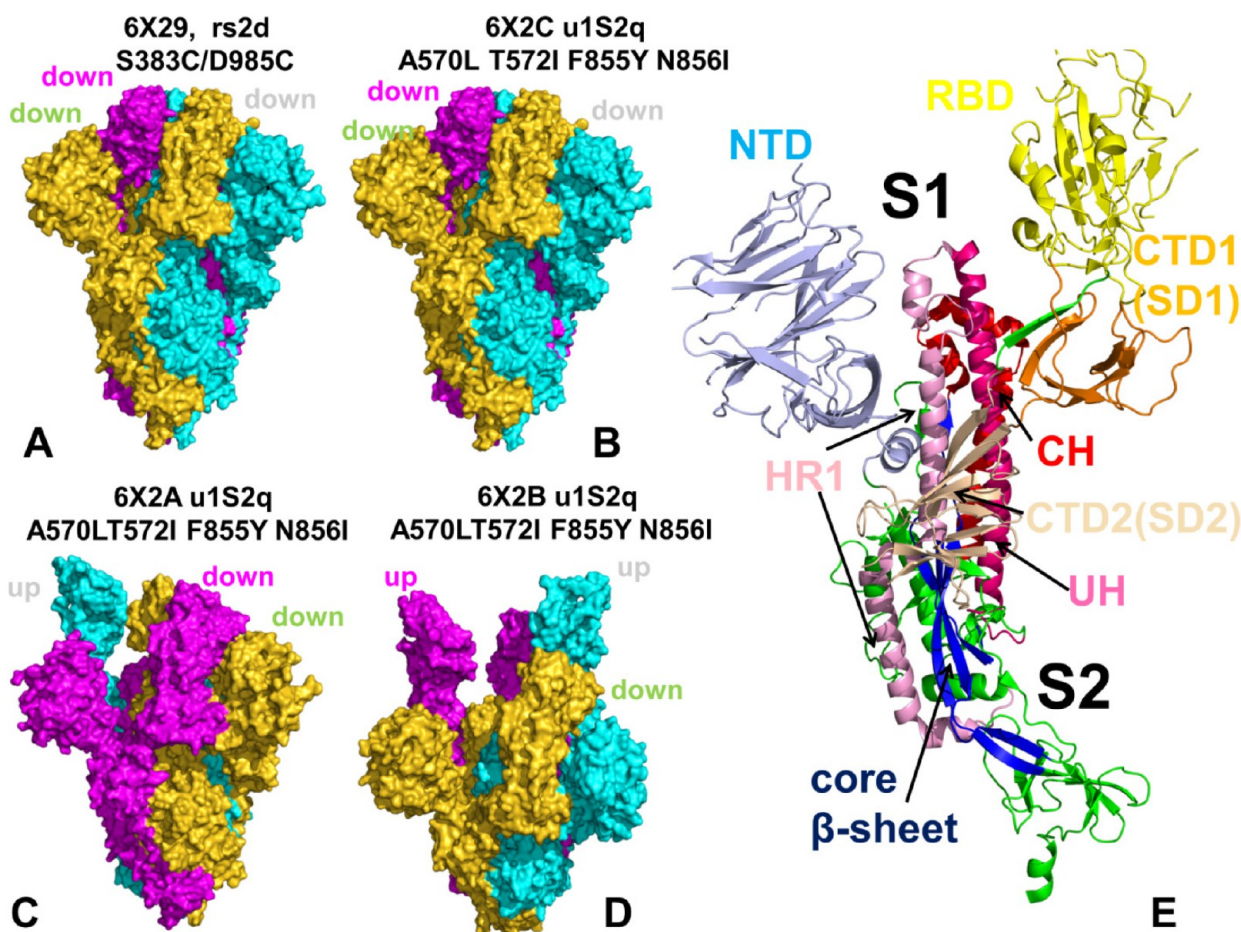


Figure 2. Cryo-EM structures of the SARS-CoV-2 spike (S) protein trimer mutants and domain organization of the SARS-CoV-2 S protein. The cryo-EM structure of the full-length S mutant trimer rS 2d S383C/D985C in the closed prefusion conformation (A), u1S2q quadruple mutant (A507L T572I F855Y N856I) in the closed conformation (B), u1S2q quadruple mutant in the partially open (1-up) conformation (C), and u1S2q quadruple mutant in the open (2-up) conformation (D). The structures are shown in surface representation with protomers in green, cyan and magenta. The down and up positions of RBDs are annotated. (E) Schematic organization of the SARS-CoV-2 S protein is shown by highlighting major domains in a single protomer. The subunit S1 regions are annotated as follows: N-terminal domain (NTD) (residues 14–306) in light blue; receptor-binding domain (RBD) (residues 331–528) in yellow; C-terminal domain 1 (CTD1) (residues 528–591) in orange; C-terminal domain 2 (CTD2) (residues 592–686) in wheat color; upstream helix (UH) (residues 736–781) in red; heptad repeat 1 (HR1) (residues 910–985) in pink; central helix (CH) (residues 986–1035) in hot pink; antiparallel core β -sheet (β_{46} – β_{49} – β_{50}) (in blue) is assembled from an N-terminal β -strand (β_{46}) (residues 711–736) preceding the UH region and a C-terminal β -hairpin (β_{49} – β_{50}) (residues 1045–1076) located downstream of the CH region.

$$k_i = \sum_j A_{ij} \quad (7)$$

$$P_i = 1 - \left(\frac{k_i}{k_{si}} \right)^2 \quad (9)$$

We previously demonstrated that spectral network clustering targets functional modules in proteins.⁸⁸ The network clustering is based on the spectral decomposition of the network Laplacian L defined as

$$L = D - A \quad (8)$$

where D is the degree matrix, a diagonal matrix whose diagonal is the degree vector, and A is the adjacency matrix. The eigenvector corresponding to the second minor eigenvalue of the Laplacian (Fiedler's vector) assigns nodes (residues) to two different clusters on the basis of the sign of its corresponding components. The method is then binary and hierarchical.

Once the network is divided into a given number of clusters (powers of 2), we define the participation coefficient, defined as

where k_i is the overall node degree, while k_{si} is the node degree including only links with nodes (residues) that belong to their own cluster. The participation coefficient P describes the propensity of residue nodes to participate into intercluster communication. Functional regions in proteins have been identified on the basis of values of non null participation coefficients, pointing to an active role of corresponding residues in signal transmission between different clusters (functional domains). We highlighted as highly active communication residues the nodes with $P > 0.75$, using our previous studies showing that such residues typically correspond to important regulatory centers of signal transmission between protein domains.⁹⁶ The proposed methodology of network clustering was implemented as a Cytoscape plugin.⁹⁹

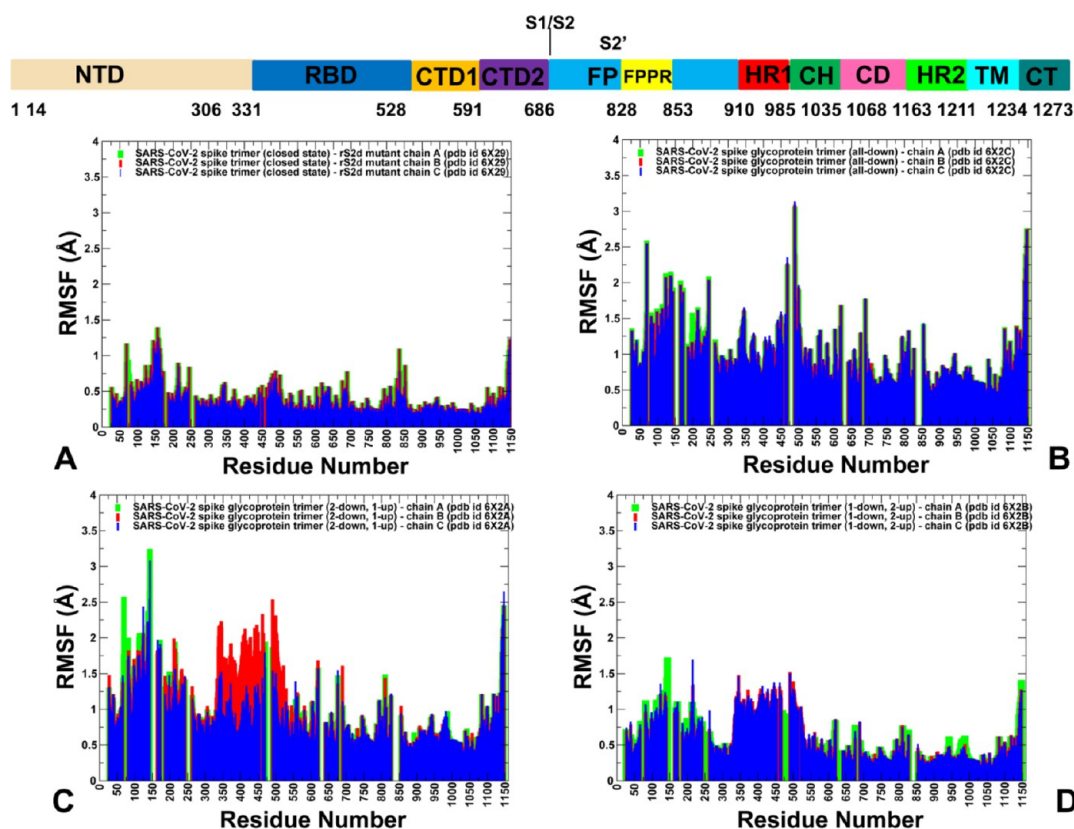


Figure 3. CABS-GG conformational dynamics of the SARS-CoV-2 spike (S) protein trimer mutants. A schematic representation of domain organization and residue range for the full-length SARS-CoV-2 spike (S) protein is shown above conformational dynamics profiles. (A) Root-mean-square fluctuations (RMSF) profiles obtained from CABS-CG simulations of the cryo-EM structure of the full-length S mutant trimer rS 2d S383C/D985C in the closed prefusion conformation. (B) RMSF profile for the u1S2q quadruple mutant (A507L T572I F855Y N856I) in the closed conformation. (C) RMSF profile for the u1S2q quadruple mutant in the partially open (1-up) conformation. (D) RMSF profile for the u1S2q quadruple mutant in the open (2-up) conformation. The profiles for protomer chains A, B, and C are shown in green, red, and blue bars, respectively.

Dynamic-Based Modeling of Residue Interaction Network and Community Analysis. A graph-based representation of protein structures^{100–102} is used to represent residues as network nodes and the *inter-residue edges* to describe residue interactions. The details of graph construction using residue interaction cutoff strength (I_{\min}) were outlined in our previous studies.^{103–105} We constructed the residue interaction networks using both dynamic correlations¹⁰² and coevolutionary residue couplings¹⁰⁵ that provide complementary descriptions of the allosteric interactions and yield robust network signatures of long-range couplings and communications. The ensemble of shortest paths is determined from the matrix of communication distances by the Floyd–Warshall algorithm.¹⁰⁶ Network graph calculations were performed using the python package NetworkX.¹⁰⁷

Using the constructed protein structure networks, we computed the residue-based betweenness parameter. The short path betweenness centrality of residue i is defined to be the sum of the fraction of shortest paths between all pairs of residues that pass through residue i :

$$C_b(n_i) = \sum_{j < k} \frac{g_{jk}(i)}{g_{jk}} \quad (10)$$

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 betweenness centrality metric is also computed by evaluating the average shortest path length (ASPL) change by systematically removing individual nodes.^{108,109} A Z-score is then computed for each node based on the change of ASPL compared to the initial one.

The Girvan–Newman algorithm^{110,111} is used to identify local communities. To characterize global bridges from a community structure, we employed the intercommunity bridgeness metric.^{112–114} The algorithmic details were specified in our recent studies.^{115,116} The bridgeness parameter uses community detection as input:

$$G(i) = \sum_{j \in J} l_{ij} \delta_{ij} \quad (11)$$

Here the sum is over communities J (different from the community of node i , denoted as I), δ_{ij} is equal 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 . It is measured by the inverse of the number of links between them. Nodes that are only linked to nodes of their own community have $G(i) = 0$, while nodes involved in bridging two (or more) communities have a positive value of the index. All parameters were computed using the python package NetworkX.¹⁰⁷ The network parameters were also evaluated by Cytoscape package for network analysis.¹¹⁷ The ModuLand program within the Cytoscape platform^{118,119} was adapted to determine a hierarchical network structure and

centralities of the identified communities in the dynamic residue interaction networks.

RESULTS AND DISCUSSION

Coarse-Grained Simulations of the SARS-CoV-2 Spike Trimers: Mutation-Induced Differential Stabilization and Unique Dynamic Signatures of Functional States.

We employed multiple CABS-CG simulations followed by atomistic reconstruction and refinement to provide a detailed comparative analysis of dynamic landscapes that are characteristic of the major functional states of the SARS-CoV-2 S trimer. Atomistic structures of the SARS-CoV-2 S prefusion trimer mutants in multiple functional states were selected for this analysis.³⁶ These mutant variants of SARS-CoV-2 spike protein were originally chosen in structural studies to probe the dynamic equilibrium between functional states by targeting contact regions between the RBD and S2, the RBD and NTD, the SD1 and S2, and the SD2 and S2 (Figures 1, 2). Simulations were performed for the closed form of a double cysteine mutant, S383C/D985C (RBD to S2 double mutant (rS 2d), pdb id 6X29), that introduced a covalent link between RBD and S2 regions at the C-terminus of HRR1 and was shown to stabilize the “all-down” conformation (Figure 2A). In addition we conducted CABS-CG simulations for the closed, partially open (1 RBD-up) and open forms (2 RBDs up) of the SARS-CoV-2 S SD1 to S2 quadruple mutant (u1S2q) (A570L T572I F855Y N856I) (Figure 2B–D). While all-atom MD simulations with the explicit inclusion of the glycosylation shield could provide a rigorous assessment of conformational landscape of the SARS-CoV-2 S proteins, such simulations remain technically challenging and computationally demanding due to the enormous size of a complete SARS-CoV-2 S system embedded onto the membrane.^{56–59} The abundance and complexity of information in all-atom simulations of glycosylated S protein may also obscure the networking signatures and global modulation effect of the glycan shield⁵⁹ on the protein dynamics, energetics and particularly long-range allosteric couplings examined in our study.

We opted to combine the efficiency and reproducibility of simplified CABS-CG simulations with more detailed energetic and network analysis of all-atom reconstructed trajectories that also incorporated the structurally resolved glycan environment in the key glycosite positions. The proposed approach was adopted to characterize the global dynamic and network signatures of the SARS-CoV-2 S structures that are generally robust to perturbations of the glycan microenvironment and driven mainly by the underlying topology and modularity of the S structures. We also used the proposed approach to identify hinge regions and key regulatory switch centers that can impose a primary control over conformational changes between distinct functional states of the SARS-CoV-2 S structures. These global dynamic and network features of the S protein landscapes are assumed to be largely determined by the underlying topology of the SARS-CoV-2 S structures and can be captured through a proposed combination of CABS-CG dynamics and subsequent atomistic refinement of trajectories with structurally resolved glycans.

Using CABS-CG simulations, we examined how mutations in the interdomain contacts could affect the global dynamic profiles of the closed, partially open, and open states revealing the important regions of flexibility (Figure 3).

We also explored connections between the intrinsic dynamics of the SARS-CoV-2 S protein and their propensity

to access up and down RBD states. A comparative analysis of the dynamics profiles showed an appreciable and consistent difference between a considerably more flexible S1 subunit and generally structurally rigid S2 subunit in all functional states (Figure 3). The S2 subunit exhibited very small thermal fluctuations with RMSF < 1.0 Å, which is consistent with a strong sequence conservation in these regions.^{120–122} Indeed, the S2 domain displays considerably higher conservation as compared to the S1 domains, particularly in the FP region (residues 812–829), the HR1 (residues 910–985), and the CH regions (residues 986–1035). The CTD1 (residues 529–591) (Figure 2E) is a structural relay between RBD and S2 regions that can arguably sense the functional movements in the S1 and S2 domains and communicate signal from and to the fusion peptide.³⁴ These segments of the S1 subunit are largely stable in all functional states of the S trimer (Figure 3A–D).

The differences in mobility of the closed-down forms of the S trimer could be clearly seen from the RMSF profile of the S383C/D985C (RBD to S2 double mutant (rS 2d) (Figure 3A) and another “all-down” spike form of the quadruple mutant (u1S2q) (A570L T572I F855Y N856I) (Figure 3B). Strikingly, a significant stabilization of the S383C/D985C closed trimer can be contrasted with the mobility of the closed form of the quadruple mutant that may be triggered by rigid body shifts of SD1 domain and displacement of the A570L + T572I-containing loop from the native position. These findings are consistent with the experimental structural data suggesting that mutation-induced structural changes could trigger the increased flexibility in both the RBD and NTD regions.³⁶ Conformational dynamics of the quadruple mutant in the partially open form with 1 up RBD (Figure 3C) was similar to that in the closed mutant state (Figure 3B), showing an appreciable degree of mobility in the S1 subunit regions and a fairly sharp contrast with structurally stable S2 regions. In particular, S2 subdomains HR1, CH, and CD showed considerable stability during simulations. As expected, the mobility of the RBD-up protomer was greater than that of the other “down” RBDs (Figure 3B,C). Both closed and partially open states featured larger thermal fluctuations in the NTD and RBD regions, particularly of the exposed RBM segment involved in recognition of the host receptor. Interestingly, even in a fully closed conformation of the SARS-CoV-2 S quadruple mutant with all RBDs in the “closed-down” position, the flexibility of the RBDs can be maintained (Figure 3B). These results are consistent with our recent studies suggesting that closed and open prefusion forms of the SARS-CoV-2 spike trimer could become more dynamic than the “locked-closed” trimer conformations in which mobility of the NTD and RBD regions are largely suppressed.⁶⁴

Strikingly, conformational dynamics of the closed state for the S383C/D985C double mutant (Figure 3A) was very similar to the dynamics of the open quadruple mutant trimer with 2 up RBDs (Figure 3D). Both structures showed relatively small fluctuations and structural stabilization of both S1 and S2 domains with RMSF < 1.0 Å for most of the protein residues. Our results are consistent with structural studies showing that quadruple mutant (u1S2q) (A570L T572I F855Y N856I) targeting the interdomain positions may alter the dynamic preferences of the SARS-CoV-2 S trimer by destabilizing the closed state while enhancing stability of the open form.³⁶ In general, the results showed that both closed and open forms of the S trimer can maintain flexibility of the S1 subunit required

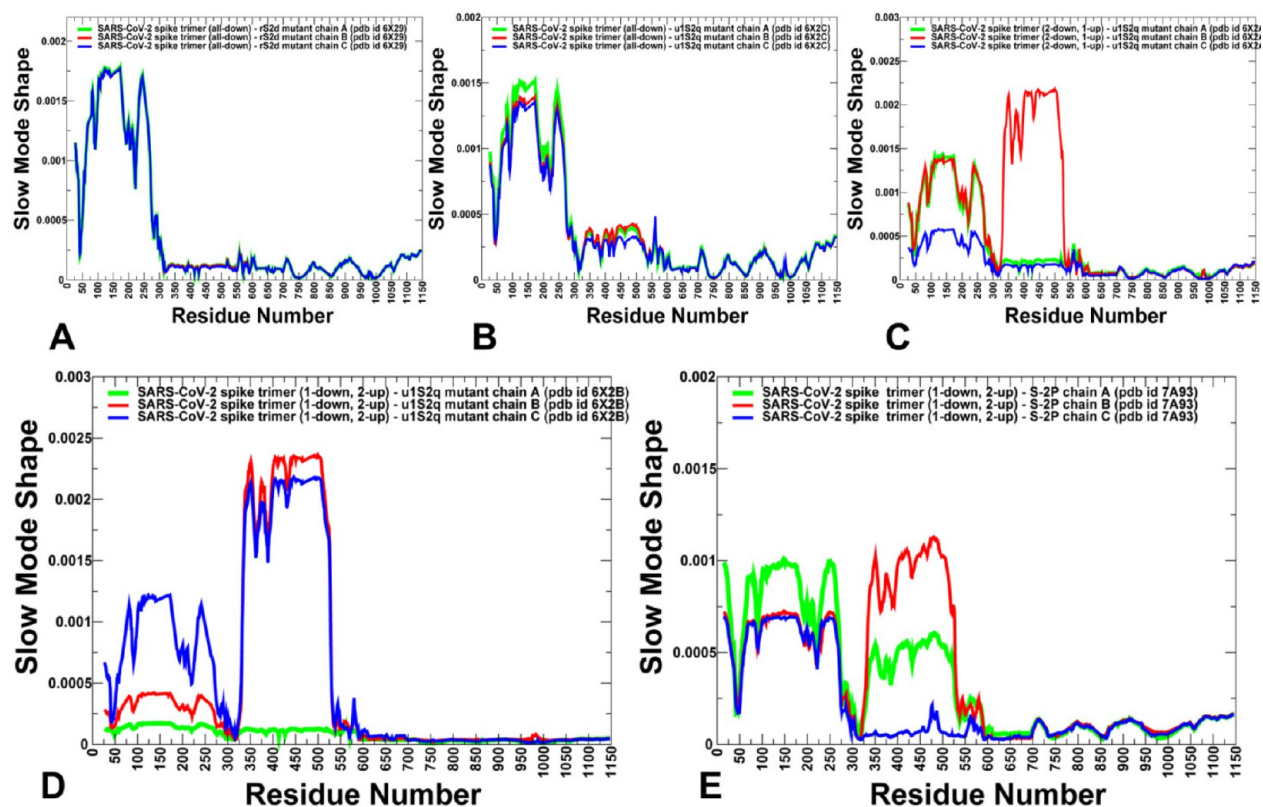


Figure 4. Collective dynamics of the SARS-CoV-2 spike mutant trimers in the prefusion form. The mean square displacements in functional motions are averaged over the three lowest frequency modes. (A) Essential mobility profiles for the structure of the S mutant rS 2d S383C/D985C (pdb id 6X29).³⁶ (B) Slow mode profiles for the cryo-EM structure of the u1S2q quadruple mutant (A507L T572I F855Y N856I) in the all-down closed conformation (pdb id 6X2C).³⁶ (C) Slow mode profiles for the cryo-EM structure of the full-length u1S2q quadruple mutant in the partially open (1-up) conformation (pdb id 6X2A).³⁶ (D) The slow mode profiles for the structure of u1S2q quadruple mutant in the open (2-up) conformation (pdb id 6X2B).³⁶ (E) Essential mobility profiles for the structure of S-2P (K986P/V987P) mutant in the open (2-up) conformation (pdb id 7A93).⁴⁵ The profiles for protomer chains A, B and C are shown in green, red, and blue lines, respectively.

for dynamic transformations of RBDs. We found that mutations of the interdomain residues can change the dynamic profiles and stability of the distinct functional states.

The central finding of this analysis is that quadruple mutation (u1S2q) may reverse the dynamic balance between closed and open states and promote stabilization of the receptor-accessible form of the SARS-CoV-2 S protein. We argue that binding of SARS-CoV-2 S quadruple mutant with ACE2 receptor may further exploit these thermodynamic preferences, triggering transition to the fully open state of spike protein and virus entry to the host receptor.

Hinge Regions in the SARS-CoV-2 Spike Structures Aligned with Mutational Sites That Drive Dynamic Shifts between Functional Forms. Here we examined the conjecture that SARS-CoV-2 S mutants can modulate the intrinsic equilibrium of the S protein by exploiting preexisting soft modes of collective motions defined by the protein topology. To characterize collective motions and determine the distribution of hinge regions in the SARS-CoV-2 S trimers, we performed PCA of trajectories derived from CABS-CG simulations and also determined the essential slow modes using ENM analysis.^{89–92} The essential profiles for the slowest modes and reported the functional dynamics profiles were averaged over the first three major low frequency modes (Figure 4). The conserved hinge regions that can regulate the interdomain movements between RBD and NTD are localized at F518 and V539 positions (Figure 3A–C). These hinges are

responsible for the coordination of the functional movements of the NTD and RBD with respect to the S2 subunit and are shared by all states. The second hinge position V539 is localized in the CTD1 region (residues 529–591) and could coordinate the concerted dynamic changes between RBD and FP regions. Importantly, several additional hinge sites in the closed and partially open forms were precisely aligned with the sites of mutations. These hinge positions correspond to residues S383, K386, D428, T547, L570, I572, S591, S750, Y855, I856, T962, and D985 (Figure 4A–C). Notably, the slow mode shape for the rS 2d S mutant showed that RBD regions (residues 331–528) may be largely restricted in collective motions as these residues corresponded to minima along the essential mobility profile (Figure 4A). This reflected the enhanced stabilization of the RBD regions induced by the introduction of the disulfide bond S383C/D985C at the critical juncture between RBD and S2 domains. In contrast, the RBD regions are more prone to functional movements in the closed form of the u1S2q S mutant (Figure 4B) as the slow mode profile revealed that RBD residues are now aligned with local maxima and correspond to moving regions. The hinge centers in the closed state of the u1S2q S mutant corresponded to small regulatory clusters near F318, V539, S70–S72, and 750–755 residues (Figure 4B). It is worth noting that L570, I572, Y855, I856, and S591 hinges are situated near the interdomain SD1-S2 interfaces that act as regulatory switch

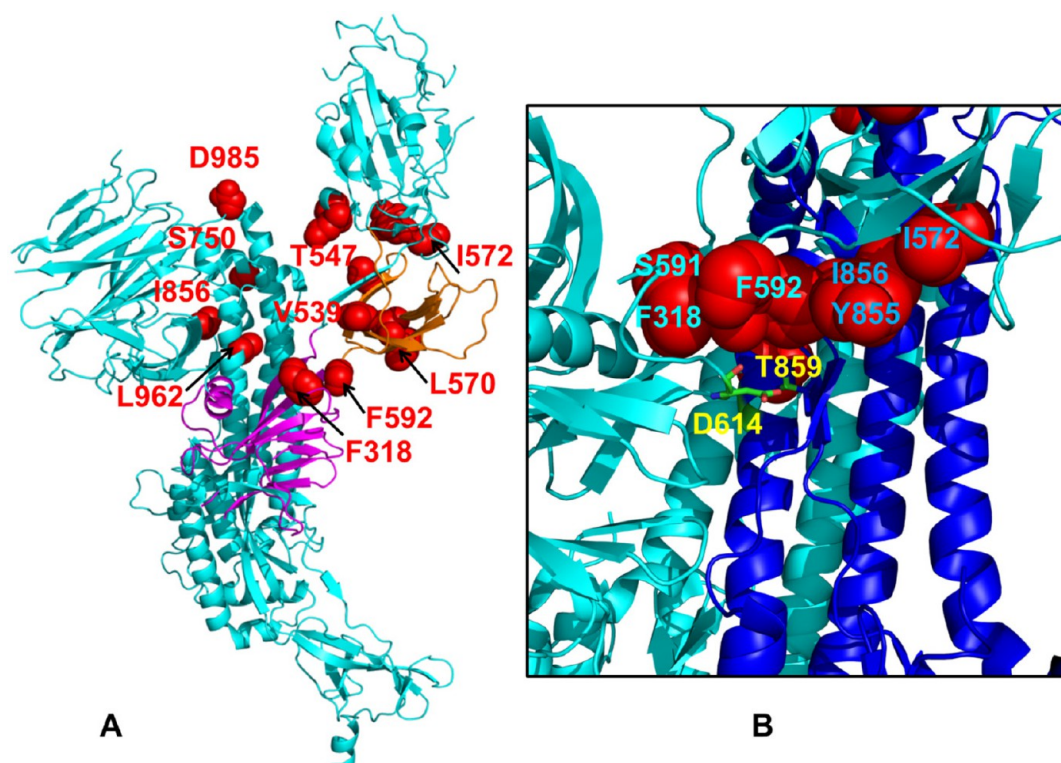


Figure 5. Structural maps of the hinge regions in the SARS-CoV-2 spike trimer mutants. (A) The major hinge sites (shown in red spheres) are mapped onto a closed-down protomer of the u1S2q S mutant in the closed state (pdb id 6X2C). These hinge centers include F318, V539, L570, I572, T547, F592, S750, I856 and L962 residues. Notably, these regulatory points are located near the interprotomer and interdomain contact interfaces and corresponded to mutational sites in the u1S2q S trimer. (B) A close-up map of hinge sites near the interprotomer interface is shown with F318, S591 and F592 on one protomer and I572, Y855, an I856 on the adjacent protomer. The key hydrogen bonding between D614 (protomer A) and T859 (protomer B) is shown with D614 and T859 in atom-colored sticks. The close proximity of the salt bridge to the network of hinge centers can be seen.

centers governing the population shifts between closed and open forms.³⁶

According to our results, hinge sites in the closed (Figure 4B) and partially open conformations (Figure 4C) of the S quadruple mutant can form a fairly broad network of hinge centers, indicating that these functional states are prone to collective movements and coordinated conformational changes facilitating transitions to the open states. In some contrast, the collective dynamics profile of the open form with 2 up RBDs was different, featuring a largely immobilized S2 subunit in the slow modes (Figure 4D). The movements of the NTDs were also partly repressed, while the RBDs corresponded to the dominantly moving regions in the slow modes. We also noticed that functional movements of the NTD regions can be progressively reduced going from the closed to a partially open and then open states of the S quadruple mutant (Figure 4B–D). At the same time, the collective movements of RBDs become more dominant in the functional dynamics of the S quadruple mutant when adopting an open form with 2 up RBDs (Figure 4D).

A direct comparison of ENM-based slow mode profiles for the quadruple mutant (u1S2q) (A570L T572I F855Y N856I) (Figure 4D) and S-2P (K986P/V987P) mutant (Figure 4E) showed that conformational plasticity of S protein and minor topological changes could lead to different collective dynamics signatures. In the S-2P mutant, we observed noticeable deviations in slow mode motions for the “standing-up” RBDs showing reduced mobility and similar amplitudes of displacements for both NTD and RBD regions. The S2 subunit in the

S-2P mutant can feature moderate slow mode motions around local hinge sites (Figure 4E), while S2 is completely immobilized in the quadruple mutant (Figure 4D). Importantly, the collective dynamics in the open S quadruple mutant trimer with 2 up RBDs can be characterized by large movements of S1 subunit and RBDs around the largely immobilized S2 regions. These functional motions can promote access and binding with the host receptor and transition to the fully open state with all RBDs in the active position. On the basis of these findings, we argue that the quadruple mutant could also prime the S trimer for efficient binding with ACE2 and initiation of the membrane fusion and transition to the postfusion form.

Structural mapping of the slow mode profiles illustrated that structurally immobilized region of functional motions in the quadruple trimer corresponded to the core of the S2 subunit, while the NTDs and RBDs in both forms are prone to large conformational changes (Figure S1, Supporting Information). In the closed mutant form, there is greater propensity for movements of NTDs while the RBDs are aligned with dominant functional movements in the open quadruple form (Figure S1, Supporting Information). Of interest are distinct structural maps of the essential profiles in the open state of the quadruple mutant and the S-2P mutant, featuring the increasing propensities for collective movements in the RBD regions of the quadruple mutant (Figure S1, Supporting Information).

By mapping major hinge centers onto the protomer conformations, we observed that regulatory centers of

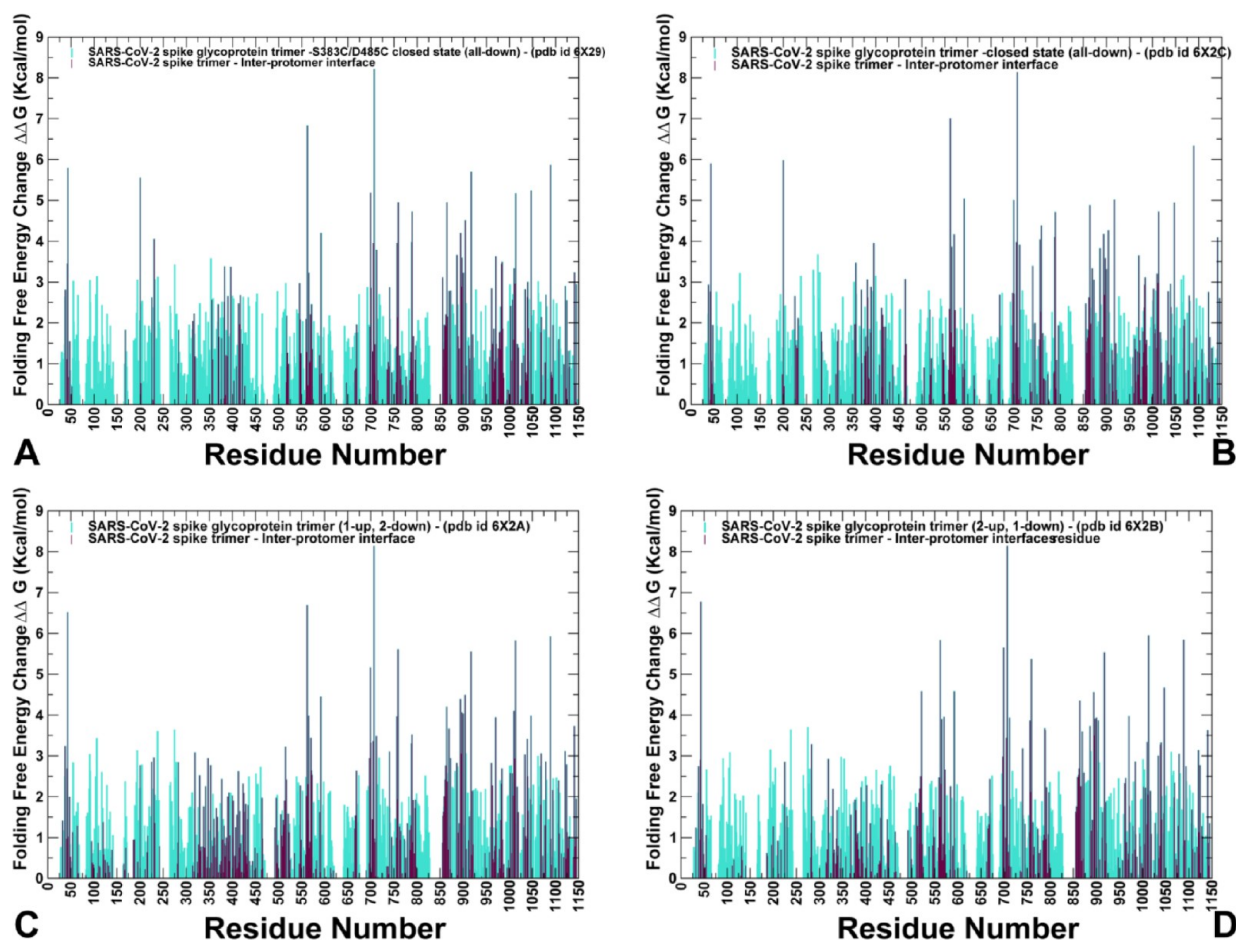


Figure 6. Systematic alanine scanning of the protein residues in the SARS-CoV-2 mutant trimers. The protein stability free energy changes upon alanine mutations for the S mutant rS 2d S383C/D485C in the all-down closed conformation (A), the cryo-EM structure of the u1S2q quadruple in the all-down closed conformation (B), the cryo-EM structure of the u1S2q quadruple mutant in the partially open (1-up) conformation (C), and the structure of u1S2q quadruple mutant in the open (2-up) conformation (D). The residue-based protein stability changes upon alanine mutations are shown in light gray bars, and the respective stability changes for residues involved in the interprotomer contacts are shown in thick maroon-colored bars. The computed values were averaged over equilibrium samples from molecular simulations.

functional motions in the closed form could coordinate rigid body rearrangements of RBDs and NTDs and global displacements of the S1 regions with respect to more rigid S2 subunit. The network of hinge sites in the closed form highlights the position of these residues at the border of SD1 and SD2 regions and between SD1 and RBD regions (Figure 5A). The functional movements of RBDs are dictated by main hinge centers located near F318, S591, F592, V539, L570, and I572 residues. Structural map of these hinge centers indicated that these sites can form a protective shield around one of the key salt bridges between D614 of one protomer and T859 of the adjacent protomer (Figure 5B).

A structural comparison of major hinge sites in the closed, partially open conformations of the S quadruple mutant (Figure S2, Supporting Information) highlighted that all forms share common regulatory points of collective motions (F318, V539, F592, L570). Moreover, it could be seen that a network of major hinge centers featured in the closed state becomes smaller in the open state where these sites are localized in the CTD1 region (Figure S2, Supporting Information). Hence, collective motions in the open state that are regulated through CTD1 hinges may couple the RBD and FP regions to facilitate allosteric couplings and potentially prime the S protein for membrane fusion. Our analysis supported the notion that the

introduction of mutations at the border between SD1 and S2 domains engineered in the quadruple mutant may promote and accelerate functional conformational changes between closed and open states in the SARS-CoV-2 S trimer, priming the spike protein for binding with the host receptor and subsequent transitions to the postfusion state.

To probe a potential role of glycosylation sites in assisting functional dynamics and collective motions in the SARS-CoV-2 mutant structures, we highlighted the structural alignments of 16 glycosylation sites 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, and N1158) projected onto conformational maps of the essential profiles in the closed and open forms (Figure S1, Supporting Information). Strikingly, this simple analysis showed a strong overlap of the glycosylation sites dispersed across both the S1 and S2 subunits with the regions of increased mobility in collective motions (Figure S1, Supporting Information). Among major conserved glycosites, eight are located in the NTD regions and two (N331 and N343) are located in the RBD but outside of the RBM region accessible to ACE2. Some of these glycosylation sites (N122, N149, N74, N165, N234, and N343) are aligned with peaks of the essential profile, especially in the open states (Figure 4C,D). In fact, the

slow mode profiles of the partially open (1RBD up) and open S states (2 RBDs up) showed that N165 and N234 corresponded almost precisely to the local maxima in the NTD regions, while N331 of RBD is aligned with the local hinge site responsible for collective movements of RBD with respect to S1 subunit (Figure 4C,D).

These findings may partly rationalize the experimental results showing that N165 and N234 glycans may not only shield the S protein from antibody-mediated immunity but also provide conformational control acting as stabilizers of the open state and inhibitors of transitions to the closed state.⁴⁹ According to our observations, by interacting with the key mobile elements of collective motions in the open state and restricting their movements, N165 and N234 glycans would likely impede conformational transitions to the closed states. Combined with the fact that the N331 glycosite is aligned with the local hinge positions, our results suggested a plausible scenario in which N331, N165, and N234 may be coupled during allosteric changes and act synchronously to modulate conformational preferences of the S protein. These findings are also consistent with microsecond all-atom MD simulations of a fully glycosylated SARS-CoV-2 S protein embedded in a viral membrane model showing a collective behavior of these glycans that could form topologically important glycan microdomains.⁵⁹ Recent experimental studies presented strong evidence for the presence of O-glycosylation at T323 position and indicated possible glycosylation at S325 of the S protein.^{122,123} A possible O-glycosylation at Thr323 of SARS-CoV-2 subunit S1 glycoprotein has been predicted by computational analysis.¹²⁴ Interestingly, our results showed a close overlap between glycosites N331, T323, and S325 and important hinge center in the RBD region anchored by F318 that is in the close proximity of these glycosites (Figure S1, Supporting Information).¹²⁵ Hence, glycans may provide an additional stabilizing contribution to immobilize the hinge site in the RBD region, thereby modulating collective conformational changes in the S protein.

Ensemble-Based Alanine Scanning and Energetic Analysis of the Interprotomer Contact Sites in the SARS-CoV-2 Spike Mutants. We employed the equilibrium ensembles generated by CABS-CG simulations of the SARS-CoV-2 mutant to perform a systematic alanine scanning of the protein residues (Figure 6). We also highlighted and examined stability and binding free energy changes for the interprotomer residues. The objective was also to identify whether mutational sites of the interdomain contacts and hinge positions in the SARS-CoV-2 trimers also corresponded to notable energetic hotspots. Although the general pattern of the residue-based alanine scanning profile in the SARS-CoV-2 S mutants is preserved, the structural differences and subtle packing alterations and rearrangements were revealed through this analysis. As expected, in the closed form of the rS 2d mutant (Figure 6A) quadruple mutant (Figure 6B), we observed the largest number of interprotomer sites that contribute significantly to protein stability. These hotspot clusters are broadly distributed, including the RBD residues and CTD1 domain (residues 529–591) but most of the hotspots are localized in the S2 subunit regions, including the HR1 (residues 910–985), CH (residues 986–1035), and UH regions (residues 736–781).

These peaks included hydrophobic residues Y789, F759, F592, F562, Y756, S383, Y396, L570, and I572 (Figure 6A,B). Some of these hotspots included mutational sites and

regulatory positions such as S383, L570, and I572. However, mutations in the key hinge positions F318 and V539 produced destabilization changes of 1.53 and 2.1 kcal/mol that are similar to a spectrum of hydrophobic tightly packed residues located away from regulatory regions. The experimental studies showed that the D614G mutation can shift the population of the SARS-CoV-2 S trimer toward open states.⁴² According to the energetic analysis, the D614G mutation yielded a fairly moderate energy loss of 0.8 kcal/mol as this important residue is relatively exposed. These results suggested that the local energetic loss caused by the D614G mutation may not be sufficient to explain radical effects of this mutation on the dynamic equilibrium of S trimer that are likely linked with allosteric changes in the residue interaction networks and structural perturbations around important hinge centers. We found that the density of energetic hotspot clusters decreases in the partially open and open states (Figure 6C,D). In the open form of the S quadruple mutant, most of the interprotomer hotspots are localized in the S2 region, which is consistent with the conformational dynamics analysis showing the increased mobility of S1 regions. Interestingly, alanine scanning showed that the interprotomer contacts in S2 subunit of the open form can be weakened as compared to the closed state (Figure 6D).

The characteristic peaks and energetic hotspots are mostly conserved, featuring L570 and I572 positions among the most important for protein stability of the interdomain and interprotomer interfaces in the quadruple S mutant. At the same time, mutations in other hinge positions F318, S383, V539, and S591 produced more moderate destabilization changes. Hence, alanine scanning analysis demonstrated that while the hinge centers and sites of engineered mutations (S383, K386, D428, T547, L570, I572, S591, S750, Y855, I856, T962, D985) can appreciably contribute to stability of the SARS-CoV-2 S trimer, only several regulatory points (L570, I572) corresponded to the energetic protein stability hotspots.

These findings suggested that mutation-induced population shifts between distinct functional states observed for rS 2d and u1S2q S proteins arise primarily not from local interaction changes at the interprotomer interfaces but are driven by long-range allosteric changes and reorganizations in the allosteric interaction networks.

Protein Contact Networks and Network Clustering in the SARS-CoV-2 Spike Mutants: Probing Interdomain Communicating Regions and Mediating Centers. We constructed protein contact network models for the SARS-CoV-2 mutant proteins using a model developed in our earlier studies.^{96–98} The protein contact network is a network whose nodes are the protein residues and links are active contacts between residues in the protein structure. For the rS 2d (S383C/D985C) mutant, structural modularity is symmetrically arranged as evidenced by the distribution of two well-defined clusters (Figure S3, Supporting Information). The active region is in the central core of the structure and points to residues crucial for structural stability. In the closed state of the quadruple mutant we observed a similar pattern of clustering partition, but some relevant differences seemed to emerge as the clusters to be more distributed and less segregated (Figure S3, Supporting Information). Network clustering for a partially open u1S2q mutant witnessed a shift in the domain partitioning, with the identified modules corresponding to the S1 region of the single “up” protomer

Table 2. Communicating Regions in the SARS-CoV-2 Spike Trimer Mutants^a

pdb code	$\langle \text{hydroph} \rangle$	communicating residues ($P > 0$)	$\langle \text{hydroph} \rangle_{\text{active}}$	highly communicating residues ($P \geq 0.75$)	$\langle \text{hydroph} \rangle_{\text{active+}}$
6X29	-0.92	260	-0.66	29	-1.10
6X2A	-0.90	87	-0.68	16	-1.72
6X2B	-0.91	241	-0.74	25	0.25
6X2C	-0.91	302	-0.73	38	-1.14

^aThe second, fourth and sixth columns report the mean hydrophobicity of the whole structures, of the communicating regions ($P > 0$), and of the highly active residues ($P \geq 0.75$), respectively. The third and the fifth columns reports the number of communicating residues with $P > 0$ and of the number of the intermodular communicating hotspots (residues with $P \geq 0.75$).

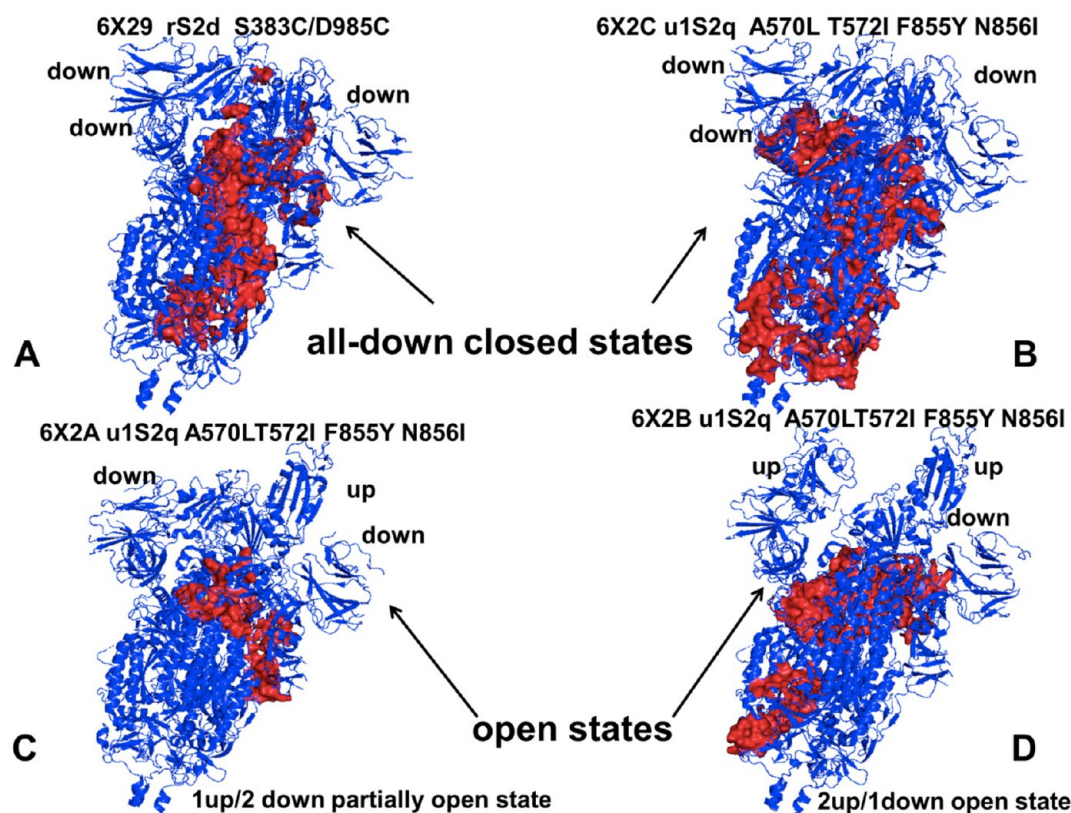


Figure 7. Structural mapping of communicating residues with participation coefficient $P > 0$ for the S mutant rS 2d S383C/D985C in the all-down closed conformation (A), the cryo-EM structure of the u1S2q quadruple in the all-down closed conformation (B), the cryo-EM structure of the u1S2q quadruple mutant in the partially open (1-up) conformation (C) and the structure of the u1S2q quadruple mutant in the open (2-up) conformation (D). The SARS-CoV-2 S mutant trimers are shown in ribbons. The predicted communicating residues are shown in red spheres.

and the other module representing the remaining part of the S trimer structure (Figure S3, Supporting Information). The presence of two activated protomers in the open mutant conformation drives back structural symmetry of cluster partitioning (Figure S3, Supporting Information). We observed that the two clusters are well distributed and strongly intermingled as compared to the rS 2d down conformation.

The participation coefficient P values describe the propensity of nodes to contribute to the intercluster communication.^{96–98} In this model, we characterized the importance of functional regions in allosteric interactions based on the average participation coefficient values of the corresponding residues involved in signal transmission between different clusters (functional domains). Table 2 reports the general features of the four structures, including the number of detected intercluster communicating residues and their hydrophobicity according to the Kyte–Doolittle scale.¹²⁶ The number of highly communicating residues ($P > 0.75$) is fairly similar in the both closed forms and open state

with 2 up RBDs, while the level of communication may drop in a partially open form of the S mutant (Table 2). These findings suggested that there may be a redistribution of local modules and partial rewiring of the protein contact network upon adopting a receptor-accessible open state in the S trimer mutants.

The participation coefficient heat maps showed several interesting patterns of the intermodular communication in the SARS-CoV-2 S trimer mutants (Figure S4, Supporting Information). The actively participating residues ($P > 0$) in both closed forms are similarly distributed across the entire trimer (Figure S4, Supporting Information). However, this distribution is more biased toward the interdomain regions between RBDs and S2 unit in the S383C/D985C double mutant, while communicating sites form more densely populated clusters in the S2 regions of the quadruple mutant (Figure S4, Supporting Information).

Structural mapping of communicating residues in the closed states showed a high density of these sites in the S2 regions

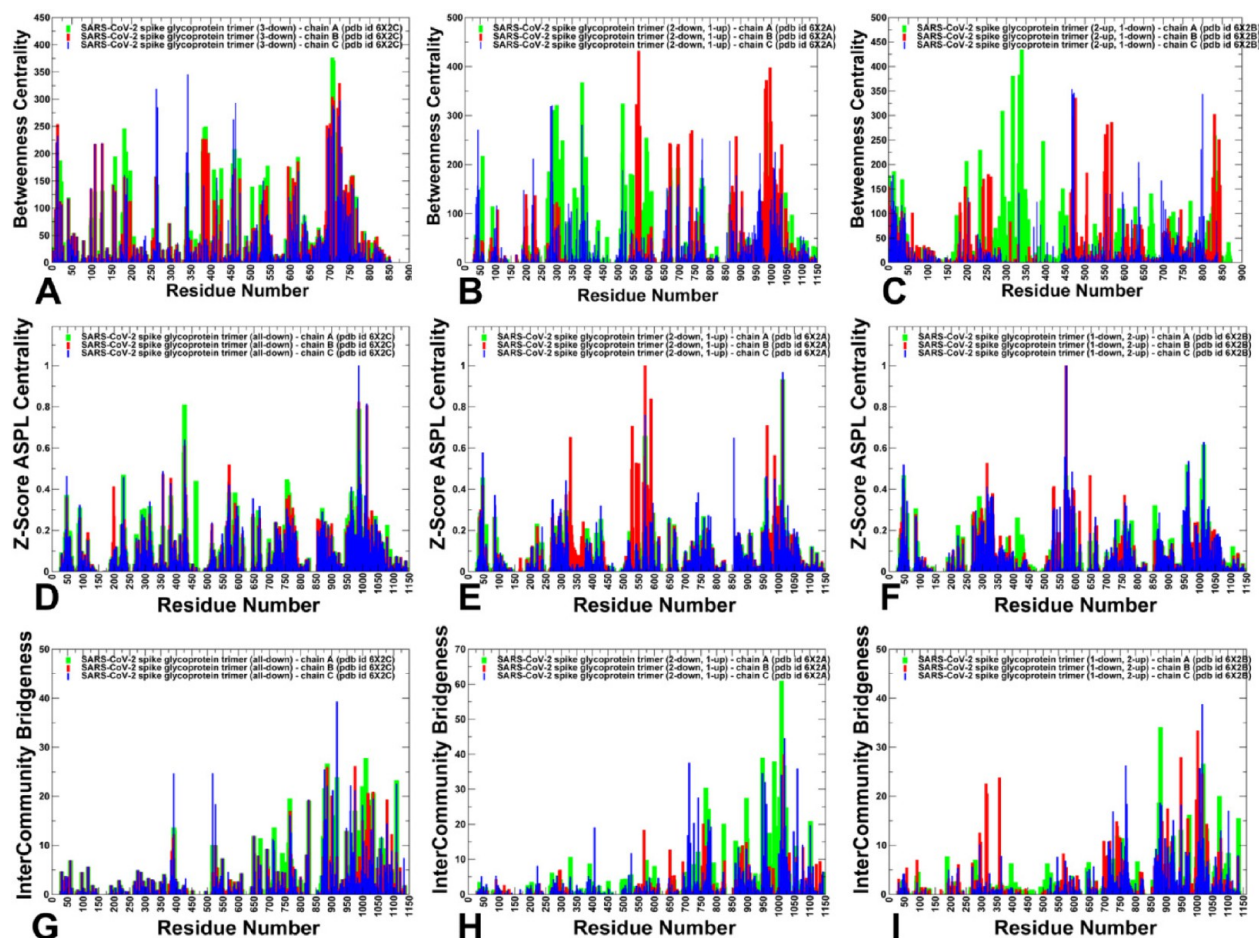


Figure 8. Network centrality analysis of the SARS-CoV-2 S mutant trimers. The short path betweenness centrality for the structure of the u1S2q quadruple in the all-down closed conformation (A), the structure of the u1S2q quadruple mutant in the partially open (1-up) conformation (B), and the structure of u1S2q quadruple mutant in the open (2-up) conformation (C). The SARS-CoV-2 S mutant trimers are shown in ribbons. The residue-based profiles are shown for protomers of the S trimer in green, red, and blue lines. The betweenness centrality profiles based on ASPL change upon removing individual nodes are shown for the structures of the u1S2q quadruple mutant in the closed, partially open, and open states in panels D, E, F respectively. The intercommunity bridgeness profiles are shown for the structures of the u1S2q quadruple mutant in the closed, partially open, and open states in panels G, H, and I respectively.

and the interdomain S1/S2 regions (Figure 7A,B). The potential communicating hotspots with high participation coefficient values ($P > 0.75$) in the closed S states included residues in the CTD1 region (N532, L533, N542, V539, T553), FP region (V826, T827), the HR1 (Q913, V915, Q926, A944, L948, C985) (Tables S1 and S2, Supporting Information). Notably, this list of communicating hotspots included C985 that is disulfide-bonded to S383 in the double mutant and residues in the FP and CTD1 regions mediating communication with the RBD regions (Table S1, Supporting Information). In the partially open state with a single protomer in the “up” position, there is a noticeable narrowing of the intermodular communicating sites that become consolidated at the border of the two major clusters located near the hinge regions between RBD and S2 (Figure 7C) including positions F318, V539, Y855, I856, N960, V963 (Table S3, Supporting Information). Some of these residues are involved in the interprotomer contacts near mutational sites (L570-V963, N960-L570). In the open form of the quadruple S mutant with 2 up RBDs, the clusters are rearranged, and clusters of communicating residues become less dispersed and more localized in the S2 regions and near the interdomain border of the S1 subunit (Figure 7D). The majority of strongly

communicating residues in the open form are located in the CH and UH regions of the S2 subunit, including S555, VL585, D586, L752, and F759 residues (Table S4, Supporting Information).

Notably, the communicating sites near S1 and S2 borders (F318, V539, L570, and I572) that featured significant participation coefficients corresponded to the key hinge centers implicated in regulation of functional movements. These findings suggested that different functional states of S mutant trimers can share the intercluster communicating sites that also correspond to major hinge centers of collective dynamics. At the same time, protein contact network analysis pointed to a partial redistribution of communication hubs in the open trimer that may reflect reorganization of the allosteric interaction networks and modulation of signal transmission mechanisms.

Dynamic Network Modeling and Centrality Analysis Identify Regulatory Regions Mediating Allosteric Interaction Networks in the SARS-CoV-2 Spike Mutants. Understanding of the interplay between population-shift mechanism governing allosteric structural changes and “dynamics-driven” mechanism in which allosteric interactions can be mediated through alterations of functional motions and

rebalancing of rigid and flexible protein regions is central objective of mechanistic models underpinning allostery.^{65–69} These models allow for quantitative analysis of allosteric molecular events in which conformational landscapes of protein systems can be remodeled by various perturbations such as mutations, ligand binding, or interactions with other proteins. In particular, conformational changes between open and closed forms of SARS-CoV-2 S protein and process of virus entry on the molecular level may exemplify a potential synergistic effect of population-shift and dynamically driven mechanisms, where allosteric structural changes between spike states are intimately coupled with an exchange of conformational mobility that accompany S protein binding with the host receptor. Structure-based network approaches have offered a simple and effective formalism for describing allosteric interactions, where the dynamic fluctuations are mapped onto a graph with nodes representing residues and edges representing weights of the measured dynamic properties.^{65–69} Network-centric models of protein structure and dynamics used in the current study can also provide a complementary perspective to physics-based landscape models and allow for quantitative analysis of allosteric changes, identification of regulatory control centers and mapping of allosteric communication pathways.^{108,127–129}

Using this framework, the residue interaction networks in the SARS-CoV-2 spike trimer structures were built using a graph-based representation of protein structures^{100–102} in which residue nodes are interconnected through both dynamic¹⁰² and coevolutionary correlations.¹⁰⁵ By employing perturbation-based network modeling, we computed ensemble-averaged distributions of several residue-based topological network metrics: the short path residue centrality (Figure 8A–C), the betweenness centrality based on average shortest path length (ASPL) change upon removing individual nodes (Figure 8D–F), and the intercommunity bridgeness (Figure 8G–I). These network metrics were explored to establish consensus in identifying mediating centers of allosteric interactions in the SARS-CoV-2 mutant structures.

We found that the high centrality residues are assembled in tight interaction clusters with the distribution of peaks that is generally similar for the closed and partially open forms of u1S2q S mutant (Figure 8A,B). In these structures, the network centrality peaks tend to align with the hinge center regions (residues 315–331, 569–572, 756–760), indicating that these key regulators of functional motions may also mediate communication in the residue interaction networks. Moreover, high centrality residues often precisely corresponded to the engineered mutational sites, suggesting that their functional role may be partly determined by global mediating propensities in the residue interaction networks (Figure 8A–C).

In addition, there are a number of local peaks in the HR1 region (residues 910–985), CH regions (residues 986–1035), and UH (residues 736–781). Notably, the distribution of high centrality peaks was broad in the closed and partially open forms, indicating a large allosteric interaction network characteristic of these states. The number of local centrality peaks noticeably reduced in the open form with 2 up RBDs, even though distribution peaks were still seen in the CH regions of S2 and the interdomain borders of S1 and S2 subunits (Figure 8C). Characteristically, a noticeable density of high centrality sites in the CH region is present and shared in

distinct conformational states of the SARS-CoV-2 spike protein (Figure 8A–C).

A similar composition of high centrality peaks was also seen using a different betweenness centrality metric (Figure 8D–F). Using this network metric, we observed consolidation of local centrality peaks into three major clusters in the open form, corresponding to interdomain borders of RBD and S2, CTD1 regions, and CH residues (Figure 8F). The centrality profiles also highlighted a gradual narrowing of the betweenness centrality distribution as we proceed from the closed state to the open forms. These observations indicated that allosteric interaction network in the open state may be partially rewired and become more localized. Another important network signature observed in the open forms was the emergence of the largest and strongest peak aligned with residues 568–572 in the CTD1 region and strong centrality density in the NTD regions near glycosite positions (Figure 8E,F). These high centrality clusters mediate allosteric couplings between RBDs and NTDs as well as between RBDs and functional FP segment in the S2 subunit. Structural maps of high centrality sites highlighted a significant density of these residues in the interprotomer and interdomain regions, primarily in the HR1, CH, and UH regions of the S2 subunit (Figure S5, Supporting Information).

Interestingly, the incorporation of glycan shield into atomistic models allowed for more detailed network analysis where the observed peaks and high centrality of glycosites in the NTD and RBD regions of the open states resulted from formation of tightly packed glycan microclusters (Figure 8A–F). Indeed, in the open states glycans at N331 and N343 positions of “up” protomer B can interact with N122, N165, and N234 glycans from the NTD regions of another open protomer C. The increased betweenness centrality (Figure 8B,C) and Z-score ASPL network centrality (Figure 8E,F) for these regions suggested that the glycan environment at glycosites N165, N234, N331, and N343 enhances the mediating role of these positions in the intramolecular communications and may contribute to stabilization of the global allosteric interaction network. These results are consistent with recent experimental⁴⁹ and computational studies^{56,57} showing that N234 and N165 glycans can act as stabilizers of the open state and that alanine mutations at these positions destabilize the RBD up conformations. Furthermore, our findings indicated that N234 and N165 glycosites acting as mediating centers may be allosterically coupled with other regulatory elements and therefore exert a partial control on conformational transitions in the SARS-CoV-2 S mutant structures. While the density of high centrality sites is reduced in the partially open and open states, we observed mediating residues near glycosites in the NTD-RBD regions as well as high centrality clusters in the CTD1 region and at the borders between RBD and S2 regions (Figure S5, Supporting Information).

Collectively, these findings provided several important insights. First, the observed narrowing in the distribution of centrality sites in the open states may reflect the mechanism of allosteric interactions in this form where signal transmission may proceed through fewer localized routes and functional centers connecting the S2 subunit with RBD regions. Second, several known regulatory sites that dictate dynamic switching between conformational states of the SARS-CoV-2 spike trimer overlap with the high centrality mediators of the interaction network. This may be important for facilitating therapeutic

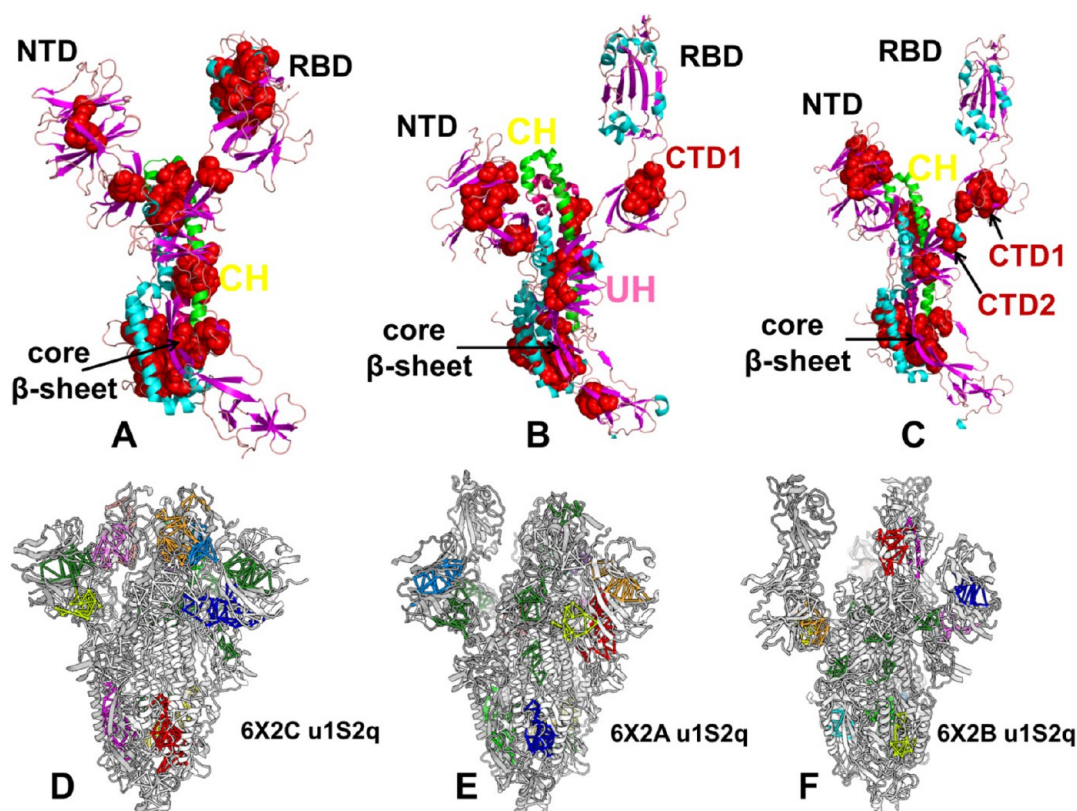


Figure 9. Network community analysis of the SARS-CoV-2 S mutant trimers. Structural mapping of key local communities (residues comprising these modules are in red spheres) in the structure of the u1S2q quadruple in the all-down closed conformation (A), the structure of the u1S2q quadruple mutant in the partially open (1-up) conformation (B), and the structure of the u1S2q quadruple mutant in the open (2-up) conformation (C). The SARS-CoV-2 S mutant protomer is shown in ribbons depicting the “down protomer” for the closed state in panel A and “up protomer” for the partially open and open states in panels B and C. Schematic representation of the distribution for major communities in the closed-down form of the u1S2q S mutant (D), partially open form with 1 up-RBD (E), and open form with 2 up-RBDs (F).

intervention through targeting specific mediating centers involved in signaling pathways of the receptor-accessible states of SARS-CoV-2 protein.

Community Analysis of Residue Interaction Networks and Intercommunity Bridgeness Analysis Suggest Principles of Signal Transmission in the SARS-CoV-2 Spike Mutants. Using community decomposition, the residue interaction networks were divided into local interaction modules in which residues are densely interconnected. A community-based model of allosteric communications is based on the notion that groups of residues that form local interacting communities are correlated and switch their conformational states cooperatively. In this model, allosteric communications can be transmitted through a chain of stable local modules connected via intercommunity bridges. By describing protein dynamics as a dynamically evolving network of interconnected modules, the topological regularities of the network structure can be identified.^{65–69}

We performed community analysis of the SARS-CoV-2 S trimers and computed the intermodular bridgeness profiles (Figure 8G–I). The bridgeness profiles reflect the differences in concentration of local communities in the S1 and S2 subunits as most of stable modules and corresponding connectors tend to be localized in the S2 regions, particularly in the HR1 (residues 910–985), CH regions (residues 986–1035), and UH (residues 736–781). Interestingly, open forms featured a number of intercommunity bridges in the UH regions (residues 736–781), suggesting that UH residues

could provide stepping stones for the global routes connecting the S2 and S1 subunits. Structural map of major communities illustrated a dense distribution of stable network modules that form hubs of an allosteric network that connects S2 and S1 subunits and mediated through the UH and CH regions of the protein core. The projection of largest and most stable communities on the protomer structures revealed the overall topological arrangement pointing to conservation of intra- and interprotomer stable modules in the S2 subunit, primarily localized in a core antiparallel β -sheet and UH and CH regions (Figure 9A–C).

A conserved 3D organization of the core β -sheet and the UH and CH regions among different coronaviruses is a salient feature of the S glycoprotein structure.²¹ The conserved communities in the S2 regions shared by all states included (Y1067-L1049-F906), (C1032-H1048-S1051-H1064), (M1029-V1033-P1053), and (F1042-E1031-R1039) that provide stabilization anchors for mediating allosteric interactions in the S2 subunit. Interestingly, the overall topology of local communities in the S2 subunit is mostly retained in all forms, suggesting that allosteric signaling mechanisms may be governed by functional regions in the core β -sheet and UH and CH regions.

We also delineated conformation-specific community rearrangements that provide some insight into changes in allosteric interaction networks between different functional forms of the SARS-CoV-2 S trimer mutants. In the closed conformation, the intramolecular communities prominently

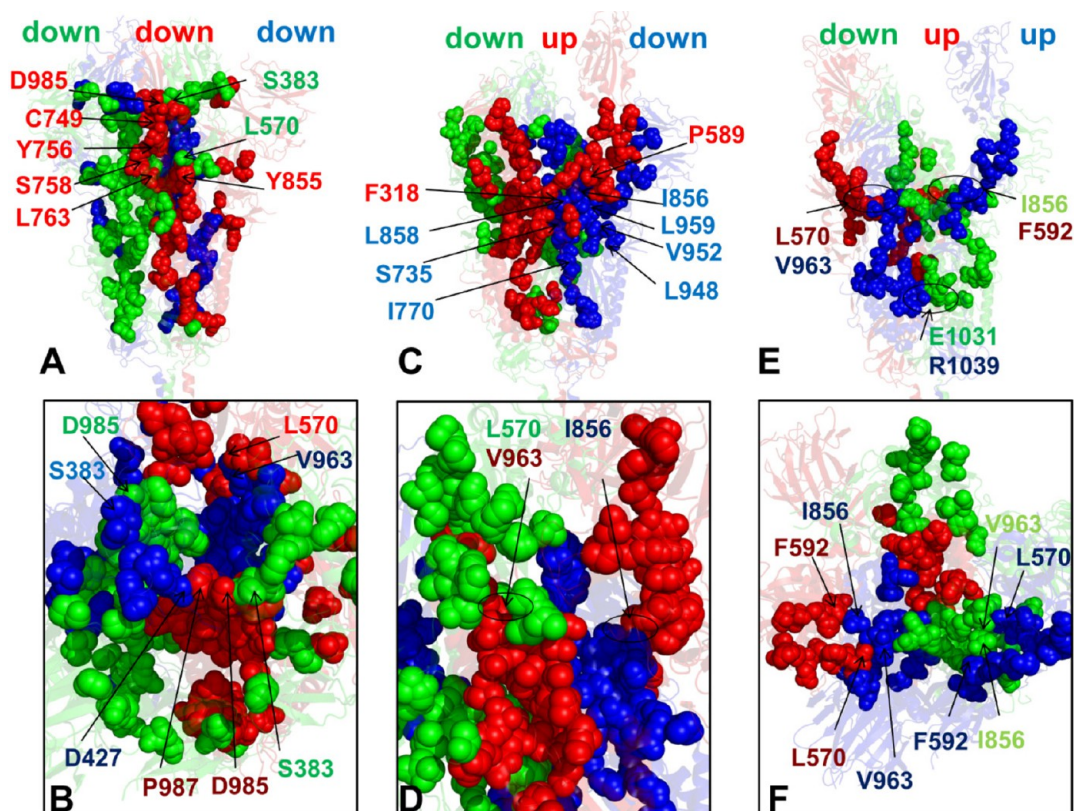


Figure 10. Molecular topography of major allosteric communication pathways in the SARS-CoV-2 S mutant trimers. (A) Structural mapping of communication paths for the u1S2q quadruple mutant in the all-down closed conformation. The residues comprising the pathways are shown in spheres colored by the respective protomer. The protomers are in green, red, and blue ribbons with the decreased transparency. (B) Close-up of the interprotomer connections in the major communication paths in the closed form. (C) Structural mapping of communication paths for the u1S2q quadruple mutant in the partially open (1-up) conformation. (D) Close-up of the interprotomer connections in the major communication paths in the partially open form. (E) Structural mapping of communication paths for the u1S2q quadruple mutant in the open (2-up) conformation. (F) Close-up of the interprotomer connections in the major communication paths in the open form.

feature residues in both the S1 and S2 subunits (Figure 9A). These modules in the NTD and RBD regions include clusters (W353-F400-Y423), (D290-F275-F58), (F342-F374-W436), (Y365-L387-F515-V382), (R509-S438-D442), (F464-F429-S514), (R355-Y396-F464-S514-F429-P426), (Y365-V395-F392). In addition, we detected a number of local communities in the CH, FP, and CTD1 regions such as (F1042-E1031-R1039), (C1032-H1048-S1064), and (Q1054-S816-E819). A detailed structural map of local communities in the closed form is highlighted for a protomer (Figure 9A) and a schematic projection of largest communities is shown for the closed state of the u1S2q mutant trimer (Figure 9D). It can be seen that stabilizing network modules are broadly distributed across the structure, indicative of a broad and large allosteric network in the closed form. By comparing the community distributions in the closed (Figure 9A) and open forms (Figure 9B,C) we found a noticeable redistribution in the local network modularity that becomes particularly apparent in the open conformation with 2 up RBDs. When the S trimer adopts a partially open conformation, the RBD in the “up” position reduces the contacts with the S2 subunit of an adjacent protomer. Our community analysis reflected these structural changes as density of local communities in the RBD and NTD regions was appreciably reduced in the “up” protomer. At the same time, stable modules are preserved in the S2 subunit and are localized primarily in the UH and CH regions (Figure 9B,E). These communities include (T791-K795-L806),

(L1049-L1067-F906-V911), (F898-F802-F797), (F906-I923-L916), and (R905-M1050-F898-M902) that responsible for stability and communication between HR1, CH, and UH regions. We noticed that the allosteric interaction network in the partially open form (Figure 9E) remained quite similar to the network of the closed form (Figure 9D), showing only decoupling of the NTDs and RBDs in the “up” protomer. In some contrast, local communities in the NBD and RBD regions in the open conformation with 2 up RBDs can be weakened or completely disassembled (Figure 9C,F). Moreover, the network of local communities is less dense in the S2 subunit as stabilizing modules become more segregated. The open form of the S trimer featured a significant population of communities in the CTD1 region (residues 529–591) which is implicated in allosteric connection between RBD and S2 regions to communicate signal from and to the fusion peptide.³⁴ These local communities in the “up” protomers included (L552-F541-I587-C589), (C538-T549-C590), and (F543-L546-V576-L585) (Figure 9C). Structural projection of dominant communities on the S trimer structures in the open state (Figure 9F) highlighted the reduced density and a partially rewired organization of stabilizing modules as compared to strongly interconnected closed form. The preservation of key stabilizing communities in the S2 subunit and CTD1 regions allows for the interprotomer connections and could arguably result in localized but still efficient signaling between S1 and S2 regions.

Allosteric Communication Pathways in the SARS-CoV-2 Spike States: Differential Modulation of Signal Transmission through Cascaded Regulatory Switching.

Another emerging concept central to computational models of allostery is identification of regulatory control points that mediate long-range communications in the system and modeling of allosteric pathways between conformational switch centers.^{65–69} The current understanding of communication paths in proteins is based on the ensemble-based statistical model that assumes existence of multiple paths where the validity and thermodynamic preferences of a particular route are often tested against experimental data using mutagenesis of key switch points that mediate specific communication.^{65–69}

In our study, we leveraged the results of community decomposition and used edge betweenness (or edge centrality) in the global interaction network as a proxy for modeling and analysis of allosteric communication pathways. This parameter is defined as the ratio of all the shortest paths passing through a particular edge to the total number of shortest paths in the network. Through structural mapping of the high centrality edges in most probable pathways of the ensembles, we obtained a global map of major communication routes in the interaction network of the SARS-CoV-2 trimer mutants (Figure 10A–C). The topography of these maps showed significant differences in the distribution and density of major communication routes connecting S1 and S2 regions in different functional states of the S trimer. The communication highway system in the closed state featured a dense interconnected network with several alternative global routings that engage molecular switches connecting the protomers in key positions to enable allosteric signal transmission (Figure 10A,B). In the closed form, the major communication routes can proceed through key S2 segments, including primarily CH regions 986–1035 (S1021, Y1007, A1015) and UH regions 736–781 (S758, F759, Y759, N751, L754, L755, C749, E748). Notably, the UH residues provide an important part of the global route connecting S2 and S1 subunits that dominates allosteric paths. The high edge centrality links in the ensemble of communication paths featured by several key interprotomers are L570(chain A)-V963(chain B), S383(chain B)-D485(chain C), D571(chain B)-S967(chain C), P589(chain A)-Y855(chain B), and P987(chain A)-D427(chain C) (Figure 10A,B). By exploring links between the UH and CH regions, the communication paths within a single protomer can reach the regulatory switch points of P987 and D485 residues that enable the interprotomer connections (Figure 10A,B). We found that at this juncture communication paths can link P987 of the protomer with D427 of the one adjacent protomer, and D485 with S383 of the other adjacent protomer. Using these switch centers, allosteric paths can connect directly and efficiently to the RBD regions of the adjacent protomers (Figure 10B). These results highlighted the important role of S383, D485, and P987 residues that can function as critical regulators of the conformational equilibrium in S-2P and rS 2d spike mutants.^{27,35–38} According to our findings, these strategically located residues can provide key mediating bridges for efficient allosteric communication paths between S2 and S1-RBD regions in the S trimer (Figure 10A,B). Moreover, these switch centers can act as hubs in the ensemble of allosteric communication paths in the closed form as a significant fraction of communication routes in the S2 subunit can utilize these hubs to reach RBD regions of the neighboring

protomers (Figure 10A,B). Molecular mapping of allosteric pathways revealed a well-defined topography and significant density of communication routes with switch residues at the interprotomer contact interfaces. The central result of this analysis is that these experimentally confirmed regulatory hotspots (S383, D485 and K987P)^{35–38} can also function as key network bridges of communication pathways and control switch points of signal transmission in the closed form of the SARS-CoV-2 spike trimer.

It is worth noting that structural architecture of the closed form and strategic location of switch contacts can preserve the overall topology of communication routes in the closed states of the wild-type and mutant S proteins. However, not only do K986P and V987P mutations shift the equilibrium away from the “all-down” form²⁷ but also the efficiency of predicted pathways in the closed conformations can be compromised due to weakening of stabilizing interactions at the switch control points. The weakening of the key interprotomer switch contacts may significantly hamper the efficiency of allosteric communication in the S closed trimers. Structural mapping of hubs in the ensemble of most probable pathways indicated that a significant number of these mediating sites are located in the S2 regions and near the interprotomer and interdomain boundaries between S1 and S2 subunits (Figure S6, Supporting Information). The density of communication hubs is high in the closed form reflecting the broader allosteric network and diverse communication routes. However, the number of hubs can be markedly reduced in the open S trimer with the RBDs in the up conformation, making allosteric signaling more dependent on the fewer interprotomer bridges governing communications. We argue that the dense allosteric network of high centrality sites and communication hubs seen in the closed form could allow for diverse alternative routes between S2 and S1 regions, thereby protecting the robustness of allosteric signaling. At the same time, allosteric communications in the open form may become more sensitive and vulnerable to targeted mutations of the key mediating centers.

In the partially open spike with one RBD in the up position (chain B), the high edge centrality links corresponded to key interprotomer contacts including L570(chain A)-V963(chain B), P589(chain A)-I856(chain B), L570(chain B)-N960(chain C), L570(chain B)-K964(chain C), and P987(chain B)-D427(chain C) (Figure 10 C,D). We found a partial redistribution of communication pathways in the partially open form, where allosteric routes in the “up” protomer proceeded via S2 subunit utilizing HR1 (residues 910–985) and passing through a hydrophobic stretch of 934–940, L948, V952, A956, V963 until connecting with the Y855/I856 switch to link to “down” protomer and reach the RBD regions (Figure 10C,D). An alternative path can pass through the S2 subunit of the “down” protomer by utilizing UH regions until reaching a key switch point Y855/I856 and connecting to the “up” protomer and its RBD regions. Hence, in the partially open state, the dominant pathways appeared to utilize Y855/I856 regulatory center to connect S2 regions with activated RBD in the “up” protomer. We also found that in the partially open state, a significant population of probable paths could proceed through HR1 region, including residues 934–940 that form a cluster targeted by naturally occurring mutations in SARS-CoV-2 protein.^{130,131} This region also undergoes significant conformational changes during postfusion transition¹³⁰ and our analysis suggested that mutations in these positions may have a stronger effect on signal transmission in the open state,

as these residues are not involved in communication paths in the closed form.

Interestingly, in the open spike form featuring 2 up RBDs the distribution of communication pathways changes even more drastically (Figure 10E,F). Among high edge centrality links are a number of the interprotomer connections that could create a cascade of switches for propagating communication signal. These links included P589(chain A)-Y855(chain B), D1041 (chain B)-S1030 (chain C), G413 (chain A)-E988-(chain C), V963(chain A)-L570(chain C), I856(chain A)-F592(chain C), K964(chain A)-L570(chain C), and D427-(chain A)-P987(chain C). Importantly, some of these critical switch links dominate the ensemble of communication paths included sites L570, F592, Y855, I856, and K964. The ensemble of dominant pathways connects the FP residues I818, E819, V826 on one of the “up” protomers to the HR1 region (residues 910–985) via V952, A956, V963, and Q965, and then through the critical switch Y855/I856, the path links to the other “up” protomer by connecting to another potential switch center F592 of the CTD1 region (Figure 10E,F). In turn, the CTD1 region provides mediating bridges for communication routes to the RBD of the “up” monomer. Another optimal path along the “up” protomer can connect FP residues via I856 to L570/I572 in the CTD1 of the “down” protomer and subsequently to the RBD regions. Strikingly, a new network of efficient routes connecting the FP regions with RBD residues can emerge that leverages several conserved switches in the receptor-accessible “up” positions. A significant fraction of communication paths navigates through switch points Y855, I856, and F592 located in the close proximity of D614 at the interprotomer interface with T859 of the adjacent protomer (Figure 5B). On the basis of this analysis, we argue that D614G mutation not only abolishes the local interprotomer contacts and weakens stability of the trimer but also could adversely affect allosteric communication routes passing through switch Y855/I856 and F592 to reach the S1 subunit regions (Figure 10E,F). This is consistent with other studies showing that a close proximity of D614G to the hinge bending region could lead to the increased switch flexibility and promote transitions to the active “up” state, making the mutant more amenable for binding with host receptor ACE2.¹³² It is worth noting that molecular mechanisms underlying the D614G effect are not well understood.¹³³ Several studies suggested that D614G may affect the intraprotomer energetics and favor a “one-up” open conformation by eliciting contract changes in the CTD1 and FP regions.^{134,135} In support of our results, recent cryo-EM structures of SARS-CoV-2 S mutants revealed an allosteric effect that seemingly small D614G changes in the SD2 domain have on RBD dynamics, suggesting that SD2 regions may be involved in modulating RBD transitions between up and down conformations.¹³⁶ Indeed, according to our analysis, the communication paths in the open form may often exploit FP and CTD1/CTD2 regions to connect S2 regions with activated RBDs in the up position (Figure 10E,F).

Strikingly, even though some of the regulatory switch points are utilized by probable paths in all states, the composition of switching cascades and topography of signaling routes can considerably change between the closed and open forms. In the closed form the broad ensemble of pathways connecting S2 and S1 regions proceed through S383, D985, P987, DS427, L570, and V963 control points (Figure 10A,B). A different cascade of regulatory control points (L570, I572, F592, Y855,

I856, and K964) may control signal transmission in the open form that featured paths navigating through FP and CTD1/CTD2 regions to connect with RBDs. In addition, the density and width of the communication path ensemble can be progressively reduced in the partially open and open states. Although, the signaling routes become more localized in the open states, allosteric interactions and cross-talk of the S1 and S2 subunits can be maintained through a narrow ensemble of communication routes that are highly dependent on efficiency of switch points in the CTD1 region, including L570, I572, and F592.

The greater diversity of allosteric communication paths in the closed-down form could ensure a proper balance of the network robustness and efficiency to maintain resilience against random mutational changes. In some contrast, efficiency of signal transmission in the open states could be more vulnerable to mutations of specific conformational switches, particularly in the CTD1 region.

CONCLUSIONS

This study systematically examined functional mechanisms of the SARS-CoV-2 S mutants through the lens of allosteric regulation. We found that mutations of the interdomain residues can change the dynamic profiles and stability of the distinct functional states. Functional dynamics analysis revealed conserved hinge regions that appeared to be aligned with engineered mutational sites driving population shifts between closed and open states. This analysis supported the notion that the introduction of mutations at the border between SD1 and S2 domains engineered in the quadruple mutant may accelerate functional conformational changes in the SARS-CoV-2 S trimer, priming spike protein for binding with the host receptor and subsequent transitions to the postfusion state. The results of this study revealed key functional regions and regulatory centers that govern collective dynamics and allosteric interactions in the SARS-CoV-2 spike proteins. Dynamic modeling of residue interaction networks and community analysis identified stabilizing structural modules in the S2 subunit acting as anchoring regions for signal transmission in the SARS-CoV-2 S mutants. The central finding of this study is that the allosteric interactions and pathways in different functional states can be modulated through a versatile network of the interprotomer regulatory centers. We found that mutations at the key interdomain positions can differentially modulate distribution of conformational states and uniquely define topography of signal propagation pathways operating through state-specific cascades of control switch points. We argue that targeting the predicted mediating centers of allosteric interactions and communication hubs in the SARS-CoV-2 S proteins with chemical probes could be beneficial for probing molecular mechanisms underlying the virus infectivity profile and discovery of specific and broad-spectrum allosteric modulators of the SARS-CoV-2 protein functions.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcc.0c10637>.

Structural maps of functional motions for the SARS-CoV-2 mutant structures (Figure S1), structural maps of the hinge regions in different functional states of SARS-

CoV-2 u1S2q quadruple mutants (Figure S2), network clustering for the SARS-CoV-2 mutant structures (Figure S3), structural mapping of the participation coefficient distributions (Figure S4, structural maps of high centrality sites in the SARS-CoV-2 mutant structures (Figure S5), structural maps of communication pathway hubs for the SARS-CoV-2 mutant structures (Figure S6), and the participation coefficients for highly communicating residues ($P > 0.75$) in the different functional forms of the SARS-CoV-2 double and quadruple mutants (Tables S1–S4) (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.jpcc.0c10637>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

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

ABBREVIATIONS

SARS, severe acute respiratory syndrome; RBD, receptor binding domain; ACE2, angiotensin-converting enzyme 2 (ACE2); NTD, N-terminal domain; RBD, receptor-binding domain; CTD1, C-terminal domain 1; CTD2, C-terminal domain 2; FP, fusion peptide; FPPR, fusion peptide proximal region; HR1, heptad repeat 1; CH, central helix region; CD, connector domain; HR2, heptad repeat 2; TM, transmembrane anchor; CT, cytoplasmic tail

REFERENCES

- (1) Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K. S. M.; Lau, E. H. Y.; Wong, J. Y.; et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **2020**, *382*, 1199–1207.
- (2) Wang, C.; Horby, P. W.; Hayden, F. G.; Gao, G. F. A novel coronavirus outbreak of global health concern. *Lancet* **2020**, *395*, 470–473.
- (3) Yi, Y.; Lagniton, P. N. P.; Ye, S.; Li, E.; Xu, R. H. COVID-19: what has been learned and to be learned about the novel coronavirus disease. *Int. J. Biol. Sci.* **2020**, *16*, 1753–1766.
- (4) Wu, A.; Peng, Y.; Huang, B.; Ding, X.; Wang, X.; Niu, P.; Meng, J.; Zhu, Z.; Zhang, Z.; Wang, J.; et al. Genome composition and

divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* **2020**, *27*, 325–328.

(5) de Wit, E.; van Doremalen, N.; Falzarano, D.; Munster, V. J. SARS and MERS: recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **2016**, *14*, 523–534.

(6) Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **2020**, *395*, 497–506.

(7) Astuti, I.; Ysrafil, I. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): An overview of viral structure and host response. *Diabetes Metab. Syndr.* **2020**, *14*, 407–412.

(8) Schoeman, D.; Fielding, B. C. Coronavirus envelope protein: current knowledge. *Virol. J.* **2019**, *16*, 69.

(9) 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.

(10) 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*, 271–280.

(11) Zhou, P.; Yang, X. L.; Wang, X. G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H. R.; Zhu, Y.; Li, B.; Huang, C. L.; et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *579*, 270–273.

(12) 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.

(13) Duan, L.; Zheng, Q.; Zhang, H.; Niu, Y.; Lou, Y.; Wang, H. The SARS-CoV-2 spike glycoprotein biosynthesis, structure, function, and antigenicity: Implications for the design of spike-based vaccine immunogens. *Front. Immunol.* **2020**, *11*, 576622.

(14) 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*, 894–904.

(15) 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.

(16) 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.

(17) 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.

(18) Chakraborti, S.; Prabakaran, P.; Xiao, X.; Dimitrov, D. S. The SARS coronavirus S glycoprotein receptor binding domain: fine mapping and functional characterization. *Virol. J.* **2005**, *2*, 73.

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

(20) 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.

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

(22) 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.

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

(24) 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*, 15092.

(25) 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.

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

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

(28) 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.

(29) Fan, X.; Cao, D.; Kong, L.; Zhang, X. Cryo-EM analysis of the post-fusion structure of the SARS-CoV spike glycoprotein. *Nat. Commun.* **2020**, *11*, 3618.

(30) 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*, 445–454.

(31) 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*, 281–292.

(32) 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.

(33) Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020**, *367*, 1444–1448.

(34) 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.

(35) 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.

(36) 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.

(37) 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.

(38) 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.

(39) Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E. E.; Bhattacharya, T.; Foley, B.; et al. Tracking changes in SARS-CoV-2 Spike: Evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **2020**, *182*, 812–827.

(40) Plante, J. A.; Liu, Y.; Liu, J.; Xia, H.; Johnson, B. A.; Lokugamage, K. G.; Zhang, X.; Muruato, A. E.; Zou, J.; Fontes-Garfias, C. R. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* **2020**, DOI: 10.1038/s41586-020-2895-3.

(41) Hou, Y. J.; Chiba, S.; Halfmann, P.; Ehre, C.; Kuroda, M.; Dinno, K. H.; Leist, S. R.; Schäfer, A.; Nakajima, N.; Takahashi, K. SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. *Science* **2020**, *370*, 1464–1468.

(42) Yurkovetskiy, L.; Wang, X.; Pascal, K. E.; Tomkins-Tinch, C.; Nyalile, T. P.; Wang, Y.; Baum, A.; Diehl, W. E.; Dauphin, A.; Carbone, C.; et al. Structural and functional analysis of the D614G SARS-CoV-2 spike protein variant. *Cell* **2020**, *183*, 739–751.

(43) 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.

(44) 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. *bioRxiv* 2020; DOI: 10.1101/2020.09.10.286948.

(45) 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.

(46) Yang, T. J.; Chang, Y. C.; Ko, T. P.; Draczkowski, P.; Chien, Y. C.; Wu, K. P.; Khoo, K. H.; Chang, H.-W.; Hsu, S. D.; et al. Cryo-EM analysis of a feline coronavirus spike protein reveals a unique structure and camouflaging glycans. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 1438–1446.

(47) Watanabe, Y.; Berndsen, Z. T.; Raghvani, 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*, 2688.

(48) 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.

(49) Henderson, R.; Edwards, R. J.; Mansouri, K.; Janowska, K.; Stalls, V.; Kopp, M.; Haynes, B. F.; Acharya, P. Glycans on the SARS-CoV-2 spike control the receptor binding domain conformation. *bioRxiv* 2020; DOI: 10.1101/2020.06.26.173765.

(50) Jo, S.; Qi, Y.; Im, W. Preferred conformations of N-glycan core pentasaccharide in solution and in glycoproteins. *Glycobiology* **2015**, *26*, 19–29.

(51) Harbison, A.; Fadda, E. An atomistic perspective on ADCC quenching by core-fucosylation of IgG1 Fc N-glycans from enhanced sampling molecular dynamics. *Glycobiology* **2020**, *30*, 407–414.

(52) Amaro, R. E.; Li, W. W. Molecular-level simulation of pandemic influenza glycoproteins. *Methods Mol. Biol.* **2012**, *819*, 575–594.

(53) Grant, O. C.; Montgomery, D.; Ito, K.; Woods, R. J. Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Sci. Rep.* **2020**, *10*, 14991.

(54) Zhao, P.; Praissman, J. L.; Grant, O. C.; Cai, Y.; Xiao, T.; Rosenbalm, K. E.; Aoki, K.; Kellman, B. P.; Bridger, R.; Barouch, D. H.; et al. Virus-receptor interactions of glycosylated SARS-CoV-2 spike and human ACE2 receptor. *Cell Host Microbe* **2020**, *28*, 586–601.

(55) 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.

(56) 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.

(57) Yu, A.; Pak, A. J.; He, P.; Monje-Galvan, V.; Casalino, L.; Gaieb, Z.; Dommer, A. C.; Amaro, R. E.; Voth, G. A. A multiscale coarse-grained model of the SARS-CoV-2 virion. *Biophys. J.* **2020**, DOI: 10.1016/j.bpj.2020.10.048.

(58) Casalino, L.; Dommer, A.; Gaieb, Z.; Barros, E. P.; Sztain, T.; Ahn, S. H.; Trifan, A.; Brace, A.; Bogetti, A.; Ma, H.; et al. AI-driven multiscale simulations illuminate mechanisms of SARS-CoV-2 spike dynamics. *bioRxiv* 2020; DOI: 10.1101/2020.11.19.390187.

- (59) Ghorbani, M.; Brooks, B. R.; Klauda, J. B. Exploring dynamics and network analysis of spike glycoprotein of SARS-CoV-2. *bioRxiv* 2020; DOI: 10.1101/2020.09.28.317206.
- (60) 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.
- (61) Ali, A.; Vijayan, R. Dynamics of the ACE2-SARS-CoV-2/SARS-CoV spike protein interface reveal unique mechanisms. *Sci. Rep.* **2020**, *10*, 14214.
- (62) Verkhivker, G. M. Coevolution, dynamics and allostery conspire in shaping cooperative binding and signal transmission of the SARS-CoV-2 spike protein with human angiotensin-converting enzyme 2. *Int. J. Mol. Sci.* **2020**, *21*, 8268.
- (63) Di Paola, L.; Hadi-Alijanvand, H.; Song, X.; Hu, G.; Giuliani, A. The discovery of a putative allosteric site in the SARS-CoV-2 spike protein using an integrated structural/dynamic approach. *J. Proteome Res.* **2020**, *19*, 4576–4586.
- (64) Verkhivker, G. M. Molecular simulations and network modeling reveal an allosteric signaling in the SARS-CoV-2 spike proteins. *J. Proteome Res.* **2020**, *19*, 4587–4608.
- (65) Guo, J.; Zhou, H. X. Protein allostery and conformational dynamics. *Chem. Rev.* **2016**, *116*, 6503–6515.
- (66) Dokholyan, N. V. Controlling allosteric networks in proteins. *Chem. Rev.* **2016**, *116*, 6463–6487.
- (67) Nussinov, R.; Tsai, C. J. Allostery without a conformational change? Revisiting the paradigm. *Curr. Opin. Struct. Biol.* **2015**, *30*, 17–24.
- (68) Tsai, C. J.; Nussinov, R. A unified view of “how allostery works. *PLoS Comput. Biol.* **2014**, *10*, No. e1003394.
- (69) Wodak, S. J.; Paci, E.; Dokholyan, N. V.; Berezovsky, I. N.; Horovitz, A.; Li, J.; Hilser, V. J.; Bahar, I.; Karanicolas, J.; Stock, G.; et al. Allostery in its many disguises: From theory to applications. *Structure* **2019**, *27*, 566–578.
- (70) Kolinski, A. Protein modeling and structure prediction with a reduced representation. *Acta Biochim. Polym.* **2019**, *51*, 349–371.
- (71) 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.
- (72) 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.
- (73) 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.
- (74) 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.
- (75) Jamroz, M.; Orozco, M.; Kolinski, A.; Kmiecik, S. Consistent view of protein fluctuations from all-atom molecular dynamics and coarse-grained dynamics with knowledge-based force-field. *J. Chem. Theory Comput.* **2013**, *9*, 119–125.
- (76) Badaczewska-Dawid, A. E.; Kolinski, A.; Kmiecik, S. Protocols for fast simulations of protein structure flexibility using CABS-Flex and SURPASS. *Methods Mol. Biol.* **2020**, *2165*, 337–353.
- (77) 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.
- (78) 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.
- (79) Hoof, R. W.; Sander, C.; Vriend, G. Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures. *Proteins: Struct., Funct., Genet.* **1996**, *26*, 363–376.
- (80) 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.
- (81) Fiser, A.; Sali, A. ModLoop: Automated modeling of loops in protein structures. *Bioinformatics* **2003**, *19*, 2500–2501.
- (82) Fernandez-Fuentes, N.; Zhai, J.; Fiser, A. ArchPred: A template based loop structure prediction server. *Nucleic Acids Res.* **2006**, *34*, W173–W176.
- (83) Ko, J.; Lee, D.; Park, H.; Coutsiaris, E. A.; Lee, J.; Seok, C. The FALC-Loop web server for protein loop modeling. *Nucleic Acids Res.* **2011**, *39*, W210–W214.
- (84) Krivov, G. G.; Shapovalov, M. V.; Dunbrack, R. L., Jr Improved prediction of protein side-chain conformations with SCWRL4. *Proteins: Struct., Funct., Genet.* **2009**, *77*, 778–795.
- (85) Rotkiewicz, P.; Skolnick, J. Fast procedure for reconstruction of full-atom protein models from reduced representations. *J. Comput. Chem.* **2008**, *29*, 1460–1465.
- (86) Lombardi, L. E.; Marti, M. A.; Capece, L. CG2AA: backmapping protein coarse-grained structures. *Bioinformatics* **2016**, *32*, 1235–1237.
- (87) 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.
- (88) 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.
- (89) Bahar, I.; Lezon, T. R.; Yang, L. W.; Eyal, E. Global dynamics of proteins: bridging between structure and function. *Annu. Rev. Biophys.* **2010**, *39*, 23–42.
- (90) Yang, L. W.; Rader, A. J.; Liu, X.; Jursa, C. J.; Chen, S. C.; Karimi, H. A.; Bahar, I. oGNM: online computation of structural dynamics using the gaussian network model. *Nucleic Acids Res.* **2006**, *34*, W24–W31.
- (91) Li, H.; Chang, Y. Y.; Lee, J. Y.; Bahar, I.; Yang, L. W. DynOmics: dynamics of structural proteome and beyond. *Nucleic Acids Res.* **2017**, *45*, W374–W380.
- (92) Eyal, E.; Lum, G.; Bahar, I. The anisotropic network model web server at 2015 (ANM 2.0). *Bioinformatics* **2015**, *31*, 1487–1489.
- (93) 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.
- (94) Dehouck, Y.; Gilis, D.; Rooman, M. A new generation of statistical potentials for proteins. *Biophys. J.* **2006**, *90*, 4010–4017.
- (95) 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.
- (96) 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.
- (97) De Ruvo, M.; Giuliani, A.; Paci, P.; Santoni, D.; Di Paola, L. Shedding light on protein-ligand binding by graph theory: The topological nature of allostery. *Biophys. Chem.* **2012**, *165*–166, 21–29.
- (98) 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.
- (99) Cumbo, F.; Paci, P.; Santoni, D.; Di Paola, L.; Giuliani, A. GIANT: A Cytoscape plugin for modular networks. *PLoS One* **2014**, *9*, No. e105001.
- (100) Brinda, K. V.; Vishveshwara, S. A network representation of protein structures: Implications for protein stability. *Biophys. J.* **2005**, *89*, 4159–4170.
- (101) Vijayabaskar, M. S.; Vishveshwara, S. Interaction energy based protein structure networks. *Biophys. J.* **2010**, *99*, 3704–3715.
- (102) 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.

- (103) Czemerer, J.; Buse, K.; Verkhivker, G. M. Atomistic simulations and network-based modeling of the Hsp90-Cdc37 chaperone binding with Cdk4 client protein: A mechanism of chaperoning kinase clients by exploiting weak spots of intrinsically dynamic kinase domains. *PLoS One* **2017**, *12*, No. e0190267.
- (104) 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.
- (105) Stetz, G.; Verkhivker, G. M. Computational analysis of residue interaction networks and coevolutionary relationships in the Hsp70 chaperones: A community-hopping model of allosteric regulation and communication. *PLoS Comput. Biol.* **2017**, *13*, No. e1005299.
- (106) Floyd, R. W. Algorithm 97: Shortest path. *Commun. ACM* **1962**, *5*, 345.
- (107) Hagberg, A. A.; Schult, D. A.; Swart, P. J. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference (SciPy2008)*, Pasadena; Varoquaux, G., Vaught, T., Millman, J., Eds.; 2008; pp 11–15.
- (108) del Sol, A.; Fujihashi, H.; Amoros, D.; Nussinov, R. Residues crucial for maintaining short paths in network communication mediate signaling in proteins. *Mol. Syst. Biol.* **2006**, *2*, 2006.0019.
- (109) Brysbaert, G.; Mauri, T.; Lensink, M. F. Comparing protein structures with RINspectator automation in Cytoscape. *F1000Research* **2018**, *7*, 563.
- (110) Girvan, M.; Newman, M. E. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 7821–7826.
- (111) Newman, M. E. Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **2006**, *74*, 036104.
- (112) Jensen, P.; Morini, M.; Karsai, M.; Venturini, T.; Vespignani, A.; Jacomy, M.; Cointet, J.-P.; Mercklé, P.; Fleury, E. Detecting global bridges in networks. *J. Complex Netw.* **2016**, *4*, 319–329.
- (113) Rao, C. R. Diversity and dissimilarity coefficients: A unified approach. *Theor. Popul. Biol.* **1982**, *21*, 24–43.
- (114) Stirling, A. A general framework for analyzing diversity in science, technology and society. *J. R. Soc., Interface* **2007**, *4*, 707–719.
- (115) 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.
- (116) 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.
- (117) 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.
- (118) 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.
- (119) Szalay-Beko, 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.
- (120) 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.
- (121) Shah, M.; Ahmad, B.; Choi, S.; Woo, H. G. Sequence variation of SARS-CoV-2 spike protein may facilitate stronger interaction with ACE2 promoting high infectivity. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 3402.
- (122) Armijos-Jaramillo, V.; Yeager, J.; Muslin, C.; Perez-Castillo, Y. SARS-CoV-2, an evolutionary perspective of interaction with human ACE2 reveals an undiscovered amino acids necessary for complex stability. *Evol. Appl.* **2020**, *13*, 2168–2178.
- (123) Zhou, D.; Tian, X.; Qi, R.; Peng, C.; Zhang, W. Identification of 22 N-glycosites on spike glycoprotein of SARS-CoV-2 and accessible surface glycopeptide motifs: implications for vaccination and antibody therapeutics. *Glycobiology* **2020**, DOI: 10.1093/glycob/cwaa052.
- (124) Shajahan, A.; Supekar, N. T.; Gleinich, A. S.; Azadi, P. Deducing the N- and O-glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology* **2020**, *30*, 981–988.
- (125) Uslupehlivan, M.; Sener, E. Glycoinformatics approach for identifying target positions to inhibit initial binding of SARS-CoV-2 S1 protein to the host cell. *bioRxiv* **2020**; DOI: 10.1101/2020.03.25.007898.
- (126) Kyte, J.; Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **1982**, *157*, 105–132.
- (127) del Sol, A.; O'Meara, P. Small-world network approach to identify key residues in protein-protein interaction. *Proteins: Struct., Funct., Genet.* **2005**, *58*, 672–82.
- (128) Bode, C.; Kovacs, I. A.; Szalay, M. S.; Palotai, R.; Korcsmaros, T.; Csermely, P. Network analysis of protein dynamics. *FEBS Lett.* **2007**, *581*, 2776–2782.
- (129) Liang, Z.; Verkhivker, G. M.; Hu, G. Integration of network models and evolutionary analysis into high-throughput modeling of protein dynamics and allosteric regulation: theory, tools and applications. *Briefings Bioinf.* **2020**, *21*, 815–835.
- (130) Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Foley, B.; Giorgi, E. E.; Bhattacharya, T.; Parker, M. D., et al. Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2. *bioRxiv* **2020**; DOI: 10.1101/2020.04.29.069054.
- (131) 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*, 1284–1294.
- (132) Laha, S.; Chakraborty, J.; Das, S.; Manna, S. K.; Biswas, S.; Chatterjee, R. Characterizations of SARS-CoV-2 mutational profile, spike protein stability and viral transmission. *Infect., Genet. Evol.* **2020**, *85*, 104445.
- (133) Jackson, C. B.; Zhang, L.; Farzan, M.; Choe, H. Functional importance of the D614G mutation in the SARS-CoV-2 spike protein. *Biochem. Biophys. Res. Commun.* **2020**, DOI: 10.1016/j.bbrc.2020.11.026.
- (134) Mansbach, R. A.; Chakraborty, S.; Nguyen, K.; Montefiori, D.; Korber, B.; Gnanakaran, S. The SARS-CoV-2 spike variant D614G favors an open conformational state. *bioRxiv* **2020**; DOI: 10.1101/2020.07.26.219741.
- (135) Xu, C.; Wang, Y.; Liu, C.; Zhang, C.; Han, W.; Hong, X.; Wang, Y.; Hong, Q.; Wang, S.; Zhao, Q., et al. Conformational dynamics of SARS-CoV-2 trimeric spike glycoprotein in complex with receptor ACE2 revealed by cryo-EM. *bioRxiv* **2020**; DOI: 10.1101/2020.06.30.177097.
- (136) Gobeil, S.; Janowska, K.; McDowell, S.; Mansouri, K.; Parks, R.; Manne, K.; Stalls, V.; Kopp, M.; Henderson, R.; Edwards, R. J., et al. D614G mutation alters SARS-CoV-2 spike conformational dynamics and protease cleavage susceptibility at the S1/S2 junction. *bioRxiv* **2020**; DOI: 10.1101/2020.10.11.335299.